ISSN 1413-3555

BJPT Brazilian Journal of Physical Therapy

2016 Mar/Apr; 20(2)

abrapg-ft
ASSOCIAÇÃO BRASILEIRA DE PESQUISA
E PÓS-GRADUAÇÃO EM FISIOTERAPIA

BJPT Brazilian Journal of Physical Therapy

FINANCIAL SUPPORT

COFFITO
Conselho Federal de Fisioterapia e Terapia Ocupacional

CREFITO-3
Conselho Regional de Fisioterapia e Terapia Ocupacional da 3ª Região

FAPEMIG

FAPESP

UFMG
UNIVERSIDADE FEDERAL DE MINAS GERAIS

ufscar

UNICID
Universidade Cidade de S. Paulo

CNPq
Conselho Nacional de Desenvolvimento Científico e Tecnológico

CAPES
Ministério da Educação
Ministério da Ciência, Tecnologia e Inovação
GOVERNO FEDERAL
BRASIL
PAÍS RICO É PAÍS SEM POBREZA

ISSN 1413-3555

BJPT Brazilian Journal of Physical Therapy

Journal affiliated with the
Associação Brasileira de Editores Científicos
ABEC Brasil

COPE COMMITTEE ON PUBLICATION ETHICS

abrapg ft
ASSOCIAÇÃO BRASILEIRA DE PESQUISA E PÓS-GRADUAÇÃO EM FISIOTERAPIA

Braz J Phys Ther. 2016 Mar-Apr; 20(2):105-190

The Brazilian Journal of Physical Therapy is published by the Associação Brasileira de Pesquisa e Pós-Graduação em Fisioterapia – ABRAPG-Ft (Brazilian Association for Research and Graduate Studies in Physical Therapy). Published since 1996, the Brazilian Journal of Physical Therapy adopts a peer review process. Each article is only published after it is accepted by the reviewers, who are maintained anonymous during the process.

The editors accept no responsibility for damage to people or property, which may have been caused by the use of ideas, techniques or procedures described in the material published by this journal.

The submission of articles presupposes that these articles, with the exception of extended summaries, have not been previously published elsewhere, nor submitted to any other publication.

The abbreviated title of the journal is **Braz. J. Phys. Ther.** and this must be used in references, footnotes and bibliographic legends.

The Brazilian Journal of Physical Therapy is freely accessible at the homepage on the web: http://www.scielo.br/rbfis.

Mission

To publish original research articles on topics related to the areas of physical therapy and rehabilitation sciences, including clinical, basic or applied studies on the assessment, prevention, and treatment of movement disorders.

Indexed in

Cataloguing Card

Brazilian Journal of Physical Therapy / Associação Brasileira de Pesquisa e Pós-Graduação em Fisioterapia. v. 1, n. 1 (1996). – São Carlos: Editora Cubo, 1996-

v. 20, n. 2 (March/April 2016).
Bimonthly
Continued Revista Brasileira de Fisioterapia
ISSN 1413-3555

1. Physical Therapy. 2. Studies. I. Associação Brasileira de Pesquisa e Pós-Graduação em Fisioterapia.

Librarian: Dormélia Pereira Cazella (CRB 8/4334)

Contact Address
Brazilian Journal of Physical Therapy
Rod. Washington Luís, Km 235,
Caixa Postal 676, CEP 13565-905
São Carlos, SP - Brasil
+55(16) 3351-8755
contato@rbf-bjpt.org.br
www.rbf-bjpt.org.br

Technical and Administrative Support
Ana Paula de Luca
Leonor A. Saidel Aizza

Desktop Publishing and Editorial Consulting
editora cubo
soluções para o universo acadêmico

Printed in acid free paper.

BJPT Brazilian Journal of Physical Therapy

Review Article

Systematic Review

Original Articles

review article

BJPT Brazilian Journal of Physical Therapy

The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument

Lidwine B. Mokkink[1], Cecilia A. C. Prinsen[1], Lex M. Bouter[1], Henrica C. W. de Vet[1], Caroline B. Terwee[1]

ABSTRACT | Background: COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) is an initiative of an international multidisciplinary team of researchers who aim to improve the selection of outcome measurement instruments both in research and in clinical practice by developing tools for selecting the most appropriate available instrument. **Method:** In this paper these tools are described, i.e. the COSMIN taxonomy and definition of measurement properties; the COSMIN checklist to evaluate the methodological quality of studies on measurement properties; a search filter for finding studies on measurement properties; a protocol for systematic reviews of outcome measurement instruments; a database of systematic reviews of outcome measurement instruments; and a guideline for selecting outcome measurement instruments for Core Outcome Sets in clinical trials. Currently, we are updating the COSMIN checklist, particularly the standards for content validity studies. Also new standards for studies using Item Response Theory methods will be developed. Additionally, in the future we want to develop standards for studies on the quality of non-patient reported outcome measures, such as clinician-reported outcomes and performance-based outcomes. **Conclusions:** In summary, we plea for more standardization in the use of outcome measurement instruments, for conducting high quality systematic reviews on measurement instruments in which the best available outcome measurement instrument is recommended, and for stopping the use of poor outcome measurement instruments.

Keywords: COSMIN; measurement properties; outcome measures; systematic reviews of instruments; outcome selection.

BULLET POINTS

- COSMIN aims to improve instrument selection in research and clinical practice.
- Description of COSMIN tools for selecting most appropriate instrument.
- Call for standardization in instrument use.
- Call for conducting high quality systematic reviews on instruments.
- Call for stopping the use of poor measurement instruments.

HOW TO CITE THIS ARTICLE

Mokkink LB, Prinsen CAC, Bouter LM, de Vet HCW, Terwee CB. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. Braz J Phys Ther. 2016 Mar-Apr; 20(2):105-113. http://dx.doi.org/10.1590/bjpt-rbf.2014.0143

Introduction

COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) is an initiative of an international multidisciplinary team of researchers with a background in epidemiology, psychometrics, qualitative research, and health care, who have expertise in the development and evaluation of outcome measurement instruments[1]. The COSMIN initiative aims to improve the selection of outcome measurement instruments both in research and in clinical practice by developing tools for selecting the most appropriate instrument. The COSMIN Steering Committee (see Appendix 1), founded in 2005, was inspired by a lack of clarity in the literature about terminology and definitions of measurement properties. Moreover, there exists an impressive amount of outcome measurement instruments and there are even many instruments measuring the same construct, developed for the same patient population, and still new ones are being developed. So researchers and clinicians

[1]Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands
Received: Nov. 13, 2015 Revised: Nov. 16, 2015 Accepted: Nov. 24, 2015

have to choose the most suitable instrument for their application.

The process of selecting outcome measures for specific purposes is complex. Choices involve conceptual considerations, such as defining the construct and population; practical aspects, such as burden for patients and raters, and costs; and quality aspects assessed by nine different measurement properties clustered in the domains reliability, validity and responsiveness[2]. Selecting unsuitable or poor quality outcome measurement instruments may introduce bias in the conclusions of studies. This may lead to a waste of resources and be unethical because participating patients contribute little or nothing to the body of knowledge but still suffer from the burdens and risks of the study[3].

An additional problem is that in systematic reviews of clinical trials the results reported cannot be compared and statistically pooled when different instruments are used to measure the same construct of interest in each study. Moreover, in clinical trials evaluating the benefits and harms of health care interventions, often a great variety of outcomes are reported. This makes it even more difficult to compare and combine results. This hampers the usefulness of clinical trial evidence to inform clinicians, at the cost of the best possible care for patients. Standardization in outcomes and outcome measurement instruments in specific areas of research is therefore highly warranted.

The COSMIN initiative wants to improve the selection of outcome measurement instruments by developing methodological guidelines based on consensus reached in a broad international panel of experts. The initial focus was on patient-reported outcome measures (PROMs). Therefore, the focus of this paper is only on PROMs.

First, some conceptual considerations concerning the selection of an outcome measurement instrument are explained. Next, the tools yielded by the COSMIN initiative will be described. Finally, we describe our future plans for research.

Conceptual considerations when selecting outcome measurement instruments

It is important to understand the distinction between an outcome and an outcome measurement instrument. An outcome refers to the construct of interest. Since we talk about patient-reported outcomes, the outcome is often a phenomenon that cannot be observed directly, for example fatigue or health-related quality of life. The outcome chosen defines *what* is being measured. An outcome measurement instrument refers to *how* the outcome is being measured. It refers to the specific outcome measurement instrument. For example, the Neurological Fatigue Index for multiple sclerosis (NFI-MS)[4] or the Skindex-29[5] to measure quality of life in dermatology.

When selecting an outcome measurement instrument for research or clinical practice, first the outcome to be measured should be clearly defined. That is, one should define what to measure. For example, when measuring a broad construct such as health-related quality of life, it should be clarified which subdomains are relevant for the target population in the specific context of interest. Sometimes several definitions exist for an outcome. There are, for instance, multiple definitions for the construct 'disability'. The World Health Organization (WHO) defines 'disability' as a broad concept: 'problems an individual may experience in functioning, namely impairments, activity limitations and participation restrictions'[6]. Nagi[7] defined disability more narrowly as 'a pattern of behaviour that evolves in situations of long-term or continued impairment that are associated with functioning limitations' (previously called 'handicap' in the International Classification of Functioning of the WHO[8]). Without explicitly defining or describing the intended outcome, people may have different ideas about it and interpret it differently.

Next, one has to choose a specific instrument. Often, for the same outcome multiple measurement instruments are available. To select the best available outcome measurement instrument the COSMIN initiative has yielded several tools.

Standardization of the selection of outcomes and outcome measurement instruments in specific areas of research will improve consistencies in reporting and decrease difficulties in comparing and combining the findings in systematic reviews and meta-analyses. This can be obtained by the development of Core Outcome Sets (COS). A COS is an agreed standardized set of outcomes that should be measured and reported, as a minimum, in all clinical trials in a specific disease or trial population (i.e. *what* to measure)[9]. Once the COS is defined, it is then important to achieve consensus on which outcome measurement instruments should be selected to measure the core outcomes, referring to Core Outcome Measurement Instruments (i.e *how* to measure)[10]. The existence or use of a core outcome set does not imply that outcomes in a particular trial should

be restricted to those in the relevant core outcome set. Rather, there is an expectation that the core outcomes will be collected and reported, making it easier for the results of trials to be compared, contrasted and combined as appropriate; while researchers continue to explore other outcomes as well[11].

COSMIN tools

The COSMIN initiative has developed the following tools to help researchers and clinicians choosing the most appropriate outcome measurement instrument:

1. COSMIN taxonomy and definitions of measurement properties;
2. COSMIN checklist to evaluate the methodological quality of studies on measurement properties;
3. Search filter for finding studies on measurement properties;
4. Protocol for systematic reviews of outcome measurement instruments;
5. Database of systematic reviews of outcome measurement instruments;
6. Guideline for selecting outcome measurement instruments for outcomes included in a Core Outcome Set.

We performed an international Delphi study aiming to develop consensus-based standards for assessing the methodological quality of studies on measurement properties[1,2,12-14]. Results from this study were the COSMIN taxonomy and definitions, and the COSMIN checklist.

COSMIN taxonomy and definitions

We first developed a taxonomy and reached consensus on definitions of the measurement properties (see Table 1)[2]. Nine measurement properties clustered within three domains, i.e. reliability, validity and responsiveness, were considered relevant in the evaluation of outcome measurement instruments (Figure 1).

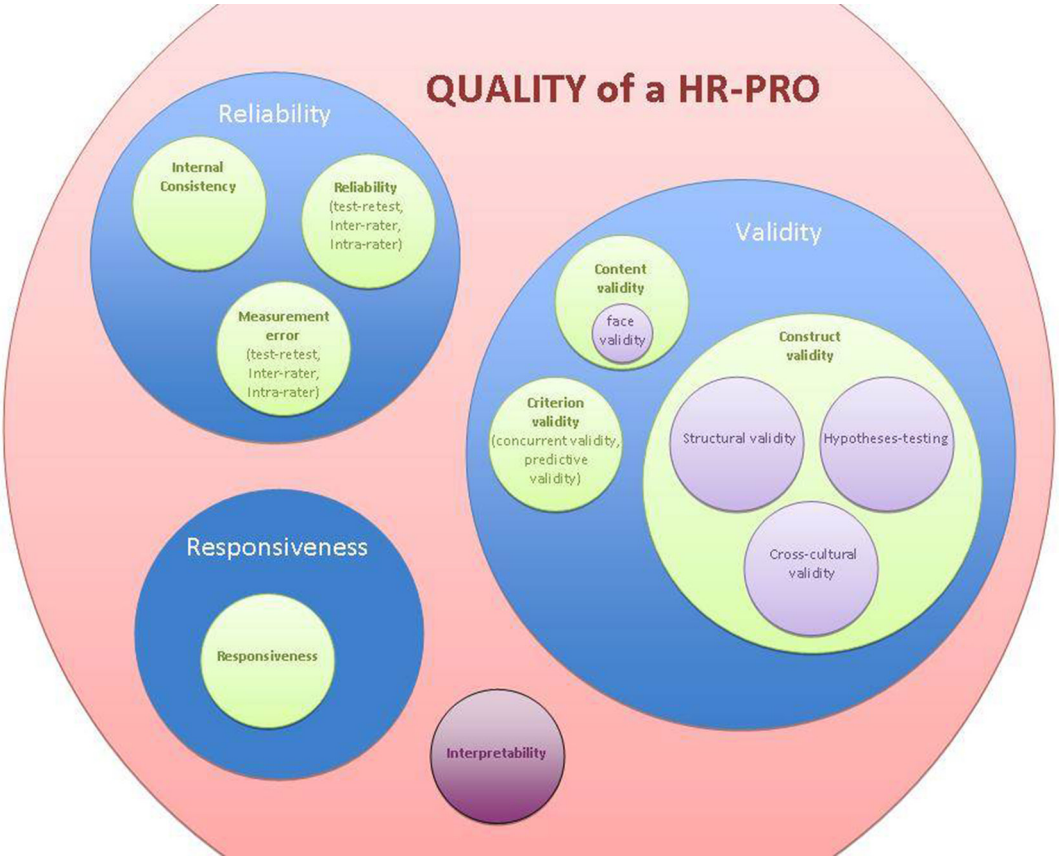

Table 1. Definitions of domains, measurement properties, and aspects of measurement properties.

Term			Definition
Domain	**Measurement property**	**Aspect of a measurement property**	
Reliability			The degree to which the measurement is free from measurement error
Reliability (extended definition)			The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: e.g. using different sets of items from the same health related-patient reported outcomes (HR-PRO) (internal consistency); over time (test-retest); by different persons on the same occasion (inter-rater); or by the same persons (i.e. raters or responders) on different occasions (intra-rater)
	Internal consistency		The degree of the interrelatedness among the items
	Reliability		The proportion of the total variance in the measurements which is due to 'true'† differences between patients
	Measurement error		The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured
Validity			The degree to which an HR-PRO instrument measures the construct(s) it purports to measure
	Content validity		The degree to which the content of an HR-PRO instrument is an adequate reflection of the construct to be measured
		Face validity	The degree to which (the items of) an HR-PRO instrument indeed looks as though they are an adequate reflection of the construct to be measured
	Construct validity		The degree to which the scores of an HR-PRO instrument are consistent with hypotheses *(for instance with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups)* based on the assumption that the HR-PRO instrument validly measures the construct to be measured
		Structural validity	The degree to which the scores of an HR-PRO instrument are an adequate reflection of the dimensionality of the construct to be measured
		Hypotheses testing	Idem construct validity
		Cross-cultural validity	The degree to which the performance of the items on a translated or culturally adapted HR-PRO instrument are an adequate reflection of the performance of the items of the original version of the HR-PRO instrument
	Criterion validity		The degree to which the scores of an HR-PRO instrument are an adequate reflection of a 'gold standard'
Responsiveness			The ability of an HR-PRO instrument to detect change over time in the construct to be measured
	Responsiveness		Idem responsiveness
Interpretability*			The degree to which one can assign qualitative meaning - that is, clinical or commonly understood connotations – to an instrument's quantitative scores or change in scores

†The word 'true' must be seen in the context of the CTT, which states that any observation is composed of two components – a true score and error associated with the observation. 'True' is the average score that would be obtained if the scale were given an infinite number of times. It refers only to the consistency of the score, and not to its accuracy[15]. *Interpretability is not considered a measurement property, but an important characteristic of a measurement instrument.

COSMIN checklist

We developed a critical appraisal tool, i.e. the COSMIN checklist[12-14], containing standards for evaluating the methodological quality of studies on the measurement properties of outcome measurement instruments. The COSMIN checklist and a supplementary manual can be optained from the COSMIN website[16]. For each measurement property a box with standards was developed. These standards describe design requirements and preferred statistical methods. For example, in a high quality study of internal consistency, first a check for the unidimensionality of the (sub)scale should be done (Box A item 5 of the COSMIN checklist)[12]. Subsequently the internal consistency statistic should be calculated for the items of this unidimensional (sub) scale (Box A item 7)[12]. Other standards concern, for instance, using an appropriate time interval between test and retest administration when investigating test-retest reliability and measurement error (Box B and C item 8)[12], or formulating a priori hypotheses for hypotheses testing (a form of construct validity) (Box F item 4)[12].

When examining the interrater reliability and agreement of the items of the COSMIN checklist, we found that the reliability of the individual items was low (i.e. only 6% of the items had a Kappa statistic above 0.75), but that the agreement between raters was appropriate for 80% of the items[17]. When using the COSMIN checklist in a systematic review, we recommend getting some prior on-the-job training and experience, completing it by two independent raters, and reaching consensus about the ratings[17]. To use the COSMIN checklist in a systematic review of measurement instruments, we developed a four-point rating system for scoring the items of the COSMIN checklist[14]. With this version it is possible to calculate overall methodological quality scores per study on a measurement property. This is useful and enlightening in systematic reviews, as it allows to present conclusions on the quality of the instruments under study accompanied by various levels of evidence[14].

Search filter for finding studies on measurement properties

To facilitate the selection of outcome measurement instruments to be included in a systematic review of measurement instruments, a search filter was developed and validated in cooperation with clinical librarians for finding studies on measurement properties in PubMed[18]. In such a review, the filter can be combined with search terms for the outcome and the population of interest. The filter for finding studies on measurement properties showed to have a sensitivity of 97.4% and a positive predictive value of 4.4%. We translated this filter for EMBASE and CINAHL, and all filters are available from the COSMIN website[16].

Protocol for systematic reviews of outcome measurement instruments

Systematic reviews of outcome measurement instruments are important for the evidence-based selection of instruments. In such a review, the measurement properties of all outcome measurement instruments for a specific construct in a specific population are described and compared according to predefined criteria, and a conclusion is drawn about the most appropriate instrument.

We developed a protocol for performing systematic reviews of measurement instruments, including a 10-step procedure (available from the COSMIN website). In this protocol we describe how the COSMIN search filter[18] can be used to identify all relevant outcome measurement instruments, as well as how the COSMIN checklist[12] can be used to assess the quality of the included studies. In addition to the search filter for studies on measurement properties, and if the review concerns PROMs, a PROM filter developed by the University of Oxford can be used (available from the COSMIN website).

In addition, we describe the method of a best evidence synthesis in which the number of studies, their quality and (consistency of) results can be combined to determine the strength of the evidence for each measurement property. For example, strong evidence for a positive reliability is obtained when consistent positive results (ICCs or Kappas >0.70) are found in at least two studies of good quality or one study of excellent quality. The procedure is similar to the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach[19] that is used in reviews of clinical trials. Previously developed cut-off values (such as ICC or Kappas >0.70) are used to determine whether an outcome measurement instrument has good measurement properties[20].

In 2009 we concluded, based on a review of systematic reviews of measurement instruments, that the quality of these reviews should and could be improved[21]. Recently, we updated this review, and concluded that the quality of published systematic reviews of measurement instruments has improved[22]. However,

there is still room for improvement with regards to the search strategy, and especially the quality assessment of the included studies and instruments as well as the data synthesis. Therefore, we are currently updating the protocol for performing systematic reviews of measurement instruments, aiming to publish it as a peer-reviewed guideline for systematic reviews of outcome measurement instruments (manuscript in preparation). In this way, we aim to contribute to the improvement of systematic reviews of measurement instruments.

Database for systematic review of outcome measurement instruments

The COSMIN initiative maintains an overview of published systematic reviews of outcome measurement instruments. This overview is presented in a searchable database available on the COSMIN website[16]. Currently, it contains 569 systematic reviews and we aim to update this overview yearly. The COSMIN database provides a good starting point to search for and select outcome measurement instruments.

Guideline for selecting outcome measurement instruments for outcomes included in a Core Outcome Set

COSMIN collaborated with the COMET (Core Outcome Measures in Effectiveness Trials) initiative to develop a guideline for the selection of outcome measurement instruments for outcomes included in a COS[10]. We reached consensus among a large group of experts on four main steps in the selection of outcome measurement instruments for COS: Step 1) conceptual considerations; Step 2) finding existing outcome measurement instruments; Step 3) quality assessment of outcome measurement instruments; and Step 4) generic recommendations on the selection of outcome measurement instruments for outcomes included in a COS. The resulting consensus-based guideline can be used by COS developers in defining *how* to measure core outcomes (submitted publication by Prinsen CA, et al. How to select outcome measurement instruments for outcomes included in a 'Core Outcome Set' – a practical guideline).

Ongoing and future studies

At the moment, we work on updating the COSMIN checklist. Over the past years, users of the COSMIN checklist have identified gaps in the available standards. Recent regulatory guidelines on outcome measurement instruments development and evaluation call for an extension of the COSMIN checklist with respect to its standards for the quality of studies on content validity within the specific context of interest. Therefore, a Delphi study is underway which aims to reach consensus on new COSMIN standards and criteria for evaluating the content validity (including face validity) of PROMs. In these new standards, the quality of the development process of PROMs will be taken into account, and criteria for what constitutes good content validity will be developed.

In addition, a shift has taken place in recent years from the use of traditional statistical methods (i.e. Classical Test Theory (CTT)) to the recommended use of newer statistical methods (e.g. Item Response Theory (IRT)[23] and Rasch Measurement Theory[24]) analyses for developing and evaluating outcome measurement instruments. This requires an extension of the COSMIN standards for studies using IRT and Rasch methods. Clear methodological advantages of using IRT or other modern test theory methods over or in addition to CTT have been described[25]. Well-developed IRT-based instruments, have probably better measurement properties than CTT-based instruments[26,27]. In addition, IRT allows for Computer Adaptive Testing (CAT), a method of questionnaire administration in which a computer algorithm iteratively selects questions based on previous answers. Questionnaires that are completed by CAT dramatically decreases the burden for patients to complete questionnaires and improving precision[28-31]. Examples of IRT-based instruments are the Patient Reported Outcomes Measurement Information System (PROMIS) instruments, which are available as CAT instruments as well as static short forms[32]. A next step to be addressed is to achieve consensus among an international group of experts on standards for the methodological quality of studies using IRT and Rasch methods for evaluating measurement properties and to operationalize these standards into a user-friendly and easily applicable checklist to be used e.g. in systematic reviews of outcome measurement instruments.

The COSMIN standards were originally developed for evaluating the quality of studies on the measurement properties of PROMs. Although the COSMIN standards have also been used in systematic reviews of other types of outcome measurement instruments, adaptations are required to use the COSMIN standards for evaluating the quality of studies on the measurement properties of other patient-centered outcome measurement instruments, such as clinician-reported outcome measure

(e.g. a goniometer to measure range of motion), or a performance based test (e.g. a six minute walk test to measure walking speed). It is our ambition to develop new standards specific for other types of instruments.

Finally, we want to develop reporting guidelines for studies on measurement properties, and for systematic reviews on measurement properties.

Need for high quality systematic reviews of outcome measurement instruments and Core Outcome Set development

By the development of the COSMIN tools described above and by generating awareness for the importance of selecting high quality instruments, COSMIN aims to accomplish that researchers and clinicians make their choices on outcomes and outcome measurement instruments more informed. We plea for more standardization in the use of outcomes and outcome measurement instruments. We support the aim of the COMET initiative to stimulate the development of COS. The use of COS will lead to more standardization in outcome reporting in specific areas of research, making it easier for the results of trials to be compared and combined as appropriate. COSMIN strongly encourages researchers to perform high quality systematic reviews of outcome measurement instruments. More high quality systematic reviews of outcome measurement instruments are needed to make an informed choice for the best instrument for a specific purpose and for stopping the use of poor outcome measurement instruments.

References

1. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. BMC Med Res Methodol. 2006;6(1):2. http://dx.doi.org/10.1186/1471-2288-6-2. PMid:16433905.
2. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737-45. http://dx.doi.org/10.1016/j.jclinepi.2010.02.006. PMid:20494804.
3. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. Lancet. 2014;383(9912):166-75. http://dx.doi.org/10.1016/S0140-6736(13)62227-8. PMid:24411645.
4. Mills RJ, Young CA, Pallant JF, Tennant A. Development of a patient reported outcome scale for fatigue in multiple sclerosis: The Neurological Fatigue Index (NFI-MS). Health Qual Life Outcomes. 2010;8(1):22. http://dx.doi.org/10.1186/1477-7525-8-22. PMid:20152031.
5. Chren MM, Lasek RJ, Quinn LM, Mostow EN, Zyzanski SJ. Skindex, a quality-of-life measure for patients with skin disease: reliability, validity, and responsiveness. J Invest Dermatol. 1996;107(5):707-13. http://dx.doi.org/10.1111/1523-1747.ep12365600. PMid:8875954.
6. World Health Organization – WHO. ICF: International Classification of Functioning, disability and health. Geneva: WHO; 2001.
7. Nagi SZ. Some conceptual issues in disability and rehabilitation. In: Sussman MB, editor. Sociology and rehabilitation. Columbus: Ohio State University Press; 1965. p. 100-13.
8. World Health Organization – WHO. International classification of International Classification of Impairments, Disabilities, and Handicaps.Geneva: WHO; 1980.
9. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. Trials. 2012;13(1):132. http://dx.doi.org/10.1186/1745-6215-13-132. PMid:22867278.
10. Prinsen CA, Vohra S, Rose MR, King-Jones S, Ishaque S, Bhaloo Z, et al. Core Outcome Measures in Effectiveness Trials (COMET) initiative: protocol for an international Delphi study to achieve consensus on how to select outcome measurement instruments for outcomes included in a 'core outcome set'. Trials. 2014;15(1):247. http://dx.doi.org/10.1186/1745-6215-15-247. PMid:24962012.
11. Boers M, Kirwan JR, Tugwell P, Beaton D, Bingham CO 3rd, Conaghan PG, et al. The OMERACT handbook. OMERACT; 2015.
12. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010;19(4):539-49. http://dx.doi.org/10.1007/s11136-010-9606-8. PMid:20169472.
13. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Med Res Methodol. 2010;10(1):22. http://dx.doi.org/10.1186/1471-2288-10-22. PMid:20298572.
14. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res. 2012;21(4):651-7. http://dx.doi.org/10.1007/s11136-011-9960-1. PMid:21732199.
15. Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use. Oxford: University Press; 2008.
16. COnsensus-based Standards for the selection of health Measurement INstruments – COSMIN [Internet]. Amsterdam: COSMIN; 2015 [cited 2015 Sept. 28]. Available from: http://www.cosmin.nl
17. Mokkink LB, Terwee CB, Gibbons E, Stratford PW, Alonso J, Patrick DL, et al. Inter-rater agreement and reliability

of the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) checklist. BMC Med Res Methodol. 2010;10(1):82. http://dx.doi.org/10.1186/1471-2288-10-82. PMid:20860789.

18. Terwee CB, Jansma EP, Riphagen II, Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res. 2009;18(8):1115-23. http://dx.doi.org/10.1007/s11136-009-9528-5. PMid:19711195.
19. GRADE working group [Internet]. 2005 [cited 2015 Sept. 28]. Available from: http://www.gradeworkinggroup.org
20. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34-42. http://dx.doi.org/10.1016/j.jclinepi.2006.03.012. PMid:17161752.
21. Mokkink LB, Terwee CB, Stratford PW, Alonso J, Patrick DL, Riphagen I, et al. Evaluation of the methodological quality of systematic reviews of health status measurement instruments. Qual Life Res. 2009;18(3):313-33. http://dx.doi.org/10.1007/s11136-009-9451-9. PMid:19238586.
22. Terwee CB, Prinsen CA, Garotti MGR, Suman A, Vet HC, Mokkink LB. The quality of systematic reviews of health-related outcome measurement instruments. Qual Life Res. 2015. http://dx.doi.org/10.1007/s11136-015-1122-4. PMid:26346986.
23. Lord FM. Applications of item response theory to practical testing problems. Mahwah: Erlbaum; 1980.
24. Andrich D. Rasch models for measurement. Beverly Hills: Sage Publications; 1988.
25. Embretson SE, Reise SP. Item response theory for psychologists. Mahwah: Lawrence Erlbaum Associates; 2000.
26. Fries JF, Bruce B, Bjorner J, Rose M. More relevant, precise, and efficient items for assessment of physical function and disability: moving beyond the classic instruments. Ann Rheum Dis. 2006;65(Suppl 3):iii16-21. http://dx.doi.org/10.1136/ard.2006.059279. PMid:17038464.
27. Lai JS, Cella D, Choi S, Junghaenel DU, Christodoulou C, Gershon R, et al. How item banks and their application can influence measurement practice in rehabilitation medicine: a PROMIS fatigue item bank example. Arch Phys Med Rehabil. 2011;92(10 Suppl):S20-7. http://dx.doi.org/10.1016/j.apmr.2010.08.033. PMid:21958919.
28. Fries JF, Krishnan E, Rose M, Lingala B, Bruce B. Improved responsiveness and reduced sample size requirements of PROMIS physical function scales with item response theory. Arthritis Res Ther. 2011;13(5):R147. http://dx.doi.org/10.1186/ar3461. PMid:21914216.
29. Thissen D, Reeve BB, Bjorner JB, Chang CH. Methodological issues for building item banks and computerized adaptive scales. Qual Life Res. 2007;16(S1 Suppl 1):109-19. http://dx.doi.org/10.1007/s11136-007-9169-5. PMid:17294284.
30. Haley SM, Ni P, Hambleton RK, Slavin MD, Jette AM. Computer adaptive testing improved accuracy and precision of scores over random item selection in a physical functioning item bank. J Clin Epidemiol. 2006;59(11):1174-82. http://dx.doi.org/10.1016/j.jclinepi.2006.02.010. PMid:17027428.
31. Bjorner JB, Chang CH, Thissen D, Reeve BB. Developing tailored instruments: item banking and computerized adaptive assessment. Qual Life Res. 2007;16(S1 Suppl 1):95-108. http://dx.doi.org/10.1007/s11136-007-9168-6. PMid:17530450.
32. PROMIS – Dynamic tools to measure health outcomes from the patient perspective [Internet]. Chapel Hill: PROMIS; 2015 [cited 2015 Sept. 28]. Available from: http://www.nihpromis.org

Correspondence
Lidwine B. Mokkink
VU University Medical Center
Institute for Health and Care Research
Department of Epidemiology and Biostatistics and EMGO
De Boelelaan 1089a, 1081 HV
Amsterdam, the Netherlands
e-mail: w.mokkink@vumc.nl

Appendix 1. COSMIN steering committee members.

Original COSMIN steering committee

The original COSMIN steering committee was responsible for the initiation and organisation of the Delphi study in which the COSMIN taxonomy and COSMIN checklist was developed. Members of the original steering committee were:

Lidwine B (Wieneke) Mokkink, Assistant professor in Clinimetrics, Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands. **Caroline B Terwee**, Associate professor in Measurement, Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands. **Jordi Alonso**, Head of the Health Services Research Group. Epidemiology and Public Health Program. Institut Municipal d'Investigació Mèdica (IMIM-Hospital del Mar), Barcelona, Spain. **Donald L Patrick**, Professor of Health Services and Director of the Social and Behavioral Sciences Program at the University of Washington, Director of the Seattle Quality of Life Group, Seattle, US. **Lex M Bouter**, Professor of Methodology and Integrity, VU University and Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands. **Henrica CW de Vet**, Professor of Clinimetrics, Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands. **Dirk L Knol**, Statistician, Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands. **Paul W Stratford**, Professor School of Rehabilitation Science, Faculty of Health Sciences, and Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada.

Current COSMIN steering committee

The current steering committee is responsible for updating the COSMIN checklist. Members of the current steering committee are:

Lidwine B (Wieneke) Mokkink, Assistant professor in Clinimetrics, Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands. **Cecilia AC (Sanna) Prinsen**, Assistant professor in Clinimetrics, Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands. **Jordi Alonso**, Head of the Health Services Research Group. Epidemiology and Public Health Program. Institut Municipal d'Investigació Mèdica (IMIM-Hospital del Mar), Barcelona, Spain. **Donald L Patrick**, Professor of Health Services and Director of the Social and Behavioral Sciences Program at the University of Washington, Director of the Seattle Quality of Life Group, Seattle, US. **Lex M Bouter**, Professor of Methodology and Integrity, VU University and Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands. **Henrica CW de Vet**, Professor of Clinimetrics, Department of Epidemiology and Biostatistics, Emgo Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands. **Marjan Westerman**, Assistan professor, expert in qualitative research, VU University, Amsterdam, the Netherlands. **Caroline B Terwee**, Associate professor in Measurement, Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands. Dirk Knol and Paul Stratford have retired from the COSMIN group in 2015.

systematic review

BJPT Brazilian Journal of Physical Therapy

Assessment of the measurement properties of the post-stroke motor function instruments available in Brazil: a systematic review

Elaine Lima[1], Luci F. Teixeira-Salmela[2], Luan Simões[1], Ana C. C. Guerra[1], Andrea Lemos[1]

ABSTRACT | Background: While there are several instruments in Brazil that measure motor function in patients after stroke, it is unknown whether the measurement properties of these instruments are appropriate. **Objective**: To identify the motor function instruments available in Brazil for patients after stroke. To assess the methodological quality of the studies and the results related to the measurement properties of these instruments. **Method**: Two independent reviewers conducted searches on PubMed, LILACS, CINAHL, Web of Science, and Scopus. Studies that aimed to cross-culturally adapt an existing instrument or create a Brazilian instrument and test at least one measurement property related to motor function in patients after stroke were included. The methodological quality of these studies was checked by the COSMIN checklist with 4-point rating scale and the results of the measurement properties were analyzed by the criteria developed by Terwee et al. **Results**: A total of 11 instruments were considered eligible, none of which were created in Brazil. The process of cross-cultural adaptation was inadequate in 10 out of 11 instruments due to the lack of back-translation or due to inappropriate target population. All of the instruments presented flaws in the measurement properties, especially reliability, internal consistency, and construct validity. **Conclusion**: The flaws observed in both cross-cultural adaptation process and testing measurement properties make the results inconclusive on the validity of the available instruments. Adequate procedures of cross-cultural adaptation and measurement properties of these instruments are strongly needed.

Keywords: stroke; validity of tests; reproducibility of results; translating; physical therapy.

BULLET POINTS

- 11 studies were found that assessed post-stroke motor function in Brazilian patients.
- Most of the cross-cultural adaptation was conducted without the target population.
- Flaws in the measurement properties made the results inconclusive.
- Caution should be taken in the selection of instruments for research and clinical practice.

HOW TO CITE THIS ARTICLE

Lima E, Teixeira-Salmela LF, Simões L, Guerra ACC, Lemos A. Assessment of the measurement properties of the post-stroke motor function instruments available in Brazil: a systematic review. Braz J Phys Ther. 2016 Mar-Apr; 20(2):114-125. http://dx.doi.org/10.1590/bjpt-rbf.2014.0144

Introduction

Various measurement instruments have been created with the objective of assessing motor function in post-stroke individuals[1-3]. These instruments aim to verify the ability to maintain or change the body's position in space, walk and move around, move and handle objects, as well as verify motor coordination and fine manual motricity[1-3]. These abilities involve aspects related to activities and participation and the structure and function of the organs and systems, as described in the International Classification of Functioning, Disability, and Health (ICF)[4].

The application of these instruments aims to measure upper limb function, trunk function, or global motor function[1-3]. Some instruments assess performance through the observation of performed activities, while others are based on questionnaires on motor function[1-3]. After stroke, motor function can present various degrees of impairment and generate social and economic loss. Therefore, it is essential to use valid instruments to achieve an effective rehabilitation[5-7].

In general, the instruments used in Brazil to assess post-stroke motor function were developed in other

[1] Departamento de Fisioterapia, Universidade Federal de Pernambuco (UFPE), Recife, PE, Brazil
[2] Departamento de Educação Física, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brazil
Received: Apr. 16, 2015 Revised: Aug. 12, 2015 Accepted: Aug. 26, 2015

countries, usually in English and, consequently, targeted to the original population[1,8]. However, before an instrument can be used in a new country, culture, and/or language, a cross-cultural adaptation process is necessary. This process requires a standardized method involving the language translation and the cross-cultural adaptation to maintain its content validity[9,10]. After this process, the new scale should be applied to the new target population and its measurement properties can be analyzed to check if the adapted instrument truly measures the construct in the new setting[9-12].

The instrument can only be considered valid and reliable for use in a new cultural-clinical context through the adequate evaluation of the measurement properties[9-11]. The objectives of this systematic review were to identify the measurement instruments of motor functions in post-stroke individuals available in Brazil, to assess the methodological quality of the studies, and to assess the results of these studies.

Method

Two independent reviewers (EL and LS) conducted searches and selected eligible studies in the PUBMED, LILACS, SCOPUS, CINAHL, and WEB OF SCIENCE databases between February and March of 2014, according to the search strategy presented in Table 1. There was no language restriction.

Either cross-cultural adaptation studies or Brazilian instruments that assessed the motor function of post-stroke individuals in at least one item were considered eligible. Furthermore, these studies had to have verified at least one measurement property of

Table 1. Research strategies for each research database.

DATABASE	RESEARCH STRATEGY
MEDLINE (PUBMED)	((“Brazil “ [Mesh]) OR Portuguese OR Brazilian) AND ((“Stroke” [Mesh]) OR (“Paresis” [Mesh])) AND ((“Questionnaires” [Mesh]) OR scale OR test OR performance based test) AND Sensitive search filter for measurement properties NOT Exclusion Filter *
LILACS	(Brazil OR Portuguese Brazilian OR) AND (Stroke OR Stroke OR paresis) AND (Trunk OR upper limbs OR lower limbs OR sensorimotor function OR motor OR Function motor activity OR mobility OR coordination balance OR instrumentation OR comparative studies OR validation studies OR translations OR translation adjustment OR cross-cultural equivalence OR Validity OR validation OR Reliability OR reproducibility OR reproducible OR psychometric tests OR psychometric properties OR clinimetric clinimetric OR property OR valuation OR inter-observer OR variation results OR Intra-examiner OR mony retest OR inter-rater OR intraobserver OR interparticipants OR intraparticipants OR internal consistency Rasch OR Effect OR Effect floor ceiling OR disability assessment OR questionnaires OR scale tests)
CINAHL AND WEB OF SCIENCE	(Brazil OR Brazilian OR Portuguese) AND (Stroke OR Paresis) AND (questionnaires OR scale OR test OR comparative studies OR validation studies OR validation OR translations OR cross cultural OR cross cultural adaptation OR cross cultural comparison OR cross-cultural equivalence OR validity OR reliability OR reproducibility OR psychometrics OR clinimetrics OR outcome assessment OR observer variation OR reproducibility of results OR internal consistency OR alpha Cronbach OR agreement OR precision OR test-retest OR interrater OR inter-rater OR intrarater OR intra-rater OR intertester OR inter-tester OR intratester OR intra-tester OR interobserver OR inter-observer OR intraobserver OR intraobserver OR intertechnician OR inter-technician OR intratechnician OR intra-technician OR interexaminer OR inter-examiner OR intraexaminer OR intra-examiner OR interassay OR inter-assay OR intraassay OR intra-assay OR interindividual OR inter-individual OR intraindividual OR intra-individual OR interparticipant OR inter-participant OR intraparticipant OR intra-participant OR kappa OR concordance OR intraclass OR dimension OR subscale OR responsivity OR ceiling effect OR floor effect OR Item response model OR IRT OR Rasch)
SCOPUS	TITLE-ABS-KEY(stroke OR paresis) AND TITLE-ABS-KEY(cross cultural adaptation OR cross cultural comparison OR translation) AND TITLE-ABS-KEY(Brazil OR Brazilian OR Portuguese)
	TITLE-ABS-KEY(stroke OR paresis) AND TITLE-ABS-KEY(validation OR validity OR validation studies) AND TITLE-ABS-KEY(Brazil OR Brazilian OR Portuguese)
	TITLE-ABS-KEY(stroke OR paresis) AND TITLE-ABS-KEY(psychometrics OR clinimetrics OR reliability OR reproducibility OR responsiveness OR internal consistency OR intra examiner OR inter examiner OR responsivity OR ceiling effect OR floor effect) AND TITLE-ABS-KEY(brazil OR brazilian OR portuguese)

*In accordance with the recommendations of the Consensus-based Standards for the selection of health Measurements Instruments-COSMIN[13].

these instruments. Studies that involved individuals with other neurological conditions were excluded.

The two reviewers (EL and LS) screened the studies by title and abstract performing a pre-selection through eligibility criteria on the computer screen. Then, they read the full text of the studies potentially eligible to confirm their inclusion. It was pre-defined that disagreements between two reviewers were arbitrated by a third reviewer (AL).

The data extraction was performed in a standardized way through a pre-established data extraction form. The following data were extracted: title, authors, year of publication, journal, study objectives, eligibility criteria of the participants, instrument objective (discriminative, predictive or evaluative)[12], number of subscales/items/domains, and domain assessed according to ICF[4].

The evaluation of the methodological quality of the included studies was performed through the COSMIN checklist with 4-point rating scale, which is a tool created through the Consensus-based Standards for the selection of health Measurements Instruments (COSMIN), with the aim of scoring and classifying the quality of the methods used for the study of each measurement property[14-16].

The COSMIN checklist with 4-point rating scale is composed of nine boxes: A- Internal consistency, B- Reliability, C- Measurement error, D- Content validity, E- Structural validity, F- Hypothesis tests, G- Cross-cultural validity, H- Criterion validity, and I- Responsiveness[14-16].

Each box includes a series of items that assess the measurement property methodology. These items are classified on a scale of 4 points: 1- Poor, 2- Fair, 3- Good, and 4- Excellent. The final classification for each box is determined by the lowest score achieved by any of the items[14-16].

In addition to the boxes mentioned above, there is still another box that should be completed for each measurement property. This box aims to identify the clinical-epidemiological profile of the population, analyzing age mean, distribution by gender, illness characteristic, country of origin, and spoken language[14-16].

For example, to assess the internal consistency, box A presents 11 items: the first 3 items assess the missing data. Item 4 assesses the sample size; items 5, 6, and 7 assess questions related to unidimensionality; item 8 verifies the presence of other methodological flaws; and the other items verify the statistical method[14-16].

Item 4 of this box assesses the sample size for internal consistency as follows: Excellent (N=100), Good (between 50 and 99 participants), Fair (between 30 and 40 participants), and Poor (less than 30 participants). Similar to item 4, the remaining items are scored on a 4-point scale, according to specific criteria. In the end, even if the instrument has obtained "excellent" classification in the other items, but in item 4 received a "good" score for having a sample size between 50 and 99 participants, the internal consistency of the instrument will be classified as having "good" internal consistency as the lowest score is used[14-16].

Furthermore, the COSMIN recommends that to complement the evaluation of an instrument, the quality criteria developed by Terwee et al.[11] should be used; these criteria classify the measurement properties as Positive (+), Negative (-), or doubtful (?) focusing on the analysis of the obtained results[11]. The use of the Terwee et al.[11] criteria complements the evaluation of the measurement properties, as the COSMIN does not determine the cut offs that are considered adequate for the statistical analysis of each measurement property. In other words, the fact that a study used Cronbach's α, one of the statistical measurements advocated by COSMIN, to verify the internal consistency does not guarantee the quality of this property, as adequate values may not have been reached[11,14-16].

For example, internal consistency receives a positive score when the unidimensionality is verified, with the participation of 100 or more individuals and through Cronbach's α (between 0.70 and 0.95). If α does not reach this interval, the score will be negative. When the unidimensionality is not verified, or if there is another methodological flaw, the score will be classified as inconclusive[11].

Results

A total of 529 studies were found, of which only 14 studies[17-30] were included through the eligibility criteria (Figure 1). Two instruments (Test Évaluant les Membres supérieurs des Personnes Âgées - TEMPA[20] and the Jebsen-Taylor Test[21]) were not specifically created for post-stroke individuals; however, they have been validated in Brazil for this population and were included in this review.

In the 14 studies included, 11 instruments were identified. Three of them (Motor Activity Log - MAL[18,19], Fugl-Meyer scale[23,24], and National Institute of Health Stroke Scale - NIHSS[25,26]) were analyzed in two studies each, and the other 8 in only one

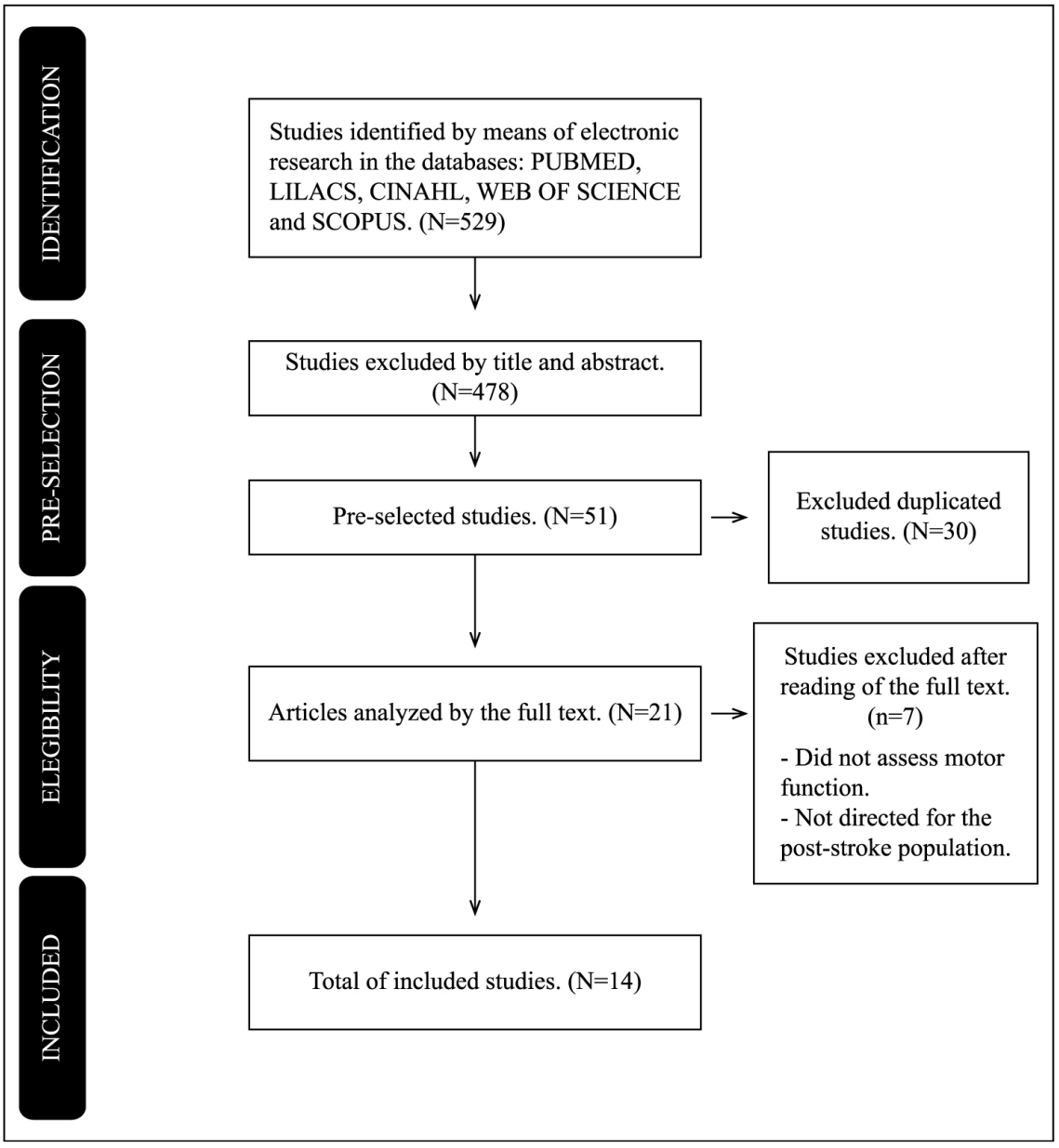

Figure 1. Identification, selection, and inclusion of the studies.

study. The characterization of the included studies is presented in Table 2.

None of the included instruments were Brazilian, therefore all had to be submitted to the cross-cultural adaptation process. The most frequently measured properties were reliability (n=11 studies), construct validity through hypothesis testing (n=6 studies), and internal consistency (n=6 studies). None of the included studies assessed responsiveness. The evaluation of the measurement properties is shown in Table 3.

Cross-cultural adaptation

Only one instrument (MAL)[18,19] was adapted to Brazilian culture in accordance with the recommended method[9,14-16]. The cross-cultural adaptation of the Rivermead Mobility Index (RMI)[22] obtained a "good" classification, given that it is not clear whether an expert committee participated or whether a pre-test was conducted[9].

Six instruments[17,21,23,25-27,30] obtained a classification considered "poor" for the cross-cultural adaptation

Table 2. Description of the instruments that assess the motor function of post-stroke individuals available in Brazil.

INSTRUMENT	AUTHOR/YEAR	CONSTRUCT	OBJECTIVE	TYPE OF INSTRUMENT	DIMENSION OF CIF	Nº OF ITEMS/ SUBSCALES	BRAZIL VERSION (AUTHOR/YEAR)
Postural Assessment Scale	Benaim et al.[31] (1999)	Maintenance and postural changes	Evaluative	Based on performance	Activity	12 items	Yoneiama et al.[28] (2008)
Trunk impairment Scale	Fujiwara et al.[32] (2004)	Trunk motor function	Evaluative	Based on performance	Structure and function	7 items	Lima et al.[30] (2008)
Trunk impairment Scale	Verheyden et al.[33] (2004)	Trunk motor function	Evaluative	Based on performance	Structure and function /Activity	17 items	Castelassi et al.[29] (2009)
Fugl-Meyer Scale	Fugl-Meyer et al.[34] (1975)	Global sensory-motor function	Evaluative	Based on performance	Structure and function /Activity	7 subscales	Maki et al.[24] (2006) Michaelsen et al.[23] (2011)
Jebsen-Taylor Test	Jebsen et al.[35] (1969)	Upper limb motor function	Evaluative	Based on performance	Activity	7 items	Ferreiro et al.[21] (2010)
Motor Assessment Scale (MAS)	Carr et al.[36] (1985)	Global motor function	Evaluative	Based on performance	Activity	8 items	Conte et al.[27] (2009)
Motor Activity Log (MAL)	Uswatte et al.[37] (2006)	Upper limb motor function	Evaluative	Questionnaire	Participation	30 items	Saliba et al.[18] (2011) Pereira et al.[19] (2012)
National Institutes of Health stroke scale (NIHSS)	Brott et al.[38] (1989)	Losses from the Stroke	Predictive	Based on performance	Activity	11 items (2 for motor function)	Cincura et al.[25] (2009) Caneda et al.[26] (2006)
Rivermead Mobility Index (RMI)	Collen et al.[39] (1991)	Global motor function	Evaluative	Questionnaire	Activity	15 items	Pavan et al.[22] (2010)
TEMPA (Test Évaluant les Membres supérieurs des Personnes Âgées)	Desrosiers et al.[40] (1993)	Upper limb motor function	Evaluative	Based on performance	Activity	9 items (8 items in the Brazilian version)	Michaelsen et al.[20] (2008)
Wolf Motor Function Test	Morris et al.[41] (2004)	Upper limb motor function	Evaluative	Based on performance	Structure and function /Activity	17 items	Pereira et al.[17] (2011)

Table 3. Evaluation of the methodological quality of the included studies through the COSMIN checklist with 4-point rating scale: Consensus-based standard for the selection of health measurement instruments.

INSTRUMENT	AUTHOR/YEAR	USED TRI	TRI	INTERNAL CONSISTENCY (A)	RELIABILITY (B)	MEASUREMENT ERROR (C)	F HYPOTHESIS TEST (F)	CROSS-CULTURAL ADAPTATION (G)	GENERALIZATION BY BOX
Postural Assessment Scale	Yoneiama et al.[28] (2008)	No	-	Poor	(Inter and test-retest) Poor	-	Poor	Fair	A Excellent B Excellent C Excellent F Excellent G Poor
Trunk Commitment Scale	Lima et al.[30] (2008)	No		Poor	(Inter and test-retest) Poor	-	Poor	Poor	A Excellent B Excellent F Excellent G Poor
Trunk Deficiency Scale	Castelassi et al.[29] (2009)	No	-	Poor	Poor	Poor	Poor	Fair	A Excellent B Excellent F Excellent G Reasonable
Fugl-Meyer Scale	Maki et al.[24] (2006)	No	-	-	(Inter and test-retest) Poor	-	-	Fair	B Excellent G Poor
Fugl-Meyer Scale (manual)	Michaelsen et al.[23] (2011)	No	-	-	(Inter) Poor	-	-	Poor	B Excellent G Poor
Jebsen-Taylor Test	Ferreiro et al.[21] (2010)	No	-	Poor	(Inter and Intra) Fair	-	-	Poor	A Excellent B Excellent G Poor
Motor Assessment Scale	Conte et al.[27] (2009)	No	-	-	(Inter and Intra) Poor	-	-	Poor	B Excellent G Poor
Motor Activity Log	Saliba et al.[18] (2011)	Yes	Good	Poor	Poor (test-retest)	-	-	Excellent	A Excellent B Excellent F Excellent G Excellent
Motor Activity Log	Pereira et al.[19] (2012)	No	-	-	(Inter and test-retest) Poor	(Inter and test-retest) Poor	Fair	-	B Excellent C Excellent F Excellent

TRI: Theory of the Answer to the TEMPA item: Test Évaluant les Membres supérieurs des Personnes Âgées. The following properties were not analyzed in the studies: D: Content Validity; E: Structural validity; H: Criteria validity; I: Responsiveness; J: Interpretability.

Table 3. Continued...

INSTRUMENT	AUTHOR/YEAR	USED TRI	TRI	INTERNAL CONSISTENCY (A)	RELIABILITY (B)	MEASUREMENT ERROR (C)	F HYPOTHESIS TEST (F)	CROSS-CULTURAL ADAPTATION (G)	GENERALIZATION BY BOX
National Institute of Health stroke scale	Cincura et al.[25] (2009)	No	-	-	Good	-	Fair	Poor	B Excellent F Excellent G Poor
National Institute of Health stroke scale	Caneda et al.[26] (2006)	No	-	-	Good	-	-	Poor	B Excellent
Rivermead Mobility Index	Pavan et al.[22] (2010)	No	-	Poor	Poor (test-retest)	-	-	Good	A Excellent B Excellent G Poor
TEMPA	Michaelsen et al.[20] (2008)	No	-	-	(Inter and test-retest) Poor	-	Poor	Poor	B Excellent F Excellent G Poor
Wolf Motor Function Test	Pereira et al.[17] (2011)	No	-	-	(Inter and Intra) Poor	(Inter and Intra) Poor	-	Poor	B Excellent C Excellent G Poor

TRI: Theory of the Answer to the TEMPA item: Test Évaluant les Membres supérieurs des Personnes Âgées. The following properties were not analyzed in the studies: D: Content Validity; E: Structural validity; H: Criteria validity; I: Responsiveness; J: Interpretability.

process. The TEMPA[20] and Motor Assessment Scale (MAS)[27] were created through simple translations into Portuguese. The NIHSS[25,26] was inadequately adapted in two studies, one of which only performed a single translation into Portuguese[25]. In the adaptation of the Wolf Motor Test (Wolf)[17], Jebsen-Taylor Test[21], and the Trunk Impairment Scale[30], pre-tests were not performed either.

For the Fugl-Meyer Scale[23,24], Posture Assessment Scale[28], and Trunk Deficiency Scale[29], the quality of the adaptation process was "fair" because it included a pre-test but did not include an adequate description of the assessed sample. However, the Fugl-Meyer Scale[23] manual, which was produced in a different study to the production of the instrument, presented a "poor" process as it included only one translation into Portuguese. In the Posture Assessment Scale[28] and Trunk impairment Scale[29], the translation and back translation were performed by only one translator.

Reliability

All of the instruments were tested for reliability. Eight (MAS[27], MAL[18,19], Wolf Motor Function Test[17], TEMPA[20], Posture Assessment Scale[28], Trunk impairment Scale[30], and the study of the Fugl-Meyer manual[24]) received a "poor" classification because they included fewer than 30 participants and used the intraclass correlation coefficient (ICC) when this was not indicated. The studies of the Fugl-Meyer Scale[23] (N=50) and RMI[22] (N=95) had good samples but were classified as "poor" for having used inadequate statistical methods (i.e. ICC and the Wilcoxon test, respectively). The Jebsen-Taylor Test[21] was considered "fair" for presenting a sample between 30 and 49 individuals (n=40). The reliability of the NIHSS was verified in two studies with "good" methodology and samples of 51 and 62 participants, respectively[25,26].

Measurement error

The measurement error was verified in three instruments (Trunk impairment Scale[29], Wolf Motor Function Test[17], and MAL[18,19]) through the Bland-Altman plot analysis; however, the methodological quality was classified as "poor" because the sample included less than 30 individuals.

Internal consistency

Six instruments[18,19,21,22,28-30] were tested for internal consistency; however, the methodological quality was classified as "poor" in all of them. In five instruments (Posture Assessment Scale[28], RMI[22], Jebsen Taylor Test[21], and Trunk impairment Scales[29,30]), the reason was the lack of factor analysis. Moreover, in the Posture Assessment Scale[28] and Trunk impairment Scales[29,30], the sample included less than 30 individuals and in the study of MAL[18,19], the sample included less than 5 individuals per item of the instrument for unidimensionality.

Construct validity

Construct validity was analyzed in six instruments (MAL[19], TEMPA[20], Posture Assessment Scale[28], NIHSS[25], and Trunk Impairment Scales[29,30]) through the hypothesis tests by correlation with the Fugl-Meyer Scale[19,20,28,29], Barthel Index[25], Berg Balance Scale, and Functional Independence Measure[30]. The study method used in four of these instruments was classified as "poor" due to inadequate sample size (n<30)[20,28-30].

The MAL[19] and NIHSS[25,26] presented "fair" methodological quality in the validity tests, as the hypotheses about the direction and magnitude of the correlation were not previously formulated or described in the study; however, it was possible to assume the expected direction for the correlation (positive or negative).

Terwee criteria

As for the evaluation of the results of the measurement property analysis using the criteria of Terwee et al.[11], the majority of the studies presented doubtful results in the study of measurement properties, with the exception of the inter-examiner reliability of the NIHSS[25,26], which presented positive results with Kappa coefficient >0.70 in items 5a, 5b, 6a, and 6b (referring to upper and lower limb motor function) (Table 4).

The results of the measurement error tests of the internal consistency and of construct validity were considered doubtful due to the methodological flaws presented, as described previously[17,19,28].

The ceiling and floor effects, which reflect interpretability, were verified in two instruments (Trunk Impairment Scale[29] and Posture Assessment Scale[28]). The percentage of individuals who reached the minimum and maximum scores was lower than 15%, but with an inadequate sample size (<50). However, other measures of interpretability like the minimum clinically important difference and minimum important difference were not analyzed. Finally, criterion validity and responsiveness were not tested in any of the eligible studies.

Table 4. Measurement properties assessment through the Terwee et al.[11] criteria (2007).

Instrument	Author/ year	Internal Consistency	Reliability	Measurement error	Content validity	Construct validity	Criteria validity	Responsiveness	Floor and ceiling effect	Interpretability
Posture Assessment Scale	Yoneima et al.[28] (2008)	?	(inter and test-retest) ?	NA	NA	?	NA	NA	?	NA
Trunk commitment Scale	Lima et al.[30] (2008)	?	?	NA	NA	?	NA	NA	NA	NA
Trunk Deficiency Scale	Castelassi et al.[29] (2009)	?	?	?	NA	?	NA	NA	?	NA
Fugl-Meyer Scale	Maki et al.[24] (2006)	NA	(Inter and test-retest) ?	NA	NA	NA	NA	NA	NA	NA
Fugl-Meyer Scale Manual	Michaelsen et al.[23] (2011)	NA	(Inter) ?	NA	NA	NA	NA	NA	NA	NA
Teste de Jebsen-Taylor	Ferreiro et al.[21] (2010)	?	(inter and intra) ?	NA	NA	NA	NA	NA	NA	NA
Motor Assessment Scale	Conte et al.[27] (2009)	NA	(inter and intra) ?	NA	NA	NA	NA	NA	NA	NA
Motor Activity Log	Saliba et al.[18] (2011)	?	(test-retest) ?	NA	NA	?	NA	NA	NA	NA
Motor Activity Log	Pereira et al.[19] (2012)	NA	(inter and test-retest) ?	(inter and test-retest) ?	NA	?	NA	NA	NA	NA
NIHSS	Cincura et al.[25] (2009)	NA	(inter) +	NA	NA	?	NA	NA	NA	NA
NIHSS	Caneda et al.[26] (2006)	NA	(inter) +	NA	NA	NA	NA	NA	NA	NA
Rivermead Mobility Index	Pavan et al.[22] (2010)	?	(test-retest) ?	NA	NA	NA	NA	NA	NA	NA
TEMPA	Michaelsen et al.[20] (2008)	NA	(inter and test-retest) ?	NA	NA	?	NA	NA	NA	NA
Wolf Motor Function Test	Pereira et al.[17] (2011)	NA	(Inter and intra) ?	(Inter and intra) ?	NA	NA	NA	NA	NA	NA

(+)= Positive; (-)= Negative; (?)= Inconclusive; (NA)= Not Assessed. TEMPA: Test Évaluant les Membres supérieurs des Personnes Âgées; NIHSS: National Institute of Health Stroke Scale.

Discussion

The results of this review showed that the available instruments in Brazil for assessing post-stroke motor function are arising from cross-cultural adaptation, not from newly developed Brazilian. However, the findings are inconclusive regarding the quality of the cross-cultural adaptation as well as from measurement properties, due to flaws with regards to methodology. The main methodological flaw observed during the cross-cultural adaptation process of the included instruments was the absence of a pre-testing of the final version[17,20-22,27,30]. Only one instrument (Motor Activity Log - MAL)[18,19] followed the recommended processes for an adequate cross-cultural adaptation.

The goal of applying the instrument in the target population (pre-test) before the measurement property analysis aims to identify possible imperfections in the interpretation of the items of an interview and the viability of the tasks proposed by the instrument for the target population. Therefore, the performance of the pre-test allows the identification of possible adjustments necessary in the instrument, based on the direct participation of the population for which it was adapted[11].

Although some instruments performed a pre-test, most of the studies did not described the sample properly[24,28,29]. To allow the generalization of the results of a cross-cultural adaptation, the COSMIN checklist recommends that the participants involved in the pre-test should be clinically and epidemiologically reported in terms of age, gender, characteristics of the illness, and source of patients (hospital, clinic, community, etc.)[14-16].

The absence of a back translation was also verified in some instruments[20,27]. This stage has the important aspect of allowing the verification of semantic equivalence between the original instrument and what was created in the new language, allowing necessary adjustments in the new version. It was also observed that, in some instruments[28,29], the stages of translation and back translation were performed by a single translator. The performance of multiple translations is recommended in the literature because it allows the interaction between specialists in the construct and in the languages involved, allowing a more adequate process of cultural adaptation and the maintenance of semantic equivalence[14-16].

Concerning the measurement properties, methodological flaws were also verified. The reliability was verified in all studies; however, in the majority of these, a sample size of less than 30 participants was selected[17-20,22,27-30]. The adequate number is at least 50 participants, and for an ideal sample, the recruitment of at least 100 participants is recommended[14-16].

In addition, the intra-lass correlation coefficient was often chosen as the statistical method when it was, in fact, inadequate. The adequate method for instruments with ordinal type scores is the Kappa coefficient[17-20,22-24,27-30]. The only instrument with an adequate study method for reliability, the NHSS, presented flaws in the cross-cultural adaptation[25,26].

For internal consistency, the majority of the studies did not report factorial analysis or unidimensionality study of the items[21,22,28-30]. These analyses are important because they intend to verify the number of dimensions into which the items are distributed and whether subscales are formed in the instrument. The only instrument to present the unidimensionality through the Rasch analysis, the (MAL)[18,19], included an inadequate number of participants[14-16].

In terms of internal consistency, a sample of 7 participants is indicated for each item of the instrument, requiring a minimum of 100. For example, for an instrument of 30 items, a sample of 210 would be indicated[14-16]. It is recommended that internal consistency should be assessed in two ways: through the classic form, or by the item response theory. First, Cronbach's alpha should be calculated after the performance of the factorial analysis, which identifies the number of subscales where the alpha must be calculated[14-16]. Second, the Rasch mathematical model is indicated to assess the unidimensionality of the items, verifying the presence of items that can be adjusted or removed from the instrument[14-16].

The flaws observed in the construct validity of the instruments[20,25,26,28-30] generate uncertainties about the degree to which the Brazilian versions of the included instruments truly measure the correct construct. It is recommended that 100 participants be assessed and that hypotheses be previously formulated about the direction and magnitude of the expected correlation between the scores of the tested instruments and the comparator instrument[14-16].The responsiveness and the criterion validity were not analyzed in any of the studies. The criterion validity is analyzed to verify the degree with which the scores of the instruments are an adequate reflection of the "gold-standard". However, for motor function measurement, no such instrument was observed in Brazil.

The absence of the responsiveness study, observed in all of the instruments, hampers the identification of the ability of these instruments to detect changes

in the assessed construct over time. Therefore, there is no evidence that it will be possible to quantify any motor function changes in post-stroke individuals in clinical research[14-16].

Finally, the interpretability of the obtained scores in these instruments still has not been clarified. Despite the fact that the ceiling and floor effects in the Posture Assessment Scale and Trunk Deficiency Scale were analyzed and had favorable results, the sample size in both studies was inadequate. None verified the minimum important change (MIC) or the minimum important difference (MID). These results are relevant because the MIC is the smallest change in the construct score the patients observe as important and the MID corresponds to the minimum difference in the construct among patients that is considered important[14-16]. None of the instruments were tested for their interpretability and responsiveness. As such, it remains unknown whether these instruments are able of measuring clinical changes over time.

Final considerations

Future studies should revise the cross-cultural adaptation processes, following all of the recommended stages (translation, synthesis of translations, back translation, expert committee, and pre-test). Moreover, the measurement properties should be analyzed with an adequate number of participants and the application of statistical methods that reflect the validity of each property. The results of this review point out that health professionals must be cautious when selecting instruments to assess post-stroke motor function for use in research and clinical practice in Brazil.

References

1. Soriano FFS, Baraldi K. Escalas de avaliação funcional aplicáveis a pacientes pós- Acidente Vascular Encefálico. Con Scientia e Saúde. 2010;9(3):521-30.
2. Pomeroy VM, King LM, Pollock A, Alison B, Peter L. Electrostimulation for promoting recovery of movement or functional ability after stroke. Cochrane Database Syst. Rev. 2006(2):1-74. http://dx.doi.org/10.1002/14651858.CD003241.pub2.
3. Pollock A, Gill B, Pomeroy VM, Langhorne P. Physiotherapy treatment approaches for the recovery of postural control and lower limb function following stroke. Cochrane Database Syst. Rev. 2007(1):1-92. http://dx.doi.org/10.1002/14651858.CD001920.pub2.
4. Leitão A. Manual da Classificação Internacional da Funcionalidade Incapacidade e Saúde (CIF). Lisboa: OMS; 2004.
5. Falcão IV, Carvalho EMF, Barreto KML, Lessa FJD, Leite VMM. Acidente Vascular Cerebral precoce: implicações para adultos em idade produtiva atendidos pelo Sistema Único de Saúde. Rev Bras Saude Mater Infant. 2004;4(1):95-102. http://dx.doi.org/10.1590/S1519-38292004000100009.
6. Ferreira LTD, Pinto EB. Perfil Funcional dos pacientes acometidos por acidente vascular encefálico de um centro de referência em Salvador. Fisioter Bras. 2008;9(2):81-5.
7. Hendricks HT, van Limbeek J, Geurts AC, Zwarts MJ. Motor Recovery after Stroke: A systematic Review of the Literature. Arch Phys Med Rehabil. 2002;83(11):1629-37. http://dx.doi.org/10.1053/apmr.2002.35473. PMid:12422337.
8. Ceccato RB, Almeida CI. O planejamento da reabilitação na fase aguda após acidente vascular encefálico. Acta Fisiátrica. 2010;17(1):33-47.
9. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976). 2000;25(24):3186-91. http://dx.doi.org/10.1097/00007632-200012150-00014. PMid:11124735.
10. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al, Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. Value Health. 2005;8(2):94-104. http://dx.doi.org/10.1111/j.1524-4733.2005.04054.x. PMid:15804318.
11. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34-42. http://dx.doi.org/10.1016/j.jclinepi.2006.03.012. PMid:17161752.
12. Kirshner B, Guyatt G. A Methodological framework for assessing health indices. J Chronic Dis. 1985;38(1):27-36. http://dx.doi.org/10.1016/0021-9681(85)90005-0. PMid:3972947.
13. Terwee CB, Jansma EP, Riphagen II, De Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurements instruments. Qual Life Res. 2009;18(8):1115-23. http://dx.doi.org/10.1007/s11136-009-9528-5. PMid:19711195.
14. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010;19(4):539-49. http://dx.doi.org/10.1007/s11136-010-9606-8. PMid:20169472.
15. Mokink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Med Res Methodol. 2010;(10):22.
16. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology and definitions of measurements properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737-45. http://dx.doi.org/10.1016/j.jclinepi.2010.02.006. PMid:20494804.
17. Pereira ND, Michaelsen SM, Menezes IS, Ovando AC, Lima RC, Teixeira-Salmela LF. Reliability of the Brazilian version of the Wolf Motor Function Test in adults with hemiparesis. Rev Bras Fisioter. 2011;15(3):257-65. http://dx.doi.org/10.1590/S1413-35552011000300013. PMid:21829991.
18. Saliba VA, Magalhães LC, Faria CDCM, Laurentino GE, Cassiano JG, Teixeira-Salmela LF. Adaptação Transcultural e análise das propriedades psicométricas da versão brasileira do instrumento Motor Activity Log. Rev Panam

Salud Publica. 2011;30(3):262-71. http://dx.doi.org/10.1590/S1020-49892011000900011. PMid:22069074.

19. Pereira ND, Ovando AC, Michaelsen SM, Anjos SM, Lima RCM, Nascimento LR, et al. Motor Activity Log-Brasil: reliability and relationships with motor impairments in individuals with chronic stroke. Arq Neuropsiquiatr. 2012;70(3):196-201. PMid:22392112.
20. Michaelsen SM, Natalio MA, Silva AG, Pagnussat AS. Reliability of the translation and adaptation of the Test d'Évaluation des Membres Supérieurs des Personnes Âgées (TEMPA) to the Portuguese language and validation for adults with hemiparesis. Rev Bras Fisioter. 2008;12(6):511-9. http://dx.doi.org/10.1590/S1413-35552008005000012.
21. Ferreiro KN, Santos RL, Conforto AB. Psychometrics properties of the Portuguese version of the Jebsen-Taylor test for adults with mild hemiparesis. Rev Bras Fisioter. 2010;14(5):377-82. http://dx.doi.org/10.1590/S1413-35552010005000018. PMid:21203696.
22. Pavan K, Cruz LC, Nunes MF, Menezes LG, Marangoni BE. Cross-cultural adaptation and validation of the Rivermead Mobility Index in stroke patients within the Brazilian cultural and language context. Arq Neuropsiquiatr. 2010;68(1):52-5. http://dx.doi.org/10.1590/S0004-282X2010000100012. PMid:20339653.
23. Michaelsen SM, Rocha AS, Knabben RJ, Rodrigues LP, Fernandes CGC. Translation, adaptation and inter-rater reliability of the administration manual for the Fugl-Meyer assessment. Rev Bras Fisioter. 2011;15(1):80-8. http://dx.doi.org/10.1590/S1413-35552011000100013. PMid:21519719.
24. Maki T, Quagliato EMAB, Cacho EWA, Paz LPS, Nascimento NH, Inoue MMEA, et al. Estudo de Confiabilidade da aplicação da escala de Fugl-Meyer no Brasil. Rev Bras Fisioter. 2006;10(2):177-83. http://dx.doi.org/10.1590/S1413-35552006000200007.
25. Cincura C, Pontes-Neto OM, Neville IS, Mendes HF, Menezes DF, Mariano DC, et al. Validation of the National Institutes of health stroke scale, Modified Rankin Scale and Barthel Index in Brazil: The role of cultural adaptation and structured interviewing. Cerebrovasc Dis. 2009;27(2):119-22. http://dx.doi.org/10.1159/000177918. PMid:19039215.
26. Caneda MAG, Fernandes JG, Almeida AG, Mugnol FE. Confiabilidade de escalas de comprometimento neurológico em pacientes com acidente vascular cerebral. Arq Neuropsiquiatr. 2006;64(3A):690-7. http://dx.doi.org/10.1590/S0004-282X2006000400034. PMid:17119821.
27. Conte ALF, Ferrari PP, Carvalho TB, Relvas PCA, Neves RCM, Rosa SF. Reliability, comprehension and acceptability of the Portuguese version of the Motor Assessment Scale in stroke patients. Rev Bras Fisioter. 2009;13(5):405-11. http://dx.doi.org/10.1590/S1413-35552009005000056.
28. Yoneiama SM, Roiz RM, Oliveira TM, Oberg TG, Lima NMFV. Validação da versão brasileira da Escala de avaliação postural para pacientes após acidente vascular encefálico. Acta Fisiatr. 2008;15(2):96-100.
29. Castelassi CS, Ribeiro EAF, Fonseca VC, Beinotti F, Oberg TD, Lima NMFV. Confiabilidade da versão brasileira da escala de deficiências de tronco em hemiparéticos. Fisioter Mov. 2009;22(2):189-99.
30. Lima NMFV, Rodrigues SY, Fillipo TM, Oliveira R, Oberg TG, Cacho EWA. Versão brasileira da Escala de Comprometimento do tronco: um estudo de validade em sujeitos pós-acidente vascular encefálico. Fisioter Pesqui. 2008;16(3):248-53. http://dx.doi.org/10.1590/S1809-29502008000300006.
31. Benaim C, Pérennou DA, Villy J, Rousseaux M, Pelissier JY. Validation of a standardized assessment of postural control in stroke patients: the Postural Assessment Scale for Stroke Patients (PASS). Stroke. 1999;30(9):1862-8. http://dx.doi.org/10.1161/01.STR.30.9.1862. PMid:10471437.
32. Fujiwara T, Liu M, Tsuji T, Sonoda S, Mizuno K, Akaboshi K, et al. Development of a new measure to assess trunk impairment after stroke (Trunk Impairment Scale): its psychometric properties. Am J Phys Med Rehabil. 2004;83(9):681-8. http://dx.doi.org/10.1097/01.PHM.0000137308.10562.20. PMid:15314532.
33. Verheyden G, Nieuwboer A, Mertin J, Preger R, Kiekens C, De Weerdt W. The trunk impairment scale: a new tool to measure motor impairment of the trunk after stroke. Clin Rehabil. 2004;18(3):326-34. http://dx.doi.org/10.1191/0269215504cr733oa. PMid:15137564.
34. Fugl-Meyer AR, Jaasko L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient: 1. A method for evaluation of physical performance. Scand J Rehabil Med. 1975;7(1):13-31. PMid:1135616.
35. Jebsen RH, Taylor N, Trieschmann RB, Trotter MJ, Howard LA. An objective and standardized test of hand function. Arch Phys Med Rehabil. 1969;50(6):311-9. PMid:5788487.
36. Carr JH, Shepherd RB, Nordholm L, Lynne D. Investigation of a new motor assessment scale for stroke patients. Phys Ther. 1985;65(2):175-80. PMid:3969398.
37. Uswatte G, Taub E, Morris D, Light K, Thompson PA. The Motor Activity Log-28: assessing daily use of the hemiparetic arm after stroke. Neurology. 2006;67(7):1189-94. http://dx.doi.org/10.1212/01.wnl.0000238164.90657.c2. PMid:17030751.
38. Brott TG, Adams HP Jr, Olinger CP, Marler JR, Barsan WG, Biller J, et al. Measurements of acute cerebral infarction: a clinical examination scale. Stroke. 1989;20(7):864-70. http://dx.doi.org/10.1161/01.STR.20.7.864. PMid:2749846.
39. Collen FM, Wade DT, Robb GF, Bradshaw CM. The Rivermead Mobility Index: a further development of the Rivermead Motor Assessment. Int Disabil Stud. 1991;13(2):50-4. http://dx.doi.org/10.3109/03790799109166684. PMid:1836787.
40. Desrosiers J, Hébert R, Dutil E, Bravo G. Development and reliability of an upper extremity function test for the elderly: the TEMPA. Can J Occup Ther. 1993;60(1):9-16. http://dx.doi.org/10.1177/000841749306000104.
41. Morris DM, Uswatte G, Crago JE, Cook EW 3rd, Taub E. The reliability of the wolf motor function test for assessing upper extremity function after stroke. Arch Phys Med Rehabil. 2001;82(6):750-5. http://dx.doi.org/10.1053/apmr.2001.23183. PMid:11387578.

Correspondence
Andrea Lemos
Universidade Federal de Pernambuco (UFPE)
Departamento de Fisioterapia
Avenida Professor Moraes Rego, 1235, Cidade Universitária
CEP 50670-901, Recife, PE, Brazil
e-mail: andrealemos4@gmail.com

original article

BJPT Brazilian Journal of Physical Therapy

Stimulus electrodiagnosis and motor and functional evaluations during ulnar nerve recovery

Luciane F. R. M. Fernandes[1], Nuno M. L. Oliveira[1], Danyelle C. S. Pelet[2], Agnes F. S. Cunha[2], Marco A. S. Grecco[3], Luciane A. P. S. Souza[1]

ABSTRACT | Background: Distal ulnar nerve injury leads to impairment of hand function due to motor and sensorial changes. Stimulus electrodiagnosis (SE) is a method of assessing and monitoring the development of this type of injury. **Objective**: To identify the most sensitive electrodiagnostic parameters to evaluate ulnar nerve recovery and to correlate these parameters (Rheobase, Chronaxie, and Accommodation) with motor function evaluations. **Method**: A prospective cohort study of ten patients submitted to ulnar neurorrhaphy and evaluated using electrodiagnosis and motor assessment at two moments of neural recovery. A functional evaluation using the DASH questionnaire (Disability of the Arm, Shoulder, and Hand) was conducted at the end to establish the functional status of the upper limb. **Results**: There was significant reduction only in the Chronaxie values in relation to time of injury and side (with and without lesion), as well as significant correlation of Chronaxie with the motor domain score. **Conclusion**: Chronaxie was the most sensitive SE parameter for detecting differences in neuromuscular responses during the ulnar nerve recovery process and it was the only parameter correlated with the motor assessment.

Keywords: chronaxie; ulnar nerve; evaluation studies; disability evaluation; rehabilitation; movement.

BULLET POINTS

- Stimulus electrodiagnosis is a reliable, noninvasive method of identifying neural regeneration.
- Chronaxie was the most sensitive parameter for assessing ulnar regeneration.
- Chronaxie and motor evaluation should be used to monitor neural regeneration.

HOW TO CITE THIS ARTICLE

Fernandes LFRM, Oliveira NML, Pelet DCS, Cunha AFS, Grecco MAS, Souza LAPS. Stimulus electrodiagnosis and motor and functional evaluations during ulnar nerve recovery. Braz J Phys Ther. 2016 Mar-Apr; 20(2):126-132. http://dx.doi.org/10.1590/bjpt-rbf.2014.0138

Introduction

Injury to the ulnar nerve is one of the most common upper limb peripheral nerve lesions. Eser et al.[1] conducted a retrospective study and found that most cases involved lesions of the ulnar nerve, with 337 cases (27%), followed by lesions of the median nerve, with 273 cases (22%). A complete and detailed evaluation of the hand is essential in order to identify a suitable treatment and achieve the best response to therapy. Rosén and Lundborg[2] developed a model for the specific evaluation of median and ulnar nerve lesions, considering three domains: motor, sensory, and pain/discomfort. In the motor domain, evaluation of the median and ulnar nerves involves testing the strength of key hand muscles, along with dynamometry measurements of handgrip strength.

Regarding functional evaluation, the DASH (Disability of the Arm, Shoulder, and Hand) questionnaire[3], which consists of three modules, employs a series of questions related to different tasks involving the upper limbs. This instrument was developed to measure dysfunction and physical symptoms in the upper limbs and to evaluate progress over time[3].

Among the various means of evaluating peripheral nerve lesions, a traditional physical therapeutic resource, which has nonetheless been infrequently utilized, is stimulus electrodiagnosis (SE). It is a reliable, noninvasive method of monitoring neural conditions and recovery progress[4,5]. In addition to its use for evaluation purposes, SE is the only resource available to establish the ideal conditions of therapeutic

[1] Departamento de Fisioterapia Aplicada, Universidade Federal do Triângulo Mineiro (UFTM), Uberaba, MG, Brazil
[2] Curso de Fisioterapia, Universidade Federal do Triângulo Mineiro (UFTM), Uberaba, MG, Brazil
[3] Departamento de Cirurgia, Universidade Federal do Triângulo Mineiro (UFTM), Uberaba, MG, Brazil
Received: Dec. 18, 2014 Revised: June 30, 2015 Accepted: Oct. 06, 2015

electrostimulation, ensuring use of the most suitable electrical pulse for treatment of a specific lesion[4,5]. SE is one of the most objective means of evaluating and monitoring the evolution of a peripheral nerve lesion[4,6,7], and of guiding the use of therapeutic electrostimulation[5]. In the present study, we attempted to test the efficiency and the importance of SE as a 'tool-of-the-trade' for physical therapists, particularly in ulnar nerve recovery.

There are no reports in literature concerning the use of SE in upper limb peripheral nerve lesions. No protocols have been defined for electrostimulation, and no studies have investigated the relationship between SE parameters and the results of physical therapeutic evaluation. If a good relationship is found, it could emphasize the importance of SE in the process of nerve lesion recovery and rehabilitation. Thus, the hypothesis of this study was that the SE parameters correlate with motor scores during peripheral ulnar lesion recovery. The objective of this study, therefore, was to identify the most sensitive parameters to use for the evaluation of ulnar nerve recovery and to correlate the SE parameter values with motor performance. The neuromuscular responses, which were obtained using electrodiagnosis during the recovery process after neurorrhaphy of the ulnar nerve, were evaluated, and the electrodiagnosis parameters were correlated with the results of motor assessments. Our objective was to determine whether the motor gains are linked to neural recovery, showing a possible new use for SE.

Method

Subjects

An observational, prospective cohort study was carried out to investigate the recovery of the ulnar nerve after neurorrhaphy. All subjects received information about the objectives and procedures of the study and signed an informed consent form, in accordance with regulation 466/12 of the Brazilian National Health Council. The study was approved by the ethics committee of Universidade Federal do Triângulo Mineiro (UFTM), Uberaba, MG, Brazil (protocol number 1663). The inclusion criteria selected patients with ulnar nerve lesions in the region of the wrist and distal forearm and who had been submitted to neurorrhaphy and then underwent physical therapy during the first three months after surgery. The exclusion criteria were refusal to participate in the study, postoperative complications (infection), failure to attend the evaluations, and absence of muscular response during the electrodiagnosis examination.

A sample size of nine patients was determined, based on the standard deviation values of 3.03 obtained for Chronaxie in a pilot project involving five patients with ulnar nerve lesions. Sample size estimation was calculated using Power and Sample v.3.0.4 software with power of 80% and α=0.05. The Chronaxie variable was selected due to its recognized importance in electrodiagnostic examinations[4,8,9].

The study was conducted with ten patients, including six men and four women with a mean age of 42 (SD=15) years. In the sample, only one subject was left-handed, and the right-hand side was more severely affected by lesions acquired in the workplace. All of the patients underwent neurorrhaphy, carried out by the same medical team, and were referred to the same physical therapy service at the UFTM, which followed the same protocol developed for the study. Nine patients underwent surgery in the first three weeks following injury.

Two evaluations of the injured limb were performed: initial (EV1) and final (EV2). EV1 was conducted during the initial phase (between 4 and 6 months post-surgery) and EV2 was performed at a later stage (between 10 and 15 months post-surgery). Evaluations (denoted EVWL) were also made on the contralateral limb (i.e. without lesion). All evaluations were conducted by the same examiner and under the same conditions.

Equipment and functional evaluation

The equipment used for the SE tests included: a) a universal pulse generator (Model Nemesys 941, Quark, Brazil); b) aluminum electrodes (10×5 cm); c) natural plant sponges (10×5 cm); d) an electrodiagnosis pen; and e) evaluation sheets (available from the instrument manual). For the motor assessment, a hydraulic dynamometer (Jamar®) was used, and the functional evaluation was conducted using the DASH questionnaire that had been translated and validated for use in Portuguese[10,11].

Electrodiagnostic testing

The parameters considered in the SE were Rheobase, Chronaxie, and Accommodation. Rheobase corresponds to the minimum stimulation intensity able to produce the smallest muscular contraction that can be perceived visually. This was achieved using a rectangular pulse with a period (T) of 1.0 s and an interval between pulses (R) of 2.0 s[8,12,13]. Chronaxie

corresponds to the shortest time necessary to produce a muscular contraction, also using a rectangular pulse with an interval of 2.0 s and an amplitude equal to two times the Rheobase obtained previously[8,12,14]. Accommodation is defined similarly to Rheobase, but the measurement is performed using an exponential pulse with a period of 1.0 s and an interval between the pulses of 2.0 s[8,14]. Both Rheobase and Accommodation are measures of intensity and are given in units of milliamps (mA), while Chronaxie is a measure of the duration or width of the pulse and is therefore given in milliseconds (ms)[8].

The subjects were placed in the seated position with the upper limb supported and maintaining the shoulder adduced, the elbow flexed at 90°, the forearm supine, and the wrist in a neutral position. First, the skin was cleansed with 70% alcohol in order to reduce its impedance. The muscle evaluated was the abductor of the fifth finger. An SMS (strong muscle stimulation) current was used to locate the motor point, employing a monopolar technique with two electrodes. One was a pen-type (active) electrode with an area sufficiently small to be able to stimulate the abductor muscle of the fifth finger. Dampened gauze was used to cover the metal tip in order to avoid direct contact with the skin. The other (passive) dispersive electrode had a greater area (in order to diminish the concentration of the electric charge on the skin) and was attached to the contralateral upper limb with an elastic band. The interface between this electrode and the skin was filled with a dampened sponge on the contractile part of the brachial biceps in order to close the circuit, following the recommendations provided in the manufacturer's manual. The stimulation electrode was positioned perpendicularly to the muscle under evaluation, and the pressure and angle of the pen were set after the motor point had been located. The intensity used was sufficient to induce a visible contraction. The SE was then initiated and the Rheobase, Chronaxie, and Accommodation values were recorded. The test was performed bilaterally, with the contralateral side used as the control.

Evaluation of motor performance

Hand muscle strength and grip strength were determined according to the standardized Rosén and Lundborg motor score procedure[2]. The muscles used to evaluate hand strength were the abductor of the fifth finger, the fourth palmar, and the first dorsal interosseous. The results were graded from zero to five, according to the Highet scale[15], and the values obtained for the three muscles were added and divided by 15 (the value for a normal individual).

The position adopted for measurements of grip strength was that recommended by the American Society of Hand Therapists (ASHT)[16]. Three measurements were performed and the arithmetic mean was calculated and divided by the mean for the healthy side.

Functional evaluation

In this study, only the module of the DASH questionnaire that evaluates functional ability was used. The score obtained varies from 0 to 100%, and the higher the score is, the greater the functional limitation[3,10,11]. The DASH test was only used in the final evaluation, in order to measure and describe functional status.

Data analysis

The normality of the SE data (Rheobase, Chronaxie, and Accommodation) was assessed using the Shapiro-Wilks test, and only the Chronaxie values were shown not to have normal distribution. Mean, standard deviation (SD), median, and maximum and minimum values were obtained through descriptive analysis. For the Chronaxie inferential analysis, the Wilcoxon matched pair (time) and Mann-Whitney U test for independent samples (side) were used. For Rheobase and Accommodation inferential analysis, the Student t test for dependent samples (time) and the Student t test for independent samples (side) were used. Finally, we calculated the correlation between the SE data and the Rosén and Lundborg[2] motor domain scores using the Spearman rank test. For all the tests, the significance level was set at 5%. The software Statistica 7 was used for all analyses.

Results

The SE values obtained were compared considering the initial (EV1) and final (EV2) evaluations and evaluations of the sides with and without lesion (EVWL) (Figures 1-3). Chronaxie was the parameter that best represented recovery of the ulnar nerve (Figure 2).

The Chronaxie values obtained for the sides without lesion were very close to zero, and similar results were obtained in the final evaluation. The minimum, mean, standard deviation, median, and maximum values for Rheobase, Chronaxie, and Accommodation obtained in the initial (EV1), final (EV2), and side without lesion (EVWL) evaluations are shown in Table 1.

According to statistical analysis, Rheobase did not present significant differences between the times of lesion EV1 and EV2 (p=0.56) or the side evaluated in EV1 (p=0.53) and in EV2 (p=0.88). Furthermore, the Accommodation values did not show any significant differences between the times of lesion (p=0.61) or between the sides (EV1: p=0.18 and EV2: p=0.56). On the other hand, the Chronaxie values were significantly different between EV1 and EV2 (p=0.01), as well as between the sides tested (EV1: p=0.00 and EV2: p=0.00).

The mean (SD) DASH values of the final evaluation were 33.1% (SD=21.3%) with minimum of 3.3% and maximum of 59.2%.

The relationship between the SE values and the Rosén and Lundborg[2] motor domain scores was investigated by calculating the Spearman correlation coefficient (r_s) (Table 2).

The Chronaxie parameter was the only parameter that showed a significant negative correlation with the Rosén and Lundborg[2] motor domain score in both the initial and final evaluations.

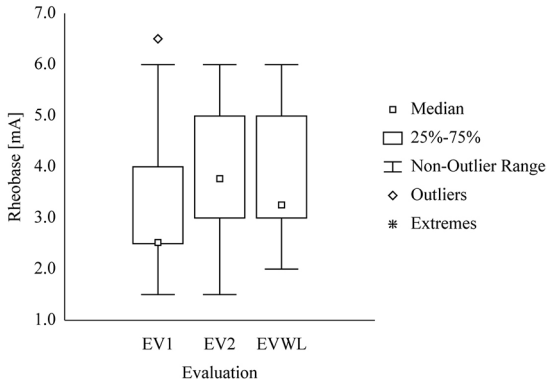

Figure 1. Boxplot of values of the initial (EV1), final (EV2), and side without lesion (EVWL) evaluations for Rheobase.

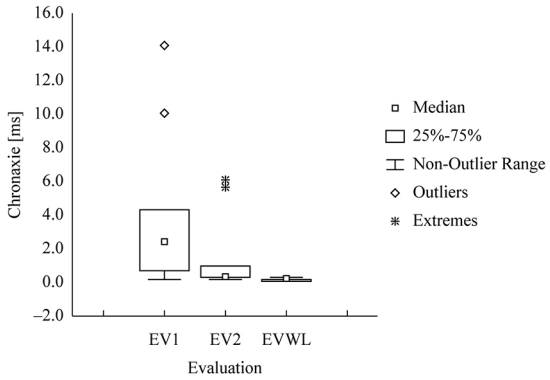

Figure 2. Boxplot of values of the initial (EV1), final (EV2), and side without lesion (EVWL) evaluations for Chronaxie.

Discussion

This study contributes to the literature concerning quantitative evaluation of recovery of the ulnar nerve, using a test that has been largely ignored in clinical

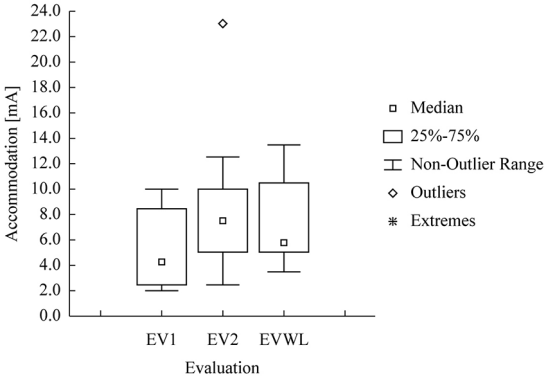

Figure 3. Boxplot of values of the initial (EV1), final (EV2), and side without lesion (EVWL) evaluations for Accommodation.

Table 1. Minimum, Mean, Median, Standard Deviation (SD), and Maximum values of Rheobase, Chronaxie, and Accommodation in the initial (EV1), final (EV2), and side without lesion (EVWL) evaluations.

	Evaluation	Minimum	Mean	SD	Median	Maximum
Rheobase [mA]	EV1	1.50	3.40	1.69	2.5	6.50
	EV2	1.50	3.75	1.37	3.75	6.00
	EVWL	2.00	3.85	1.43	3.25	6.00
Chronaxie [ms]	EV1	0.20	3.93	4.56	2.40	14.00
	EV2	0.20	1.51	2.28	0.30	6.00
	EVWL	0.10	0.18	0.07	0.20	0.30
Accommodation [mA]	EV1	2.00	5.3	3.16	4.25	10.00
	EV2	2.50	8.7	5.94	7.50	23.00
	EVWL	3.50	7.4	3.49	5.75	15.50

Table 2. Correlations between the values of the electrodiagnosis parameters and motor domain scores).

		r_s	*P*-value
EV1	Motor domain X Rheobase	0.313	0.38
	Motor domain X Chronaxie	–0.757	0.01*
	Motor domain X Accommodation	0.396	0.26
EV2	Motor domain X Rheobase	0.355	0.31
	Motor domain X Chronaxie	–0.794	0.01*
	Motor domain X Accommodation	0.247	0.49

EV1=Initial Evaluation. EV2=Final Evaluation. r_s=Spearman correlation coefficient. *significant at $p<0.05$.

practice in recent years. SE is a resource that can be used to aid diagnosis, evaluate the stage of a lesion and nerve recovery, and define the parameters used in electrostimulation[5]. In addition to SE, the Rosén and Lundborg[2] motor domain score and functional evaluation with the DASH questionnaire[3,10,11] were used.

The profile of the assessed patients with nerve lesions was as follows: mostly males (60%); average age of 42 years; right-hand limb most commonly affected (60% of cases), and most common cause of lesion was workplace accident (50%). A similar patient profile was reported by Eser et al.[1], who conducted a study using the data for 938 patients evaluated using electromyography and diagnosed with peripheral nerve lesions located in the upper limbs and the lower limbs. In that study, 71% of patients were male, the average age was 38 years, and the right-hand side was most frequently affected (55%). Most of the lesions (77%) were located in the upper limbs, and the main cause was car accidents (26.9%).

All patients were submitted to the same surgical procedure, which was performed by the same medical team. However, there were differences in clinical scenarios, including scarring processes and recovery times. Furthermore, the individuals showed differences in terms of both regeneration and sensitivity thresholds. Nonetheless, even considering these factors, Chronaxie proved to be sufficiently sensitive for detecting differences between the lesion phases.

It is important to emphasize that, although the SE method was described many years ago, it is rarely used in clinical practice despite the advantages described above and still requires further scientific investigation. It is likely that, in addition to a lack of information in the literature, its poor use could be related to difficulties encountered during application of the procedure. The test is detailed and requires an experienced physical therapist for its application and interpretation of the results. Another difficulty is related to the equipment required, because there are currently few options commercially available.

However, according to the initial hypothesis, the Chronaxie parameter correlates with motor scores during the recuperation of a peripheral ulnar lesion. Rheobase and Accommodation did not demonstrate any correlation. Chronaxie was the only parameter that showed significant differences between times of lesion and side with and without lesion. If only the Chronaxie test (based on Rheobase studies) was performed, which is relatively easy, it would be possible to understand the lesion and predict the motor behavior.

Chronaxie and Rheobase was first defined more than one hundred years[12] ago, and since that time, various researchers have studied these parameters in cases of peripheral nerve lesion[4,8,9,17,18]. They found that Chronaxie, which provides a measure of the neuromuscular electrical excitation threshold, was the most sensitive parameter for use in detection of nerve lesions. In this study, Chronaxie was also found to be the most sensitive parameter for use in lesion diagnosis and assessment of recovery of the ulnar nerve. The behavior of this parameter during the recovery/regeneration process was similar for all the patients, with high values during the initial phase and low values during the recovery phase. In the latter case, the values were very close to those obtained for the side without lesion.

There have been few reports of Chronaxie values for patients with peripheral nerve lesion. Licht et al.[19] associated Chronaxie values with the type and severity of lesion. The lesions were classified using six levels of severity, and the Chronaxie values were: 30-60 ms (neurotmesis); 20-30 ms (total axon degeneration); ~20 ms (partial axon degeneration); 10-20 ms (neuropraxy); 1-10 ms (moderate neuropraxy); and <1 ms (mild neuropraxy). The authors did not provide any information concerning the sample population. In this study, the Chronaxie values obtained in the initial evaluation correspond to normal physiology and moderate denervation. In the final evaluation, the values correspond to light denervation[19].

Ervilha and Araújo[4] conducted a study of Chronaxie using healthy individuals and individuals who had shown peripheral nerve lesions for more than eight months and less than two years. Seven muscles of the upper limb were evaluated and three Chronaxie value intervals were defined, depending on the severity of the lesion: the first (0.13 ms, SD=0.80 ms) was

representative of normal individuals, the second (1.5-20 ms) reflected moderate peripheral lesion, and the third (>30 ms) indicated severe lesion and a poor prognosis. Although the duration of the lesion was considered and different muscles were evaluated with the aim of classifying all of the nerves of the upper limb, there were gaps remaining between the established intervals where Chronaxie values were not associated with lesion severity.

Therefore, Chronaxie was found to be a sensitive and useful parameter that could be used to evaluate the process of recovery/regeneration following ulnar nerve lesion. A reduction in Chronaxie towards normal values is indicative of reinnervation or the avoidance of further degeneration of the muscle fiber[8]. Here, the Chronaxie values obtained in EV2 were very close to the values obtained for the side without lesion, for all but two of the patients.

The Rheobase and Chronaxie parameters were studied by Lee et al.[20] in patients suffering from encephalopathy after cerebrovascular accident. The results obtained for the paretic and non-paretic sides were compared, showing that the Rheobase and Chronaxie values were significantly higher for the paretic side. It could be inferred that reduction in muscular activity in cases of paresis or peripheral nerve lesion contributed to the need for greater stimulation in terms of both intensity and duration.

In the present study, the Rheobase values showed no similarity between patients or during the phases of lesion. High Rheobase values were found for both the side without lesion and in the final evaluation of some of the patients. The Accommodation parameter has not been the target of scientific studies in patients with peripheral nerve lesion, although studies have been conducted with animals[8,9,21]. Comparisons with the present study were, therefore, not possible.

Comparisons between SE parameters and clinical data for peripheral nerve lesions could not be found in literature. In the present study, Chronaxie showed a significant negative correlation with the values obtained for the Rosén and Lundborg[2] motor domain score. The correlation was negative because, in the process of neural regeneration, the Chronaxie values tended to diminish towards 0.2 ms while the motor domain score increased towards 1.0. The other parameters did not show any correlation with the motor domain score.

In terms of clinical applications, the results of this study reinforce the need for detailed, quantitative and carefully directed evaluation of patients with peripheral nerve lesions. The findings also indicate the desirability to reinstate electrodiagnostic evaluation in clinical practice. Chronaxie, especially, is a valuable parameter that can be used in assessments of the recovery/regeneration process. The Chronaxie value is extremely useful for determination of the duration of the electrical impulse used for muscle stimulation and helps in the application of stimulations that are more comfortable[22]. The optimum duration of an impulse is equal to the Chronaxie of the muscle that it aims to stimulate[23]. Because the Chronaxie test requires the Rheobase value, we also have an indication of a possible stimulation intensity value.

A limitation of this study was that we did not construct the quadratic and triangular pulse graphics, which might have given us a better idea of the recuperation process. The construction of such graphics should be included in future studies. The use of DASH at baseline could also provide information about functional gain during nerve recuperation.

In conclusion, stimulus electrodiagnosis is a quantitative, noninvasive technique for neuromuscular evaluation and can be used to accompany recovery following neurorrhaphy of the ulnar nerve. The Chronaxie parameter proved to be most sensitive for identifying differences between initial and final evaluations of the limb on the lesion side, as well as between the sides with and without lesion. This parameter also presented correlation with the results of clinical motor domain assessment. The renewed use of electrodiagnosis should therefore be encouraged, and the technique should be included in both clinical practice and academic courses.

Acknowledgements

Financial disclosure statements have been obtained, and no conflicts of interest have been reported by the authors or by any individuals in control of the content of this article. The authors would like to thank the patients who participated in this study. The project was financed by FAPEMIG and FUNEPU.

References

1. Eser F, Aktekin LA, Bodur H, Atan C. Etiological factors of traumatic peripheral nerve injuries. Neurol India. 2009;57(4):434-7. http://dx.doi.org/10.4103/0028-3886.55614. PMid:19770544.
2. Rosén B, Lundborg G. A new model instrument for outcome after nerve repair. Hand Clin. 2003;19(3):463-70. http://dx.doi.org/10.1016/S0749-0712(03)00003-9. PMid:12945644.
3. Hudak PL, Amadio PC, Bombardier C, Beaton D, Cole D, Davis A, et al. Development of an upper extremity outcome

measure: the DASH (disabilities of the arm, shoulder and hand). Am J Ind Med. 1996;29(6):602-8. http://dx.doi.org/10.1002/(SICI)1097-0274(199606)29:6<602::AID-AJIM4>3.0.CO;2-L. PMid:8773720.

4. Ervilha UF, Araujo RC. Estudo sobre a frequência de distribuição da cronaxia e a sua correlação com distintos graus de lesões nervosas periféricas. Rev Bras Fisioter. 1997;1(2):45-50.
5. Coutinho E L, Brasileiro OSM, Parizotto NA, Carmo JM, Carrinho PM. Fisioterapia-abordagem clínica e terapêutica das lesões nervosas periféricas. In: Tatagiba M, Mazzer N, Aguiar PHP, Pereira CU. Nervos periféricos: diagnóstico e tratamento clínico e cirúrgico. Rio de Janeiro: Revinter; 2003. p. 210-25.
6. Parry GJ. Electrodiagnostic studies in the evaluation of peripheral nerve and brachial plexus injuries. Neurol Clin. 1992;10(4):921-34. PMid:1331738.
7. Koka R, Hadlock TA. Quantification of functional recovery following rat sciatic nerve transection. Exp Neurol. 2001;168(1):192-5. http://dx.doi.org/10.1006/exnr.2000.7600. PMid:11170734.
8. Russo TL, França C, Castro C, Salvini TF. Alterations of chronaxie, rheobase and accommodation in denervated skeletal muscle submitted to electrical stimulation. Rev Bras Fisioter. 2004;8(2):169-75.
9. Russo TL, Peviani SM, Freria CM, Gigo-Benato D, Geuna S, Salvini TF. Electrical stimulation based on chronaxie reduces atrogin-1 and myoD gene expressions in denervated rat muscle. Muscle Nerve. 2007;35(1):87-97. http://dx.doi.org/10.1002/mus.20668. PMid:17034040.
10. Orfale AG, Araújo PMP, Ferraz MB, Natour J. Translation into Brazilian Portuguese, cultural adaptation and evaluation of the reliability of the disabilities of the arm, shoulder and hand questionnaire. Braz J Med Biol Res. 2005;38(2):293-302. http://dx.doi.org/10.1590/S0100-879X2005000200018. PMid:15785841.
11. Cheng HMS, Sampaio RF, Mancini MC, Fonseca ST, Cotta RMM. Disabilities of the arm, shoulder and hand (DASH): factor analysis of the version adapted to Portuguese/Brazil. Disabil Rehabil. 2008;30(25):1901-9. http://dx.doi.org/10.1080/09638280701749342. PMid:19061116.
12. Irnich W. The terms "chronaxie" and "rheobase" are 100 years old. Pacing Clin Electrophysiol. 2010;33(4):491-6. http://dx.doi.org/10.1111/j.1540-8159.2009.02666.x. PMid:20132498.
13. Ashley Z, Sutherland H, Lanmuller H, Unger E, Li F, Mayr W, et al. Determination of the chronaxie and rheobase of denervated limb muscles in conscious rabbits. Artif Organs. 2005;29(3):212-5. http://dx.doi.org/10.1111/j.1525-1594.2005.29037.x. PMid:15725219.
14. Paternostro-Sluga T, Schuhfried O, Vacariu G, Lang T, Fialka-Moser V. Chronaxie and accommodation index in the diagnosis of muscle denervation. Am J Phys Med Rehabil. 2002;81(4):253-60. http://dx.doi.org/10.1097/00002060-200204000-00003. PMid:11953542.
15. Seftchick JL, Detullio LM, Fedorczyk JM, Aulicino PL. Clinical examination of the hand. In: Skirven TM, Osterman AL, Fedorczyk JM. Rehabilitation of the hand and upper extremity. 6th ed. St Louis: Mosby; 2011.
16. Fess EE. The need for a consistent international language in hand rehabilitation. J Hand Ther. 1993;6(4):329. http://dx.doi.org/10.1016/S0894-1130(12)80337-1. PMid:8124449.
17. Fernández AM. Electrodiagnóstico y electroestimulación de músculos denervados. Fisioterapia. 2001;23:23-35. http://dx.doi.org/10.1016/S0211-5638(01)72970-7.
18. Robinson AJ, Snyder-Mackler L. Eletrofisiologia clínica: eletroterapia e teste eletrofisiológico. São Paulo: Art Med; 2002.
19. Licht S, Jornet A, Garcia-Alsina J. Electrodiagnóstico y electromiografía. Barcelona: Jims; 1970.
20. Lee W-D, Kim J-H, Lee J-U, Kim M-Y, Lee L-K, Yang S-M, et al. Differences in rheobase and chronaxie between the paretic and non-paretic sides of hemiplegic stroke patients: a pilot study. J Phys Ther Sci. 2013;25(6):717-9. http://dx.doi.org/10.1589/jpts.25.717. PMid:24259837.
21. Alves JSM, Leal-Cardoso JH, Santos-Junior FFU, Carlos PS, Silva RC, Lucci CM, et al. Limb immobilization alters functional electrophysiological parameters of sciatic nerve. Braz J Med Biol Res. 2013;46(8):715-21. http://dx.doi.org/10.1590/1414-431X20132626. PMid:23969978.
22. Seitz O, Gillert O. Klinischer Beitrag zur modernen Reizstromdiagnostik. Biomed Tech Eng. 1957;2(6):182-8.
23. Spielhoz N. Estimulação elétrica do músculo desnervado. In: Nelson RM, Hayes KW, Currier DP, editors. Eletroterapia clínica. Barueri: Manole; 2002. p. 411-46.

Correspondence
Luciane Fernanda Rodrigues Martinho Fernandes
Universidade Federal do Triângulo Minerio
Departamento de Fisioterapia Aplicada
Av. Frei Paulino, 30, Bairro Abadia
CEP 38025-180, Uberaba, MG, Brasil
e-mail: lfrm@terra.com.br, fernandes.luciane72@gmail.com

original article

BJPT Brazilian Journal of Physical Therapy

Wound healing treatment by high frequency ultrasound, microcurrent, and combined therapy modifies the immune response in rats

Raciele I. G. Korelo[1], Marcelo Kryczyk[2], Carolina Garcia[2], Katya Naliwaiko[3], Luiz C. Fernandes[4]

ABSTRACT | **Background:** Therapeutic high-frequency ultrasound, microcurrent, and a combination of the two have been used as potential interventions in the soft tissue healing process, but little is known about their effect on the immune system. **Objective:** To evaluate the effects of therapeutic high frequency ultrasound, microcurrent, and the combined therapy of the two on the size of the wound area, peritoneal macrophage function, $CD4^+$ and $CD8^+$, T lymphocyte populations, and plasma concentration of interleukins (ILs). **Method:** Sixty-five Wistar rats were randomized into five groups, as follows: uninjured control (C, group 1), lesion and no treatment (L, group 2), lesion treated with ultrasound (LU, group 3), lesion treated with microcurrent (LM, group 4), and lesion treated with combined therapy (LUM, group 5). For groups 3, 4 and 5, treatment was initiated 24 hours after surgery under anesthesia and each group was allocated into three different subgroups (n=5) to allow for the use of the different therapy resources at on days 3, 7 and 14 Photoplanimetry was performed daily. After euthanasia, blood was collected for immune analysis. **Results:** Ultrasound increased the phagocytic capacity and the production of nitric oxide by macrophages and induced the reduction of $CD4^+$ cells, the $CD4^+/CD8^+$ ratio, and the plasma concentration of IL-1β. Microcurrent and combined therapy decreased the production of superoxide anion, nitric oxide, $CD4^+$-positive cells, the $CD4^+/CD8^+$ ratio, and IL-1β concentration. **Conclusions:** Therapeutic high-frequency ultrasound, microcurrent, and combined therapy changed the activity of the innate and adaptive immune system during healing process but did not accelerate the closure of the wound.

Keywords: wound healing; ultrasonic therapy; electric stimulation therapy; physical therapy specialty.

BULLET POINTS

- Ultrasound therapy and microcurrent changed the activity of the innate and adaptive immune system.
- Combined therapy further accelerated the response of the adaptive immune system.
- Ultrasound, microcurrent, and combined therapy did not accelerate the closure of acute wounds.

HOW TO CITE THIS ARTICLE

Korelo RIG, Kryczyk M, Garcia C, Naliwaiko K, Fernandes LC. Wound healing treatment by high frequency ultrasound, microcurrent, and combined therapy modifies the immune response in rats. Braz J Phys Ther. 2016 Mar-Apr; 20(2):133-141. http://dx.doi.org/10.1590/bjpt-rbf.2014.0141

Introduction

Wound healing is a process in which the body tissue repairs itself through the coordinated action of extra- and intracellular events[1]. The wound-healing process leads to a formation of new tissue that is structurally and functionally identical to its previous state[2]. Biomarkers of inflammation can be monitored, and their persistence in the blood has been associated with disturbances in the healing process. This source of the inflammation has been attributed to macrophages[3].

Macrophages are cells from the innate immune system that phagocytose and kill pathogenic organisms[2]; they also produce proinflammatory cytokines (e.g.tumor necrosis factor [TNF]-α, interleukin [IL]-1β, and IL-6) and chemically reactive oxygen and nitrogen molecules. Forty eight hours after tissue injury, macrophages also release IL-10, an anti-inflammatory cytokine which initiates the remodeling of tissues in conjunction with growth factors[4]. Following the

[1] Programa de Pós-graduação em Educação Física, Coordenação do Curso de Fisioterapia, Universidade Federal do Paraná (UFPR), Curitiba, PR, Brazil
[2] Faculdade Dom Bosco, Curitiba, PR, Brazil
[3] Departamento de Biologia, UFPR, Curitiba, PR, Brazil
[4] Programa de Pós-graduação em Educação Física, Departamento de Fisiologia, UFPR, Curitiba, PR, Brazil
Received: Dec. 24, 2014 Revised: June 03, 2015 Accepted: Sept. 11, 2015

migration of macrophages to injured tissue, subsets of lymphocytes do the same. For example, $CD4^+$ T lymphocytes are anti-inflammatory while $CD8^+$ T lymphocytes are proinflammatory, and the $CD4^+/CD8^+$ ratio is reduced as the healing process progresses[5].

The management of wound preparation must follow the principle of tissue infection moisture edge (TIME), meaning that the microenvironment must be maintained to support conditions favorable to the healing process[6]. For the treatment of wounds, non-focused high-frequency ultrasound (HFU)[7-11] and microcurrent electrical stimulation (MET)[12-18] have been used as adjuvant therapies for wound healing. Several studies have reported the positive effects of these approaches based on different mechanisms[7,8,10,12,15,16,18] in the wound-healing process. However, the effects of HFU and MET on the immune system are not fully known, particularly in terms of macrophages and subsets of lymphocyte cells.

The combination of these two therapeutic modalities could amplify the effect of one modality alone. Combined therapy in the context of this paper consists of the therapeutic application of an electrical current through the ultrasonic transducer, providing sound pulses and electrical current flow simultaneously. This association saves time in the application and it is believed that the ultrasound increases the permeability of cell membranes, favoring the action of electrical currents on the nervous tissue[19]. However, research is scarce on the benefits of this synergy, and to the authors knowledge, the combined therapy of the two modalities has not been studied in relation to its effect on wound healing. Therefore, it was hypothesized that the use of physical therapy modalities (i.e. HFU and MET) applied alone would promote beneficial changes in the activities of the innate and adaptive immune system, and that the combined therapy would result in further acceleration of the wound-healing process over that which was believed to occur when each modality was applied as a treatment in isolation.

In this study, the effects of HFU, MET, and combined therapy on wound area, peritoneal macrophage function, lymphocyte immune functions, and the plasma concentration of ILs were investigated.

Method

Animal model

A total of 65 male Wistar rats were used (7 weeks old, weighing about 312.7 g, standard error of the mean [SEM]=11.6 g). The rats were kept in an environment that had a constant temperature (23±1 °C) under a light/dark cycle (12/12 h), with free access to food and water. This study followed the ethical rules established by Brazilian Law 11.794/08 and the recommendations of the Brazilian College of Animal Experimentation, Moreover, it was approved by the Animal Research Ethics Committee of the Universidade Federal do Paraná (CEUA-UFPR), Curitiba, Paraná, Brazil (protocol number: 561/2012).

Experimental design

The animals (n=65) were randomly allocated into five groups as follows: control (C, group 1, n=5), lesion and no treatment (L, group 2, n=15), lesion treated with ultrasound therapy (LU, group 3, n=15), lesion treated with microcurrent therapy (LM, group 4, n=15), and lesion treated with combined therapy (LUM, group 5, n=15). With the exception of the control group, all groups were subdivided into three subgroups (n=5 each) according to the number of days of treatment. The treatment was conducted once a day everyday until exsanguinated at 3, 7, and 14 days of treatment. The euthanasia of animals was carried out by decapitation following the ethical rules of the Resolution of the CFMV (Federal Board of Veterinary Medicine) number 1000/12. Two animals of group 2 were lost – during the surgical procedure to inflict the original injury. In order to solve this drawback these two animals were replaced at the time of surgery.

Surgical procedure

The rats were anesthetized with xylazine (0.05 mL, 10 mg/mL, IM/intramuscular) and ketamine (0.15 mL, 50 mg/mL, IM) diluted in 9% potassium chloride (0.25 mL). After the hair was removed from the back of each rat, the surgical procedure was carried out in which an area of skin was excised by one of the investigators (Figure 1), as described in details elsewhere[20]. After creating the wound, the lesion was cleaned with sterile gauze soaked in saline and received a primary cover (circular self-adhesive dressing of 1.5 cm) maintained for 4 hours after the excision. The animals were kept at room temperature under a warm heating pad to prevent hypothermia until full recovery. The animals were then kept in the room noted above and were allowed to have social interaction with the other four animals of the same subgroup.

Treatment protocol

The treatment was initiated 24 hours after surgery (Figure 2) and carried out once a day everyday, following the schedule (3, 7, or 14 days), before

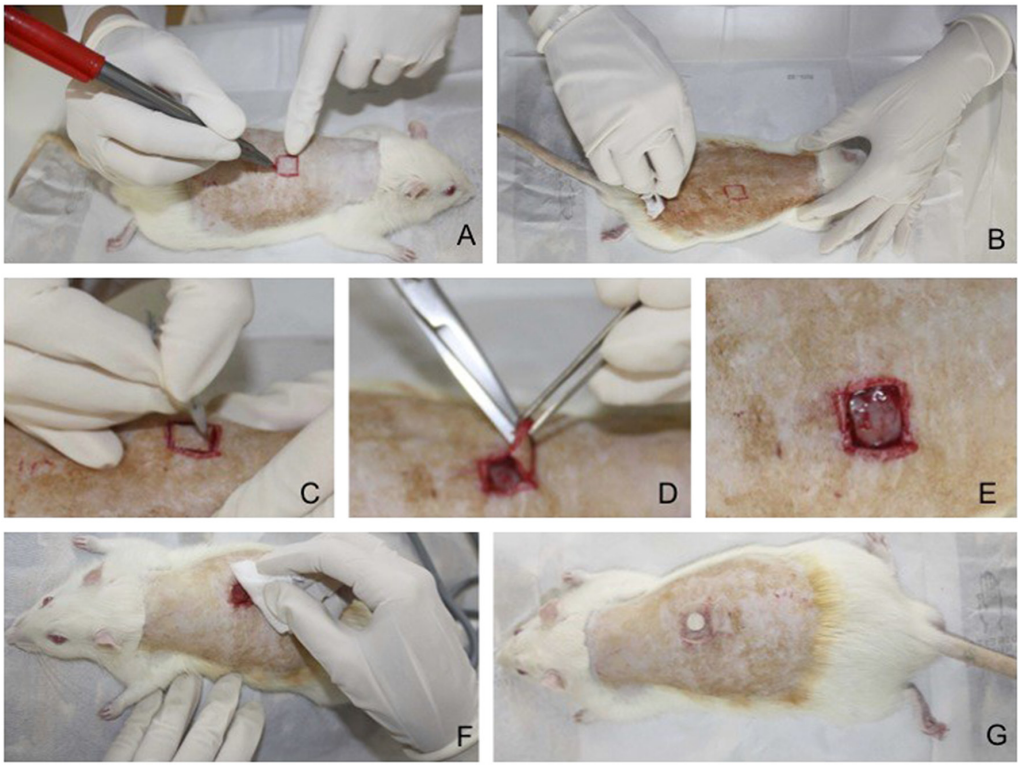

Figure 1. Surgical procedure for creating the lesion. (A) After hair removal, skin-marking excision wound with plastic mold of 1 cm^2 to mark the skin, located on the dorsal medial line of the animal, using 1 cm below the transverse line connecting the lower angle of the scapula as the cranial limit. (B) Asepsis with topical povidone-iodine. (C) Incision with a scalpel blade around the marked tissue. (D) Dissection of the excised skin in the suprafascial plane with tweezers and Mayo scissors, respecting the muscular fascia (2 mm deep). (E) Resection of the skin segment demarcated. (F) Cleaning of excision wound with sterile gauze soaked in saline solution. (G) Primary coverage placement with self-adhesive dressing.

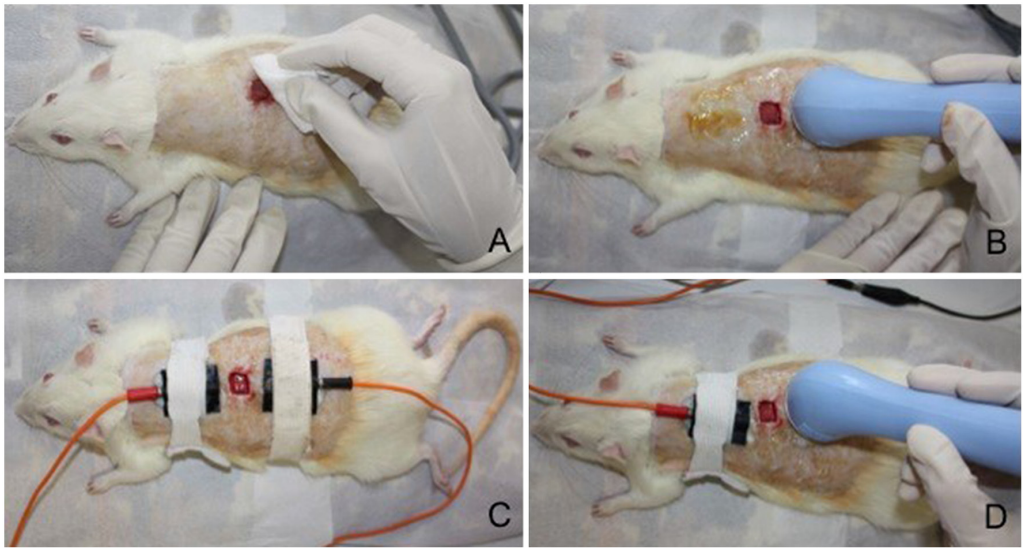

Figure 2. Intervention procedures in the different groups. (A) Cleaning the wound with sterile gauze and saline solution made for the L group (lesion with no treatment group) and prior to the application of the other three forms of intervention. (B) Application of therapeutic ultrasound in animals from the LU group (lesion treated with ultrasound group). (C) Application of microcurrent in animals from the LM group (lesion treated with microcurrent group). (D) Application of combined therapy in animals from the LUM group (lesion treated with ultrasound and microcurrent group).

euthanasia of animals. With the exception of the control group, all groups had their wounds cleaned daily, under anesthesia, prior to the application of the therapeutic resources. This was done with sterile gauze soaked in saline, and primary coverage was not returned after cleaning. The equipment used underwent assessment and received calibration certification. For the dosage, each parameter was chosen based on previous scientific evidence[8,13,17]. For the LU group (group 3), the HFU (Sonopulse III® IBRAMED portable equipment) was set using a transducer at 3 MHz with the following parameters: effective radiating contact area of 3.5 cm^2; pulsed mode (100 Hz); 50% duty cycle; spatial and temporal average intensity of 1.0 W/cm^2 and 0.5 W/cm^2, respectively; and therapeutic dose of 33.6 J/cm^2. The therapy was carried out for 8 minutes via direct contact with water soluble gel around the wound area (25 cm^2). For the LM group (group 4), a Neurodyn Esthetic® IBRAMED was used to deliver MET. Conductive silicone electrodes (30 mm × 50 mm) with water soluble gel were wrapped around the wound and the following parameters were used: intensity 300 μA, pulse frequency 100 Hz, monophasic rectangular pulse with polarity inversion at every 3 seconds, performed for 8 minutes. For the LUM group (group 5), the two modalities were interconnected using a cable specifically so that the ultrasonic transducer could apply both the sound and electrical pulse. Both modalities were parameterized and the procedures were the same as those described for the LU & LM groups (groups 3 & 4); the application was carried out with a static silicon-carbon electrode positioned at the cranial edge of the lesion. The ultrasonic transducer applied by direct contact by manual sliding the transducer around the free borders of the wound.

Percentage reduction of wound area-photoplanimetry

Wound contraction was evaluated daily until 14 days after surgery and therapies by photographic documentation. For standardization of images, a fixation system was used for the camera with the distance set at 15 cm; the light was controlled using two cold lamps with a power of 45 W. Evaluation wound contraction was analyzed by measuring the area of wound (transition of regular scar tissue skin on the photo), calculated using ImageJ software 1.47t[15], calibrated by the caliper used adjacent to the edge of wound[13], and expressed in percentage. The percentage reduction in the area of the wound was calculated using the following formula[21]:

$$\text{Percentage reduction of wound area } (\%) = \frac{(\text{Initial} - \text{Final wound area}) \times 100}{\text{Initial wound area}}$$

Interleukin plasma concentration

The measurements of IL-1β, 6, and 10 and TNF-α in the plasma were performed following the instructions of the ELISA kit manufacturer (Boster Immunoleader®).

Macrophage immune parameters

Resident macrophages were obtained by intraperitoneal lavage from the groups treated for 14 days to verify the functional parameters. Phagocytic capacity, neutral red retention, superoxide anion, hydrogen peroxide (H_2O_2), and nitric oxide (NO) production were performed in quadruplicate as described elsewhere[22].

T lymphocyte subpopulations $CD4^+/CD8^+$

Flow cytometry was used to measure the presence of $CD4^+$ and $CD8^+$, markers of helper and cytotoxic T lymphocytes, respectively, on the surface of freshly prepared blood lymphocytes, as has been described elsewhere[23]. About 10^6 cells/mL in phosphate buffered solution (PBS) supplemented with 0.1% fetal bovine serum (wt/v) and 0.05% sodium azide (wt/v) were incubated for 20 min at 4 °C in the dark with fluorescently labeled monoclonal antibodies to $CD4^+$ (clone OX-38) and $CD8^+$ (clone OX-8). After staining, cells were washed twice with PBS and immediately examined for fluorescence using a Becton Dickinson FACSCalibur.

Statistical analysis

Data are presented as mean±SEM and were subjected to normal analysis and homogeneity of variance using the Shapiro-Wilk test and Levene's test, respectively. For analysis of the percentage reduction in wound area, analysis of variance (ANOVA) with repeated mixed measures was used following a design 2 (reduction of percentage wound area: pre vs. post intervention) × 4 (treatment: lesion vs. ultrasound vs. microcurrent vs. combined therapy) × 3 (time: 3 days vs. 7 days vs. 14 days), with a significance level of $p<0.05$. For the other variables (IL plasma concentration, functional parameters of macrophage and lymphocyte subpopulations), two-way ANOVA between groups was used. Post hoc tests with Bonferroni's correction were employed for multiple comparisons, with a statistical significance of $p<0.05$. The effect size was determined by calculating omega as follows: $\omega=0.01$ was considered a small effect, $\omega=0.06$ was considered

an average effect, and values of ω above 0.14 were considered a large effect. Analyses were performed with IBM-SPSS software version 19 and the graphics prepared with GraphPad PRISM® software version 5.0 for Windows.

Results

Percentage reduction of wound area

There was a significant reduction in the wound area pre- and post-therapy intervention (Wilks's lambda=0.042, $F_{1,47}$=1076.7, p=0.00, partial eta squared=0.958). In addition, there was a positive interaction between time (3, 7 and 14 days) and pre-and post-intervention in the reduction of wound area (Wilks's lambda=0.096, $F_{2,47}$=221.1, p=0.00, partial eta squared=0.904). There was no significant interaction between the type of treatment and the percentage reduction in wound area pre- and post-intervention (Wilks's lambda=0.881, $F_{3,47}$=2.118, p=0.12, partial eta squared=0.119). Finally, no significant interaction was observed between the types of treatment for percentage reduction in wound area pre- and post-intervention at different times (Wilks's lambda=0.882, $F_{6,47}$=1.047, p=0.41, partial eta squared=0.118).

Interleukin concentrations

After 3 days (Table 1), only IL-1β was significantly increased (p=0.04 vs. C). The interventions in all three treatment groups (i.e.LU, LM, LUM [groups 3, 4 & 5]) induced a marked reduction in IL-1β (p=0.04, p=0.02, and p=0.04, respectively vs. L). HFU therapy induced a twofold increase in the concentration of IL-6 when compared to the control group (group 1) (p=0.04). At days 7 and 14, the concentration of such cytokines was undetectable.

Macrophage immune parameters

Table 2 shows the functional parameters of peritoneal macrophages and lymphocytes after surgery at 14 days of therapy. Phagocytosis and neutral red uptake were different between groups ($F_{4,39}$=4.095, p=0.03, ω=0.22 and $F_{4,39}$=7.390, p=0.00, ω=0.38, respectively). Phagocytosis and neutral red uptake in the L group (group 2) were reduced by 47% (p=0.03 vs. C) and 22.5% (p=0.08 vs. C), respectively. The HFU treatment

Table 1. Plasma concentration of interleukin (IL)-1β, IL-6, IL-10, and tumor necrosis factor (TNF)-α for the experimental period of 3 days after the excision wound model between the different intervention groups (n=5 for each group; C: control; L: lesion without treatment; LU: lesion treated with ultrasound; LM: lesion treated with microcurrent; LUM: lesion treated with combined therapy). The data represent the mean (minimum - maximum), expressed as pg/mL. Each experiment was performed in triplicate.

	C	L	LU	LM	LUM
IL-1β	nd	45 (27-64)[a]	2.6 (0-37)[b]	nd[b]	13 (6-52)[b]
TNF-α	nd	nd	nd	nd	nd
IL-6	nd	3.2 (0.2-6.1)	8.9 (3.5-11)[a]	1.2 (0.1-5.1)	nd[c]
IL-10	nd	nd	nd	nd	nd

[a]p<0.05 compared to C; [b]p<0.05 compared to L; [c]p<0.05 compared to LU; nd (not detectable).

Table 2. Functional parameters from peritoneal macrophages and blood $CD4^+$ and $CD8^+$ T lymphocytes 14 days after excision in the groups (n=5 for each group; C: control; L: lesion with no treatment; LU: lesion treated with ultrasound; LM: lesion treated with microcurrent; LUM: lesion treated with combined therapy). Macrophage data represent mean±SEM, expressed as absorbance/10^6 cells. Experiment was performed in octuplicate. Data from lymphocytes are expressed as percentage and the experiments were performed in duplicate.

	C	L	LU	LM	LUM
Macrophages Phagocytic capacity	1.91±0.21	1.0±0.20[a]	1.9±0.28[b]	1.74±0.35	1.58±0.24
Red retention	0.40±0.02	0.31±0.01[a]	0.32±0.01	0.26±0.01[a]	0.29±0.01[a]
Superoxide anion	2.56±0.19	2.12±0.15	1.71±0.16[a]	1.41±0.75[a,b]	1.44±0.15[a,b]
Hydrogen peroxide	1.60±0.07	0.93±0.06[a]	0.85±0.03[a]	1.09±0.08[a]	0.82±0.03[a,c]
Lymphocytes T $CD4^+$	21.2±0.4	22.6±0.2[a]	19.0±0.2[a,b]	19.4±0.3[a,b]	12.0±1.9[a,b,c,d]
T $CD8^+$	19.0±0.2	14.9±0.3[a]	19.0±1.9	20.3±0.8[b]	17.6±5.8
$TCD4^+/CD8^+$	1.1±0.04	1.5±0.01[a]	1.0±0.11[b]	0.9±0.05[b]	0.8±0.17[b]

[a]p<0.05 compared to C; [b]p<0.05 compared to L; [c]p<0.05 compared to LU; [d]p<0.05 compared to LM.

(LU group) [group 3] showed only the recovery of phagocytosis capacity (p=0.04 vs. the lesion without treatment group (group 2)). Neither microcurrent therapy (group 4) nor combined therapy (group 5) reversed the effect caused by excision (p>0.05). Superoxide and hydrogen peroxide production were different between groups ($F_{4,39}$=9.807, p=0.00, ω=0.46 and $F_{4,39}$=25.786, p=0.00, ω=0.78, respectively). Superoxide production did reduce in group 2 (lesion without treatment group) but was not different from the control group (p=0.56) [group 1]. HFU therapy (group 3) caused a further reduction in superoxide production but the reduction was not different from the lesion without treatment group (p=0.72) [group 2]. In contrast, the microcurrent therapy (group 4) and combined therapy (LUM) (group 5) caused a significant reduction (p<0.05 vs. group 2) in the superoxide production (33%) when compared to the Lesion without treatment group (group 2). For hydrogen peroxide, the lesion without treatment group (group 2) showed a reduction of H_2O_2 production of 42% (p=0.00 vs. the control group (group 1)). The different modalities (groups 3, 4, & 5) did not modify the H_2O_2 production when compared to the Lesion without treatment group (p>0.05) (group 2).

The production of NO is shown in Figure 3. This was different between groups ($F_{4,39}$=15.451, p=0.00, ω=0.74). The basal concentration of NO production

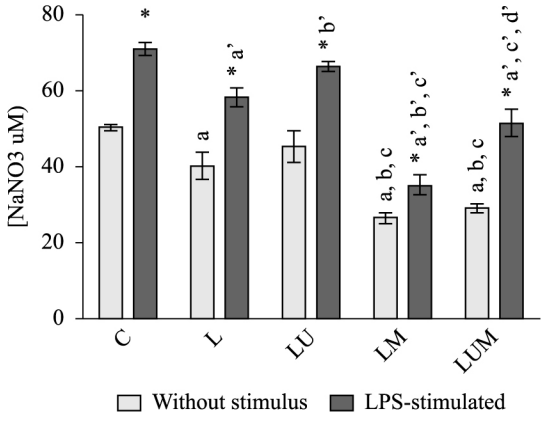

Figure 3. Nitric oxide production by peritoneal macrophages 14 days after excision model without stimulus and LPS-stimulated in the groups (C: control; L: lesion with no treatment; LU: lesion treated with ultrasound; LM: lesion treated with microcurrent; LUM: lesion treated with combined therapy). The data represent the mean±SEM, expressed as µmol/L. Each experiment was performed in quadruplicate. *p<0.05 relative to its respective group; [a]p<0.05 compared to C; [b]p<0.05 compared to L; [c]p<0.05 compared to LU; [d]p<0.05 compared to LM.

by peritoneal macrophages was reduced in the lesion without treatment group (group 2) (p=0.04 vs. the control group (group 1)). HFU therapy (group 3) did reverse the effect of the lesion (p=0.39 vs. the control group [group 1]). In contrast, for the microcurrent therapy alone group (LM group or group 4) and in combination with the HFU (the LUM group or group 5), the production was about 50% of the basal state. Under lipopolysaccharide (LPS) stimulation, it was also statistical different between the groups ($F_{4,19}$=30.875, p=0.00, ω=0.85). Under LPS stimulation, peritoneal macrophages increased in about 30% from control (p=0.00 *vs* non-stimulated). The same was seen in the L (group 2) and LU (group 3) groups (p<0.01 *vs.* non-stimulated). In the LM group (group 4), LPS stimulation increased the nitric oxide production by ~20% (p=0.02 *vs.* non-stimulated). In the presence of LPS, the combined therapy (group 5) increased the NO production by ~40% (p<0.01 *vs.* non-stimulated).

T lymphocyte subpopulations $CD4^+/CD8^+$

The excised lesion (group 2) increased (Table 2) the $CD4^+$ T lymphocyte subset (6.19%) as compared to control group (group 1) (p=0.03). The therapy with HFU (group 3) or microcurrent (group 4) significantly decreased (~15%) the population of $CD4^+$ T lymphocytes when compared to the Control (group 1) and L (group 2) groups (p=0.02 and p=0.04, respectively). The combined therapy (group 5) showed an even further reduction 45%; (p=0.04 vs. control (group 1), p=0.03 vs. lesion without treatment (group 2), p=0.02 vs. LU (group 3) and p=0.02 vs. LM (group 4)). $CD8^+$ T lymphocytes in the lesion group (group 2) was reduced by 22% (p=0.02 vs. control group (group 1)). The different therapies (groups 3 & 4) and their combination (group 5) recovered the CD8+ T lymphocyte population to control levels (p>0.05 vs. control group (group 1)). The $CD4^+/CD8^+$ ratio increased in the lesion group (group 2) (p=0.03 vs. control group (group 1)). The different therapies (groups 3 & 4) and their combination (group 5) reversed the effects caused by the excision lesion to control level (p=0.04 vs. LU (group 3), p=0.03 vs. LM (group 4), p=0.02 vs. LUM (group 5)).

Discussion

In this study, the effect of HFU, microcurrent, and combined therapies on the immune system and healing process in a wound-induced excision model was investigated. Contrary to the findings of this

study, other studies have shown acceleration of the wound-healing process with HFU[8,10,20] and MET[12,16]; however, this previous research did not explore the immune system as was done in the present study. Inflammation is a key factor in the healing process, followed by cell proliferation and maturation[2-4]. Our results showed that the two different physical therapy modalities were able to significantly modulate macrophages immune parameters, decrease the expression of pro-inflammatory cytokines, and decrease the expression of $CD4^+$-positive cells in association with a reduced $CD4^+/CD8^+$ ratio. In addition, the combination of both modalities (combined therapy) further decreased the expression of $CD4^+$-positive cells and the $CD4^+/CD8^+$ ratio. These changes, which were brought about by the combined therapy, suggest that the rats immune system was attempting to solve the healing process. However, although the different approaches resulted in a significant reduction in the wound area, their use for this purpose was not supported, as there was no statistically significant difference compared to untreated excised lesions.

Because these therapies were able to change the immune parameters, helping to modulate the healing process, and given the fine line of percentage change between the different interventions, a change in wound area would also be expected. However, the present study did not detect such a result. This result might have been affected by the dosage parameters of the modalities used, although based on previous evidence[8,13,17], the dosage used in this study was not enough to promote the acceleration of wound healing. It must be pointed out that in the present study, only a specific dosages were evaluated; thus, the results cannot be extrapolated to conditions in which other parameters and/or different forms of electrical stimulation may be used. Further research is needed to determine the ideal dose-response treatment for the different stages of wound healing.

It has been shown[24] that different doses of HFU and microcurrents increase the tensile strength of tissue being repaired, which is important for tissue quality. As a matter of fact, the cytokine network is important for tissue quality[2,4]; the present study showed that different physical therapy modalities are able to alter the plasma concentrations of ILs.

HFU reduced the concentration of IL-1β (known as pro-inflammatory) along with an increase in IL-6 (in the early phase of inflammation, this is a chemoattractant cytokine for keratinocytes) 3 days after the lesion was generated. This was seen with both modalities for IL-1β concentration. The concentrations of IL-6 and TNF-α increased markedly and IL-1β and IL-10 were observed in lower amounts[24]. The methods used in this study somehow contributed to the delicate balance between these cytokines, attenuating the inflammatory response because IL-1β is a potent inducer of the transduction signaling cascade for growth factors involved in the migration of inflammatory cells and the production of prostaglandins for nociception[25].

An in vitro[26] study showed that HFU therapy stimulated the release of IL-1β at a low rate due to three possible mechanisms: an increase in cell membrane permeability; changes in the signal transduction that regulated gene expression; and alterations in the cytoskeleton, affecting cell metabolism and gene expression. However, the precise mechanism is still unknown. The reduction of IL-1β by the microcurrent therapy may have lead to the possible involvement of nuclear factor κB (NFκB) and mitogen-activated protein kinase (MAPK)[27].

Fourteen days after wound-induced excision, HFU increased the phagocytosis capacity and the NO production in the presence of the challenger LPS. The authors are not aware of any study that investigated the innate immune system following physical therapy treatment and the authors hypothesize that this increase was due to the capacity of the ultrasound therapy to induce a stable cavitation related to the transitory increment of cell permeability to calcium influx[7,11]. Along with increased phagocytic capacity, HFU also increased NO production in the presence of LPS at the same level of control group. This increase may have been caused by higher intracellular calcium concentration[28,29], revealing its antioxidant potential and resulted in low levels of the other two redox molecules (i.e. superoxide anion and H_2O_2). Indeed, the formation of redox by ultrasound is related to the sonolysis of water generated by the cavitation[11]; however, the acoustic parameters used in the present experiment were insufficient to induce their formation, perhaps because ultrasonic frequencies up to 2 MHz were not able to sonolyze water[30].

Interestingly, the microcurrent decreased superoxide anion production concurrently with increased production of NO in the presence of LPS, confirming its antimicrobial power, but with lower expression when compared to the lesion without treatment group (group 2). An excess of redox molecules can damage tissues and amplify the pro-inflammatory response, perhaps leading to a chronic stage[31]. Microcurrent therapy increases the adenosine triphosphate (ATP)

concentration, where ATP acts as an antioxidant to stabilize mitochondrial function[32]. In fact, low-grade NO production at the end of the healing process restores collagen concentration to physiological levels, but the mechanisms for this are not known[33]. Curiously, combined therapy reduced H_2O_2, superoxide anion, and NO production. NO production has been proposed as a common mechanism for both therapeutic approaches[34], however, this still needs to be proven. The present data does not support this suggestion. The authors think that different energy forms–whether mechanical or electrical–trigger a sequence of events to stimulate or inhibit cell processes which leads to the wound-healing process.

Finally, both modalities alone and in combination reduced the $CD4^+/CD8^+$ ratio, suggesting the resolution of the inflammatory process[5]. The $CD8^+$ T cell population was maintained at control levels, resulting in better tissue quality, although this was only applicable to microcurrent therapy because $CD4^+$ cells are considered up-regulatory, leading to the formation of fragile scar tissue and $CD8^+$ down-regulation, which makes the scar tissue more denserigid[5]. In addition, $CD8^+$ cells regulate the production of cytokines through $CD4^+$, and consequently, result in a lower level of cytokines[5]. Different organisms with high regenerative capacity (e.g. zebrafish, salamander, and human fetuses) have lower power for stimuli to activate the signaling cascade of the healing process and present with an absence of inflammatory cells at the site of injury, which might be a prerequisite for better repair and tissue quality and/or complete regeneration[3]. This suggests that the modalities used in this study could be useful in modulating the immune system, thereby helping in the formation of a new tissue.

In future studies, the quality of scar tissue should also be measured and local biomarkers should be investigated. In the present study, only circulating markers were measured because the aim of the study was to look at the healing phenomenon from a systemic perspective. The present work sheds light on the therapeutic approaches and the effect of two modalities used alone or together on the immune system. An environment of excessive inflammation may lead to inadequate healing, thereby delaying the healing process and increasing the amount of scar tissue. This means that it might be feasible to manipulate the immune response in the healing process, especially if the patient was willing to sacrifice the wound area reduction rate for a better esthetic result. The physical therapy modalities studied here could contribute to such applications.

Conclusions

High frequency ultrasound, microcurrent and combining the two modalities were able to modulate the activity of the innate and adaptive immune system, improving the inflammatory environment but not accelerating the wound-healing process. More studies are needed to understand the mechanisms involved in the modulation of wound healing using different physical agents, as well as the relation of these different physical modalities with the quality of the newly formed tissue.

References

1. Ennis WJ, Sui A, Bartholomew A. Stem cells and healing: impact on inflammation. Adv Wound Care. 2013;2(7):369-78. http://dx.doi.org/10.1089/wound.2013.0449. PMid:24587974.
2. Janis JE, Harrison B. Wound healing: part I. basic science. Plast Reconstr Surg. 2014;133(2):199e-207e. http://dx.doi.org/10.1097/01.prs.0000437224.02985.f9. PMid:24469191.
3. Eming SA, Hammerschmidt M, Krieg T, Roers A. Interrelation of immunity and tissue repair or regeneration. Semin Cell Dev Biol. 2009;20(5):517-27. http://dx.doi.org/10.1016/j.semcdb.2009.04.009. PMid:19393325.
4. Strbo N, Yin N, Stojadinovic O. Innate and adaptive immune responses in wound epithelialization. Adv Wound Care. 2014;3(7):492-501. http://dx.doi.org/10.1089/wound.2012.0435. PMid:25032069.
5. Boyce DE, Jones WD, Ruge F, Harding KG, Moore K. The role of lymphocytes in human dermal wound healing. Br J Dermatol. 2000;143(1):59-65. http://dx.doi.org/10.1046/j.1365-2133.2000.03591.x. PMid:10886136.
6. Leaper DJ, Schultz G, Carville K, Fletcher J, Swanson T, Drake R. Extending the TIME concept: what have we learned in the past 10 years? Int Wound J. 2012;9(Suppl 2):1-19. http://dx.doi.org/10.1111/j.1742-481X.2012.01097.x. PMid:23145905.
7. Ennis WJ, Lee C, Plummer M, Meneses P. Current status of the use of modalities in wound care: electrical stimulation and ultrasound therapy. Plast Reconstr Surg. 2011;127(Suppl 1):93S-102S. http://dx.doi.org/10.1097/PRS.0b013e3181fbe2fd. PMid:21200278.
8. Taradaj J, Franek A, Brzezinska-Wcislo L, Cierpka L, Dolibog P, Chmielewska D, et al. The use of therapeutic ultrasound in venous leg ulcers: a randomized, controlled clinical trial. Phlebology. 2008;23(4):178-83. http://dx.doi.org/10.1258/phleb.2008.008015. PMid:18663117.
9. Cullum N, Al-Kurdi D, Bell-Syer SE. Therapeutic ultrasound for venous leg ulcers. Cochrane Database Syst Rev. 2010;6(6):CD001180. http://dx.doi.org/10.1002/14651858.CD001180.pub3. PMid:20556749.
10. Polak A, Franek A, Blaszczak E, Nawrat-Szoltysik A, Taradaj J, Wiercigroch L, et al. A prospective, randomized, controlled, clinical study to evaluate the efficacy of high-frequency ultrasound in the treatment of stage II and stage III pressure ulcers in Geriatric patients. Ostomy Wound Manage. 2014;60(8):16-28. PMid:25105475.

11. Furusawa Y, Hassan MA, Zhao QL, Ogawa R, Tabuchi Y, Kondo T. Effects of therapeutic ultrasound on the nucleus and genomic DNA. Ultrason Sonochem. 2014;21(6):2061-8. http://dx.doi.org/10.1016/j.ultsonch.2014.02.028. PMid:24657073.
12. Bayat M, Asgari-Moghadam Z, Maroufi M, Rezaie FS, Bayat M, Rakhshan M. Experimental wound healing using microamperage electrical stimulation in rabbits. J Rehabil Res Dev. 2006;43(2):219-26. http://dx.doi.org/10.1682/JRRD.2005.05.0089. PMid:16847788.
13. Balakatounis KC, Angoules AG. Low-intensity electrical stimulation in wound healing: review of the efficacy of eternally applied currents resembling the current of injury. Eplasty. 2008;8:283-91. PMid:18552975.
14. Poltawski L, Watson T. Bioelectricity and microcurrent therapy for tissue healing: a narrative review. Phys Ther Rev. 2009;14(2):104-14. http://dx.doi.org/10.1179/174328809X405973.
15. Korelo RIG, Valderramas S, Ternoski B, Medeiros DS, Andres LF, Adolph SMM. Microcurrent application as analgesic treatment in venous ulcer: a pilot study. Rev Lat Am Enfermagem. 2012;20(4):753-60. http://dx.doi.org/10.1590/S0104-11692012000400016. PMid:22990161.
16. Thakral G, LaFontaine J, Najafi B, Talal TK, Kim P, Lavery LA. Electrical stimulation to accelerate wound healing. Diabet Foot Ankle. 2013;4(0). http://dx.doi.org/10.3402/dfa.v4i0.22081. PMid:24049559.
17. Bélanger A-Y. Therapeutic electrophysical agents: evidence behing practice. 3rd ed. Baltimore: Lippincott Williams & Wilkins; 2014.
18. Torkaman G. Electrical stimulation of wound healing: a review of animals experimental evidence. Adv Wound Care (New Rochelle). 2014;3(2):202-18. http://dx.doi.org/10.1089/wound.2012.0409. PMid:24761359.
19. Moretti FA, Marcondes FB, Provenza JR, Fukuda TY, Vasconcelos RA, Roizenblatt S. Combined therapy (ultrasound and interferential current) in patients with fibromyalgia: once or twice in a week? Physiother Res Int. 2012;17(3):142-9. http://dx.doi.org/10.1002/pri.525. PMid:22114059.
20. Altomare M, Nascimento AP, Romana-Souza B, Amadeu TP, Monte-Alto-Costa A. Ultrasound accelerates healing of normal wounds but not of ischemic ones. Wound Repair Regen. 2009;17(6):825-31. http://dx.doi.org/10.1111/j.1524-475X.2009.00542.x. PMid:19821959.
21. Rijswijk L. Measuring wounds to improve outcomes. Am J Nurs. 2013;113(8):60-1. http://dx.doi.org/10.1097/01.NAJ.0000432967.20419.1b. PMid:23884002.
22. Bonatto SJR, Folador A, Aikawa J, Yamazaki RK, Pizatto N, Oliveira HH, et al. Lifelong exposure to dietary fish oil alters macrophage responses in Walker 256 tumor-bearing rats. Cell Immunol. 2004;231(1-2):56-62. http://dx.doi.org/10.1016/j.cellimm.2004.12.001. PMid:15919370.
23. Jeffery NM, Newsholme EA, Calder PC. Level of polyunsaturated fatty acids and the n-6 to n-3 polyunsaturated fatty acid ratio in the rat diet alter serum lipid levels and lymphocyte functions. Prostaglandins Leukot Essent Fatty Acids. 1997;57(2):149-60. http://dx.doi.org/10.1016/S0952-3278(97)90005-3. PMid:9250698.
24. Ng GY. Comparing therapeutic ultrasound with microamperage stimulation therapy for improving the strength of Achilles tendon repair. Connect Tissue Res. 2011;52(3):178-82. http://dx.doi.org/10.3109/03008207.2010.500752. PMid:20672987.
25. Grimstad Ø, Sandanger Ø, Ryan L, Otterdal K, Damaas JK, Pukstad B, et al. Cellular sources and inducers of cytokines presents in acute wound fluid. Wound Repair Regen. 2011;19(3):337-47. http://dx.doi.org/10.1111/j.1524-475X.2011.00668.x. PMid:21371163.
26. Doan N, Reher P, Meghji S, Harris M. In vitro effects of therapeutic ultrasound on cell proliferation, protein synthesis, and cytokine production by human fibroblasts, osteoblasts, and monocytes. J Oral Maxillofac Surg. 1999;57(4):409-19, discussion 420. http://dx.doi.org/10.1016/S0278-2391(99)90281-1. PMid:10199493.
27. Kim JH, Choi H, Suh MJ, Shin JH, Hwang MH, Lee HM. Effect of biphasic electrical current stimulation on IL-1β stimulated annulus fibrosus cells using *in vitro* microcurrent generating chamber system. Spine (Phila Pa 1976). 2013;38(22):E1368-76. http://dx.doi.org/10.1097/BRS.0b013e3182a211e3. PMid:23823576.
28. Wink DA, Hines HB, Cheng RY, Switzer CH, Flores-Santana W, Vitek MP, et al. Nitric oxide and redox mechanisms in the immune response. J Leukoc Biol. 2011;89(6):873-91. http://dx.doi.org/10.1189/jlb.1010550. PMid:21233414.
29. Bernatchez SF, Menon V, Stoffel J, Walters SA, Lindroos WE, Crossland MC, et al. Nitric oxide levels in wound fluid may reflect the healing trajectory. Wound Repair Regen. 2013;21(3):410-7. http://dx.doi.org/10.1111/wrr.12048. PMid:23627618.
30. Lifka J, Ondruschka B, Hofmann J. The use of ultrasound for the degradation of pollutants in water: aquasononolysis: a review. Eng Life Sci. 2003;3(6):253-62. http://dx.doi.org/10.1002/elsc.200390040.
31. Guo S, Dipietro LA. Factors affecting wound healing. J Dent Res. 2010;89(3):219-29. http://dx.doi.org/10.1177/0022034509359125. PMid:20139336.
32. Lee BY, Al-Waili N, Stubbs D, Wendell K, Butler G, Al-Waili T, et al. Ultra-low microcurrent in the management of diabetes mellitus, hypertension on chronica wounds: report of twelve cases and discussion of mechanism of action. Int J Med Sci. 2009;7(1):29-35. http://dx.doi.org/10.7150/ijms.7.29. PMid:20046232.
33. Park JE, Abrams MJ, Efron PA, Barbul A. Excessive nitric oxide impairs wound collagen accumulation. J Surg Res. 2013;183(1):487-92. http://dx.doi.org/10.1016/j.jss.2012.11.056. PMid:23290597.
34. Ennis WJ, Lee C, Meneses P. A biochemical approach to wound healing through the use of modalities. Clin Dermatol. 2007;25(1):63-72. http://dx.doi.org/10.1016/j.clindermatol.2006.09.008. PMid:17276203.

Correspondence
Raciele I. Guarda Korelo
Universidade Federal do Paraná
Coordenação do Curso de Fisioterapia
Avenida Coronel Francisco Heráclito dos Santos, 210
CEP 81531-980, Curitiba, PR, Brazil
e-mail: raciele@ufpr.br

original article

BJPT Brazilian Journal of Physical Therapy

Test-retest, inter- and intra-rater reliability of the flexicurve for evaluation of the spine in children

Juliana A. Sedrez[1], Cláudia T. Candotti[1], Maria I. Z. Rosa[2], Fernanda S. Medeiros[2], Mariana T. Marques[1], Jefferson F. Loss[1]

ABSTRACT | Introduction: The early evaluation of the spine in children is desirable because it is at this stage of development that the greatest changes in the body structures occur. **Objective:** To determine the test-retest, intra- and inter-rater reliability of the Flexicurve instrument for the evaluation of spinal curvatures in children. **Method:** Forty children ranging from 5 to 15 years of age were evaluated by two independent evaluators using the Flexicurve to model the spine. The agreement was evaluated using Intraclass Correlation Coefficients (ICC), Standard Error of the Measurement (SEM), and Minimal Detectable Change (MDC). **Results:** In relation to thoracic kyphosis, the Flexicurve was shown to have excellent correlation in terms of test-retest reliability ($ICC_{2,2}$=0.87) and moderate correlation in terms of intra-($ICC_{2,2}$=0.68) and inter-rater reliability ($ICC_{2,2}$=0.72). In relation to lumbar lordosis, it was shown to have moderate correlation in terms of test-retest reliability ($ICC_{2,2}$=0.66) and intra- ($ICC_{2,2}$=0.50) and inter-rater reliability (ICC=0.56). **Conclusion:** This evaluation of the reliability of the Flexicurve allows its use in school screening. However, to monitor spinal curvatures in the sagittal plane in children, complementary clinical measures are necessary. Further studies are required to investigate the concurrent validity of the instrument in order to identify its diagnostic capacity.

Keywords: child; spinal curvatures; reliability studies; physical therapy; rehabilitation.

BULLET POINTS

- The Flexicurve has test-retest, intra- and inter-rater reliability confirmed.
- The Flexicurve can be used for evaluating spinal curvatures in children.
- The Flexicurve can be used in school screening to detect postural alterations.

HOW TO CITE THIS ARTICLE

Sedrez JA, Candotti CT, Rosa MIZ, Medeiros FS, Marques MT, Loss JF. Test-retest, inter- and intra-rater reliability of the flexicurve for evaluation of the spine in children. Braz J Phys Ther. 2016 Mar-Apr; 20(2):142-147. http://dx.doi.org/10.1590/bjpt-rbf.2014.0139

Introduction

Spine evaluation is essential both for monitoring[1] and diagnosing vertebral alterations, with radiography being the most appropriate method for both processes[2]. However, radiography entails exposure to undesirable radiation levels, so non-invasive methods are of great benefit[3] because of their lower cost, fewer technical difficulties, and the absence of exposure to ionizing radiation[4].

Among the non-invasive methods is the Flexicurve, a flexible ruler that was first described by Takahashi and Atsumi[5]. The Flexicurve allows measurements in the sagittal plane and can be used in several surroundings[6]. The psychometric properties of this instrument have been described for use with adults[7], and it is seen as a low-cost and quick evaluation instrument[8].

School-age children have a significant prevalence of postural imbalances[9], and the early detection of postural changes can be important. The Flexicurve instrument can be a screening tool because it is easily accessible for the school environment. This is even more important when considering that schoolchildren are likely to exhibit poor daily posture[10], modifying it over the years. In other words, in seeking body balance, students' posture adapts to their lifestyle practices, and proper or improper posture habits lead to repercussions in adulthood[11].

[1] Programa de Pós-graduação em Ciência do Movimento Humano, Escola de Educação Física, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil
[2] Departamento de Fisioterapia, Escola de Educação Física, UFRGS, Porto Alegre, RS, Brazil
Received: Jan. 15, 2015 Revised: July 27, 2015 Accepted: Oct. 13, 2015

However, the use of any alternative postural evaluation instrument requires an evaluation of its psychometric properties[12]. Although the Flexicurve instrument is known for its use in the adult population[7], it is necessary to verify the instrument's test-retest, intra-and inter-rater reliability prior to its use with children, because this population has distinct characteristics, such as thorax size and the sagittal curvatures of the spine. Therefore, the present study aimed to determine the test-retest, intra- and inter-rater reliability of the Flexicurve instrument for the evaluation of spinal curvatures in children.

Method

Sample

The sample size was calculated according to Walter et al.[13] and Donner and Eliasziw[14], assuming: the null hypothesis value of Intraclass Correlation Coefficient (ICC) to be 0.40 (e.g. on the basis that any value lower than .40 might be considered clinically "unacceptable"); 80% of power; two replicated measurements (one for each evaluator or twice by the same evaluator); and a significance level of 95% to detect an ICC value of .70 (based on previous literature[7]), a minimum of 33 participants was found. Allowing for losses, 40 children who had undergone X-ray examination in a hospital in Porto Alegre were invited to participate in the study. The inclusion criteria were children of both sexes, ranging in age from 5 to 15 years old. Children who had previous surgery or congenital deformity in spinal structures were excluded. This study was approved by the Research Ethics Committee of Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil (number 19685), and the children's guardians signed the informed consent form.

Evaluation protocol

The evaluation consisted of modeling the spine with the Flexicurve instrument, which provided the Flexicurve angles (FA) for the thoracic and lumbar spine. The same protocol was repeated in each evaluation: (1) on the child's bare back, the spinous processes (SP) of the C7, T1, T12, L1, L5, and S1 vertebrae were palpated and marked with stickers; (2) the child was in the standing position with normal posture; (3) elbows and shoulders were flexed to 90° and supported on the wall; (4) while remaining motionless; (5) the Flexicurve was molded to the child's back over the spine; (6) the Flexicurve was removed from the child's back and placed on graph paper, where the curvature was drawn and the SP marked; and (7) the FA was obtained using Biomec-FLEX free software (www.ufrgs.br/biomec), in which the input data consisted of the coordinate values representing the thoracic and lumbar curvatures, and the output data consisted of the curvature angles in the sagittal plane. The procedures (steps 1 to 7) were performed in accordance with the literature[7].

Design procedures

The spine postural evaluations using the Flexicurve instrument were performed by two previously trained independent evaluators (Ev1 and Ev2), with each subject being evaluated four times in two days. On the first day, there were two successive evaluations (Measure 1 and 2) by the same evaluator (Ev1) and a third evaluation (Measure 3) by a second evaluator (Ev2). After a seven-day interval, the children were re-evaluated (Measure 4) by one evaluator (Ev1). Both evaluators had at least two years' experience with postural evaluation of the spine and received 20 hours training in the use of the Flexicurve, which consisted of palpation, molding, transfer to paper, and analysis using the software.

For the test-retest reliability evaluation, data from the first evaluation (Measure 1) and from the second evaluation (Measure 2), performed successively by the same evaluator (Ev1), were used[15]. For the intra-rater reliability evaluation, data from the first evaluation (Measure 1) and from the evaluation performed by the same examiner (Ev1) seven days after (Measure 4) were used[15]. For the inter-rater reliability evaluation data from the first evaluation by Ev1 (Measure 1) and from the evaluation performed on the same day by Ev2 (Measure 3) were used[15].

Statistical analysis

The statistical analysis was performed with SPSS 17.0. Initially, a data descriptive analysis was carried out using descriptive statistics. The data normality was confirmed using the Shapiro-Wilk test. To verify the test-retest, intra- and inter-rater reliability, the Intra-Class Coefficient ($ICC_{2,2}$), the Standard Error Measurement (SEM), and Minimum Detectable Change (MDC) were calculated. $ICC_{2,2}$ was based on a 2-way (random effects) repeated-measures analysis of variance model with absolute agreement. The values found in the ICC were classified according to literature[16] as weak (ICC<0.40), moderate (ICC between 0.40 and 0.75), and excellent (ICC>0.75). The Standard Error of the Measurement (SEM) was estimated using the following

formula: SEM = SD $\sqrt{1-ICC_{2,2}}$, where SD is the standard deviation of the measurements. The Minimum Detectable Change (MDC) was estimated based on a 95% confidence interval, where MDC=1.96 * SEM. The level of significance adopted for all tests was 0.05.

Results

Twenty-five (25) boys and 15 girls were evaluated (Table 1). The results for the test-retest reliability of thoracic kyphosis and lumbar lordosis angles expressed by ICC values were excellent, with SEM values less than 4.5° and MDC values less than 8.5° (Table 2).

For the evaluation of intra-rater reliability, the results obtained by the ICC showed excellent and moderate levels for the angles of thoracic kyphosis and lumbar lordosis, respectively, with SEM values less than 6.0° and MDC values less than 11.5° (Table 2).

For the evaluation of inter-rater reliability, the results obtained by the ICC demonstrated excellent and moderate levels for the thoracic kyphosis and lumbar lordosis angles, respectively, with SEM values less than 6.0° and MDC values less than 11.5° (Table 2).

Discussion

This study aimed to determine the test-retest, intra-and inter-rater reliability of the Flexicurve instrument for evaluating spinal curvatures in children. Acceptable correlation levels were found for the angles of thoracic kyphosis and lumbar lordosis. These results differ from those of Teixeira and Carvalho[6], which showed only excellent levels of both inter- (ICC=0.94) and intra-rater reliability (ICC=0.87) for thoracic kyphosis in an adult population, and those of Oliveira et al.[7], which also showed excellent levels of inter-rater reliability (ICC=0.94 for thoracic kyphosis; ICC=0.83 for lumbar lordosis) and intra-rater reliability (ICC=0.83 for thoracic kyphosis; ICC=0.78 for lumbar lordosis) in adults. Both studies used the Flexicurve instrument.

However, Letafatkar et al.[17] evaluated the lumbar region with the Flexicurve instrument and found results

Table 1. Anthropometric data of the sample (mean±SD).

Sample	Age (years)	Body mass (kg)	Height (cm)	BMI (kg/cm²)
Total (n=40)	10.2±2.8	39.3±12.9	1.4±0.2	19.5±2.8
Boys (n=25)	11.0±2.3	42.9±12.0	1.4±0.1	20.0±2.5
Girls (n=15)	9.1±3.3	33.5±12.5	1.3±0.2	18.7±3.2

BMI: Body mass index; SD: standard deviation.

Table 2. Results for test-retest, intra- and inter-rater reliability.

		Mean±SD(°)	$ICC_{2,2}$ (95% CI)	*p* (ICC)	SEM(°)	MDC(°)
Test-retest reliability						
Kyphosis (n=40)	**Measure 1**	37.5±9.3	0.93 (0.87-0.96)	<0.01	2.5	4.9
	Measure 2	37.5±9.5				
Lordosis (n=40)	**Measure 1**	26.0±9.5	0.80 (0.61-0.89)	<0.01	4.3	8.4
	Measure 2	26.2±9.5				
Intra-rater reliability						
Kyphosis (n=38)	**Measure 1**	36.0±9.9	0.82 (0.65-0.91)	<0.01	4.1	8.1
	Measure 4	36.4±9.5				
Lordosis (n=38)	**Measure 1**	24.8±9.5	0.67 (0.36-0.83)	<0.01	5.7	11.2
	Measure 4	26.0±10.4				
Inter-rater reliability						
Kyphosis (n=40)	**Measure 1**	35.7±9.0	0.83 (0.68-0.91)	<0.01	4.1	8.0
	Measure 3	37.6±10.8				
Lordosis (n=40)	**Measure 1**	25.2±9.5	0.72 (0.47-0.85)	<0.01	5.7	11.2
	Measure 3	27.2±12.0				

ICC: Intraclass Correlation Coefficient; SEM: Standard Error Measurement; MDC: Minimum Detectable Change; SD: standard deviation.

that corroborate those of the present study, showing moderate intra-rater reliability (ICC ranging from 0.62 to 0.69) and inter-rater reliability (ICC=0.54). Lovell et al.[18] also evaluated the lumbar region and found that the intra-rater reliability ranged from moderate to excellent (ICC 0.73 to 0.94) in addition to moderate inter-rater reliability, with ICC values of 0.41 and 0.50, which suggested that the evaluation of the lumbar region in adults with the Flexicurve may be viable if performed by the same person, but the degree of reproducibility may vary from evaluator to evaluator.

Dunleavy et al.[19], who also investigated the inter-and intra-evaluator reliability of the Flexicurve using variable spine length and width, found that the measurements of the total length of the spine showed good intra-rater reliability (ICC=0.93), but the evaluation of the thoracic length, lumbar length, chest width, and lumbar width showed moderate intra-rater reliability (ICC=0.61-0.80). In addition, the inter-rater reliability for all measures was moderate (ICC=0.58 to 0.72), and the mean lengths indicated significant differences among the evaluators.

It is worth noting that, in all the evaluations of test-retest, intra- and inter-rater reliability of this study, the correlations were always lower in the lumbar region than in the thoracic region, demonstrating an inherent difficulty in evaluating this region. Previous studies have demonstrated the difficulty in evaluating the lumbar region. For example, Hinman[20] evaluated the inter-rater reliability of the Flexicurve instrument by novice evaluators and found excellent correlation levels for thoracic kyphosis indices (ICC 0.93 and 0.94) and lower levels of correlation for lumbar lordosis (ICC of 0.60 and 0.73). The author noted that the difficulty in molding the Flexicurve to regions of smaller curvature or even concave features of the lumbar spine might have caused the greatest variability in lumbar measurements. Thus, Hinman's results[20] are in accordance with the present study in that there is difficulty when evaluating smaller curvatures with the Flexicurve, particularly in the lower back. Another difficulty in evaluating the lumbar region is related to the palpation of anatomical landmarks in this region since the characteristics of the lumbar vertebrae hamper the identification and location of the spinal process[21].

In the literature, several non-invasive instruments for exclusively evaluating thoracic kyphosis are described, a fact that is probably related to difficulties in the non-invasive evaluation of lumbar lordosis. For example, we can cite studies such as the one by Perriman et al.[22], which verified the concurrent validity of the flexible electrogoniometer; the study by D'Osualdo et al.[23], which validated the arcometer; and the study by Lewis and Valentine[24], which determined the test-retest reliability of the inclinometer. All of these studies exclusively evaluated thoracic kyphosis.

In addition to the adequate levels of test-retest, intra- and inter-rater reliability obtained in this study, it is important to point out the variability inherent in the measurement in order to facilitate the correct interpretation of the results during clinical follow-up. Thus, the SEM values reflect the precision of the measurement, which in this study vary from 2.5° to 5.7°, depending on the region and the analysis conducted. These values can be considered clinically acceptable, since the variability found in the gold standard technique for measuring spinal curvature, the Cobb angle, is from 5° to 10°, for both intra- and inter-rater reliability[25]. Furthermore, it is also important to know what magnitude of the change in the measurement would be necessary to determine the existence of a real change rather than a mere measurement error[26]. Based on MDC values, it can be concluded that, when Flexicurve is used by the same evaluator or different evaluators during follow-up procedures in children, there needs to be a minimum of 8° to be considered a real change in the thoracic curvature and 11° in the lumbar curvature (Table 2).

MDC values around 10° suggest the instrument has poor sensitivity, which may be considered as a limitation for its use. On the other hand, it is important to point out that the SEM and MDC values are dependent on the variability among the subjects in the sample, which in this study was between 14° to 57° for thoracic curvature and 6° to 46° for lumbar curvature. Hence, in this study, the SEM and MDC values do not necessarily reflect the reliability of the measurements, but rather the variability of the sample. Nevertheless, restricted variability among the sample could also represent a limitation. Only two children in the sample presented increased thoracic curvature (over 50°)[27,28], while none were found to have increased lumbar curvature (over 66.8°)[29], which restricts the use of the Flexicurve in children in this range.

In summary, the test-retest, intra- and inter-rater reliability of the results presented in this study show that the Flexicurve instrument can be used in the initial evaluation of the spinal curvatures in children. However, for an instrument to be used for the purposes of diagnosis, it should be subjected to a validation

process, and most studies on instrument validation are conducted with adult subjects, not children. Therefore, the Flexicurve still lacks concurrent validation in relation to gold standard technique for the detection of spinal alteration in order to ensure its diagnostic capacity.

Conclusion

The Flexicurve evaluation method has test-retest, intra- and inter-rater reliability for the population of children between 5 and 15 years of age, which would allow its use in school screening to detect postural alterations at an early stage. However, to monitor spinal curvatures in the sagittal plane in children, complementary clinical measures are necessary.

References

1. Vrtovec T, Pernus F, Likar B. A review of methods for quantitative evaluation of spinal curvature. Eur Spine J. 2009;18(5):593-607. http://dx.doi.org/10.1007/s00586-009-0913-0. PMid:19247697.
2. Jefferson RJ, Weisz I, Turner-Smith AR, Harris JD, Houghton GR. Scoliosis surgery and its effect on back shape. J Bone Joint Surg Am. 1988;70(2):261-6. PMid:3346300.
3. Turner-Smith AR, Harris JD, Houghton GR, Jefferson RJ. A method for analysis of back shape in scoliosis. J Biomech. 1988;21(6):497-509. http://dx.doi.org/10.1016/0021-9290(88)90242-4. PMid:3209594.
4. Chen Y-L, Lee Y-H. A non-invasive protocol for the determination of lumbosacral vertebral angle. Clin Biomech. 1997;12(3):185-9. http://dx.doi.org/10.1016/S0268-0033(97)00076-4. PMid:11415692.
5. Takahashi E, Atsumi H. Age differences in thoracic form as indicated by thoracic index. Hum Biol. 1955;27(2):65-7. PMid:14381054.
6. Teixeira FA, Carvalho GA. Reliability and validity of thoracic kyphosis measurements using the Flexicurve method. Rev Bras Fisioter. 2007;11(3):199-204. http://dx.doi.org/10.1590/S1413-35552007000300005.
7. Oliveira TS, Candotti CT, La Torre M, Pelinson PPT, Furlanetto TS, Kutchak FM, et al. Validity and reproducibility of the measurements obtained using the Flexicurve instrument to evaluate the angles of thoracic and lumbar curvatures of the spine in the sagittal plane. Rehabil Res Pract. 2012;2012:186156. http://dx.doi.org/10.1155/2012/186156. PMid:22619723.
8. Hinman MR. Comparison of thoracic kyphosis and postural stiffness in younger and older women. Spine (Phila Pa 1976). 2004;4(4):413-7. http://dx.doi.org/10.1016/j.spinee.2004.01.002. PMid:15246302.
9. Vasconcelos GAR, Fernandes PRB, Oliveira DA, Cabral ED, Silva LVC. Avaliação postural da coluna vertebral em escolares surdos de 7-21 anos. Fisioter Mov. 2010;23(3):371-80. http://dx.doi.org/10.1590/S0103-51502010000300004.
10. Rosa FN. Avaliação postural em escolares de 1ª a 4ª série do 1º Grau. Revista Brasileira de Ciência e Movimento. 1991;5(2):7-10.
11. Ferronatto A, Candotti CL, Silveira RP. A incidência de alterações do equilíbrio estático da cintura escapular em crianças entre 7 a 14 anos. Movimento. 1998;5(9):24-30.
12. Thomas JR, Nelson JK, Silverman AJ. Research methods in physical activity. 5th ed. Champaign: Human Kinetics; 2005.
13. Walter DS, Eliasziw M, Donner A. Sample size and optimal design for reliability studies. Stat Med. 1998;17(1):101-10. http://dx.doi.org/10.1002/(SICI)1097-0258(19980115)17:1<101::AID-SIM727>3.0.CO;2-E. PMid:9463853.
14. Donner A, Eliasziw M. Sample size requirements for reliability studies. Stat Med. 1987;6(4):441-8. http://dx.doi.org/10.1002/sim.4780060404. PMid:3629046.
15. Joint Committee for Guides in Metrology – JCGM. JCGM 200:2008. International vocabulary of metrology: basic and general concepts and associated terms (VIM). France: Working Group 2; 2008.
16. Fleiss RL. The design and analysis of clinical experiments. New York: John Wiley & Sons; 1986.
17. Letafatkar A, Amirsasan R, Abdolvahabi Z, Hadadnezhad M. Reliability and validity of the AutoCAD software method in lumbar lordosis measurement. J Chiropr Med. 2011;10(4):240-7. http://dx.doi.org/10.1016/j.jcm.2011.02.003. PMid:22654681.
18. Lovell FW, Rothstein JM, Personius WJ. Reliability of clinical measurements of lumbar lordosis taken with a flexible rule. Phys Ther. 1989;69(2):96-102. PMid:2521531.
19. Dunleavy K, Mariano H, Wiater T, Goldberg A. Reliability and minimal detectable change of spinal length and width measurements using the Flexicurve for usual standing posture in healthy young adults. J Back Musculoskeletal Rehabil. 2010;23(4):209-14. PMid:21079300.
20. Hinman MR. Interrater reliability of Flexicurve postural measures among novice users. J Back Musculoskeletal Rehabil. 2004;17(1):33-6.
21. Whittle MW, Levine D. Measurement of lumbar lordosis as a component of clinical gait analysis. Gait Posture. 1997;5(2):101-7. http://dx.doi.org/10.1016/S0966-6362(96)01079-X.
22. Perriman DM, Scarvell JM, Hughes AR, Ashman B, Lueck CJ, Smith PN. Validation of the flexible electrogoniometer for measuring thoracic kyphosis. Spine. 2010;35(14):E633-40. PMid:20505565.
23. D'Osualdo F, Schierano S, Iannis M. Validation of clinical measurement of kyphosis with a simple instrument, the arcometer. Spine (Phila Pa 1976). 1997;22(4):408-13. http://dx.doi.org/10.1097/00007632-199702150-00011. PMid:9055369.
24. Lewis JS, Valentine RE. Clinical measurement of the thoracic kyphosis. A study of the intra-rater reliability in subjects with and without shoulder pain. BMC Musculoskelet Disord. 2010;11(39):1-7. PMid:20193055.
25. Harrison DE, Harrison DD, Cailliet R, Janik TJ, Holland B. Radiographic analysis of lumbar lordosis: centroid, Cobb, TRALL, and Harrison posterior tangent methods. Spine. 2001;26(11):E235-42. http://dx.doi.org/10.1097/00007632-200106010-00003. PMid:11389407.
26. Hicks GE, George SZ, Nevitt MA, Cauley JA, Vogt MT. Measurement of lumbar lordosis: inter-rater reliability,

minimum detectable change and longitudinal variation. J Spinal Disord Tech. 2006;19(7):501-6. http://dx.doi.org/10.1097/01.bsd.0000210116.94273.ad. PMid:17021414.

27. Bernhardt M, Bridwell KH. Segmental analysis of the sagittal plane alignment of the normal thoracic and lumbar spines and thoracolumbar junction. Spine (Phila Pa 1976). 1989;14(7):717-21. http://dx.doi.org/10.1097/00007632-198907000-00012. PMid:2772721.

28. Goh S, Price RI, Leedman PJ, Singer KP. A comparison of three methods for measuring thoracic kyphosis: implications for clinical studies. Rheumatology. 2000;39(3):310-5. http://dx.doi.org/10.1093/rheumatology/39.3.310. PMid:10788541.

29. Lee CS, Noh H, Lee DH, Hwang CJ, Kim H, Cho SK. Analysis of sagittal spinal alignment in 181 asymptomatic children. J Spinal Disord Tech. 2012;25(8):E259-63. http://dx.doi.org/10.1097/BSD.0b013e318261f346. PMid:22688614.

Correspondence
Jefferson Fagundes Loss
Universidade Federal do Rio Grande do Sul
Escola de Educação Física
Rua Felizardo, 750
CEP 90690-200, Porto Alegre, RS, Brazil
e-mail: jefferson.loss@ufrgs.br

original article

BJPT Brazilian Journal of Physical Therapy

Development of a first-contact protocol to guide assessment of adult patients in rehabilitation services networks

Mariana A. P. Souza[1], Fabiane R. Ferreira[1], Cibele C. César[2], Sheyla R. C. Furtado[3], Wendy J. Coster[4], Marisa C. Mancini[5], Rosana F. Sampaio[3]

ABSTRACT | Objective: This paper describes the development of the Protocol for Identification of Problems for Rehabilitation (PLPR), a tool to standardize collection of functional information based on the International Classification of Functioning, Disability and Health (ICF). **Development of the protocol**: The PLPR was developed for use during the initial contact with adult patients within a public network of rehabilitation services. Steps to develop the protocol included: survey of the ICF codes most used by clinical professionals; compilation of data from functional instruments; development and pilot testing of a preliminary version in the service settings; discussion with professionals and development of the final version. The final version includes: user identification; social and health information; brief functional description (BFD); summary of the BFD; and PLPR results. Further testing of the final version will be conducted. **Conclusions:** The protocol standardizes the first contact between the user and the rehabilitation service. Systematic use of the protocol could also help to create a functional database that would allow comparisons between rehabilitation services and countries over time.

Keywords: rehabilitation; assessment; patient-centered care; international classification of functioning; disability and health.

BULLET POINTS

- Rehabilitation treatment should focus on the patient functional demands.
- The PLPR standardizes the data collected at the beginning of rehabilitation.
- Thus, it improves communication among professionals, services, and patients.
- It includes minimal sets of ICF codes, relevant for people with disabilities.
- The ICF codes will allow comparisons between services and locations over time.

HOW TO CITE THIS ARTICLE

Souza MAP, Ferreira FR, César CC, Furtado SRC, Coster WJ, Mancini MC, et al. Development of a first-contact protocol to guide assessment of adult patients in rehabilitation services networks. Braz J Phys Ther. 2016 Mar-Apr; 20(2):148-157. http://dx.doi.org/10.1590/bjpt-rbf.2014.0137

Introduction

Over the last centuries, the world has faced a global demographic transition, with an increase in life expectancy and in chronic health conditions, resulting in the emergence of new, increasingly complex, disability-causing illnesses, either transient or permanent[1,2]. These changes have been challenging health systems by increasing the demand for rehabilitation services[2]. The situation presents an ideal opportunity for the development of a consistent model of rehabilitative care that integrates these services across a continuum of care in the health system.

In Brazil, the strategy employed to deal with this new demand was to create multidisciplinary teams and to structure public rehabilitation services in an integrated network organized across three levels of care. Basic care is supported by Family Health Care Centers, whose services are delivered in the community, close to the family's residence and, in some cases, in the patient's home[3,4]. Specialized care is offered at Specialized Rehabilitation Centers, which are responsible for treatments that require higher technological support[3]. Finally, hospital care

[1] Programa de Pós-graduação em Ciências da Reabilitação, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brazil
[2] Departamento de Estatística, UFMG, Belo Horizonte, MG, Brazil
[3] Departamento de Fisioterapia, UFMG, Belo Horizonte, MG, Brazil
[4] Departament of Occupational Therapy, Boston University (BU), Boston, MA, United States
[5] Departamento de Terapia Ocupacional, UFMG, Belo Horizonte, MG, Brazil
Received: Jan. 21, 2015 Revised: June 30, 2015 Accepted: Sept. 22, 2015

is responsible for handling persons with disabilities in urgent and emergency situations as well as assignment to in-patient rehabilitation beds[3].

Given the diversity of services and professionals, these multidisciplinary teams need to have competencies beyond their specific professional skills. These competencies include good communication skills, the use of appropriate protocols and procedures that reflect the goals of the service, and an integrated focus on the needs of patients[5]. At the start of this project, the work of the rehabilitation networks was often marked by poor systematization in collecting and sharing information on the target population. Furthermore, the information gathered did not always reflect patients' functional condition, preventing construction of a database that would support proper administrative, organizational, and financial planning of rehabilitation services.

In order to overcome these inadequacies, beginning in 2012, the Brazilian Ministry of Health recommended the use of the International Classification of Functioning, Disability and Health (ICF) as a clinical and statistical tool in health services[6]. To implement the ICF in everyday services, professionals must adopt the biopsychosocial model as a guide for their actions and use evaluation and functioning data collection protocols that are consistent with the model[7-9]. Thus, a systematized approach to patients from the first contact with a rehabilitation service is important to enable correct identification of the patient's limitations in functioning.

In the context of public health in Brazil, the first contact of the individual with the health service represents a strategic moment as it guides the organization of health units and the work process. The first contact is the moment when the patient or family member first seeks the health service due to a specific health complaint – a health professional must listen to the patient's complaint and establish a therapeutic alliance. The goal is to guarantee admittance to everyone who seeks services from the public health system and to understand the needs of the individual so that each case is addressed in the most suitable way[10]. The first contact process is followed throughout the Unified Health System (*Sistema Único de Saúde – SUS*). At this first contact, collaboration is established between the patient and the health team, which brings the patient into the center of his or her own therapeutic process[11,12]. The therapeutic process can be understood as all of the treatments available to the patients through SUS (medical appointments, exams, medication, and others). The inclusion of the patient and his or her family in decisions concerning the therapeutic process has been associated with higher autonomy and accountability of the patient, increased compliance, and satisfaction with the treatment[13].

Because the beginning of the rehabilitation treatment should focus on identifying the problems and needs of the individual, effective communication among all persons involved in the process is important to ensure a complete understanding of the patient's situation[8]. The first contact seems to be the right moment to use tools that help overcome professional boundaries and incorporate different perspectives that contribute to the improvement of shared decision-making in the rehabilitation treatment. The purpose of this paper is to describe the development process of an ICF-based protocol for collecting information during the initial contact with adult patients in rehabilitation services networks.

Method

Development of the protocol for identification of problems for rehabilitation

The Protocol for Identification of Problems for Rehabilitation (Protocolo de Levantamento de Problemas para a Reabilitação – PLPR) was developed through a partnership between researchers from Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brazil and professionals representing public rehabilitation services in Belo Horizonte in the year 2012. Belo Horizonte is one of the largest cities in Brazil[14]. The city has an extensive network of rehabilitation services across the three levels of care laid out in the legislation, which includes 58 community service centers, three centers of specialized care, and beds in 33 hospitals. These services involve multidisciplinary teams that include physical therapists, occupational therapists, speech pathologists, nutritionists, psychologists, social workers, pharmacists, and physical educators. In total, there are more than 500 professionals involved in these services[15].

The PLPR was developed after a series of meetings that included 61 rehabilitation professionals and rehabilitation managers from the public rehabilitation services of Belo Horizonte, as well as rehabilitation researchers from UFMG. Each professional who participated in this group was selected by his or her immediate manager. The main goal of these meetings was to re-design the model of care of the public rehabilitation network of Belo Horizonte, in an attempt

to follow the guidelines from the Ministry of Health[3,6]. The gatherings also provided an opportunity for this group to discuss and develop the protocol to facilitate the implementation of the proposed new model.

Development of the PLPR involved a series of steps: 1) survey of the ICF codes most frequently used by professionals in the public rehabilitation services; 2) compilation of information contained in functional instruments available in the literature; 3) development of a preliminary version of the protocol; 4) pilot testing of the preliminary version; 5) discussion with rehabilitation professionals; and 6) development of the final version of the protocol.

For the survey of ICF codes, the professionals were asked to assemble a list of the codes that were most frequently used in rehabilitation services in their workplace. To create this list, the professionals were instructed to confer with their colleagues (other rehabilitation professionals in the services) and to select the second-level ICF codes most often used in their clinical practice.

In addition, the following sources were also analyzed and reviewed in order to guide the selection of items to be included in the protocol: ICF Checklist; ICF Core Sets (e.g. Chronic Widespread Pain; Low Back Pain; and Stroke)[16]; functional evaluations already used by the professionals in the services (e.g. Functional Independence Measure – FIM; Visual Analogue Scale for Pain; and Medical Outcomes Study Short Form 36 – SF-36)[17]; and the *Minimal Generic Set* and the *Disability Set*, which are considered relevant to persons with disabilities[18]. The functional measures used as references to develop the protocol have already been linked to the ICF or were developed using the ICF model. To define which questions and codes should be included in the protocol, rehabilitation professionals, managers, and researchers were guided by daily practice in the services, including the most frequent patient requirements for rehabilitation, questions considered essential to decide which service and what kind of treatment the patient needs, and the protocol's feasibility (time to complete). After discussing these issues and analyzing all of the material, the codes that would comprise the PLPR were defined and guiding questions were written, resulting in a preliminary version of the instrument.

Initial pilot testing

To check the feasibility of the protocol, a pilot test was carried out across the community service centers and centers of specialized care. The rehabilitation professionals who were participating in the development process were asked to apply the draft protocol to all patients seeking treatment at the rehabilitation services for one month. After that, the professionals were asked to meet with their team partners to discuss and record their experiences when using this version. Any concerns were then discussed among the researchers and professionals. According to the professionals, the time to complete the protocol varied from 15 to 30 minutes, decreasing as professionals became accustomed to using it. After this initial testing, revisions were made to the preliminary version of the protocol and a final version was developed. Further testing of the final version will be conducted and presented in subsequent studies.

Results

Final protocol for identification of problems for rehabilitation

The final version of the protocol consisted of four parts: 1) user identification; 2) social and health information; 3) brief functional description (BFD) and summary of the BFD; and 4) results. The user identification section included information such as name and health unit.

The social and health information section includes questions concerning risk factors and self-perception of emotional and physical health. This part also includes information about ICF environmental factors such as employment status, use of prosthetics and/or orthotics, the need for assistance from others to perform daily tasks, and ongoing health treatments.

The BFD was created based on sets of ICF codes considered relevant for people with a medical condition that causes disability or poses a risk for developing disability[18]. These sets of items are called the *Minimal Generic Set* (MGS) and the *Disability Set* (DS). The MGS corresponds to a set of seven codes proposed by the WHO to be used in surveys regarding disabilities and health. The DS is a set of 22 ICF codes, including the seven codes from the MGS and 15 more related to body function, activity, and participation. The DS codes are proposed as good descriptors of disability situations and are included in a project by the WHO and the World Bank[18].

Based on the codes obtained from the rehabilitation professionals and on the professionals' experiences in the pilot test, the minimal sets proposed by WHO were expanded. Content considered important to the performance of different professionals in the

multidisciplinary teams was added and codes thought to be less relevant to the reality of the services were removed. For example, we decided to remove the code "Sexual Functions" (b640) because it was found to be an uncomfortable question to ask the patients in the initial contact – most of them could not answer properly. On the other hand, we decided to add questions regarding communication skills (d330–Speaking; d350–Conversation) to identify whether the patient needed to be seen by a speech pathologist. Table 1 compares the codes in the sets suggested by the WHO and those in the preliminary and final versions of the BFD of the PLPR. The final version of the BFD comprises 25 codes distributed in 10 domains (Table 1).

In order to standardize the use of the BFD codes, a guiding question was created for each of the 25 codes, based on the description in the ICF manual for each second-level code included in the BFD and their higher codes (see Appendix 1). For example, to create the guiding question for code b455 (Exercise tolerance functions), the ICF manual was consulted,

Table 1. Comparison of the ICF codes in the Minimal Generic Set and Disability Set and in the versions of the Brief Functional Description.

ICF codes in each domain	MGS and Disability Set (WHO)	BFD – Preliminary Version (PLPR)	BFD – Final Version (PLPR)
Mobility			
b455 Exercise tolerance functions	✓	✓	✓
b710 Mobility of joint functions	✓	✓	✓
b730 Muscle power functions	✓	-	-
d410 Changing basic body position	-	-	✓
d450 Walking (G)	✓	✓	✓
d455 Moving around (G)	✓	✓	-
d470 Using transportation	✓	✓	✓
Communication			
d330 Speaking	-	✓	✓
d350 Conversation	-	✓	✓
Eutrophy			
b510 Ingestion functions	-	✓	✓
b530 Weight maintenance functions	-	✓	✓
Self-care			
d510 Washing oneself	✓	✓	✓
d530 Toileting	-	-	✓
d540 Dressing	✓	✓	✓
d570 Looking after one's health	✓	✓	✓
Pain			
b280 Sensation of pain (G)	✓	✓	✓
Interpersonal activity			
d710 Basic interpersonal interactions	✓	✓	✓
d920 Recreation and leisure	✓	✓	✓
Energy and Sleep			
b130 Energy and drive functions (G)	✓	✓	✓
b134 Sleep functions	✓	✓	✓
Affect			
b152 Emotional functions (G)	✓	✓	✓
b640 Sexual functions	✓	✓	-
d240 Handling stress and other psychological demands	✓	✓	✓
d770 Intimate relationships	✓	✓	✓
General tasks and demands			
d230 Carrying out daily routine (G)	✓	✓	✓
d640 Doing housework	✓	✓	✓ *
d660 Assisting others	✓	✓	✓ *
Remunerative employment			
d850 Remunerative employment (G)	✓	✓	✓ +

MGS: minimal generic set; DS: disability set; BFD: brief functional description. **(G) Codes in bold** represent the seven codes of the MGS. ✓ Code present. - Code absent. * After discussions with professionals, it was decided to transfer the codes d640 (Doing housework) and d660 (Assisting others) from the "Affect" domain to the "General tasks and demands" domain on the final version of the PLPR. + To better identify difficulties in performing tasks related to the remunerative work, it was decided to transfer the code d850 (Remunerative employment) from the "General tasks and demands" domain and create a "Remunerative employment" domain.

and the descriptions of codes b455, b4550, b4551, and b4552 were analyzed, resulting in the question "*When engaging in physical effort, do you feel tired or short of breath?*".

To describe the individual's level of function and disability for each of the BFD codes, we used the ICF qualifiers. Thus, after asking the reference question for a particular BFD code, if the patient reported some difficulty in the situation represented by the code, he/she would be asked to quantify the difficulty on a 5-point scale from 0 (no disability or difficulty) to 4 (complete disability or difficulty). Therefore, it is the patient or his/her proxy who quantifies the extension of the problems in the questions of the BFD.

After initial pilot testing, the professionals reported that patients had difficulty comprehending the ordinal 0-4 rating scale. Consequently, a scale from 0 (no disability or difficulty) to 10 (complete disability or difficulty) was used instead, as patients were more familiar with this type of scale. Transformations of original ICF qualifiers to the 11-point scale were conducted using a conversion table present in the manual[19]. A visual analog scale was created for patients who reported difficulty understanding the BFD questions. This figure includes a graded color code and a numerical 0-10 scale grouped according to ICF qualifiers, with descriptive words for each qualifier. The professional chooses the format most suited to the patient's understanding in order to quantify the severity of his/her problem. After completion of the PLPR form, the professional fills out a table on the front page of the protocol (summary of the BFD), coloring in the spaces relative to the qualifier for each BFD question, which results in a graphic representation of the patient's main functional limitations (Figure 1).

A final BFD score summarizes the functional information reported by the patient. It varies from 0 to 100 points based on the sum of normalized sub-scores from each of the 10 BFD domains. Higher final scores represent lower functional level. Each BFD item is rated on a five-point scale according to the following ICF qualifiers: (0) no impairment or difficulty; (1) mild impairment or difficulty; (2) moderate impairment or difficulty; (3) severe impairment or difficulty; or (4) complete impairment or difficulty[19]. The two qualifiers (8) non-specified and (9) non-applicable receive a score of zero.

In order to normalize each BFD domain by their number of items and keep the same maximum score (10) across domains, a weight was created for each domain. For example, the mobility domain has 5 items and its raw score can vary from 0 to 20. By attributing

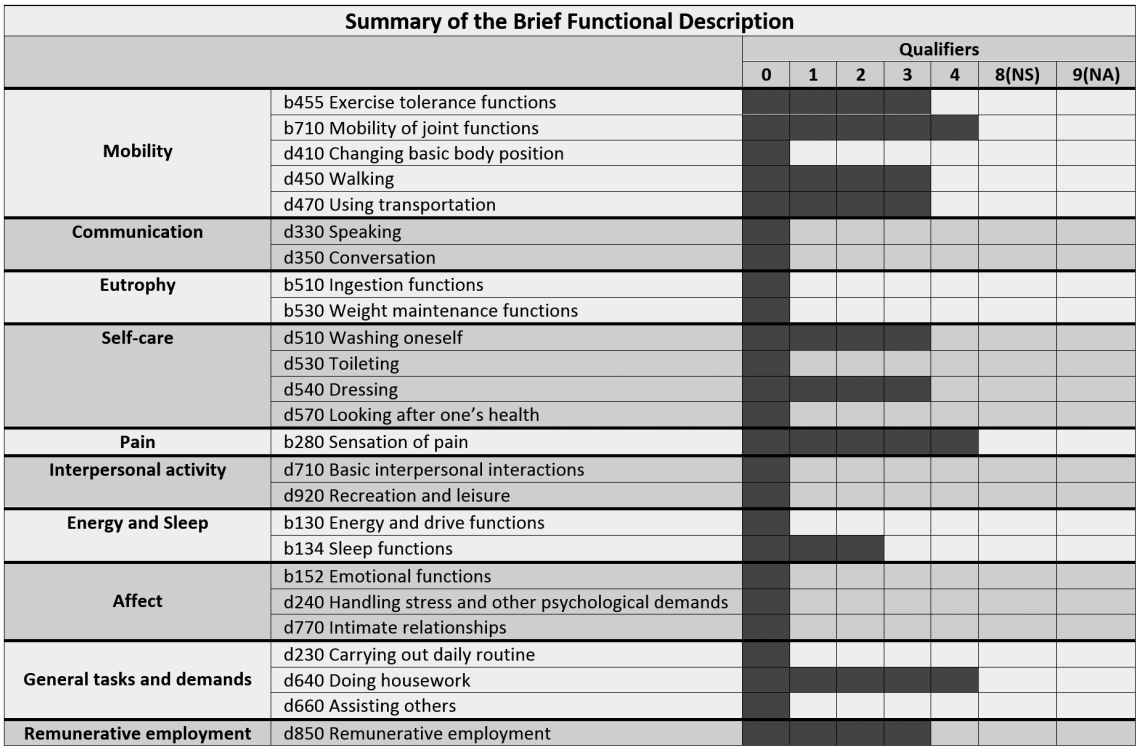

Figure 1. Example of summary of the brief functional description.

a weight of 0.5 to this domain, its maximum raw score becomes 10. Furthermore, the pain domain, which has only one item that can be scored on a 0 to 4 scale, receives a weight of 2.5. The weights were created to normalize the impact of each domain on the protocol's final score. Table 2 shows the weights attributed to each BFD domain and gives an example of final score computation.

The PLPR final score may help guide the planning of actions for the rehabilitation services network,

Table 2. Example of scoring in the Brief Functional Description.

BFD Domains	Qualifier reported	Weight in the domain	Total score in the domain (sum of qualifiers in the domain × weight)
Mobility			
b455 Exercise tolerance functions	3		
b710 Mobility of joint functions	4		
d410 Changing basic body position	0	0.5	13 × 0.5 = **6.5**
d450 Walking	3		
d470 Using transportation	3		
Communication			
d330 Speaking	0	1.25	0 × 1.25 = **0**
d350 Conversation	0		
Eutrophy			
b510 Ingestion functions	0	1.25	0 × 1.25 = **0**
b530 Weight maintenance functions	0		
Self-care			
d510 Washing oneself	3		
d530 Toileting	0	0.625	6 × 0.625 = **3.75**
d540 Dressing	3		
d570 Looking after one's health	0		
Pain			
b280 Sensation of pain	4	2.5	4 × 2.5 = **10**
Interpersonal activity			
d710 Basic interpersonal interactions	0	1.25	0 × 1.25 = **0**
d920 Recreation and leisure	0		
Energy and Sleep			
b130 Energy and drive functions	0	1.25	2 × 1.25 = **2.5**
b134 Sleep functions	2		
Affect			
b152 Emotional functions	0		
d240 Handling stress and other psychological demands	0	0.833	0 × 0.833 = **0**
d770 Intimate relationships	0		
General tasks and demands			
d230 Carrying out daily routine	0		
d640 Doing housework	4	0.833	4 × 0.833 = **3.33**
d660 Assisting others	0		
Remunerative employment			
d850 Remunerative employment	3	2.5	3 × 2.5 = **7.5**
TOTAL SCORE			**33.58**

contributing to the identification of intervention priorities for each patient and the necessary level of complexity.

The PLPR result is provided by the rehabilitation professional who completed the protocol (i.e. who received the patient in his/her first contact with the service). Based on the data collected with the PLPR, the professional identifies the patient's "primary need" for rehabilitation, the "professional indicated to coordinate the case" in the beginning of the rehabilitation process, and the "place to begin care" (i.e. in which service of the rehabilitation network the patient will start treatment). The rehabilitation location is determined primarily by the needs of the patients, availability of services, and professionals in a specific area of need, as well as issues such as the patient's ability to use public transportation safely. The coordinator of the case is responsible for optimizing patient flow across all points of the healthcare continuum, not only in rehabilitation. It is expected that rehabilitation professionals will be trained to apply the protocol and to use their clinical reasoning and experience to interpret the information gathered and make the best decision for each individual patient.

Discussion

As in other countries, the public rehabilitation network in Brazil organizes its services across different levels of care, aiming to deliver integrated assistance to patients with diverse requirements[3,20-22]. In a truly integrated system, these services work together to organize efficient and effective patient flow. For this purpose, the services should work in an integrated manner with the existing health system. One of the key points of this model of care is that, although patients may need to access different services as they progress, their transition between sites should be optimized by communication and exchange of information between services so that patients can progress in an uninterrupted continuum of assistance across different levels of care[5,20]. A model of care for rehabilitation services should consider that this is not a linear process, and that the patients often need to visit and re-visit different points of the network as their recovery progresses and new challenges are faced. This requires integrated evaluations and a care coordinator to improve efficiency of the services and support achievement of positive patient outcomes[20,23].

Another important issue is the difficulties encountered when introducing a new conceptual model to guide the actions of health services and adoption of these innovations in the daily routine of the services[24]. This is a challenge that requires considerable effort from professionals who usually must adapt to these changes without interrupting the care of patients under their responsibility. The PLPR was designed during meetings that aimed to re-design the model of care of a public rehabilitation network, and it is one of the strategies proposed for practical implementation of the transformations that result from adoption of the biopsychosocial ICF model[25].

We expect that the PLPR can contribute to improving communication among professionals and services and guiding the patient's pathway throughout the rehabilitation network. The use of this protocol systematizes the information collected in the initial contact, ensuring that this information is available online to be accessed by professionals anywhere in the network. This standardization saves time and effort of professionals and patients. Based on the identified problems and needs, patients move more quickly to advanced stages of the rehabilitation process such as the use of specific evaluations after admission for treatment at the location indicated in the PLPR[8,25]. Thus, it is expected that more equitable access can be achieved in proper locations and in a timely fashion, contributing to greater effectiveness and efficiency of rehabilitation services.

Since the release of ICF, there has been considerable research focusing on its use in several contexts such as policies, statistics, and especially in the development of ICF-based assessment tools for clinical applications[26-28]. Apart from its specific application to intervention, availability of information about functioning is essential to policy planning, service planning, and investments in rehabilitation[29-32]. The PLPR aims to meet those needs by 1) identifying the functional needs of patients in a more systematic and informative way and 2) guiding the organization of services and the planning of rehabilitation actions.

The focus of the protocol on identifying the functioning concerns reported by the patient in the initial contact with the service is crucial. The use of this protocol from the start gives the patient and/or his/her family the opportunity to report his/her functional needs and expectations regarding the rehabilitation process. This promotes active participation in the patient's own treatment planning[13]. Hence, the PLPR has great potential to improve the organization of services by increasing patient compliance, as it considers the preferences of the patient and his/her functional needs to assist in the

selection of the most appropriate professionals and services to initiate care from first contact.

In the development of the protocol, it is important to highlight the use of the minimal sets of ICF codes proposed by the WHO, as well as the participation of professionals from rehabilitation services. The inclusion of the minimal sets in the BFD will allow the comparison between the data collected with the PLPR and data from other services, locations, and at different times, as the WHO proposes the wide use of these minimal code sets in disability and health surveys[18]. Furthermore, by maintaining the majority of codes from those clusters in the protocol, it will be possible to merge databases based on the PLPR with other function-focused databases using specific statistical techniques (e.g. Item Response Theory)[33].

The active participation of professionals in the construction of the PLPR led to a protocol that is in line with the reality of rehabilitation services and increased professional compliance when applied to the daily routines of services. However, innovations that require changes are not always easy, especially when they involve clinical practice, better collaboration among disciplines, or changes in the organization of care. Studies show that behavioral changes in clinical practice are possible, but require a comprehensive approach at different levels (hospitals, ambulatory, primary care) and adaptation to specific locations and groups, similar to what has been done in the development of the PLPR[24].

In addition to the possibilities already described, the PLPR has also proved to be efficient in identifying patients who do not have a specific need for individualized care and who could take part in different group activities, undergo vocational guidance, and receive follow-up. Based on better identification of the functional needs of patients, as well as the best location to start the rehabilitation process, one might expect the use of the protocol to contribute to reducing the waiting list for rehabilitation care and the number of inappropriate transfers between services. These issues should be investigated in future studies.

References

1. Lee R. The demographic transition: three centuries of fundamental change. J Econ Perspect. 2003;17(4):167-90. http://dx.doi.org/10.1257/089533003772034943.
2. World Health Organization – WHO. World report on disability. Geneva; 2011.
3. Brasil. Ministério da Saúde. Portaria GM/MS nº 793, de 24 de Abril de 2012. Institui a Rede de Cuidados à Pessoa com Deficiência no âmbito do Sistema Único de Saúde. Diário Oficial da União; Brasília; 25 abr. 2012.
4. Brasil. Ministério da Saúde. Portaria GM/MS nº 154, de 24 de Janeiro de 2008. Cria os Núcleos de Apoio à Saúde da Família – NASF. Diário Oficial da União; Brasília; 25 jan. 2008.
5. Nancarrow SA, Booth A, Ariss S, Smith T, Enderby P, Roots A. Ten principles of good interdisciplinary team work. Hum Resour Health. 2013;11(1):19. http://dx.doi.org/10.1186/1478-4491-11-19. PMid:23663329.
6. Brasil. Ministério da Saúde. Resolução nº 452, de 10 de Maio de 2012. Estabelece que a Classificação Internacional de Funcionalidade, Incapacidade e Saúde – CIF seja utilizada no Sistema Único de Saúde, inclusive na Saúde Suplementar. Diário Oficial da União; Brasília; 6 jun. 2012.
7. Sampaio RF, Luz MT. Funcionalidade e incapacidade humana: explorando o escopo da classificação internacional da Organização Mundial da Saúde. Cad Saude Publica. 2009;25(3):475-83. http://dx.doi.org/10.1590/S0102-311X2009000300002. PMid:19300836.
8. Steiner WA, Ryser L, Huber E, Uebelhart D, Aeschlimann A, Stucki G. Use of the ICF model as a clinical problem-solving tool in physical therapy and rehabilitation medicine. Phys Ther. 2002;82(11):1098-107. PMid:12405874.
9. Rauch A, Cieza A, Stucki G. How to apply the International Classification of Functioning, Disability and Health (ICF) for rehabilitation management in clinical practice. Eur J Phys Rehabil Med. 2008;44(3):329-42. PMid:18762742.
10. Brasil. Ministério da Saúde. Acolhimento nas práticas de produção de saúde. 2nd ed. Brasília; 2006.
11. Franco TB, Bueno WS, Merhy EE. O acolhimento e os processos de trabalho em saúde: o caso de Betim, Minas Gerais, Brasil. Cad Saude Publica. 1999;15(2):345-53. http://dx.doi.org/10.1590/S0102-311X1999000200019. PMid:10409787.
12. Mitre SM, Andrade EIG, Cotta RMM. Progress and challenges facing user acceptance in the implementation and qualification of the Unified Health System in Primary Healthcare: a review of the bibliographical output in Brazil. Cien Saude Colet. 2012;17(8):2071-85. http://dx.doi.org/10.1590/S1413-81232012000800018. PMid:22899148.
13. Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. med Care Res Rev. 2013;70(4):351-79. http://dx.doi.org/10.1177/1077558712465774. PMid:23169897.
14. Instituto Brasileiro de Geografia e Estatística – IBGE. Censo Demográfico 2010. Rio de Janeiro; 2010.
15. Belo Horizonte. Gerência de Reabilitação. Profissionais de Reabilitação nos Serviços da Secretaria Municipal de Saúde de Belo Horizonte. Belo Horizonte: Secretaria Municipal de Saúde; 2013.
16. ICF Research Branch. ICF Core Sets Projects [Internet]. Nottwil: Swiss Paraplegic Research; 2013. [cited 2014 June 20]. Available from: http://www.icf-research-branch.org/icf-core-sets-projects-sp-1641024398.
17. Rehabilitation Institute of Chicago. Rehabilitation Measures Database [Internet]. Chicago; 2010 [cited 2014 June 20]. Available from: http://www. rehabmeasures.org/.
18. Cieza A, Oberhauser C, Bickenbach J, Chatterji S, Stucki G. Towards a minimal generic set of domains of functioning and health. BMC Public Health. 2014;14(1):218. http://dx.doi.org/10.1186/1471-2458-14-218. PMid:24588794.

19. World Health Organization – WHO. International Classification of Functioning, Disability and Health: ICF. 1st ed. Geneva; 2001.

20. New SWH. Rehabilitation Redesign Project Final Report [Internet]. Chatswood: Agency for Clinical Innovation; 2011 [cited 2014 July 2]. Available from: http://www.aci.health.nsw.gov.au/__data/assets/pdf_file /0020/192323/Final _Report_Model_of_Care.pdf.

21. Holdsworth LK, Webster VS, McFadyen AK. What are the costs to NHS Scotland of self-referral to physiotherapy? Results of a national Trial. Physiotherapy. 2007;93(1):3-11. http://dx.doi.org/10.1016/j.physio.2006.05.005.

22. Nancarrow S, Moran A, Freeman J, Enderby P, Dixon S, Parker S, et al. Looking inside the black box of community rehabilitation and intermediate care teams in the United Kingdom: an audit of service and staffing. Qual Prim Care. 2009;17(5):323-33. PMid:20003718.

23. Ireland H. Rehabilitation following acquired brain injury: a headway review of guidelines and evidence [Internet]. Dublin: Headway Ireland; 2012 [cited 2014 July 2]. Available from: http://www.headway.ie/download/pdf/rehabilitation_review.pdf.

24. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patient's care. Lancet. 2003;11(9391):1225-30. http://dx.doi.org/10.1016/S0140-6736(03)14546-1. PMid:14568747.

25. Sampaio RF, Ferreira FR, Souza MAP. Reorientação do Modelo Assistencial da Rede de Reabilitação SUS. Belo Horizonte: Secretaria Municipal de Saúde; 2014. Caderno de Apoio.

26. Cerniauskaite M, Quintas R, Boldt C, Raggi A, Cieza A, Bickenbach JE, et al. Systematic literature review on ICF from 2001 to 2009: its use, implementation and operationalisation. Disabil Rehabil. 2011;33(4):281-309. http://dx.doi.org/10.3109/09638288.2010.529235. PMid:21073361.

27. Ruaro JA, Ruaro MB, Souza DE, Frez AR, Guerra RO. An overview and profile of the ICF's use in Brazil: a decade of history. Rev Bras Fisioter. 2012;16(6):454-62. http://dx.doi.org/10.1590/S1413-35552012005000063. PMid:23250191.

28. Sampaio RF, Mancini MC, Gonçalves GGP, Bittencourt NFN, Miranda AD, Fonseca ST. Aplicação da Classificação Internacional de Funcionalidade, Incapacidade e Saúde (CIF) na prática clínica do fisioterapeuta. Rev Bras Fisioter. 2005;9(2):129-36.

29. United Nations Economic and Social Commission for Asia and the Pacific – UNESCO. World Health Organization – WHO. Training Manual on Disability Statistics. Bangkok; 2008.

30. O'Donovan MA, Doyle A, Gallagher P. Barriers, activities and participation: Incorporating ICF into service planning datasets. Disabil Rehabil. 2009;31(25):2073-80. http://dx.doi.org/10.3109/09638280902918738. PMid:19888837.

31. Mueller M, Lohmann S, Strobl R, Boldt C, Grill E. Patients' functioning as predictor of nursing workload in acute hospital units providing rehabilitation care: a multi-centre cohort study. BMC Health Serv Res. 2010;10(1):295. http://dx.doi.org/10.1186/1472-6963-10-295. PMid:21034438.

32. Disler PB, Roy CW, Smith BP. Predicting hours of care needed. Arch Phys Med Rehabil. 1993;74(2):139-43. [Abstract] PMid:8431096.

33. Jette AM, Haley SM. Contemporary measurement techniques for rehabilitation outcomes assessment. J Rehabil Med. 2005;37(6):339-45. http://dx.doi.org/10.1080/16501970500302793. PMid:16287664.

Correspondence
Rosana Ferreira Sampaio
Universidade Federal de Minas Gerais
Escola de Educação Física, Fisioterapia e Terapia Ocupacional
Departamento de Fisioterapia
Avenida Antonio Carlos, 6627, Campus Universitário, Pampulha
CEP 31270-401, Belo Horizonte, MG, Brazil
e-mail: rosana.sampaio@pq.cnpq.br

Appendix 1. Example of how the Brief Functional Description questions were developed.

	ICF Codes (Second Level)	Description	Description of the ICF Third and Forth Levels Codes
Mobility	b455	**Exercise tolerance functions** Functions related to respiratory and cardiovascular capacity as required for enduring physical exertion. **Inclusions:** functions of physical endurance, aerobic capacity, stamina and fatigability. **Exclusions:** functions of the cardiovascular system (b410-b429); hematological system functions (b430); respiration functions (b440); respiratory muscle functions (b445); additional respiratory functions (b450).	**b4550 General physical endurance** Functions related to the general level of tolerance of physical exercise or stamina. **b4551 Aerobic capacity** Functions related to the extent to which a person can exercise without getting out of breath. **b4552 Fatigability** Functions related to susceptibility to fatigue at any level of exertion.
MOBILITY		b455.____ Exercise tolerance functions	When engaging in physical effort, do you feel tired or short of breath?

original article

BJPT Brazilian Journal of Physical Therapy

Effects of chest wall compression on expiratory flow rates in patients with chronic obstructive pulmonary disease

Masafumi Nozoe[1], Kyoshi Mase[1], Tomoyuki Ogino[2], Shigefumi Murakami[2], Sachie Takashima[1], Kazuhisa Domen[3]

ABSTRACT | Background: Manual chest wall compression (CWC) during expiration is a technique for removing airway secretions in patients with respiratory disorders. However, there have been no reports about the physiological effects of CWC in patients with chronic obstructive pulmonary disease (COPD). **Objective:** To compare the effects of CWC on expiratory flow rates in patients with COPD and asymptomatic controls. **Method:** Fourteen subjects were recruited from among patients with COPD who were receiving pulmonary rehabilitation at the University Hospital (COPD group). Fourteen age-matched healthy subjects were also consecutively recruited from the local community (Healthy control group). Airflow and lung volume changes were measured continuously with the subjects lying in supine position during 1 minute of quiet breathing (QB) and during 1 minute of CWC by a physical therapist. **Results:** During CWC, both the COPD group and the healthy control group showed significantly higher peak expiratory flow rates (PEFRs) than during QB (mean difference for COPD group 0.14 L/sec, 95% confidence interval (CI) 0.04 to 0.24, $p<0.01$, mean difference for healthy control group 0.39 L/sec, 95% CI 0.25 to 0.57, $p<0.01$). In the between-group comparisons, PEFR was significantly higher in the healthy control group than in the COPD group (–0.25 L/sec, 95% CI –0.43 to –0.07, $p<0.01$). However, the expiratory flow rates at the lung volume at the PEFR during QB and at 50% and 25% of tidal volume during QB increased in the healthy control group (mean difference for healthy control group 0.31 L/sec, 95% CI 0.15 to 0.47, $p<0.01$: 0.31 L/sec, 95% CI 0.15 to 0.47, $p<0.01$: 0.27 L/sec, 95% CI 0.13 to 0.41, $p<0.01$, respectively) but not in the COPD group (0.05 L/sec, 95% CI –0.01 to 0.12: –0.01 L/sec, 95% CI –0.11 to 0.08: 0.02 L/sec, 95% CI –0.05 to 0.90) with the application of CWC. **Conclusion:** The effects of chest wall compression on expiratory flow rates was different between COPD patients and asymptomatic controls.

Keywords: physical therapy modalities; chronic obstructive pulmonary disease; peak expiratory flow rate**.**

Clinical Trials Identifier: UMIN000018923

BULLET POINTS

- We compared the effects of CWC on expiratory flow rates in patients with COPD and asymptomatic controls.
- It was confirmed that PEFR increased during CWC in COPD patients; however, PEFR changes during CWC were lower in COPD patients than in healthy subjects.
- Although CWC appears to be less effective in increasing absolute expiratory flow rates in COPD patients than in healthy subjects, the PEFR, which is more effective for removing secretions, did increase in the COPD group.

HOW TO CITE THIS ARTICLE

Nozoe M, Mase K, Ogino T, Murakami S, Takashima S, Domen K. Effects of chest wall compression on expiratory flow rates in patients with chronic obstructive pulmonary disease. Braz J Phys Ther. 2016 Mar-Apr; 20(2):158-165. http://dx.doi.org/10.1590/bjpt-rbf.2014.0145

Introduction

Manual chest wall compression (CWC) during expiration is a technique for removing airway secretions[1,2]. It is known that the main physiological mechanism for removing secretions is increased expiratory flow rates due to increased pleural pressure[3-5] and stretching of intercostal muscles by means of manual thoracic compression applied during exhalation[6]. Several studies have also shown that CWC increased expiratory flow rates, improved removal of airway secretions, and improved gas exchange and pulmonary mechanics[2,6-8] in patients on mechanical ventilation[6-11], as well as in patients with cystic fibrosis[3], in animal models[8,12].

[1]Department of Physical Therapy, Faculty of Nursing and Rehabilitation, Konan Women's University, Morikita-machi, Higashinada-ku, Kobe, Japan
[2]Department of Rehabilitation, Hyogo College of Medicine Sasayama Medical Center, Kurooka-cho, Sasayama-shi, Hyogo, Japan
[3]Department of Physical Medicine and Rehabilitation, Hyogo College of Medicine, Mukogawa-cho, Nishinomiya-shi, Hyogo, Japan
Received: Mar. 07, 2015 Revised: July 21, 2015 Accepted: Oct. 15, 2015

Some studies showed positive effects[8,13], but another study showed negative effects on clinical outcomes with CWC, such as expiratory flow limitation (EFL)[6].

It has also been reported that many physical therapists often use manual chest physical therapy techniques in patients with acute exacerbations of chronic obstructive pulmonary disease (COPD)[14]. It is great of importance for physical therapists to know the physiological effects of CWC in patients with COPD. It may also be harder to increase expiratory flow rates with CWC in COPD patients, because many COPD patients often show difficulty increasing expiratory flow rates because of airflow limitation[15]. However, there have been no reports of the physiological effects of CWC in patients with COPD.

The purpose of this study was to compare the effects of CWC on expiratory flow rates in patients with COPD and asymptomatic controls. We hypothesized that expiratory flow rates during CWC are harder to increase in COPD patients.

Method

Participants

This study was approved by the Hyogo College of Medicine, Nishinomiya-shi, Hyogo, Japan Ethics Committee (approval number 1189). Written, informed consent was obtained from all eligible participants. The study included fourteen clinically stable patients with stage I to IV COPD according to the Global Initiative for Chronic Obstructive Lung Disease guidelines (COPD group) who were receiving pulmonary rehabilitation at the Hyogo College of Medicine Sasayama Medical Center between October 2012 and September 2013 and who could perform spirometric testing according to the ATS/ERS Task Force Guidelines[16]. Forced expiratory volume in the first second (FEV_1) and forced vital capacity (FVC) were expressed as predicted percentage values for age, sex, and height, established by the Japanese Respiratory Society[17]. The patients were clinically stable for ≥4 weeks. Exclusion criteria were suspected asthma, other systemic conditions that could contribute to dyspnea or exercise limitation (e.g. heart failure or metabolic disorders), and neuromuscular comorbidity limiting all measurements or non-approval for study participation. All COPD patients continued their regular treatment (all COPD patients used inhaled long-acting beta 2-agonists or long-acting muscarinic antagonists, and one patient used an inhaled corticosteroid). No changes in the medications were made for the purpose of the study. Fourteen age-matched healthy subjects were also consecutively recruited from the local community (healthy control group).

Measurement procedures

All measurements were collected with the subjects in supine position. Airflow rates and lung volume changes during 1 minute of quiet breathing and during 1 minute of CWC were measured via a mouth filter (PIF-1A; MINATO Medical Science, Osaka, Japan) with a heated pneumotachograph (AE300-s; MINATO Medical Science). A mouth filter was used to avoid cross-infection. CWC was applied to each subject during expiration by a male physical therapist with 8 years of clinical experience in pulmonary physical therapy. The physical therapist stood on the right side of the subject and placed both hands on the subject's upper rib cage (upper part from the sixth rib). CWC was started from the beginning of expiration to the end of expiration. The highest tolerable level of force was applied to the subject's chest wall then released as soon as the subject began inspiration. The subject was asked to avoid actively expiring during the application of CWC but to inspire freely. All subjects also performed the inspiratory capacity (IC) maneuver at the start during quiet breathing (QB) and at the end during CWC to correct the volume measuring errors ("drift")[15].

Airflow and lung volume data were examined using an analysis system (PowerLab, ADInstruments, Castle Hill, NSW, Australia). The last five breaths during QB and CWC were analyzed, and the mean values for Ti (inspiratory time), Te (expiratory time), Ttot (total breathing cycle time), Ti/Ttot (duty cycle), Vt (tidal volume), PIFR (peak inspiratory flow rate), PEFR (peak expiratory flow rate), PEFR/PIFR, Vt/Ti (mean inspiratory flow rate), and Vt/Te (mean expiratory flow rate) were obtained for each subject.

Flow volume (FV) curve analysis was also performed by calculating the average FV curves from the last five breaths during QB and CWC[18]. The same analysis system (PowerLab, ADInstruments) was also used for averaging FV curves. The expiratory flow rate changes during QB and CWC were then examined at the same lung volume (PEFR during QB and at 50% and 25% of tidal volume during QB) and any overlap in the regions of the FV curves was determined. The presence of overlap was defined as a difference in the airflow rates within 5% between the two FV curves at the same lung volume.

Sample size calculation

The effects of CWC in COPD patients have not been reported. The sample size was calculated using the differences in PEFR with the application of CWC in the first seven subjects in each group. The mean difference in the between-group comparison was 0.195 L/sec (standard deviation (SD) 0.183). A sample size of 14 subjects per group was thus required for this study to have 80% power with alpha of 0.05.

Statistical analysis

The results are shown as means±SD. Normality of the measurement data was examined using the Kolmogorov-Smirnov test. The unpaired Student's *t*-test or Mann-Whitney's *U*-test was used to compare demographic characteristics and lung function. The between-group sex distribution was compared using chi-square analysis. Within-group comparisons by CWC intervention were done using the paired Student's *t*-test or Wilcoxon's rank-sum test, and between-group comparisons by CWC intervention were done using the unpaired Student's *t*-test or Mann-Whitney's *U*-test. A chi-square analysis was performed to compare the ratios of subjects who showed overlapping regions between the COPD group and the healthy control group. All tests were performed at a significance level of $P<0.05$. Analyses were performed with statistical software (SPSS 20, SPSS, Chicago, IL, USA).

Results

Table 1 shows the baseline characteristics of the study participants. There were significant differences between the groups in pulmonary function.

Table 2 shows the breathing pattern and lung volume changes during QB and CWC. Ti, Te, Ttot, Vt, IC, PIFR, PEFR, and Vt/Ti increased significantly, and Ti/Ttot decreased significantly more during CWC than during QB, both in the COPD group and in the healthy control group, but Vt/Te increased significantly more during CWC than during QB only in the healthy control group. As a result, PIFR/PEFR was not different during QB and CWC in both groups. In the between-group comparisons, Vt, PIFR, PEFR, Vt/Ti, and Vt/Te were significantly higher in the healthy control group than in the COPD group.

Table 3 shows the expiratory flow rates at the same lung volume during QB and CWC. The expiratory flow rates at the lung volume at PEFR during QB and at 50% and 25% of tidal volume during QB were higher only in the healthy control group, but not in the COPD group with CWC.

Figure 1 shows representative FV curves during QB and CWC in the COPD group and the healthy control group. The left FV curves in the healthy control group show increased expiratory flow rates at all lung volumes. In contrast, the right FV curves in the COPD group show increased expiratory flow rates during CWC at the beginning of expiration, but not at the end of expiration, and the FV curves had overlapping regions. Less than half of the subjects in the healthy control group (6/14 subjects; 43%) had overlapping regions in the FV curves, but almost all subjects in the COPD group had overlapping regions (13/14 subjects; 93%, $P<0.05$). All overlapping regions were seen at the end of expiration.

Table 1. Baseline characteristics of the study participants.

	Healthy control group (n=14)	COPD group (n=14)	p-value
Female, n (%)	6 (43%)	5 (36%)	0.70
Age (year)	77 (7)	80 (8)	0.21
Weight (kg)	56 (11)	48 (9)	0.059
Height (cm)	160 (10)	157 (7)	0.41
BMI (kg/m^2)	22 (3)	19 (3)	0.05
FEV_1 (L)	2.0 (0.5)	1.1 (0.6)	<0.0001
%FEV_1 (% predicted)	92 (10)	52 (20)	<0.0001
FEV/FVC (%)	78 (8)	56 (12)	<0.0001
FVC (L)	2.7 (0.6)	1.9 (0.8)	0.007
%FVC (%predicted)	95 (10)	71 (20)	0.0004

BMI: body mass index; FEV_1: forced expiratory volume on the first second; FVC: forced vital capacity. Continuous data is expressed as mean (SD), categorical data is expressed as number (%).

Table 2. Changes in breathing patterns and lung volumes during QB and CWC in both groups.

	Groups				Within-Group differences (95% CI)		Between-Group differences (95% CI)
	QB		CWC		CWC minus QB		CWC minus QB
	COPD (n=14)	Healthy (n=14)	COPD (n=14)	Healthy (n=14)	COPD	Healthy	COPD minus Healthy
Ti (sec)	1.45	1.59	1.70	1.85	0.25*	0.26*	–0.01
	(0.45)	(0.38)	(0.61)	(0.53)	(0.34)	(0.41)	(–0.31 to 0.30)
Te (sec)	2.44	2.37	3.65	3.18	1.21**	0.81**	0.40
	(0.99)	(0.87)	(1.27)	(0.80)	(0.80)	(0.73)	(–0.22 to 1.02)
Ttot (sec)	3.89	3.96	5.35	5.03	1.46**	1.07**	0.39
	(1.37)	(1.20)	(1.74)	(1.17)	(0.94)	(0.99)	(–0.39 to 1.16)
Ti/Ttot	0.38	0.41	0.32	0.37	–0.06**	–0.04*	–0.02
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.07)	(–0.07 to 0.03)
Vt (L)	0.61	0.68	0.86	1.22	0.25**	0.55**	–0.30†
	(0.22)	(0.17)	(0.29)	(0.46)	(0.17)	(0.39)	(–0.54 to –0.05)
IC (L)	1.54	2.17	1.68	2.39	0.14**	0.22**	–0.08
	(0.71)	(0.58)	(0.77)	(0.55)	(0.14)	(0.16)	(–0.20 to 0.04)
PIFR (L/sec)	0.57	0.56	0.76	0.93	0.19**	0.37**	–0.18†
	(0.13)	(0.15)	(0.24)	(0.28)	(0.19)	(0.25)	(–0.36 to –0.01)
PEFR (L/sec)	0.56	0.60	0.70	0.99	0.14**	0.39**	–0.25‡
	(0.21)	(0.20)	(0.24)	(0.36)	(0.16)	(0.27)	(–0.43 to –0.07)
PEFR/PIFR	0.97	1.05	0.97	1.06	0	0.01	–0.01
	(0.27)	(0.15)	(0.24)	(0.18)	(0.30)	(0.21)	(–0.22 to 0.19)
Vt/Ti (L/sec)	0.45	0.44	0.52	0.67	0.09**	0.23**	–0.14‡
	(0.11)	(0.12)	(0.16)	(0.17)	(0.10)	(0.14)	(–0.23 to –0.04)
Vt/Te (L/sec)	0.26	0.31	0.24	0.40	–0.03	0.09	–0.12‡
	(0.08)	(0.10)	(0.06)	(0.15)	(0.06)	(0.12)	(–0.19 to –0.04)

Results expressed as mean (standard deviation) and mean difference and 95% confidence intervals (CI) between measurement conditions. COPD: chronic obstructive pulmonary disease group; Healthy, healthy control group; QB: quiet breathing; CWC: chest wall compression; Ti: inspiratory time; Te: expiratory time; Ttot: total breathing cycle time; Ti/Ttot: duty cycle; Vt: tidal volume; IC: inspiratory capacity; PIFR: peak inspiratory flow rate; PEFR: peak expiratory flow rate; Vt/Ti: mean inspiratory flow rate; Vt/Te: mean expiratory flow rate. *, **: Significant change between QB and CWC (P<0.05, P<0.01).†, ‡: Significant change between COPD group and healthy control group (P<0.05, P<0.01).

Table 3. Changes in expiratory flow rates at the same lung volume during QB and CWC in both groups.

	Groups				Within-Group differences		Between-Group differences
	QB		CWC		CWC minus QB		CWC minus QB
	COPD (n=14)	Healthy (n=14)	COPD (n=14)	Healthy (n=14)	COPD	Healthy	COPD minus Healthy
EFR at PEFR (L/sec)	0.56	0.60	0.62	0.91	0.05	0.31**	–0.26‡
	(0.21)	(0.20)	(0.23)	(0.35)	(0.11)	(0.27)	(–0.42 to –0.09)
EFR at 50% Vt (L/sec)	0.44	0.53	0.43	0.84	–0.01	0.31**	–0.32‡
	(0.20)	(0.19)	(0.19)	(0.34)	(0.16)	(0.27)	(–0.50 to –0.15)
EFR at 25% Vt (L/sec)	0.28	0.40	0.29	0.67	0.02	0.27**	–0.25‡
	(0.09)	(0.15)	(0.17)	(0.30)	(0.12)	(0.34)	(–0.40 to –0.10)

Results expressed as mean (standard deviation), and mean difference and 95% confidence intervals (CI) between measurement conditions. COPD: chronic obstructive pulmonary disease group; Healthy: healthy control group; QB: quiet breathing; CWC: chest wall compression; EFR: expiratory flow rates; PEFR: peak expiratory flow rates; Vt: tidal volume. **: Significant change between QB and CWC (P<0.01). ‡: Significant change between COPD group and healthy control group (P<0.01).

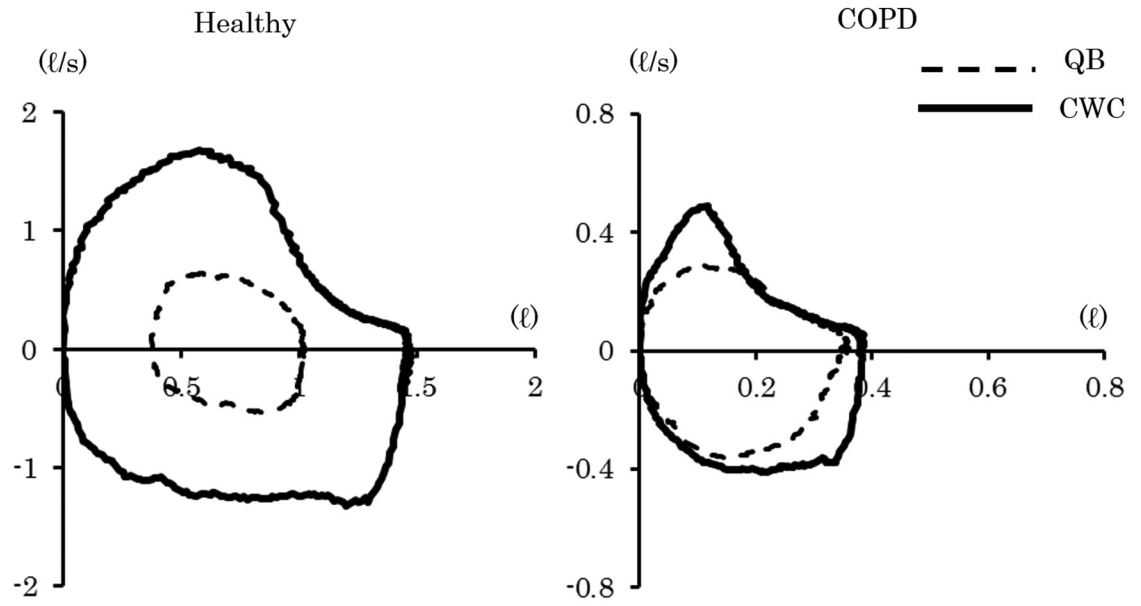

Figure 1. Representative flow-volume curves during QB and CWC. The left side shows a healthy subject, and the right side shows a COPD patient. QB: quiet breathing; CWC: chest wall compression; COPD: chronic obstructive pulmonary disease.

Discussion

In this study, the effects of CWC were examined in COPD patients, and it was confirmed that PEFR increased during CWC in these patients. However, PEFR changes during CWC were lower in COPD patients than in healthy subjects. The reason for the difference may have been the presence of expiratory flow limitation (EFL) in the COPD patients. With EFL, the expiratory flow rates never increase with increased pleural pressure, and this is often seen in patients with severe COPD[19]. Ninane et al.[20] reported assessment of EFL by studying whether the expiratory flow rates could be increased by increased pleural and abdominal pressures with abdominal compression during expiration. In the present study, CWC also increased pleural pressure by upper rib cage compression, but the participants whose expiratory flow rates did not increase during CWC probably had EFL, though the area of compression differed from that of Ninane et al.[20]. In fact, almost all COPD patients showed overlapping regions in the FV curves during QB and CWC, which suggests that it is difficult for COPD patients to increase expiratory flow rates during CWC. One more reason to consider for the discrepancy in PEFR changes between healthy and COPD participants is the increase in Vt with CWC. Generally, increasing Vt provides higher elastic forces at the start of exhalation, therefore healthy subjects have a tendency to increase expiratory flow rates more than COPD patients do.

In this study, expiratory flow rate changes during CWC were also examined using FV curves. PEFR and Vt have often been measured during various chest physical therapy techniques[2,4,7], but FV curves during these interventions have not been previously studied, except in a recent study on patients on mechanical ventilation[6]. The advantage of using FV curve assessment is to examine the expiratory flow changes at absolute volume in peripheral airway regions[3]. The expiratory flow rates at the lung volume at PEFR during QB and at 50% and 25% of tidal volume during QB increased only in the healthy control group, but not in the COPD group. Moreover, almost all COPD patients showed overlapping regions in the FV curves. This showed that it is difficult for COPD patients to increase expiratory flow in peripheral airways with CWC.

The present results showed that PEFR/PIFR did not change with CWC in the COPD group or in the healthy control group. McCarren and Alison[3] reported that PEFR/PIFR increased during vibration to 1.51 in CF patients. It was concluded that vibration was useful for CF patients because the PEFR/PIFR value needs to increase to more than 1.1 to remove secretions[3]. However, there were some differences between their methods and the present methods. They asked their subjects to inspire to total lung capacity as slowly as possible[3]. On the other hand,

they did not ask subjects to inspire slowly in their other study of normal healthy subjects[5]. In the present study, the participants were instructed to inspire freely, therefore it was not possible to determine inspiratory lung volume and breathing speed. McCarren et al.[5] reported that not only PEFR but also PIFR increased during CWC in normal subjects, so that PEFR/PIFR decreased more during CWC (to 0.64) than during QB (to 0.72). These results suggest that CWC must be combined with slow deep inspiration for effective removal of secretions. In the present results, Ti, Te, and Ttot also increased during CWC in both groups. Such changes were also seen during pursed-lips breathing (PLB) in COPD patients[21,22] because of the increased inspiratory effort during PLB. We hypothesized that increased inspiratory effort was the result of maintaining ventilation, because decreased respiratory frequency was seen not only during PLB but also CWC. Furthermore, we believe that the reason for the greater PIFR and Vt/Ti changes with CWC in the healthy control group compared to the COPD group was the presence of hyperinflation in COPD patients. It is known that lung hyperinflation leads to decreased inspiratory flow reserve by the increased elastic recoil pressure of the lung or the decreased inspiratory muscle-generated force[15].

In the present results, tidal volume was increased with CWC in both groups, but the increase was greater in the healthy control group than in the COPD group. We thought this difference was also explained by the differences in Vt/Ti between the groups. In contrast, IC increased with CWC in both groups. It is known that IC changes are associated with changes in end-expiratory lung volume, so this technique may reduce hyperinflation, such as that shown in PLB[21,22]. In fact, CWC has been used as a technique to reduce dyspnea by reducing hyperinflation, particularly in Japanese clinical practice[23].

Limitations of the study

CWC effects were measured in the supine position, but this position may affect the expiratory flow changes. Koulouris et al.[24] reported that COPD patients often have EFL, particularly in the supine position. The decreased functional residual capacity (FRC) in the supine position decreases elastic lung recoil and expiratory flow reserve. On the other hand, McCarren et al.[4] applied CWC to their subjects in the lateral recumbent position. FRC values are higher in the lateral recumbent position than in the supine position[25], so these differences may have affected the present results. However, the supine position was chosen because it provides more powerful chest compression, as shown by Toussaint et al.[26]. Moreover, CWC was applied to the upper rib cage in the present study, even though it was applied to the lower rib cage in other studies[2-5]. It is more difficult to apply CWC to the lower rib cage due to the presence of breasts in women; since there were subjects of both sexes in this sample, applying CWC to the upper rib cage was an alternative to achieve better standardization. Some of the healthy subjects in the present study also had overlapping regions in their FV curves. This may be due to age-related changes. The present subjects included many elderly persons, and their mean age was also very advanced. Since it is known that elastic lung recoil pressure and FRC decrease with age[27], some of the present healthy subjects may have had EFL. However, expiratory flow rates and tidal volume were different between the COPD group and the healthy control group. Thus, there appears to be a difference in the effects of CWC depending on the presence of COPD, even if age-related effects are excluded.

Another possible limitation was that changes in the amounts of secretions with CWC intervention and the longitudinal effects of CWC were not measured. Since the subjects were not asked to inspire to total lung capacity, the lung volume effects could not be controlled. It has been reported that chest physical therapy technique has low reliability[28]. In the present study, the effects of CWC performed by one physical therapist were examined, so that the same results may not be generalized to other physical therapists. Moreover, either the physical therapist or the assessor were not blinded. This lack of blinding may also affect the results of the study.

Clinical implications

In this study, peak expiratory flow rates, but not mean expiratory flow rates, increased in COPD patients. Volpe et al.[29] reported that peak expiratory and inspiratory flows are the key factors in secretion removal, not mean expiratory and inspiratory flows. Therefore, we believe that it would be effective to use CWC in patients with COPD to remove secretions when combined with slow inspiration, even though it would be less effective than in healthy subjects. We also showed that IC increased with CWC, so this technique may reduce hyperinflation and dyspnea.

Conclusion

The present results showed higher PEFR during CWC than during QB, both in COPD patients and in healthy subjects, but absolute expiratory flow rates during CWC increased only in healthy subjects, not in COPD patients. As a result, PEFR and Vt/Te increased with CWC more in healthy subjects than in COPD patients. Although CWC appears to be less effective in increasing absolute expiratory flow rates in COPD patients than in healthy subjects, the PEFR, which is more effective for removing secretions, did increase in the COPD group.

References

1. Branson RD. Secretion management in the mechanically ventilated patient. Respir Care. 2007;52(10):1328-42, discussion 1342-7. PMid:17894902.
2. Unoki T, Kawasaki Y, Mizutani T, Fujino Y, Yanagisawa Y, Ishimatsu S, et al. Effects of expiratory rib-cage compression on oxygenation, ventilation, and airway-secretion removal in patients receiving mechanical ventilation. Respir Care. 2005;50(11):1430-7. PMid:16253149.
3. McCarren B, Alison JA. Physiological effects of vibration in subjects with cystic fibrosis. Eur Respir J. 2006;27(6):1204-9. http://dx.doi.org/10.1183/09031936.06.00083605. PMid:16455834.
4. McCarren B, Alison JA, Herbert RD. Manual vibration increases expiratory flow rate via increased intrapleural pressure in healthy adults: an experimental study. Aust J Physiother. 2006;52(4):267-71. http://dx.doi.org/10.1016/S0004-9514(06)70006-X. PMid:17132121.
5. McCarren B, Alison JA, Herbert RD. Vibration and its effect on the respiratory system. Aust J Physiother. 2006;52(1):39-43. http://dx.doi.org/10.1016/S0004-9514(06)70060-5. PMid:16515421.
6. Guimarães FS, Lopes AJ, Constantino SS, Lima JC, Canuto P, Menezes SL. Expiratory rib cage compression, secretion clearance and respiratory mechanics in mechanically ventilated patients: A randomized crossover trial. Respir Care. 2014;59(5):678-85. http://dx.doi.org/10.4187/respcare.02587. PMid:24106324.
7. Genc A, Akan M, Gunerli A. The effects of manual hyperinflation with or without rib-cage compression in mechanically ventilated patients. Italian Journal of Physiotherapy. 2011;1(2):48-54.
8. Martí JD, Li Bassi G, Rigol M, Saucedo L, Ranzani OT, Esperatti M, et al. Effects of manual rib cage compressions on expiratory flow and mucus clearance during mechanical ventilation. Crit Care Med. 2013;41(3):850-6. http://dx.doi.org/10.1097/CCM.0b013e3182711b52. PMid:23314585.
9. MacLean D, Drummond G, Macpherson C, McLaren G, Prescott R. Maximum expiratory airflow during chest physiotherapy on ventilated patients before and after the application of an abdominal binder. Intensive Care Med. 1989;15(6):396-9. http://dx.doi.org/10.1007/BF00261500. PMid:2808898.
10. Gregson RK, Shannon H, Stocks J, Cole TJ, Peters MJ, Main E. The unique contribution of manual chest compression-vibrations to airflow during physiotherapy in sedated, fully ventilated children. Pediatr Crit Care Med. 2012;13(2):e97-102. http://dx.doi.org/10.1097/PCC.0b013e3182230f5a. PMid:21666530.
11. Unoki T, Kawasaki Y, Mizutani T, Fujino Y, Yanagisawa Y, Ishimatsu S, et al. Effects of expiratory rib-cage compression on oxygenation, ventilation, and airway-secretion removal in patients receiving mechanical ventilation. Respir Care. 2005;50(11):1430-7. PMid:16253149.
12. Wong WP, Paratz JD, Wilson K, Burns YR. Hemodynamic and ventilatory effects of manual respiratory physiotherapy techniques of chest clapping, vibration, and shaking in an animal model. J Appl Physiol. 2003;95(3):991-8. http://dx.doi.org/10.1152/japplphysiol.00249.2003. PMid:12754172.
13. Berti JS, Tonon E, Ronchi CF, Berti HW, Stefano LM, Gut AL, et al. Manual hyperinflation combined with expiratory rib cage compression for reduction of length of ICU stay in critically ill patients on mechanical ventilation. J Bras Pneumol. 2012;38(4):477-86. http://dx.doi.org/10.1590/S1806-37132012000400010. PMid:22964932.
14. Osadnik CR, McDonald CF, Holland AE. Airway clearance techniques in acute exacerbations of COPD: a survey of Australian physiotherapy practice. Physiotherapy. 2013;99(2):101-6. http://dx.doi.org/10.1016/j.physio.2012.01.002. PMid:23219648.
15. Johnson BD, Weisman IM, Zeballos RJ, Beck KC. Emerging concepts in the evaluation of ventilatory limitation during exercise: the exercise tidal flow-volume loop. Chest. 1999;116(2):488-503. http://dx.doi.org/10.1378/chest.116.2.488. PMid:10453881.
16. Miller MR, Crapo R, Hankinson J, Brusasco V, Burgos F, Casaburi R, et al. General considerations for lung function testing. Eur Respir J. 2005;26(1):153-61. http://dx.doi.org/10.1183/09031936.05.00034505. PMid:15994402.
17. Hanamoto S, Ohsuji T, Tsuyuguchi I, Kawabata S, Kimura K. Prediction formulas for pulmonary function tests expressed in linear and exponential form for healthy Japanese adults. Nihon Kyobu Shikkan Gakkai Zasshi. 1992;30(12):2051-60. PMid:1289624.
18. Nozoe M, Mase K, Murakami S, Okada M, Ogino T, Matsushita K, et al. Relationship between spontaneous expiratory flow-volume curve pattern and air-flow obstruction in elderly COPD patients. Respir Care. 2013;58(10):1643-8. http://dx.doi.org/10.4187/respcare.02296. PMid:23571516.
19. Koulouris NG, Kaltsakas G, Palamidas AF, Gennimata SA. Methods for assessing expiratory flow limitation during tidal breathing in COPD patients. Pulm Med. 2012; 2012: 1-8.
20. Ninane V, Leduc D, Kafi SA, Nasser M, Houa M, Sergysels R. Detection of expiratory flow limitation by manual compression of the abdominal wall. Am J Respir Crit Care Med. 2001;163(6):1326-30. http://dx.doi.org/10.1164/ajrccm.163.6.2004150. PMid:11371396.
21. Bianchi R, Gigliotti F, Romagnoli I, Lanini B, Castellani C, Binazzi B, et al. Patterns of chest wall kinematics during volitional pursed-lip breathing in COPD at rest. Respir Med. 2007;101(7):1412-8. http://dx.doi.org/10.1016/j.rmed.2007.01.021. PMid:17350815.

22. Bianchi R, Gigliotti F, Romagnoli I, Lanini B, Castellani C, Grazzini M, et al. Chest wall kinematics and breathlessness during pursed-lip breathing in patients with COPD. Chest. 2004;125(2):459-65. http://dx.doi.org/10.1378/chest.125.2.459. PMid:14769725.

23. Yoshimi K, Ueki J, Seyama K, Takizawa M, Yamaguchi S, Kitahara E, et al. Pulmonary rehabilitation program including respiratory conditioning for chronic obstructive pulmonary disease (COPD): Improved hyperinflation and expiratory flow during tidal breathing. J Thorac Dis. 2012;4(3):259-64. PMid:22754664.

24. Koulouris NG, Valta P, Lavoie A, Corbeil C, Chassé M, Braidy J, et al. A simple method to detect expiratory flow limitation during spontaneous breathing. Eur Respir J. 1995;8(2):306-13. http://dx.doi.org/10.1183/09031936.95.08020306. PMid:7758567.

25. Lumb AB, Nunn JF. Respiratory function and ribcage contribution to ventilation in body positions commonly used during anesthesia. Anesth Analg. 1991;73(4):422-6. http://dx.doi.org/10.1213/00000539-199110000-00010. PMid:1897767.

26. Toussaint M, Boitano LJ, Gathot V, Steens M, Soudon P. Limits of effective cough-augmentation techniques in patients with neuromuscular disease. Respir Care. 2009;54(3):359-66. PMid:19245730.

27. Janssens JP, Pache JC, Nicod LP. Physiological changes in respiratory function associated with ageing. Eur Respir J. 1999;13(1):197-205. http://dx.doi.org/10.1183/09031936.99.14614549. PMid:10836348.

28. Shannon H, Gregson R, Stocks J, Cole TJ, Main E. Repeatability of physiotherapy chest wall vibrations applied to spontaneously breathing adults. Physiotherapy. 2009;95(1):36-42. http://dx.doi.org/10.1016/j.physio.2008.08.004. PMid:19627684.

29. Volpe MS, Adams AB, Amato MBP, Marini JJ. Ventilation patterns influence airway secretion movement. Respir Care. 2008;53(10):1287-94. PMid:18811989.

Correspondence
Masafumi Nozoe
Konan Women's University
Faculty of Nursing and Rehabilitation
Department of Physical Therapy
6-2-23, Morikita-machi, Higashinada-ku
Kobe, Japan
e-mail: nozoe@konan-wu.ac.jp

original article

BJPT Brazilian Journal of Physical Therapy

Characteristics and prevalence of musculoskeletal injury in professional and non-professional ballet dancers

Michelle S. S. Costa[1], Arthur S. Ferreira[1], Marco Orsini[1], Elirez B. Silva[2], Lilian R. Felicio[3]

ABSTRACT | **Background:** Ballet is a high-performance activity that requires an advanced level of technical skills. Ballet places great stress on tendons, muscles, bones, and joints and may act directly as a trigger of injury by overuse. **Objectives:** 1) to describe the main types of injuries and affected areas related to classical ballet and 2) to compare the frequency of musculoskeletal injuries among professional and non-professional ballet dancers, considering possible gender differences among the professional dancers. **Method:** A total of 110 questionnaires were answered by professional and non-professional dancers. The questionnaire contained items related to the presence of injury, the regions involved, and the mechanism of the injury. **Results:** We observed a high frequency of musculoskeletal injuries, with ankle sprains accounting for 69.8% of injuries in professional dancers and 42.1% in non-professional dancers. Pirouettes were the most frequent mechanism of injury in professional dancers, accounting for 67.9% of injuries, whereas in the non-professional dancers, repetitive movement was the most common mechanism (28.1%). Ankle sprains occurred in 90% of the women's injuries, and muscle sprains occurred in 54.5% of the men's injuries. The most frequent injury location was the ankle joint in both sexes among the professional dancers, with 67.6% in women and 40.9% in men. **Conclusions:** The identification of the mechanism of injury and time of practice may contribute to better therapeutic action aimed at the proper function of the dancers' bodies and improved performance by these athletes.

Keywords: ballet dancers; prevalence; injuries; movement.

BULLET POINTS

- Pirouettes, repetitive movements, and landing jumps were the common mechanism of injury.
- Ankle sprains are the main injuries in professional and non-professional ballet dancers.
- Prevention programs should be conducted, with emphasis on exercises related to motor skills.

HOW TO CITE THIS ARTICLE

Costa MSS, Ferreira AS, Orsini M, Silva EB, Felicio LR. Characteristics and prevalence of musculoskeletal injury in professional and non-professional ballet dancers. Braz J Phys Ther. 2016 Mar-Apr; 20(2):166-175. http://dx.doi.org/10.1590/bjpt-rbf.2014.0142

Introduction

Dance is a series of movements in which the person moves in space and time to the rhythm of music. It is also defined as a specific expression of human motor behavior[1]. Ballet is a high-performance dance that requires an advanced level of technical skills. Ballet frequently places great stress on tendons, muscles, bones, and joints and thus may trigger injuries[2]. Ballet dancers are described as athletes because they can perform complex, physically demanding routines and are subjected to long periods of coaching[3], currently being compared to top athletes[2].

The need to understand the extent of injuries in dance has been a challenge due to methodological weaknesses involving the injuries and characteristics of the assessed individuals[3]. Sports literature has reports related to the prevalence and incidence of injuries, but there are few studies related to ballet[3-5]. Epidemiological studies of injuries in classical ballet indicate the length of performance as a major cause of injury, accounting for 40% to 80% of the lesions; however, these studies were based on amateur dancers and there was no control of the hours of daily practice[6-8]. In 1989, Bowling[4] observed that professional ballet

[1] Programa de Pós-graduação em Ciências da Reabilitação, Centro Universitário Augusto Motta (UNISUAM), Rio de Janeiro, RJ, Brazil
[2] Programa de Pós-graduação em Exercício e Ciência do Esporte, Universidade Gama Filho (UGF), Rio de Janeiro, RJ, Brazil
[3] Curso de Fisioterapia, Universidade Federal de Uberlândia (UFU), Uberlândia, MG, Brazil
Received: Mar. 20, 2015 Revised: Aug. 13, 2015 Accepted: Sept. 22, 2015

dancers had predominantly chronic injuries and that the cervical, lumbar, and ankle regions were the most affected, accounting for up to 2/3 of soft-tissue injuries. Most studies that investigated the incidence and prevalence of injuries in dance pointed to classical ballet as the dance modality with the highest technical demand[9-11]. Besides being the foundation for the performance of other forms of dance, it is the technique that has the highest rate of injuries[3,12-19].

In classical ballet, the occurrence of ankle sprains by trauma has been reported as the most frequent injury. Picon et al.[20] and Arendt and Kerschbaumer[21] reported that these lesions are related to the movements in which the dancers remain in the tip position, when high loads on the joints are required, especially during jumping and landing. According to Picon et al.[20] and Grego et al.[22], approximately 86% to 97.48% of injuries in classical ballet are related to the lower limb, especially in the joints of the foot and ankle, and 64% of these lesions were caused by micro-trauma repetition[12,20,22]. Furthermore, according to Monteiro and Grego[7], the knee and the hip correspond to 20% of regions with injury. Several factors are related to the onset and frequency of injuries in dance. However, muscle fatigue caused by overtraining, shows, and competitions seems to be one of the main triggering factors of injuries[12,13,15].

Because of the high number of injuries suffered by dancers and the recognized need to direct attention to the involvements posed by the practice of ballet, the study of characterization and frequency of injuries becomes an important ally to health professionals, identifying the main mechanisms of injury, besides collaborating in methods of injury prevention in this population. Thus, this study aims to 1) describe the main types of injuries and affected regions related to classical ballet and 2) to compare the frequency of musculoskeletal injuries among professional and non-professional ballet dancers, considering possible gender differences among the professional ones.

Method

Sample

This was a retrospective descriptive study. One hundred and ten classical dancers were evaluated in action in the State of Rio de Janeiro, Brazil. The sample included 22 men and 88 women (17.6±9.3 years of professional experience in classical ballet), of which 53 were professionals and 57 were non-professionals. The study was approved by the ethics committee of the institution (CAAE: 0030.0.307.000-11, Centro Universitário Augusto Motta-UNISUAM, Rio de Janeiro, RJ, Brazil) and all ballet dancers signed a consent form after being informed of the design and objectives of the study.

Questionnaire for prevalence of lesion

All participants were interviewed and completed a questionnaire containing items about how the injury occurred and its location. The items were selected from a self-administered questionnaire (Appendix A) based on Brooks et al.[23,24] and Allen et al.[3]. Three hundred questionnaires in total were distributed in three schools and one ballet company in Rio de Janeiro, and in two schools and one ballet company in Niterói, RJ, Brazil. 110 questionnaires were returned completely answered.

Statistical analysis

In the present study, the sample size was not calculated, and all of the 110 completely answered questionnaires were used. Descriptive data is presented using absolute and relative frequencies of categorical variables (gender; dominant body side; injury prevalence; types of injuries; location of injuries; mechanism of injury), grouped by involvement with ballet (professional and non-professional) and by gender in the professional group (male professionals and female professionals). Continuous variables (age; body mass index; hours/days of practice; years of practice; time off after injury) were presented as mean±SD because the null hypothesis of normality of distribution was not rejected using the Kolmogorov-Smirnov/Shapiro-Wilk test. For the comparative analysis, Student's t test and the Fisher test were conducted to compare averages and proportions, respectively, by adopting the value of $p<0.05$ as the significance level. The 95% confidence intervals were calculated to provide the precision of the statistical estimates. The software BioEstat version 5.3 was used for analysis.

Results

Professional versus non-professional ballet dancers

Anthropometric data and data from the training routine of male and female professional and non-professional dancers are described in Table 1. In the group of professional dancers, the mean age, hours of training, and practice time were higher than in the non-professional group. Regarding the

Table 1. Anthropometric data and data from routine of practices of professional and non-professional *ballet* dancers, as well as female and male professional dancers.

Group	Professional dancers			Non-professional dancers	P-value
Subgroup	Men	Women	All	All	
Sample size	22	31	53	57	-
Age (years)[C]	34.1±7.1 [31.2; 37.1]	34.2±6.3 [32.0; 36.4]	34.2±6.6 [32.4; 35.9]	25.0±9.4 [22.5; 27.4]	<0.001[A]; 0.990[B]
Body mass index (kg/m²)[C]	23.6±1.1 [23.1; 24.1]	19.5±1.1 [19.1; 19.9]	21.2±2.3 [20.6; 21.8]	20.7±2.5 [20.0; 21.3]	0.250[A]; <0.001[B]
Right dominance[D]	19 (86.4) [72.0; 100.7]	28 (90.3) [79.9; 100.7]	47 (88.7) [80.1; 97.2]	55 (96.5) [91.7; 101.3]	0.151[A]; 0.683[B]
Practicing time (hours/day)[C]	5.8±0.9 [5.4; 6.2]	5.9±1.0 [5.5; 6.2]	5.8±1.0 [5.6; 6.1]	2.6±1.7 [2.2; 3.1]	<0.001[A]; 0.713[B]
Practice experience (years)[C]	18.1±5.4 [15.8; 20.4]	26.3±6.5 [24.0; 28.6]	22.9±7.3 [20.9; 24.9]	12.8±8.7 [10.6; 15.1]	<0.001[A]; <0.001[B]

Data shown as mean±SD or n (%) with 95% confidence interval [lower; upper]. [A]Comparison between professional and non-professional ballet dancers. [B]Comparison between women and men professional ballet dancers. [C]Two-sample student's t-test, unequal variance (two-tailed). [D]Fisher's exact test.

professional group, the difference happens in the Body Mass Index (BMI) and years of practice, whereas the BMI was higher in male dancers and practicing time higher in females.

Regarding the most frequent types of injuries, ankle sprains correspond to 69.8% of injuries in the professional dancers and 42.1% in the non-professional ones, followed by muscle contractures in both groups. On the other hand, dislocations and subluxations were the least common lesions in both groups (Table 2).

In relation to the location of the injury, the ankle corresponded to 56.6% in the professional dancers and 35.1% in the non-professionals. The knee was the second most affected region in the group of professional dancers, and the region of the hip and leg in non-professional dancers (Table 2).

Among the mechanisms of injury, the *pirouettes* were the most common in professional dancers, but in the non-professional ones, repetitive movements. However, the group of professional dancers had a higher frequency of this mechanism when compared to the non-professional group (Table 2).

In the present study, 36.4% (n=16) of the non-professional dancers and 83% (n=44) of professionals received some kind of physical therapy.

Female versus male professional ballet dancers

All of the evaluated professional dancers, regardless of sex, had already suffered some injury as a result of ballet practice, however women have an average of post-injury off sport greater than men (p=0.009). Regarding the main types of injuries, sprains are more common in women (p=0.002) and muscle strains and sprains in men (p=0.001). The most affected region in both groups was the ankle, however women have a higher rate of injury than men in this region. Women have a high frequency of involvement in the knee, but no difference between groups was observed, since men have a high frequency of injury in the lumbar region when compared to women (Table 3).

Among the mechanisms of injury, *pirouettes* were more frequent in women, while in men, repetitive movements were the most common mechanism (Table 3).

Among the professional dancers who had injuries due to the practice of ballet, 90% (n=27) of women and 91.3% (n=21) of professional men received some kind of physical therapy.

Discussion

This study described the characteristics and frequency of injuries occurring in professional and non-professional ballet dancers, taking into consideration the characteristics and mechanisms of injury by sex and length of practice, in addition to the injured area.

The time in the profession with a company can provide challenges for ballet dancers, including the need to develop technical knowledge as well as achieve strength and required levels of technical execution. These demands may contribute to the increased rate of injuries in these ballet dancers compared to

Table 2. Data about the kind and location of the lesion between professionals and non-professionals.

Group (n)	Professionals (n=53)	Non professionals (n=57)	P-value
Dancers with injuries caused by the dance	53 (100.0) [100.0; 100.0]	44 (77.2) [66.3; 88.1]	<0.001[B]
Time off after injury - days (mean / SD)	73.0±108.3 [43.8; 102.1]	26.1±71.1 [7.7; 44.6]	0.009[A]
Type of lesion			
Sprains	37 (69.8) [67.3; 82.2]	24 (42.1) [29.3; 54.9]	0.004[B]
Muscle contractures	19 (35.8) [33.2; 48.8]	17 (29.8) [17.9; 41.7]	0.546[B]
Others (e.g. muscle strain. direct trauma)	15 (28.3) [16.2; 40.4]	4 (7.0) [0.4; 13.6]	0.005[B]
Fractures	8 (15.1) [5.5; 24.7]	6 (10.5) [2.6; 18.5]	0.572[B]
Luxations	3 (5.7) [–0.6; 11.9]	3 (5.3) [–0.5; 11.1]	1.000[B]
Subluxations	3 (5.7) [–0.6; 11.9]	0 (0.0) [0.0; 0.0]	0.109[B]
Location of Lesions			
Ankle	30 (56.6) [43.3; 69.9]	20 (35.1) [22.7; 47.5]	0.035[B]
Knee	14 (26.4) [14.5; 38.3]	4 (7.0) [0.4; 13.6]	0.009[B]
Thigh and Leg	12 (22.6) [11.4; 33.9]	9 (15.8) [6.3; 25.3]	0.468[B]
Lumbar	12 (22.6) [11.4; 33.9]	3 (5.3) [–0.5; 11.1]	0.011[B]
Foot	8 (15.1) [5.5; 24.7]	7 (12.3) [3.8; 20.8]	0.783[B]
Shoulder	4 (7.5) [0.4; 14.7]	1 (1.8) [–1.7; 5.2]	0.194[B]
Hip and Pelvis	4 (7.5) [0.4; 14.7]	7 (12.3) [3.8; 20.8]	0.530[B]
Cervical	3 (5.7) [–0.6; 11.9]	1 (1.8) [–1.7; 5.2]	0.350[B]
Face	2 (3.8) [–1.4; 8.9]	0 (0.0) [0.0; 0.0]	0.230[B]
Elbow, forearm, wrist and hand	3 (5.7) [–0.6; 11.9]	0 (0.0) [0.0; 0.0]	0.109[B]
Thorax and abdomen	0 (0.0) [0.0; 0.0]	0 (0.0) [0.0; 0.0]	1.000[B]
Mechanism of injury			
Recurrence of injury	42 (79.2) [68.3; 90.2]	14 (24.6) [13.4; 35.7]	<0.001[B]
Spins (*pirouette*)	36 (67.9) [55.4; 80.5]	13 (22.8) [11.9; 33.7]	<0.001[B]
Repetitive movement	23 (43.4) [30.1; 56.7]	16 (28.1) [16.4; 39.7]	0.112[B]
Fall after jump	10 (18.9) [8.3; 29.4]	6 (10.5) [2.6; 18.5]	0.282[B]
Fall	3 (5.7) [–0.6; 11.9]	2 (3.5) [–1.3; 8.3]	0.671[B]

Data shown as mean±SD or n (%) with 95% confidence interval [lower; upper] ($p<0.05$). [A]Two-sample student's t-test, unequal variance (two-tailed). [B]Fisher's exact test.

Table 3. Data about the type and location of the lesion of professional female and male dancers.

Group (n)	Men (n=22)	Women (n=31)	P-value
Dancers with dancing injuries	22 (100.0) [100.0; 100.0]	31 (100.0) [100.0; 100.0]	1.000[B]
Time off after injury - days (mean / SD)	43.0±57.6 [19.0; 67.1]	94.2±129.9 [48.5; 139.9]	0.009[A]
Type of lesion			
Sprains	10 (45.5) [41.3; 66.3]	27 (87.1) [84.7; 98.9]	0.002[B]
Muscle contractures	11 (50.0) [45.7; 70.9]	8 (25.8) [22.7; 41.2]	0.088[B]
Others (e.g. muscle strain. direct trauma)	12 (54.5) [33.7; 75.4]	3 (9.7) [–0.7; 20.1]	0.001[B]
Fractures	2 (9.1) [–2.9; 21.1]	6 (19.4) [5.4; 33.3]	0.445[B]
Luxations	0 (0.0) [0.0; 0.0]	3 (9.7) [–0.7; 20.1]	0.258[B]
Subluxations	1 (4.5) [–4.2; 13.2]	2 (6.5) [–2.2; 15.1]	1.000[B]
Location of Lesions			
Ankle	9 (40.9) [20.4; 61.5]	21 (67.7) [51.3; 84.2]	0.091[B]
Knee	4 (18.2) [2.1; 34.3]	10 (32.3) [15.8; 48.7]	0.348[B]
Thigh and Leg	7 (31.8) [12.4; 51.3]	5 (16.1) [3.2; 29.1]	0.202[B]
Lumbar	10 (45.5) [24.6; 66.3]	2 (6.5) [–2.2; 15.1]	0.002[B]
Foot	1 (4.5) [–4.2; 13.2]	7 (22.6) [7.9; 37.3]	0.120[B]
Shoulder	2 (9.1) [–2.9; 21.1]	2 (6.5) [–2.2; 15.1]	1.000[B]
Hip and Pelvis	1 (4.5) [–4.2; 13.2]	3 (9.7) [–0.7; 20.1]	0.633[B]
Cervical	2 (9.1) [–2.9; 21.1]	1 (3.2) [–3.0; 9.4]	0.563[B]
Face	0 (0.0) [0.0; 0.0]	2 (6.5) [–2.2; 15.1]	0.505[B]
Elbow, forearm, wrist, and hand	1 (4.5) [–4.2; 13.2]	2 (6.5) [–2.2; 14.9]	1.000[B]
Thorax and abdomen	0 (0.0) [0.0; 0.0]	0 (0.0) [0.0; 0.0]	1.000[B]
Mechanism of injury			
Recurrence of injury	15 (68.2) [48.7; 87.6]	27 (87.1) [75.3; 98.9]	0.168[B]
Spins (*pirouette*)	9 (40.9) [20.4; 61.5]	27 (87.1) [75.3; 98.9]	0.001[B]
Repetitive movement	16 (72.7) [54.1; 91.3]	7 (22.6) [7.9; 37.3]	0.001[B]
Fall after jump	4 (18.2) [2.1; 34.3]	6 (46.2) [5.4; 33.3]	1.000[B]
Fall	2 (9.1) [–2.9; 21.1]	1 (3.2) [–3.0; 9.4]	0.563[B]

Data shown as mean±SD or n (%) with 95% confidence interval [lower; upper] ($p<0.05$). [A]Two-sample student's t-test, unequal variance (two-tailed). [B]Fisher's exact test.

non-professionals ones[3,25,26]. Our data agree with these authors, because it was observed that professional ballet dancers with 5.6 hours of training/day presented higher frequency of injury in relation to the non-professional group with less training time.

Our data agree with Ekegren et al.[27], who compared 266 elite ballet dancers with varied hours of daily training. These authors reported that, with increasing training duration, there was also an increased risk of injury, particularly due to overuse[27]. However, they did not assess the relationship between the training duration and region or type of injuries. Our data show the same behavior in professional ballet dancers in spite of the overall low frequency of injuries and high frequency of both sprains and muscle injury to the lower limbs. Altogether, these characteristics suggest that these regions are mainly subjected to overload, regardless of the training duration of dancers.

Like most sports, ballet is a strenuous activity and requires good athletic condition. However, in ballet, the movement is the final objective and requires exhaustive repetition, which increases the risk of musculoskeletal injury[17,28]. Due to the determination of ballet dancers, choreographers, and ballet masters to achieve "perfection" in the movements during training, many dancers suffer injuries prior to presentations and performances.

In the present study, we found a higher frequency of injury due to repetitive movements in professional dancers compared to non-professionals. In the group of professionals, men had a higher rate of injury with this mechanism because of the lifts and wide movements and jumps that they perform[29]. According to Guimarães and Simas[13], both in professional and non-professional ballet dancers, the repetition of a particular part of choreography or isolated movement performed even after fatigue is responsible for several injuries.

Among the movements performed during training and performances, we highlight the *pointe*, *demi-pointe*, and *en dehors*. These positions can cause joint overload, as well as ligament and muscle micro-traumas, especially in the ankle and knee regions, increasing the likelihood of injury[30], as observed in this study. Thus, our data strongly suggest the inclusion of exercise programs to prevent injury, including exercises to improve the muscle strength of the lower limbs as well as the back muscles, which are the most frequently affected regions.

The various *pointe* (tiptoe) positions in ballet vary according to the position of the feet and reduce the base of support, which requires great muscle and neurophysiological effort[30]. Added to these positions, the impact of falls after jumps and *pirouettes* is related to severe ligament and musculoskeletal injuries, especially to the lower limbs[31].

The present study indicates that, among the regions affected in professional dancers, the ankle is the most common. In addition, professional women are the most affected. Our results agree with those observed by other authors; however, these authors reported about 86% of lower limb injuries, not discriminating the location of injuries in the lower limb[7,9,13,20]. Studies carried out in international dance companies confirm the results of this study, with ankle sprain as the highest prevalence of injuries in dancers[17-19,32].

The frequency found for ankle sprains occurred in dancers aged 34.7±6.9 years of age. Paradoxically, in a study conducted in a Swiss ballet company, 75% of ankle sprains occurred in young ballet dancers under 26 years of age, showing that the more experienced the dancer is, the better the technique and the lower the risk of injury[17]. This fact was not observed in our study, in which 69.8% of sprains occurred in professional dancers with an average of 22.9 years of practice. These differences may be related to the presence of specialized medical teams in sport and the presence of injury prevention programs in ballet companies overseas, which does not happen in national companies. In addition, the specific gesture movements related to the practice of ballet should also be assessed to verify possible relationships between posture misalignments during performance and specific injuries.

Hillier et al.[19] reported that men are more affected by trauma and women by repetitive stress, especially in the foot and ankle; however, our data do not agree with these authors, as professional women suffered more injuries during the *pirouette* when compared to men. Generally, the behavior of the characteristics and the frequency of the lesion between men and women are similar.

Our data indicate that, among professionals, 18.9% of the injury mechanisms are related to landing from jumps, with no differences between the sexes. As in the non-professional ballet dancers, injuries related to landing correspond to 10.5%. The jumps are highly complex movements that are performed most often by professional dancers.

Dore and Guerra[12] evaluated musculoskeletal symptoms and observed a dominant prevalence of low back pain, followed by knee pain. Our results disagree with those reported by these authors, since we observed a low frequency of lumbar complaint

among professionals. On the other hand, our data show the knee as the second affected region among the professionals. Guimarães and Simas[13] described several factors that could contribute to the overloading of the knee joint, particularly improper training, repetitive jumps, spins, and the *plié* position, causing valgus knee and excessive hyperextension. The authors also reported that both the ankle joint and the knee suffer high stress during the practice of ballet.

Gamboa et al.[5] also found a high frequency of ankle and knee injuries among professional ballet dancers, which is in agreement with our results. Moreover, these authors found an association between pronation of the subtalar joint and the prevalence of injury, overload to the knee, and risk of sprain of the anterior cruciate ligament. In spite of the homogeneity of our sample of professional ballet dancers, sports literature on this subject points to a large heterogeneity in type and location of injuries. Thus, ballet dancers should be screened with specific functional tests related to their performance, contributing to personalized prevention of musculoskeletal injuries.

Furthermore, our study showed that 90% of the female and 91.3% of the male professional dancers had undergone some kind of physical therapy, despite a high frequency of recurrence of injuries. Although we did not assess the kind of physical therapy treatment each dancer received, our data suggest the implementation of not only rehabilitative treatment, but also preventive treatment. Therefore, the therapist who is familiar with the prevalence and the mechanism of injury is better equipped to assess the type of injury involved and its extension and to design interventions that are more suitable for this type of athlete[3].

The interpretation of the data from this study should take into consideration the absence of the sample size calculation; however, the specificity and technical quality of the sample should be considered. Moreover, the frequency of the injuries was determined according to the self-report of the dancers, which could affect comparisons between studies that considered the diagnosis given by expert professionals.

Based on these findings, we can conclude that non-professional and professional ballet dancers have a high frequency of musculoskeletal injuries, with the ankle being the most affected joint due to spinning and repetitive movements. We can also conclude that gender did not affect the frequency of lesions. Therefore, technical training sessions should be emphasized to promote more appropriate movements from the biomechanical point of view, which is expected to decrease the risk of injury by repetitive movements. Moreover, physical therapy follow-up should be considered in the rehabilitation process of ballet dancers to account for their specific motor skills (e.g. jumping, *pirouettes*) and to improve muscle function in body regions with increased risk of injuries (e.g. lumbar spine and lower limbs).

In summary, our data supports better therapeutic action in the intervention of injuries, such as prevention programs, taking into consideration the dancer's movements.

References

1. Hugel F, Cadopi M, Kohler F, Perrin P. Postural control of ballet dancers: a specific use of visual input for artistic purposes. J Sports Med. 1999;20(2):86-92. PMid:10190767.
2. Batista CG, Martins EO. The prevalence of pain in classical ballet dancers. J Health Sci Inst. 2010;28(1):47-9.
3. Allen N, Nevill A, Brooks J, Koutedakis Y, Wyon M. Ballet Injuries: injury incidence and severity over 1 year. J Orthop Sports Phys Ther. 2012;42(9):781-90. http://dx.doi.org/10.2519/jospt.2012.3893. PMid:22814244.
4. Bowling A. Injuries to dance: prevalence, treatment, and perceptions of causes. BMJ. 1989;298(6675):731-4. http://dx.doi.org/10.1136/bmj.298.6675.731. PMid:2496824.
5. Gamboa JM, Roberts LA, Maring J, Fergus A. Injury patterns in elite preprofessional ballet dancers and the utility of screening programs to identify risk characteristics. J Orthop Sports Phys Ther. 2008;38(3):126-36. http://dx.doi.org/10.2519/jospt.2008.2390. PMid:18383646.
6. Camargo HCF, Ghirotto FMS. Uma visão da dança e suas lesões. Rev Bras Ciências da Saúde. 2003;1:32-5.
7. Monteiro HL, Grego LG. As lesões na dança: conceitos, sintomas, causa situacional e tratamento. Motriz. 2003;9:63-71.
8. Wiesler ER, Hunter DM, Martin DF, Curl WW, Hoen H. Ankle flexibility and injury patterns in dancers. Am J Sports Med. 1996;24(6):754-7. http://dx.doi.org/10.1177/036354659602400609. PMid:8947396.
9. Picon AP, Franchi SS. Análise antropométrica dos pés de praticantes de ballet clássico que utilizam sapatilhas de ponta. Revista Uniara. 2007;20:177-88.
10. Aquino CF, Cardoso VA, Machado NC, Franklin JS, Augusto VG. Análise da relação entre dor lombar e desequilíbrio de força muscular em bailarinas. Fisioter Mov. 2010;23(3):399-408. http://dx.doi.org/10.1590/S0103-51502010000300007.
11. Simões RD, Anjos AFP. O ballet clássico e as implicações anatômicas e biomecânicas de sua prática para os pés e tornozelo. Revista da Faculdade de Educação Física da UNICAMP. 2010;8:117-132.
12. Dore BF, Guerra RO. Sintomatologia dolorosa e fatores associados em bailarinos profissionais. Rev Bras Med Esporte. 2007;13(2):77-80. http://dx.doi.org/10.1590/S1517-86922007000200002.
13. Guimarães ACA, Simas JPN. Injuries in classical ballet. Revista de Educação Física. 2001;12:89-96.
14. Lin CF, Lee IJ, Liao JH, Wu HW, Su FC. Comparison of Postural Stability Between Injured and Uninjured Ballet

Dancers. Am J Sports Med. 2011;39(6):1324-31. http://dx.doi.org/10.1177/0363546510393943. PMid:21335350.

15. Solomon R, Micheli LJ, Solomon J, Kelley T. The "cost" of injuries in a professional ballet company: anatomy of a season. Med Probl Perform Art. 1995;10:3-10.
16. Nilsson C, Leanderson J, Wykman A, Strender LE. The injury panorama in a Swedish professional ballet company. Knee Surg Sports Traumatol Arthrosc. 2001;9(4):242-6. http://dx.doi.org/10.1007/s001670100195. PMid:11522083.
17. Bronner S, Ojofeitimi S, Mayers L. Comprehensive surveillance of dance injuries: a proposal for uniform reporting guidelines for professional companies. J Dance Med Sci. 2006;10:69-80.
18. Byhring S, Bo K. Musculoskeletal injuries in the Norwegian National Ballet: a prospective cohort study. Scand J Med Sci Sports. 2002;12(6):365-70. http://dx.doi.org/10.1034/j.1600-0838.2002.01262.x. PMid:12453164.
19. Hillier JC, Peace K, Hulme A, Healy JC. MRI features of foot and ankle injuries in ballet dancers. Br J Radiol. 2004;77(918):532-7. http://dx.doi.org/10.1259/bjr/60447506. PMid:15151980.
20. Picon AP, Lobo da Costa PH, Sousa F, Sacco ICN, Amadio AC. Biomecânica e "ballet" clássico: uma avaliação de grandezas dinâmicas do "sauté" em primeira posição e da posição "en pointe" em sapatilhas de ponta. Rev Paul Educ Fis. 2002;16:53-60.
21. Arendt YD, Kerschbaumer F. Injury and overuse pattern in professional ballet dancers. Z Orthop Ihre Grenzgeb. 2003;141(3):349-56. PMid:12822086.
22. Grego LG, Monteiro HL, Padovani CR, Gonçalves A. Lesões na dança: estudo transversal híbrido em academias da cidade de Bauru-SP. Rev Bras Med Esporte. 1999;5(2):47-54. http://dx.doi.org/10.1590/S1517-86921999000200003.
23. Brooks JH, Fuller CW, Kemp SP, Reddin DB. Epidemiology of injuries in english professional rugby union: part 1 match injuries. Br J Sports Med. 2005;39(10):757-66. http://dx.doi.org/10.1136/bjsm.2005.018135. PMid:16183774.
24. Brooks JH, Fuller CW, Kemp SP, Reddin DB. Epidemiology of injuries in english professional rugby union: part 2 training injuries. Br J Sports Med. 2005;39(10):767-75. http://dx.doi.org/10.1136/bjsm.2005.018408. PMid:16183775.
25. Grego LG, Monteiro HL, Gonçalves A, Aragon FF, Padovani CR. Agravos musculoesqueléticos em bailarinas clássicas, não clássicas e praticantes de educação física. Arq Ciênc Saúde. 2006;13:153-6.
26. Hamilton WG, Hamilton LH, Marshall P, Molnar M. A profile of the musculoskeletal characteristics of professional ballet dancers. Am J Sports Med. 1992;20(3):267-73. http://dx.doi.org/10.1177/036354659202000306. PMid:1636856.
27. Ekegren CL, Quested R, Brodrick A. Injuries in pre-professional ballet dancers: Incidence, characteristics and consequences. J Sci Med Sport. 2014;17(3):271-5. http://dx.doi.org/10.1016/j.jsams.2013.07.013. PMid:23988783.
28. Toledo SD, Akuthota V, Drake DF, Nadler SF, Chou LH. Sports and performing arts medicine: issues relating to dancers. Arch Phys Med Rehabil. 2004;85(3 Suppl 1):75-8. http://dx.doi.org/10.1053/j.apmr.2003.12.004. PMid:15034860.
29. Golomer E, Crémieux J, Dupui P, Isableu B, Ohlmann T. Visual contribution to self-induced body sway frequencies and visual perception of male professional dancers. Neurosci Lett. 1999;267(3):189-92. http://dx.doi.org/10.1016/S0304-3940(99)00356-0. PMid:10381008.
30. Costa PHL, Nora FGSA, Vieira MF, Bosch K, Rosenbaum D. Single leg balancing in ballet: effects of shoe conditions and poses. Gait Posture. 2013;37(3):419-23. http://dx.doi.org/10.1016/j.gaitpost.2012.08.015. PMid:22989743.
31. Barcellos C, Imbiriba LA. Alterações posturais e do equilíbrio corporal na primeira posição em ponta do balé clássico. Rev Paul Educ Fís. 2002;16:43-52.
32. Solomon R, Brown T, Gerbino PG, Micheli LJ. The young dancer. Clin Sports Med. 2000;19(4):717-39. http://dx.doi.org/10.1016/S0278-5919(05)70234-9. PMid:11019737.

Correspondence
Lilian Ramiro Felicio
Universidade Federal de Uberlândia
Curso de Fisioterapia
Rua Benjamim Constant, 1286
CEP 38400-678, Uberlândia, MG, Brazil
e-mail: lilianrf@uol.com.br

Appendix A. Questionnaire used to collect information of the study.

I-Identification

Name: _____

Sex: F ( ) M ( ) Ethnicity _____Marital Status_____

Date of birth: ___/____/_____ Age:_____

Weight:_____ Height: _____

Educational history: _____ Profession:_____

Address:_____

Telephone number: _____

II- Aspects related to dance

Dance company: _____

Age when you started dancing: _____

Age when you turned professional:_____

How many days per week and daily hours do you devote to practice?_____

Type of class: ( ) Ballet ( ) Contemporary ( ) Stretching

( ) Relaxing ( ) Others _____

If ballet, with or without pointe shoe? _____

Duration of classes:_____

Kind of rehearsal: ( ) Classic ( ) Contemporary

( ) Others_____

Duration of rehearsal: _____

Are you dancing at the moment?

( ) Yes ( ) No Reason_____

Do you practice any other physical activity?

( ) Yes ( ) No Which one? _____

If yes, specify the days of the week and daily hours devoted to the activity_____

How long have you practiced this activity? _____

III- Aspects related to injury.

Have you ever suffered any injury? Which one?

1. ( ) Fracture 2. ( ) Strain 3. ( ) Subluxation 4. ( ) Sprain

5. ( ) Muscular strain Location_____

6. ( ) Others _____

Have you had an injury during dance practice?

( ) Yes. Which one?_____

( ) No

If yes, how long ago?

( ) 6 months

( ) 6 months - 1 year

( ) 1 - 5 years

( ) Over 5 years

How did the injury happen? (movement performed)

_____

According to the table below, specify the location of injuries:_____

Did these lesions happen again?

Yes ( ) No ( )

Specify the number of times (each one): _____

How long were you away because of each injury?

_____

Have you had any kind of physical therapy for this lesion?
( ) Yes ( ) No.

Which one? _____
For how long?

_____

After injury, how long was it before you resumed the dance practice?

_____

Do you feel any kind of musculoskeletal pain during dance practice?

_____

IV- Check the table below for the locations and level of your pain.

		Not present	Apparent			Moderate			Severe		Unbearable	
A	Face	0	1	2	3	4	5	6	7	8	9	10
B	Neck	0	1	2	3	4	5	6	7	8	9	10
C	Abdomen	0	1	2	3	4	5	6	7	8	9	10
D	Right shoulder/Upper chest.	0	1	2	3	4	5	6	7	8	9	10
E	Left shoulder/ Upper Chest	0	1	2	3	4	5	6	7	8	9	10
F	Right elbow	0	1	2	3	4	5	6	7	8	9	10
G	Left elbow	0	1	2	3	4	5	6	7	8	9	10
H	Right forearm	0	1	2	3	4	5	6	7	8	9	10
I	Left forearm	0	1	2	3	4	5	6	7	8	9	10
J	Hand/right fist	0	1	2	3	4	5	6	7	8	9	10
K	Hand/left fist	0	1	2	3	4	5	6	7	8	9	10
L	Lumbar spine	0	1	2	3	4	5	6	7	8	9	10
M	Pelvic region	0	1	2	3	4	5	6	7	8	9	10
N	Buttocks	0	1	2	3	4	5	6	7	8	9	10
O	Hip/right thigh	0	1	2	3	4	5	6	7	8	9	10
P	Hip/right thigh	0	1	2	3	4	5	6	7	8	9	10
Q	Right Knee	0	1	2	3	4	5	6	7	8	9	10
R	Left knee	0	1	2	3	4	5	6	7	8	9	10
S	Right leg	0	1	2	3	4	5	6	7	8	9	10
T	Left leg	0	1	2	3	4	5	6	7	8	9	10
U	Right ankle	0	1	2	3	4	5	6	7	8	9	10
V	Left ankle	0	1	2	3	4	5	6	7	8	9	10
W	Right foot	0	1	2	3	4	5	6	7	8	9	10
X	Left foot	0	1	2	3	4	5	6	7	8	9	10

original article

BJPT Brazilian Journal of Physical Therapy

Reliability of measuring pectoralis minor muscle resting length in subjects with and without signs of shoulder impingement

Dayana P. Rosa[1], John D. Borstad[2], Elisa D. Pires[3], Paula R. Camargo[3]

ABSTRACT | **Background:** Pectoralis minor adaptive shortening may change scapula resting position and scapular kinematics during arm elevation. A reliable and clinically feasible method for measuring pectoralis minor length will be useful for clinical decision making when evaluating and treating individuals with shoulder pain and dysfunction. **Objectives:** To evaluate intrarater, interrater, and between-day reliability of a pectoralis minor (PM) muscle length measurement in subjects with and without signs of shoulder impingement. **Method:** A convenience sample of 100 individuals (50 asymptomatic and 50 symptomatic) participated in this study. Intra- and interrater reliability of the measurement was estimated in 50 individuals (25 asymptomatic and 25 symptomatic), and between-day reliability of the measurement repeated over an interval of 7 days was estimated in an independent sample of 50 additional participants. Pectoralis minor length was measured using a flexible tape measure with subjects standing. **Results:** Intraclass correlation coefficients ($ICC_{3,k}$) for intrarater and interrater reliability ranged from 0.86-0.97 and 0.95 for between-day reliability in both groups. Standard error of measurements (SEM) ranged from 0.30-0.42 cm, 0.70-0.84 cm, and 0.40-0.41 cm for intrarater, interrater, and between-day reliability, respectively, across the sample. The minimal detectable change (MDC) for between-day measurements ranged from 1.13-1.14 cm for both groups. **Conclusions:** In asymptomatic individuals and in those with signs of shoulder impingement, a single rater or pair of raters can measure pectoralis minor muscle length using a tape measure with very good reliability. This measurement can also be reliably used by the same rater over a seven day interval.

Keywords: measurement; scapula; physical therapy; tightness.

BULLET POINTS

- Pectoralis Minor length can be reliably measured using a tape measure.
- The presence of shoulder pain does not decrease reliability.
- Reliability is very good for measurements taken over a seven-day interval.
- Reliability estimates should not be generalized to different shoulder conditions.

HOW TO CITE THIS ARTICLE

Rosa DP, Borstad JD, Pires ED, Camargo PR. Reliability of measuring pectoralis minor muscle resting length in subjects with and without signs of shoulder impingement. Braz J Phys Ther. 2016 Mar-Apr; 20(2):176-183. http://dx.doi.org/10.1590/bjpt-rbf.2014.0146

Introduction

The pectoralis minor (PM) muscle attaches to the coracoid process of the scapula and inserts on ribs three, four, and five near the costosternal junction. It is the only scapulothoracic muscle with an anterior thoracic attachment[1,2]. It has been theorized that a habitual forward shoulder posture causes an adaptive length decrease of the pectoralis minor, which may subsequently contribute to movement alterations and/or shoulder pain[3-5]. Repetitive use of the upper extremity for activities that protract and downwardly rotate the scapula may also contribute to adaptive shortening[6,7]. The muscles' orientation determines that it will produce scapular downward rotation, anterior tilt, and internal rotation when it activates, and it is therefore an antagonist to upward rotation, posterior tilt, and external rotation, which are considered to be normal during arm elevation[8]. In support of this construct, pectoralis minor adaptive shortening has been associated with changes in the resting position of the scapula[3] and altered scapular kinematics during

[1] Programa de Pós-gradução em Fisioterapia, Universidade Metodista de Piracicaba (UNIMEP), Piracicaba, SP, Brazil
[2] Department of Physical Therapy, The Ohio State University (OSU), Columbus, OH, USA
[3] Departamento de Fisioterapia, Universidade Federal de São Carlos (UFSCar), São Carlos, SP, Brazil
Received: Apr. 02, 2015 Revised: Oct. 01, 2015 Accepted: Nov. 11, 2015

arm elevation[1]. Specifically, a group of asymptomatic subjects with relatively short pectoralis minor muscle resting length had decreased scapular posterior tilting and external rotation during arm elevation when compared to those with a relatively long muscle resting length[1]. With PM resting length identified as a potential contributor to detrimental shoulder kinematics, a reliable clinical assessment of resting length will be valuable for clinicians as they plan interventions and assess the effect of those interventions.

Measuring PM muscle length with an electromagnetic motion capture system using the coracoid process and the fourth rib as origin-insertion landmarks has shown excellent validity[9] and is considered the "gold standard" method. However, the electromagnetic system is time-consuming, expensive, not typically available to clinicians, and mainly used for research purposes. Thus, there is a need for a more clinically feasible instrument to assess PM length in subjects with postural deviations or shoulder dysfunction[3]. A tape measure and caliper both demonstrated good reliability with the electromagnetic system (Intraclass Correlation Coefficient – $ICC_{3,1}$ ranging from 0.82-0.87) to measure pectoralis minor muscle length within the same day by the same rater[1,9]. Although both tools had good reliability with the electromagnetic system, a tape measure is more readily available and easily manipulated in clinical practice.

Struyf et al.[10] recently used a caliper to measure the length of the pectoralis minor and demonstrated excellent ($ICC_{2,1}$ ranging from 0.87-0.93) and good ($ICC_{2,1}$ ranging from 0.76-0.87) intrarater reliability when reporting the Pectoralis Minor Index (PMI) in subjects with shoulder pain and asymptomatic subjects, respectively[10]. Moderate interrater reliability was demonstrated in both groups ($ICC_{2,1}$ ranging from 0.64-0.72). The PMI used by Struyf et al.[10] normalizes resting PM length to subject height[1], which is therefore not a direct assessment of absolute measurement reliability. The PMI was first proposed by Borstad and Ludewig[1] to classify people into relatively short and long PM groups and evaluate the effect of PM length on scapula kinematics. However, for assessing an individual patient and to evaluate the effectiveness of interventions to lengthen PM, a direct measurement of the muscle is more clinically relevant. In addition, as there are currently no normative values of the PMI reported in the literature, a direct measurement is more useful for making clinical decisions about an individual patient.

Another important variable that is missing in the literature is a reliability estimate of measuring PM length over time. This is a critical research gap because it limits a clinician's knowledge of how consistent the PM measurement is when used in the same patient over the course of their intervention program. Between-day reliability estimates are needed to provide clinicians with the ability to assess pectoralis minor length change over time due to treatment effects or other influences such as work or postural habits. These between-day reliability estimates are even more valuable if they have assessed a time interval that represents how the measurement will typically be used in the clinic.

The purpose of this study was to evaluate the intrarater, interrater, and between-day reliability of using a tape measure to assess pectoralis minor resting length in asymptomatic individuals and individuals with signs of shoulder impingement. Measurements of agreement, such as the minimal detectable change (MDC), standard error of measurement (SEM) of the measurement, and Bland Altman plots, were also determined in order to facilitate clinical interpretation of change over time.

Method

A convenience sample of 100 individuals (50 asymptomatic and 50 symptomatic) participated in this study. Individuals were recruited by means of fliers and direct contact from a local university setting and the community. Symptomatic subjects were recruited from a physical therapy waiting list at the clinic of Universidade Federal de São Carlos, (UFSCar), São Carlos, SP, Brazil and orthopedic clinics. All subjects were screened for eligibility by the first author and were required to be between 18 and 35 years of age. Asymptomatic individuals were included if they had no history of shoulder or cervical pathology. Because of the proposed relationship between pectoralis minor length, scapula kinematic alterations, and subacromial impingement syndrome[1,3], individuals with signs and symptoms consistent with impingement were targeted for inclusion.

The diagnosis for shoulder impingement was based on a clinical examination and self-reported history. To be classified as having shoulder impingement, subjects had to present with at least three[7,11-16] of the following: positive Neer[17] test, positive Hawkins test, positive Jobe and Moynes[18] test, pain with passive or isometric resisted shoulder lateral rotation[8,19], pain with active shoulder elevation[20], pain with palpation of rotator cuff tendons, and anterolateral shoulder pain. Individuals were excluded if they were pregnant; had ligamentous laxity based on positive Sulcus

test[21]; had apprehension during Apprehension test[22]; had history of clavicle, scapula, or humerus fracture; had systemic illnesses; or had received any treatment for shoulder pain in the last 6 months.

The study was approved by the Ethics Committee of the Universidade Metodista de Piracicaba (UNIMEP), Piracicaba, SP, Brazil (protocol number 100/12). All subjects gave their written and informed consent to participate in this study, which was conducted according to the Declaration of Helsinki.

Only the symptomatic shoulder was evaluated in the symptomatic individuals. For the asymptomatic participants, the side evaluated was randomly determined with a randomization list created by a computer program. Half of the sample (25 asymptomatic and 25 symptomatic; Sample 1) was used to determine intra- and interrater reliability, while the other half (Sample 2) was used to determine between-day reliability. Two independent samples were used to reduce the potential for bias that may have occurred with multiple measurements, such as postural adjustments by subjects or familiarity with subjects' previous measurements by raters.

Pectoralis minor resting length was measured with a tape measure with 0.10 cm resolution. To determine the interrater reliability for sample 1, all measurements were taken by two investigators who were blinded to all measures. To prevent bias, one side of the tape measure was covered with adhesive tape, blinding the examiners to the value measured. A third examiner read and recorded the measurements. To standardize the measurement technique, the primary and secondary investigators underwent training, which included studying relevant anatomy, systematically locating and palpating the coracoid process and the fourth rib landmarks, and then practicing all procedures on 10 healthy individuals. It is estimated that the training totaled approximately 6 hours.

Two bony landmarks representing the insertion and origin of the muscle were palpated, marked with a pencil, and used to represent pectoralis minor length: the caudal edge of the fourth rib at the sternum and the inferomedial aspect of the coracoid process (Figure 1A)[1]. The distance between these landmarks was estimated using the tape measure for two trials with two minutes between each trial. The caudal edge of the fourth rib was located by identifying the sternal part of the clavicle and counting down the intercostal spaces until fourth rib. The coracoid process landmark was located by palpating just distal to the concave region of the acromial end of the clavicle. During the measurements, participants were asked to remain in a standing and relaxed posture with their arms at their sides in a neutral position and to avoid postural correction (Figure 1B). Subjects were also instructed to exhale completely during the measure. The pencil marks were removed after each measurement. The same procedure was followed by the primary and secondary investigators. Following one evaluator's two measurements, subjects were allowed to rest comfortably for 5 minutes prior to the two measurements by the other evaluator. A randomization list was used to determine which evaluator measured first and second.

A test-retest design was used for sample 2 with one rater performing all measurements. Participants from sample 2 were measured in two sessions separated by an interval of seven days. This interval was selected to

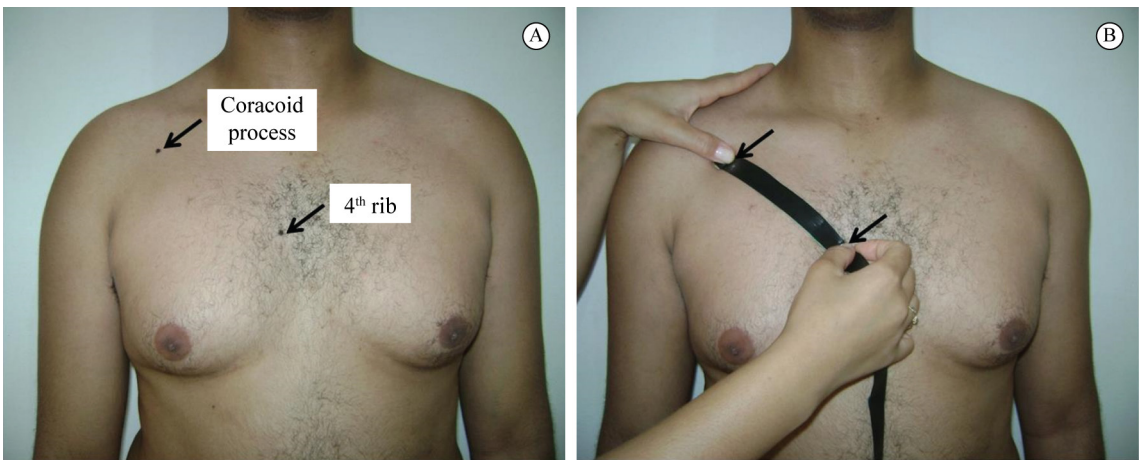

Figure 1. Bony landmarks used to represent pectoralis minor length (A) and measurement of pectoralis minor length taken with a tape measure (B).

represent the timeframe over which the measurement is likely to be used by clinicians to assess change in PM length.

The data were analyzed using the SPSS statistical package (17.0 Version). Data were normally distributed ($p>0.05$) as verified by the Kolmogorov–Smirnov test. A one-way ANOVA (Analysis of Variance) and an independent *t*-test were used to determine if baseline differences existed between groups for demographic variables and duration of pain, respectively. The relative reliability was determined by calculating the ICC for intrarater ($ICC_{3,1}$), between-day ($ICC_{3,2}$) and interrater ($ICC_{3,2}$) reliability[23]. The intrarater reliability was evaluated by separately comparing the two measurement trials from the first and second sessions. The interrater reliability was estimated using the mean of the two tape measure trials of each evaluator. The between-day reliability was estimated by comparing the mean of the two trials from each session. The ICC values were considered poor when below 0.20; fair from 0.21 to 0.40; moderate from 0.41 to 0.60; good from 0.61 to 0.80; and very good from 0.81 to 1.00[24]. The absolute reliability was defined as the SEM and MDC using the following formulas:

$SEM = \sqrt{WMS}$, where WMS (within mean square) is the within subjects mean square error term from a one-way ANOVA with subjects as the independent variable[25]; and $MDC_{95} = SEM \times \sqrt{2} \times 1.96$, for 95% confidence interval (CI)[26,27].

The SEM provides a value for measurement error for any given trial (intrarater reliability), any test occasion (between-day reliability), and any evaluator (interrater reliability)[25,28]. The MDC is an estimate of the smallest amount of change between repeated measures that can be considered to be a true change beyond measurement error[27-29] The MDC represents an outer limit of the amount of random variation that 95% of stable subjects will demonstrate when measures are collected on separate occasions.

Bland-Altman plots[30] were constructed to allow visual examination of the tape measure agreement between-days. The plots were constructed using MedCalc Software (Mariakerke, Belgium).

Results

Table 1 shows no differences among the groups in the descriptive data ($p>0.05$).

Intrarater reliability

ICC and SEM values for intrarater reliability ranged from 0.95-0.97 and 0.30-0.42 cm, respectively, for both groups (Table 2).

Interrater reliability

Table 2 also shows the interrater reliability data for both groups. ICC values for asymptomatic and symptomatic groups were 0.86 and 0.87, respectively. SEM values for the asymptomatic and symptomatic groups were 0.70 and 0.84 cm, respectively.

Between-day reliability

Table 3 reports the between-day reliability data. ICC values for both groups were 0.95 SEM values for the asymptomatic and symptomatic groups were 0.40 and 0.41 cm, respectively. MDC values for the asymptomatic and symptomatic groups were 1.13 and 1.14 cm, respectively.

Figure 2 presents the Bland-Altman plots for asymptomatic and symptomatic groups. Visual inspection of the plots for between-day reliability

Table 1. Descriptive data of the subjects.

	Sample 1		Sample 2		
	Asymptomatic (n=25)	**Symptomatic (n=25)**	**Asymptomatic (n=25)**	**Symptomatic (n=25)**	**p value**
Age (years)*	25.72±3.52	25.52±3.72	25.76±6.95	26.96±5.79	0.75
Gender	13 women; 12 men	12 women; 13 men	13 women; 12 men	14 women; 11 men	___
Weight (kg)*	67.22±10.62	70.22±15.71	64.12±10.76	67.54±9.68	0.36
Height (m)*	1.70±0.08	1.73±0.09	1.69±0.08	1.69±0.07	0.23
Evaluated shoulder	10 dominant; 15 non-dominant	10 dominant; 15 non-dominant	13 dominant; 12 non-dominant	17 dominant; 8 non-dominant	___
Duration of pain (months)*	____	41.28±37.28	_____	49.12±86.92	0.68

*Values are mean±standard deviation.

Table 2. Intrarater and interrater reliability for assessing the pectoralis minor length with the tape measure in asymptomatic and symptomatic individuals.

	Trial 1*†	Trial 2*†	$ICC_{3,1}$ (95%CI)	SEM *
Asymptomatic group (n=25)				
Rater 1	16.42±1.46	16.17±1.42	0.96 (0.92-0.98)	0.32
Rater 2	16.44±1.42	16.38±1.47	0.95 (0.90-0.98)	0.30
Symptomatic group (n=25)				
Rater 1	16.74±1.65	16.58±1.65	0.97 (0.93-0.98)	0.31
Rater 2	17.12±1.86	16.81±1.77	0.95 (0.91-0.98)	0.42
	Rater 1*†	**Rater 2*†**	**$ICC_{3,2}$ (95% CI)**	**SEM***
Asymptomatic group (n=25)	16.30±1.43	16.42±1.43	0.86 (0.68-0.94)	0.70
Symptomatic group (n=25)	16.66±1.64	16.97±1.80	0.87 (0.70-0.94)	0.84

*All units are in centimeters. † Values are mean±standard deviation. ICC: Intraclass correlation coefficients; SEM: Standard error of measurement.

Table 3. Between-day reliability for assessing the pectoralis minor length with the tape measure in asymptomatic and symptomatic individuals.

	Day 1*†	Day 2*†	$ICC_{3,2}$ (95% CI)	SEM*	MDC_{95}*
Asymptomatic group (n=25)	15.85±1.23	16.06±1.38	0.95 (0.90-0.98)	0.40	1.13
Symptomatic group (n=25)	16.23±1.55	16.27±1.68	0.95 (0.89-0.98)	0.41	1.14

*All units are in centimeters. † Values are mean±standard deviation. ICC: Intraclass correlation coefficients; SEM: Standard error of measurement; MDC: Minimal detectable change.

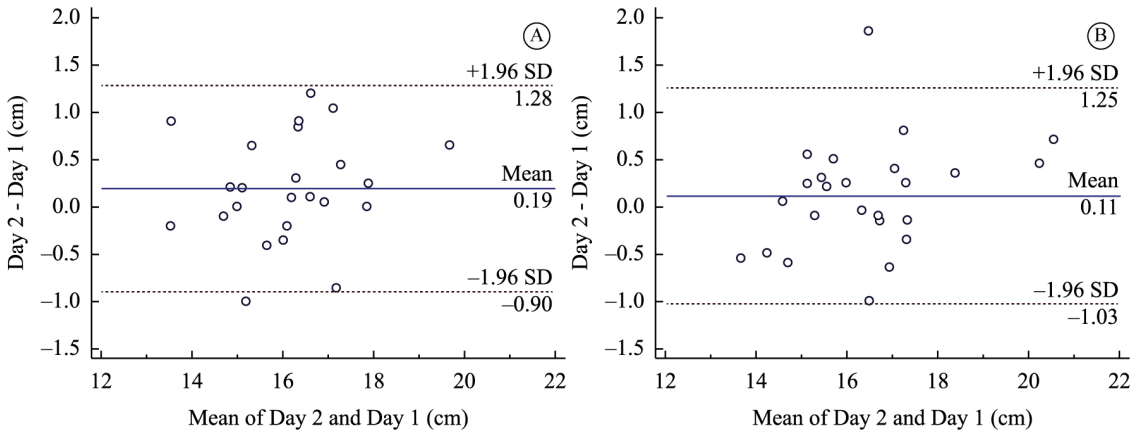

Figure 2. Bland-Altman plots for between-day agreement of measuring pectoralis minor length with a tape measure in asymptomatic (A) and symptomatic (B) groups.

revealed that all mean differences were close to zero. No systematic biases were observed. The plots show a random scatter of points above and below the mean difference line, thus showing good agreement.

Discussion

These results are similar to the measurement reliability estimates reported when using a caliper to assess PMI[10]. The current study adds to the literature by supporting the use of a tape measure to reliably estimate pectoralis minor resting length and is, to our knowledge, the first study to report the between-day reliability of this measurement. Our results suggest that the tape measure method demonstrates very good reliability for a single rater and for different raters to measure the length of this muscle within the same day. Importantly, the study also provides estimates that this measurement has very good reliability for assessing pectoralis minor resting length over a seven-day interval. As one of the purposes for any clinical measure is to reliably quantify variables over

time, the between-day reliability estimates are both practical and valuable to clinicians. The stability of these measurements over seven days in the absence of an intervention provides clinicians with a way to document a real change in PM length when assessing for treatment effects.

Pectoralis minor muscle resting length can also be assessed by measuring the linear distance from the treatment table to the posterior aspect of the acromion with the patient in the supine position, as proposed by Kendall and Provance[4]. This method was later suggested to be useful to determine shortening of the pectoralis minor[5]. The reliability of this measurement was evaluated in subjects with and without shoulder symptoms, and excellent clinical intrarater reliability (ICC=0.92-0.97) was estimated[23]. It is important to note that this is an indirect measurement of pectoralis minor muscle length that demonstrated poor correlation with the PMI[3]. The PMI is calculated by dividing pectoralis minor resting length by subject height and multiplying by 100[1]. The poor correlation between these measurements[3] likely reflects the fact that the table-to-acromion measurement is an indirect estimate of pectoralis minor length that can be influenced by scapula position changes from table and thorax forces and by altered body orientation to gravity. Conversely, the PMI uses the distance between the origin and insertion of the muscle in its calculation. Using this direct estimate between landmarks is advantageous when relying on a measurement to make treatment decisions within an individual patient.

Very good and good intrarater reliability was demonstrated when calculating the PMI in individuals with and without shoulder pain, respectively[10]. However, PM length was measured in the supine position in the previous study. The supine position is prone to misrepresenting true PM length for several reasons. First, the effect of gravity on the shoulder complex is changed, modifying the typical forces acting on the shoulder complex. Second, the plinth and weight of the thorax modify scapular position and PM length. In addition, typical functional activities of the upper extremity are done in standing, not supine. Finally, it has been shown that in supine the PM length measurement is influenced by the position of the upper extremity,[3] with full internal rotation (palm down as used in Struyf et al.[10]) resulting in higher length estimates than either neutral or full external rotation. We contend that, in standing, the normal and constant influences on the shoulder (e.g. gravity or posture) are accounted for in the measurement and therefore make it more practical, functional, and reflective of the patient's true condition.

One may also argue if scapula dyskinesis could influence the results of the present study. We believe that dyskinesis would only interfere with the measurement if there was a dynamic component to the measure. Because the measurement is taken only in resting position, this is not a factor that influenced the study results.

As stated before, the PMI is not currently useful in clinical practice because normative values have not yet been established in the literature. Estimating PMI for each patient, based on a group of similar subjects, is not feasible to clinicians. As such, information about the direct muscle length may be more applicable in the clinical practice to identify individual subject changes after intervention.

The values for intrarater reliability observed for both symptomatic and asymptomatic groups were similar in the present study. In contrast, Struyf et al.[10] reported higher intrarater reliability estimates in their group of symptomatic subjects. They explained these results as a learning effect of the examiners and due to the low variability of healthy controls. However, the asymptomatic group in their study was younger (~20 years) than the symptomatic group (~50 years), and the confidence intervals for ICC values were not provided, limiting full interpretation of the data. Our symptomatic group estimates also demonstrate that pain does not negatively influence the between-day reliability of pectoralis minor muscle length because of the very similar results found in both groups. However, it is important to note that the mean duration of symptoms is quite long for this sample and cannot be generalized to individuals with acute pain.

Interrater reliability showed wide confidence intervals for both groups (0.68-0.94), despite the very good ICC and low SEM values that represent small variability of the measure. The sample size could have contributed to these wide confidence intervals, which lead to uncertainty about the point estimate, with the true reliability potentially being anywhere within the confidence interval. Consequently, clinicians should be cautious when interpreting pectoralis minor muscle length measured by more than one rater.

Our between-day measures showed a very good reliability over time (ICC=0.95), with a small variability (CI=0.89-0.98) in both groups. Regarding visual inspection of Bland-Altman plots, good agreement can be observed, reflecting the consistency in measuring pectoralis minor length using a tape measure on

different days. The plots suggest a slightly greater length measurement on day 2 as compared with day 1, but the distribution of the difference scores indicates that there is no systematic bias. This information is important to clinical practice, because it shows stability of the measure in subjects with and without signs of impingement, and suggests that time does not influence the muscle length when no intervention is done.

Furthermore, the between-day MDC calculations suggest that a muscle length change greater than 1 cm is needed to identify a real change in pectoralis minor length on the same day and over time. However, a resting length change greater than 1 cm may not be possible given that the largest alteration in muscle length during passive pectoralis minor stretching was 0.77 cm, when performed in a similar position of the present study, with subjects in a sitting position without scapular stabilization[31]. It is not currently known how much muscle length change is possible or is associated with shoulder pathology, which makes interpreting the MDC a challenge. However, because the subjects in the present study were not evaluated for pectoralis minor shortening, our calculated MDC may not be directly applicable. It is possible that the MDC for those with a relatively shorter pectoralis minor length will be smaller because mean length and variability estimates may also be reduced. Similarly, individuals who demonstrate adaptive muscle shortening may also have more potential for an increase in muscle length than those without adaptive shortening. Moreover, variability in anthropometric characteristics may lead to differing MDC values so these results should be used with caution and applied only to individuals similar to the population used in this study. Additionally, studies that determine the minimal clinically important difference (MCID) for pectoralis minor length are needed to establish the amount of change that is meaningful and beneficial for the health status of the patient.

As the present study was conducted using only young subjects, our results cannot be generalized to older individuals or to those with other shoulder conditions.

Conclusion

This study provides additional information about intra- and interrater reliability and important new knowledge of between-day reliability using a tape measure to assess pectoralis minor resting length in asymptomatic individuals and in individuals with signs of shoulder impingement. A single rater or different raters can reliably measure pectoralis minor within the same day, and a single rater can reliably use the measurement over a seven-day interval.

Acknowledgements

The authors would like to thank Rafael Foroni Luchesi and Lívia Silveira Pogetti for helping with data collection. This study was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brazil (2012/20305-8 and 2013/50363-2).

References

1. Borstad JD, Ludewig PM. The effect of long versus short pectoralis minor resting length on scapular kinematics in healthy individuals. J Orthop Sports Phys Ther. 2005;35(4):227-38. http://dx.doi.org/10.2519/jospt.2005.35.4.227. PMid:15901124.
2. Muraki T, Aoki M, Izumi T, Fujii M, Hidaka E, Miyamoto S. Lengthening of the pectoralis minor muscle during passive shoulder motions and stretching techniques: a cadaveric biomechanical study. Phys Ther. 2009;89(4):333-41. http://dx.doi.org/10.2522/ptj.20080248. PMid:19246556.
3. Borstad JD. Resting position variables at the shoulder: evidence to support a posture-impairment association. Phys Ther. 2006;86(4):549-57. PMid:16579671.
4. Kendall FPME, Provance PG. Muscles: testing and function. 4th ed. Baltimore: Williams & Wilkins; 1993.
5. Sahrmann S. Diagnosis and treatment of movement impairment syndromes. St Louis: Mosby; 2002.
6. Lukasiewicz AC, McClure P, Michener L, Pratt N, Sennett B. Comparison of 3-dimensional scapular position and orientation between subjects with and without shoulder impingement. J Orthop Sports Phys Ther. 1999;29:574-583.
7. Ludewig PM, Cook TM. Alterations in shoulder kinematics and associated muscle activity in people with symptoms of shoulder impingement. Phys Ther. 2000;80(3):276-91. PMid:10696154.
8. Phadke V, Camargo P, Ludewig P. Scapular and rotator cuff muscle activity during arm elevation: A review of normal function and alterations with shoulder impingement. Rev Bras Fisioter. 2009;13(1):1-9. http://dx.doi.org/10.1590/S1413-35552009005000012. PMid:20411160.
9. Borstad JD. Measurement of pectoralis minor muscle length: validation and clinical application. J Orthop Sports Phys Ther. 2008;38(4):169-74. http://dx.doi.org/10.2519/jospt.2008.2723. PMid:18434665.
10. Struyf F, Meeus M, Fransen E, Roussel N, Jansen N, Truijen S, et al. Interrater and intrarater reliability of the pectoralis minor muscle length measurement in subjects with and without shoulder impingement symptoms. Man Ther. 2014;19(4):294-8. http://dx.doi.org/10.1016/j.math.2014.04.005. PMid:24815592.
11. Haik MN, Alburquerque-Sendin F, Camargo PR. Reliability and minimal detectable change of 3-dimensional scapular orientation in individuals with and without shoulder

impingement. J Orthop Sports Phys Ther. 2014;44(5):341-9. http://dx.doi.org/10.2519/jospt.2014.4705. PMid:24673448.

12. Haik MN, Alburquerque-Sendin F, Silva CZ, Siqueira-Junior AL, Ribeiro IL, Camargo PR. Scapular kinematics pre- and post-thoracic thrust manipulation in individuals with and without shoulder impingement symptoms: a randomized controlled study. J Orthop Sports Phys Ther. 2014;44(7):475-87. http://dx.doi.org/10.2519/jospt.2014.4760. PMid:24853923.
13. Camargo PR, Avila MA, Alburquerque-Sendin F, Asso NA, Hashimoto LH, Salvini TF. Eccentric training for shoulder abductors improves pain, function and isokinetic performance in subjects with shoulder impingement syndrome: a case series. Rev Bras Fisioter. 2012;16(1):74-83. http://dx.doi.org/10.1590/S1413-35552012000100013. PMid:22441232.
14. Camargo PR, Avila MA, Asso NA, Salvini TF. Muscle performance during isokinetic concentric and eccentric abduction in subjects with subacromial impingement syndrome. Eur J Appl Physiol. 2010;109(3):389-95. http://dx.doi.org/10.1007/s00421-010-1365-2. PMid:20131062.
15. Ribeiro IL, Camargo PR, Alburquerque-Sendín F, Madeleine P, Fernandez-de-las-Penas, Salvini TF. Topographical pressure pain sensitivity maps of the shoulder region in individuals with subacromial pain syndrome. Man Ther. 2016;21:134-43. http://dx.doi.org/10.1016/j.math.2015.07.002.
16. Lawrence RL, Braman JP, Laprade RF, Ludewig PM. Comparison of 3-dimensional shoulder complex kinematics in individuals with and without shoulder pain, part 1: sternoclavicular, acromioclavicular, and scapulothoracic joints. J Orthop Sports Phys Ther. 2014;44:636-645.
17. Neer CS 2nd. Anterior acromioplasty for the chronic impingement syndrome in the shoulder: a preliminary report. J Bone Joint Surg Am. 1972;54(1):41-50. PMid:5054450.
18. Jobe FW, Moynes DR. Delineation of diagnostic criteria and a rehabilitation program for rotator cuff injuries. Am J Sports Med. 1982;10(6):336-9. http://dx.doi.org/10.1177/036354658201000602. PMid:7180952.
19. Tate AR, McClure PW, Young IA, Salvatori R, Michener LA. Comprehensive impairment-based exercise and manual therapy intervention for patients with subacromial impingement syndrome: a case series. J Orthop Sports Phys Ther. 2010;40(8):474-93. http://dx.doi.org/10.2519/jospt.2010.3223. PMid:20710088.
20. Hung CJ, Jan MH, Lin YF, Wang TQ, Lin JJ. Scapular kinematics and impairment features for classifying patients with subacromial impingement syndrome. Man Ther. 2010;15(6):547-51. http://dx.doi.org/10.1016/j.math.2010.06.003. PMid:20615748.
21. Neer CS 2nd, Foster CR. Inferior capsular shift for involuntary inferior and multidirectional instability of the shoulder. A preliminary report. J Bone Joint Surg Am. 1980;62(6):897-908. PMid:7430177.
22. Rowe CR, Zarins B. Recurrent transient subluxation of the shoulder. J Bone Joint Surg Am. 1981;63(6):863-72. PMid:7240326.
23. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 3rd ed. Upper Saddle River: Pearson/Prentice Hall; 2009.
24. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med. 2001;134(8):663-94. http://dx.doi.org/10.7326/0003-4819-134-8-200104170-00012. PMid:11304107.
25. Lexell JE, Downham DY. How to assess the reliability of measurements in rehabilitation. Am J Phys Med Rehabil. 2005;84(9):719-23. http://dx.doi.org/10.1097/01.phm.0000176452.17771.20. PMid:16141752.
26. Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL. Smallest real difference, a link between reproducibility and responsiveness. Qual Life Res. 2001;10(7):571-8. http://dx.doi.org/10.1023/A:1013138911638. PMid:11822790.
27. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res. 2005;19(1):231-40. PMid:15705040.
28. Donoghue D, Physiotherapy Research and Older People (PROP) group, Stokes EK. How much change is true change? The minimum detectable change of the Berg Balance Scale in elderly people. J Rehabil Med. 2009;41(5):343-6. http://dx.doi.org/10.2340/16501977-0337. PMid:19363567.
29. Wu CY, Chuang LL, Lin KC, Lee SD, Hong WH. Responsiveness, minimal detectable change, and minimal clinically important difference of the Nottingham Extended Activities of Daily Living Scale in patients with improved performance after stroke rehabilitation. Arch Phys Med Rehabil. 2011;92(8):1281-7. http://dx.doi.org/10.1016/j.apmr.2011.03.008. PMid:21807147.
30. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999;8(2):135-60. http://dx.doi.org/10.1191/096228099673819272. PMid:10501650.
31. Borstad JD, Ludewig PM. Comparison of three stretches for the pectoralis minor muscle. J Shoulder Elbow Surg. 2006;15(3):324-30. http://dx.doi.org/10.1016/j.jse.2005.08.011. PMid:16679233.

Correspondence
Paula Rezende Camargo
Universidade Federal de São Carlos
Departamento de Fisioterapia
Rodovia Washington Luís, Km 235
CEP 13565-905, São Carlos, SP, Brazil
e-mail: prcamargo@ufscar.br

original article

BJPT Brazilian Journal of Physical Therapy

Inductive plethysmography potential as a surrogate for ventilatory measurements during rest and moderate physical exercise

Ramona Cabiddu[1], Camila B. F. Pantoni[2], Renata G. Mendes[1], Renata Trimer[1], Aparecida M. Catai[2], Audrey Borghi-Silva[1]

ABSTRACT | Background: Portable respiratory inductive plethysmography (RIP) systems have been validated for ventilatory assessment during resting conditions and during incremental treadmill exercise. However, in clinical settings and during field-based exercise, intensity is usually constant and submaximal. A demonstration of the ability of RIP to detect respiratory measurements accurately during constant intensity conditions would promote and validate the routine use of portable RIP devices as an alternative to ergospirometry (ES), the current gold standard technique for ventilatory measures. **Objective:** To investigate the agreement between respiratory variables recorded by a portable RIP device and by ES during rest and constant intensity exercise. **Method:** Tidal volume (V_T), respiratory rate (RR) and minute ventilation (V_E) were concurrently acquired by portable RIP and ES in seven healthy male volunteers during standing rest position and constant intensity treadmill exercise. **Results:** Significant agreement was found between RIP and ES acquisitions during the standing rest position and constant intensity treadmill exercise for RR and during the standing rest position for V_E. **Conclusion:** Our results suggest that portable RIP devices might represent a suitable alternative to ES during rest and during constant submaximal exercise.

Keywords: respiratory inductive plethysmography; movement; respiratory rate; minute ventilation; standing rest position; constant intensity exercise.

BULLET POINTS

- We compared respiratory inductive plethysmography (RIP) and ergospirometry (ES).
- RIP advantages over ES include portability and no need for facial apparatuses.
- Agreement was found between RIP and ES for respiratory rate and ventilation.
- RIP might represent an alternative to ES during rest and submaximal exercise.

HOW TO CITE THIS ARTICLE

Cabiddu R, Pantoni CBF, Mendes RG, Trimer R, Catai AM, Borghi-Silva A. Inductive plethysmography potential as a surrogate for ventilatory measurements during rest and moderate physical exercise. Braz J Phys Ther. 2016 Mar-Apr; 20(2):184-188. http://dx.doi.org/10.1590/bjpt-rbf.2014.0147

Introduction

Reliable measurement of ventilatory parameters is essential to support research in respiratory physiology and medicine[1]. The gold standard technique is ergospirometry (ES), which provides continuous and breath-by-breath ventilatory measures. ES systems, however, involve the use of mouthpieces, which increase dead space[2] and may be uncomfortable for the subjects[3]. In addition, they are expensive and require highly trained staff.

Respiratory inductive plethysmography (RIP) detects changes in the cross-sectional area of the rib cage and abdomen using inductive belts[4]. Respiratory volumes and timing can subsequently be obtained. Portable RIP systems incorporate belts in an elasticized vest and allow measurements to be easily obtained without a mouthpiece[5]. Evidence shows that RIP can be used to investigate respiratory mechanics in normal and symptomatic subjects[3,6,7] during rest and incremental exercise[5,6,8]. However, during rehabilitation or field-based exercise, intensity is usually submaximal and constant over a given time period[6]. A demonstration of the ability of RIP to detect respiratory measurements accurately during constant intensity conditions would promote and validate the routine use of portable RIP devices, which have the advantage of being non-invasive and easy to move

[1] Laboratório de Fisioterapia Cardiopulmonar, Departamento de Fisioterapia, Universidade Federal de São Carlos (UFSCar), São Carlos, SP, Brazil
[2] Laboratório de Fisioterapia Cardiovascular, Departamento de Fisioterapia, Universidade Federal de São Carlos (UFSCar), São Carlos, SP, Brazil
Received: Apr. 17, 2015 Revised: July 15, 2015 Accepted: Oct. 06, 2015

in clinical settings. We aim to investigate whether agreement exists between respiratory variables, including tidal volume (V_T), respiratory rate (RR) and minute ventilation (V_E), recorded by a portable RIP device and by ES in healthy male subjects in the resting standing position and during constant-intensity treadmill exercise.

Method

Subjects

Seven apparently healthy male subjects were included. Individuals were considered healthy based on an anamnesis that included a questionnaire to record demographic data, work and health status, previous surgeries, and physical activity level. All subjects were sedentary. None of them were smokers nor used medications that might affect the measurements. A visual examination was conducted to identify thoracoabdominal alterations that might alter respiratory dynamics. Anthropometric data, including weight and height, were collected. This study was approved by the Human Research Ethics Committee of Universidade Federal de São Carlos (UFSCar), São Carlos, SP, Brazil (Process no. 145/2006), and the subjects signed an informed consent form.

Protocol

Tests were performed in a laboratory with controlled temperature and humidity, always between 8 a.m. and 12 p.m. The day before and the day of the test volunteers were instructed to avoid stimulating drinks, refrain from physical exercise, and have an adequate night's sleep.

V_T [mL], RR [bpm], and V_E [L/min] were recorded simultaneously using a wearable RIP system (LifeShirt®, Vivometrics Inc., Ventura, CA, USA.) and an ergospirometer (CPX-D/BreezeSuite 6.4.1, Medical Graphics, St Paul, MN, USA). For the RIP system calibration, participants were asked to breathe seven times into an 800 mL plastic bag attached to a mouthpiece, filling and emptying it completely with each breath[9]. For the ES system, the carbon dioxide (CO_2) and oxygen (O_2) analyzers were calibrated before and after each test using a two-point measure: a calibration gas (5% CO_2, 12% O_2, and balance nitrogen) and a reference gas (room air after ATPS [ambient temperature and pressure saturated] to STPD [standard temperature and pressure, dry] correction).

Heart rate (HR) was measured while standing before and after exercise. Ventilatory variables were collected using ES and RIP during 5 minutes of standing rest. Afterwards, subjects started the treadmill exercise. Speed was increased at 30-second intervals until the HR was 20 beats faster than the resting value. This speed was maintained and ventilatory variables were collected using ES and RIP during 6 minutes.

Data analysis

An *a posteriori* power analysis (G*Power, F. Faul, Universität Kiel, Germany) was performed. Considering a p=0.05, the statistical power of this study is 60%.

Data analysis (MATLAB®, The Mathworks Inc., Natick, MA, USA) was performed on 4.5 minutes of signal acquired during rest (excluding the first 30 seconds) and on 5.5 minutes of signal acquired during exercise (excluding the first 30 seconds). Time accordance between RIP and ES breath-by-breath values was verified.

For each parameter and for each condition, the seven signals coming from all volunteers were averaged. Bland-Altman plot and Spearman correlation analyses were performed between RIP and ES values. Normality of distribution was verified using the Kolmogorov-Smirnov test. Significant correlations (p-value ≤ 0.05) were considered weak when $0 \leq r < 0.3$, moderate when $0.3 \leq r < 0.7$, and strong when $r \geq 0.7$.

Results

Mean age was 25 years (SD=3), mean weight was 73 Kg (SD=12), and mean height was 177 cm (SD=6). Average ventilatory parameters obtained during rest and exercise using ES and RIP are presented in Table 1.

Average Bland-Altman and correlation results obtained for the whole group, during rest and exercise, are shown in Figure 1 and in Figure 2, respectively.

Table 1. Average values of ventilatory parameters obtained during resting standing position and treadmill exercise using ES and RIP.

	Resting standing position		Treadmill exercise	
	ES	**RIP**	**ES**	**RIP**
V_T [mL]	487.0±34.9	471.0±131.4	1030.5±134.5	1017.2±173.1
RR [bpm]	16.2±3.2	16.4±3.7	23.3±3.8	24.8±2.6
V_E [L/min]	7.8±1.4	9.2±5.2	23.7±5.3	24.3±2.8

Data are presented as mean±SD. V_T: tidal volume; RR: respiratory rate; V_E: minute ventilation.

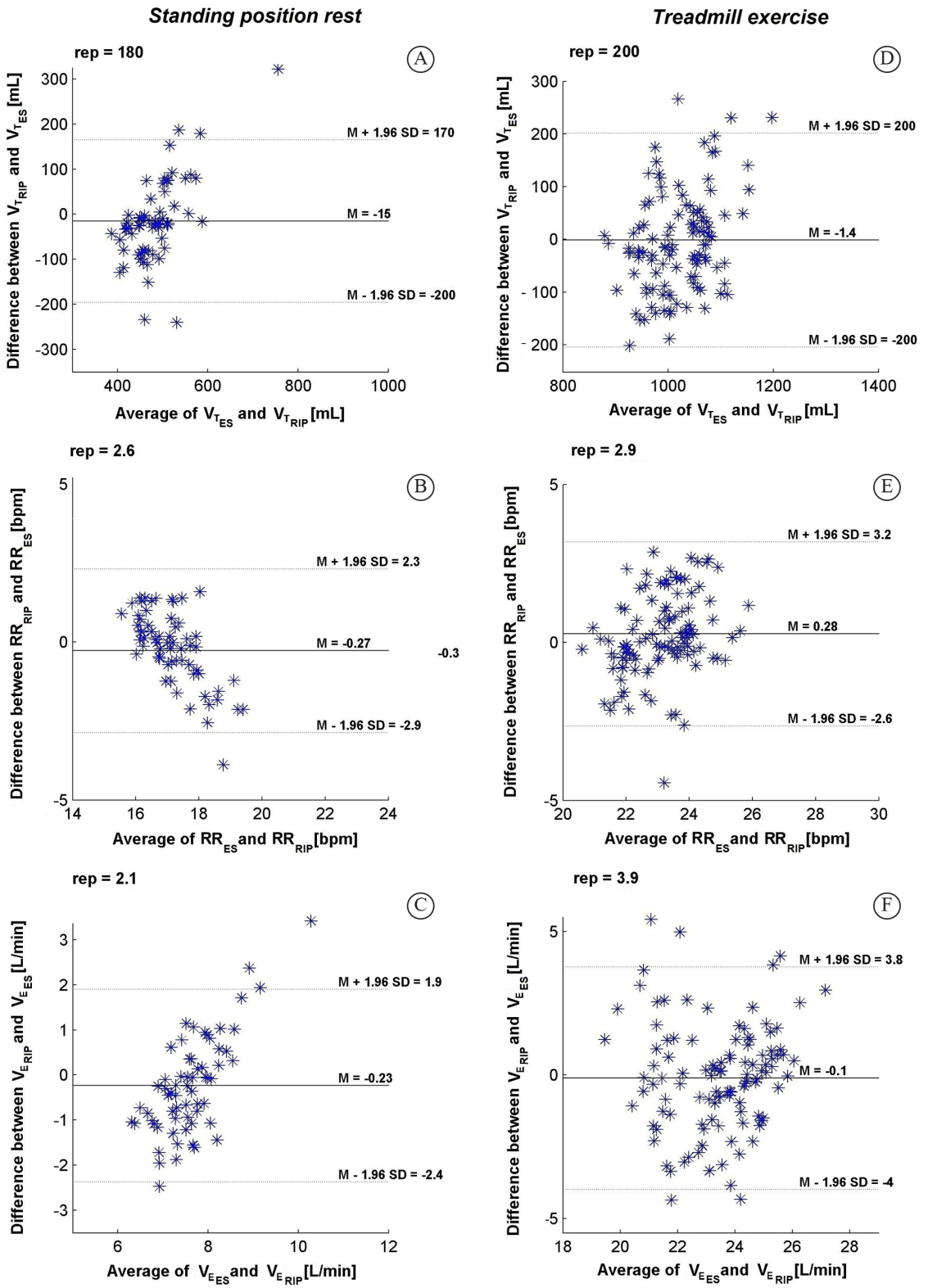

Figure 1. Bland-Altman plots showing agreement of the mean differences between (A) V_T, (B) RR, and (C) V_E average signals calculated over the signals measured using ES and RIP during the resting standing position and between (D) V_T, (E) RR, and (F) V_E average signals calculated over the signals measured using ES and RIP during the treadmill exercise. Rep: reproducibility index; M: mean; SD: standard deviation.

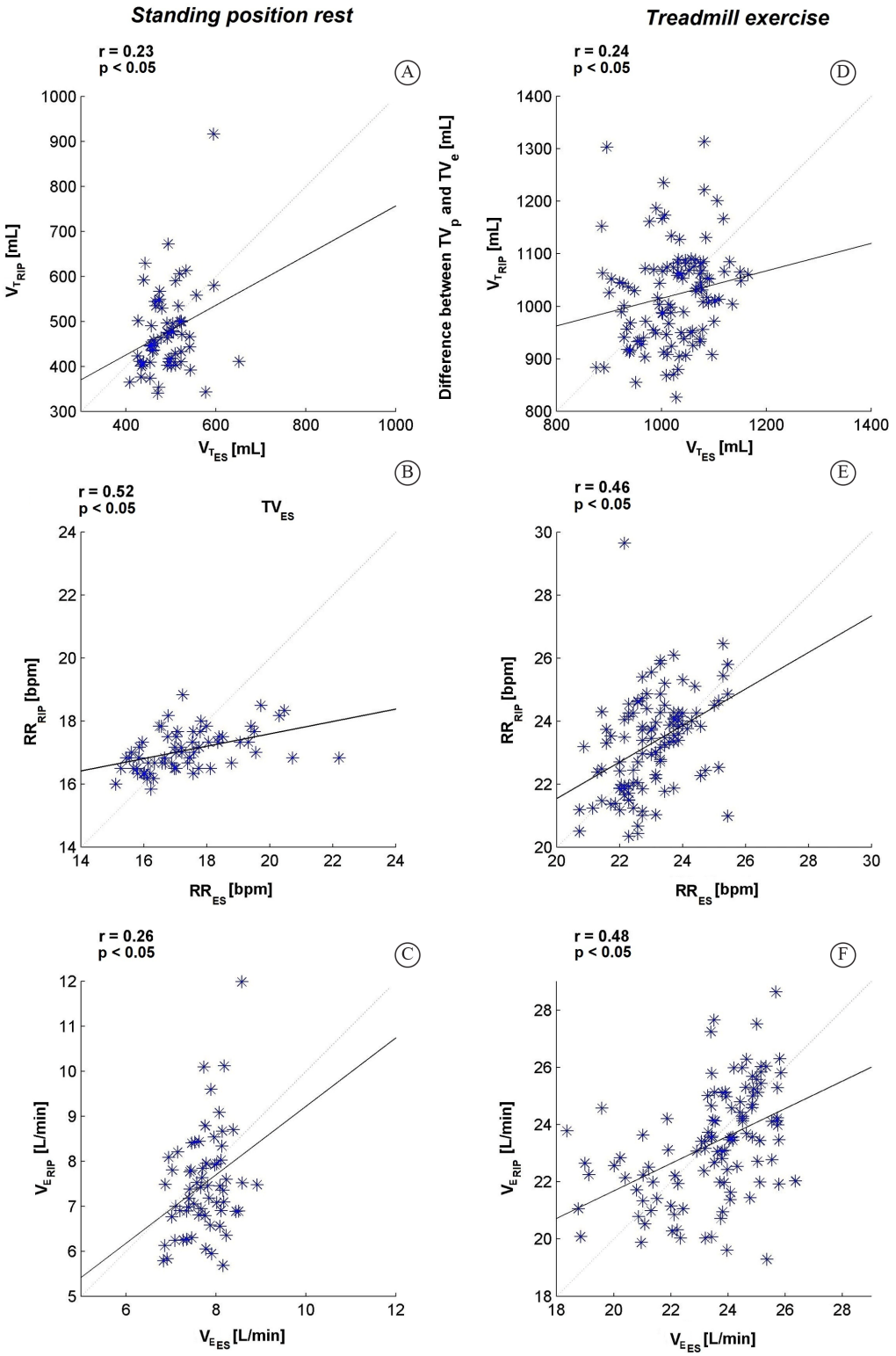

Figure 2. Correlation plots between (A) V_T, (B) RR, and (C) V_E average signals calculated over the signals measured using ES and RIP during the standing rest position and between (D) V_T, (E) RR and (F) V_E average signals calculated over the signals measured using ES and RIP during the treadmill exercise. r, correlation coefficient. For each correlation, regression lines are presented.

Discussion

The present study aimed to evaluate the potential of portable RIP as an alternative to ES during rest and steady-state treadmill exercise. Significant correlations were found between RIP and ES acquisitions; good agreement was found for RR, during rest and exercise, and V_E, during rest.

V_T values recorded using RIP and ES presented low agreement (rest: bias=15 mL; rep=180 mL; exercise: bias=1.4 mL; rep=200 mL); however, significant but weak correlations were found (rest: r=0.23; exercise: r=0.25). This is in line with Hollier et al.[1], who performed RIP and ES to measure ventilatory variables in sitting position and during two breathing tests on untreated obesity hypoventilation syndrome patients and controls. Our results suggest that RIP translates into a qualitative measurement related to ES, despite a reduced agreement between the methods. RIP-ES agreement was acceptable for RR both during rest (r=0.52; bias=0.27 bmp; rep=2.6 bpm) and exercise (r=0.46; bias=0.28 bmp; rep=2.9 bpm). Correlation between RIP and ES for V_E values was significant during rest (r=0.23) and exercise (r=0.45). Agreement was acceptable during rest (bias=0.23 L/min; rep=2.1 L/min) and low during exercise (bias=1 L/min; rep=3.9 L/min).

In summary, our results confirm that significant, quantitative agreement exists between RIP and ES acquisitions for RR, during rest and constant intensity treadmill exercise, and for V_E, during rest. To choose between one or the other, health professionals should consider what measurements are needed and that, due to the absence of airway instrumentation, RIP will probably be better tolerated[3].

Limitations of this study include a low number of subjects, which might help explain why partial disagreement was observed with Clarenbach et al.[5], who showed that RIP-ES agreement was similar for all indices in healthy volunteers and cardiorespiratory patients during progressive treadmill exercise to exhaustion. Implementation of a different kind of exercise (constant, submaximal intensity conditions, which increase the respiratory pattern variability[10]) might also have influenced results. It is also worth noticing that differences in the system calibration might have influenced measurements as well. Moreover, our investigation is limited to the study of healthy, young men during standing position and constant treadmill exercise and may not apply to other populations, as elderly people or women. In conclusion, our results suggest that RIP and ES can be used interchangeably in healthy, young male subjects to evaluate RR quantitatively during rest and constant intensity treadmill exercise and V_E during resting conditions.

References

1. Hollier CA, Harmer AR, Maxwell LJ, Menadue C, Willson GN, Black DA, et al. Validation of respiratory inductive plethysmography (LifeShirt) in obesity hypoventilation syndrome. Respir Physiol Neurobiol. 2014;194:15-22. http://dx.doi.org/10.1016/j.resp.2014.01.014. PMid:24468468.
2. Perez W, Tobin MJ. Separation of factors responsible for change in breathing pattern induced by instrumentation. J Appl Physiol. 1985;59(5):1515-20. PMid:4066581.
3. Bloch KE, Li Y, Sackner MA, Russi EW. Breathing pattern during sleep disruptive snoring. Eur Respir J. 1997;10(3):576-86. PMid:9072988.
4. Cohn MA, Rao AS, Broudy M, Birch S, Watson H, Atkins N, et al. The respiratory inductive plethysmograph: a new non-invasive monitor of respiration. Bull Eur Physiopathol Respir. 1982;18(4):643-58. PMid:7116012.
5. Clarenbach CF, Senn O, Brack T, Kohler M, Bloch KE. Monitoring of ventilation during exercise by a portable respiratory inductive plethysmograph. Chest. 2005;128(3):1282-90. http://dx.doi.org/10.1378/chest.128.3.1282. PMid:16162719.
6. Witt JD, Fisher JR, Guenette JA, Cheong KA, Wilson BJ, Sheel AW. Measurement of exercise ventilation by a portable respiratory inductive plethysmograph. Respir Physiol Neurobiol. 2006;154(3):389-95. http://dx.doi.org/10.1016/j.resp.2006.01.010. PMid:16503424.
7. Grossman P, Wilhelm FH, Brutsche M. Accuracy of ventilatory measurement employing ambulatory inductive plethysmography during tasks of everyday life. Biol Psychol. 2010;84(1):121-8. http://dx.doi.org/10.1016/j.biopsycho.2010.02.008. PMid:20176075.
8. Eberhard A, Calabrese P, Baconnier P, Benchetrit G. Comparison between the respiratory inductance plethysmography signal derivative and the airflow signal. Adv Exp Med Biol. 2001;499:489-94. http://dx.doi.org/10.1007/978-1-4615-1375-9_79. PMid:11729931.
9. Cancelliero-Gaiad KM, Ike D, Pantoni CB, Borghi-Silva A, Costa D. Respiratory pattern of diaphragmatic breathing and pilates breathing in COPD subjects. Braz J Phys Ther. 2014;18(4):291-9. http://dx.doi.org/10.1590/bjpt-rbf.2014.0042. PMid:25075999.
10. Pfaltz MC, Grossman P, Michael T, Margraf J, Wilhelm FH. Physical activity and respiratory behavior in daily life of patients with panic disorder and healthy controls. Int J Psychophysiol. 2010;78(1):42-9. http://dx.doi.org/10.1016/j.ijpsycho.2010.05.001. PMid:20472006.

Correspondence
Ramona Cabiddu
Universidade Federal de São Carlos
Departamento de Fisioterapia
Laboratório de Fisioterapia Cardiopulmonar
Rodovia Washington Luís, Km 235, SP 310
CEP 13565-905, São Carlos, SP, Brazil
e-mail: ramona.cabiddu@gmail.com

BJPT Brazilian Journal of Physical Therapy

Erratum

In the article **Effects of diaphragmatic control on the assessment of sniff nasal inspiratory pressure and maximum relaxation rate**, DOI: http://dx.doi.org/10.1590/bjpt-rbf.2014.0101, published in the Brazilian Journal of Physical Therapy, Volume 20, Number 1, page 96-103, on page 98 and 100, it reads:

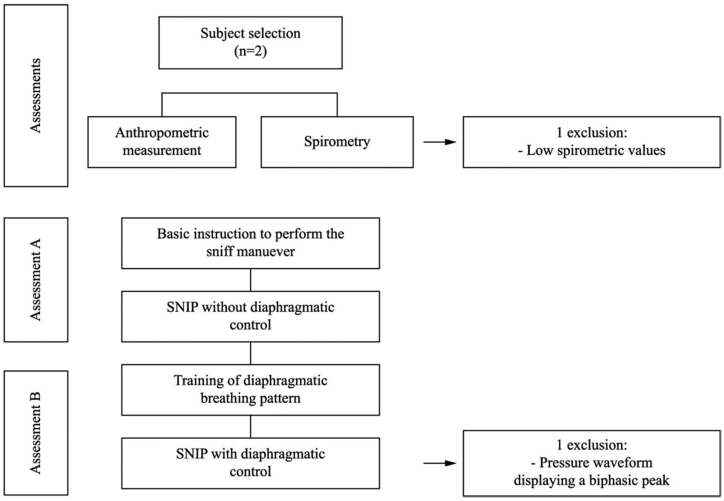

Figure 1. Study design.

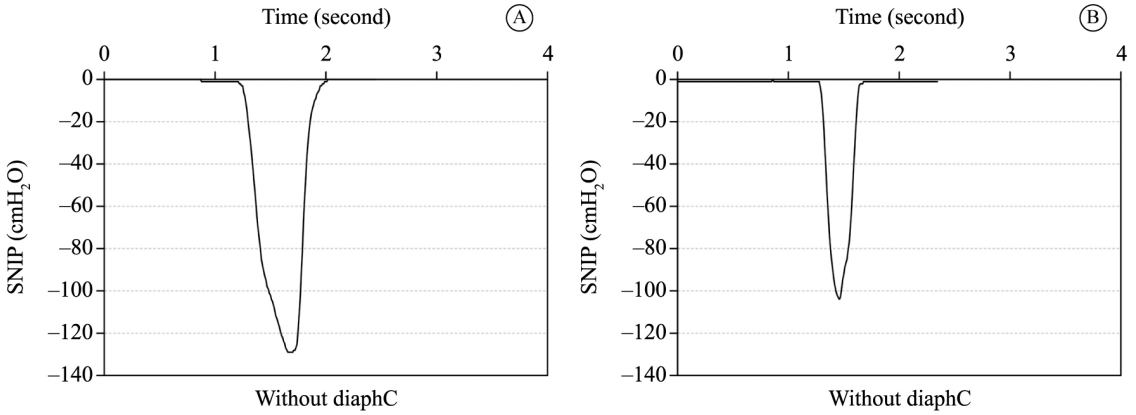

Figure 3. Graphic representation of Sniff Nasal-Inspiratory Pressure (SNIP) kinetics. Figure (A) without diaphragmatic control and (B) with diaphragmatic control in the same subject. On the left, there is a peak pressure of ~130 cmH2O and a total duration time of sniff ~630 milliseconds; on the right, there is a peak pressure of ~103 cmH2O and a total duration time of ~350 milliseconds. diaphC: diaphragmatic activation control.

It should read:

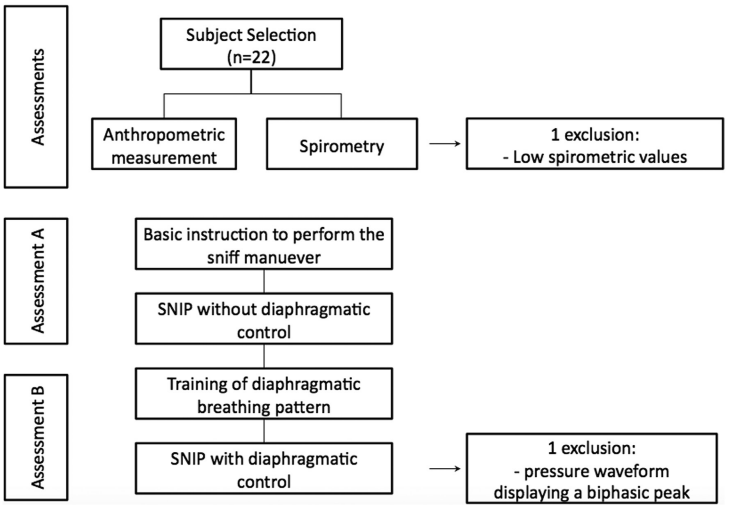

Figure 1. Study design.

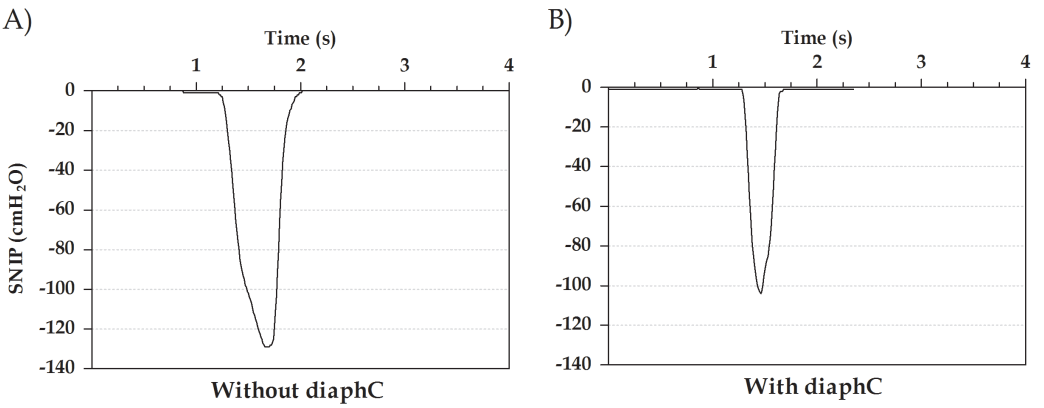

Figure 3. Graphic representation of Sniff Nasal-Inspiratory Pressure (SNIP) kinetics. Figure (A) without diaphragmatic control and (B) with diaphragmatic control in the same subject. On the left, there is a peak pressure of ~130 cmH2O and a total duration time of sniff ~630 milliseconds; on the right, there is a peak pressure of ~103 cmH2O and a total duration time of ~350 milliseconds. diaphC: diaphragmatic activation control.

 http://dx.doi.org/10.1590/bjpt-rbf.2014.0101

BJPT Brazilian Journal of Physical Therapy

INSTRUCTIONS TO AUTHORS

Scope and policies

The Brazilian Journal of Physical Therapy (BJPT) publishes original research articles, reviews, and brief communications on topics related to the professional activity of physical therapy and rehabilitation, including clinical, basic or applied studies on the assessment, prevention, and treatment of movement disorders. Our Editorial Board is committed to disseminating quality scientific investigations from many areas of expertise.

The BJPT follows the principles of publication ethics included in the code of conduct of the Committee on Publication Ethics (COPE).

All submissions will be screened for plagiarism using IThenticate software.

Open Access Policy - The BJPT is published under the Open Access model, therefore its articles are free to read, download, copy, and disseminate, provided it is for educational purposes.

Authors will not be charged any fees for the editorial processing and publication of accepted articles.

The BJPT accepts the following types of study, which must be directly related to the journal's scope and expertise areas:

a) **Experimental studies:** studies that investigate the effect(s) of one or more interventions on outcomes directly related to the BJPT's scope and expertise areas.

The World Health Organization defines a clinical trial as "any research study that prospectively allocates human participants or groups of humans to one or more health-related interventions to evaluate the effect(s) on health outcome(s)". Clinical trials include single-case experimental studies, case series, nonrandomized clinical trials, and randomized clinical trials. Randomized controlled trials (RCTs) must follow the CONSORT (Consolidated Standards of Reporting Trials) recommendations, which are available at: http://www.consort-statement.org/consort-statement/overview0/.

The CONSORT checklist and Statement Flow Diagram, available at http://www.consort-statement.org/consort-statement/flow-diagram,must be completed and submitted with the manuscript.

Clinical trials must provide registration that satisfies the requirements of the International Committee of Medical Journal Editors (ICMJE), e.g. ttp://clinicaltrials.gov/ and/or http://www.anzctr.org.au. The complete list of all clinical trial registries can be found at: http://www.who.int/ictrp/network/primary/en/index.html

We suggest that all authors register clinical trials prospectively via the website www.clinicaltrials.gov.

b) **Observational studies:** studies that investigate the relationship(s) between variables of interest related to the BJPT's scope and expertise areas without direct manipulation (e.g. intervention). Observational studies include cross-sectional studies, cohort studies, and case-control studies.

c) **Qualitative studies:** studies that focus on understanding needs, motivations, and human behavior. The object of a qualitative study is guided by in-depth analysis of a topic, including opinions, attitudes, motivations, and behavioral patterns without quantification. Qualitative studies include documentary and ethnographic analysis.

d) **Systematic reviews:** studies that analyze and/or synthesize the literature on a topic related to the scope and expertise areas of the BJPT. Systematic reviews that include meta-analysis will have priority over other systematic reviews. Those that have an insufficient number of articles or articles with low quality in the Methods section and do not include an assertive and valid conclusion about the topic will not be considered for peer-review analysis. The authors must follow the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist to format their systematic reviews. The checklist is available at http://www.prisma-statement.org/PRISMAStatement/Default.aspx and must be filled in and submitted with the manuscript. Potential authors are encouraged to read the paper Mancini MC, Cardoso JR, Sampaio RF, Costa LCM, Cabral CMN, Costa LOP. Tutorial for writing systematic reviews for the Brazilian Journal of Physical Therapy (BJPT). Braz J Phys Ther. 2014 Nov-Dec; 18(6):471-480. http://dx.doi.org/10.1590/bjpt-rbf.2014.0077.

e) **Studies on the translation and cross-cultural adaptation of questionnaires or assessment tools:** studies that aim to translate into and/or

cross-culturally adapt foreign questionnaires to a language other than that of the original version of existing assessment instruments. The authors must use the checklist (Appendix) to format this type of paper and adhere to the other recommendations of the BJPT. The answers to the checklist must be submitted with the manuscript. At the time of submission, the authors must also include written permission from the authors of the original instrument that was translated and/or cross-culturally adapted.

f) **Methodological studies**: studies centered on the development and/or evaluation of clinimetric properties and characteristics of assessment instruments. The authors are encouraged to use the Guidelines for Reporting Reliability and Agreement Studies (GRRAS) to format methodological papers, in addition to following BJPT instructions.

Important: Studies that report electromyographic results must follow the Standards for Reporting EMG Data recommended by ISEK (International Society of Electrophysiology and Kinesiology), available at http://www.isek.org/wp-content/uploads/2015/05/Standards-for-Reporting-EMG-Data.pdf.

g) **Clinical trial protocols**: The BJPT welcomes the publication of clinical trial protocols. We only accept trial protocols that are substantially funded, have ethics approval, and have been prospectively registered. Authors should use the SPIRIT statement while formatting the manuscript (http://www.spirit-statement.org).

h) **Short communications:** The BJPT will publish one short communication per issue (up to six a year) in a format similar to that of the original articles, containing 1200 words and up to two figures, one table, and ten references.

The following types of study will be considered as low priority for publication:

- narrative reviews;
- case studies.

Ethical and legal aspects

Submitting a manuscript to the BJPT implies that the paper has not been submitted simultaneously to another journal. The papers published in the BJPT are free access and distributed under the terms of Creative Commons Attribution, Non-Commercial License (http://creativecommons.org/licenses/by/3.0/deed), which allows free non-commercial use, distribution, and reproduction into any means, as long as the original format is maintained. The reproduction of part of a manuscript, even partially, including translation to another language, requires prior authorization from the editor.

The authors must cite the corresponding credits. Ideas, data or phrases from other authors without the appropriate citations and with hints of plagiarism will be subject to penalties according to the COPE code of conduct.

If part of the material has been presented in a preliminary format (at a symposium, conference, etc.), the reference of the presentation must be cited as a footnote in the title page.

The use of patient initials, names or hospital registration numbers must be avoided. Patients must not be identified in photographs, except with their express written consent attached to the original article at the time of submission.

Studies in humans must be in agreement with COPE ethical standards and must be approved by the institution's ethics committee.

Animal experiments must comply with international guidelines (such as those of the Committee for Research and Ethical Issues of the International Association for the Study of Pain, published in Pain, 16:109110, 1983).

The BJPT reserves the right not to publish manuscripts that do not adhere to the legal and ethical rules for human and animal research.

Authorship criteria

The BJPT accepts submissions of manuscripts with up to six (6) authors. The BJPT's authorship policy follows ICMJE requirements for Manuscripts Submitted to Biomedical Journals (www.icmje.org), which state that "authorship credit should be based on 1) substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; 2) drafting the article or revising it critically for important intellectual content; and 3) final approval of the version to be published." Conditions 1, 2, and 3 should all be met simultaneously. Grant acquisition, data collection, and/or general supervision of a research group do not justify authorship and must be recognized in the acknowledgements.

In exceptional cases, the editors may consider a request for submission of a manuscript with more than six (6) authors. The criteria for analysis include

the type of study, potential for citation, quality, and methodological complexity, among other things. In these exceptional cases, each author's contribution must be described at the end of the text, after the Acknowledgements and right before the References as recommended by the ICMJE. All authors are solely responsible for the content of the submitted manuscripts. All published material becomes property of the BJPT, which will retain the copyrights. Therefore, no material published in the BJPT may be reproduced without written permission from the editors. All authors of the submitted manuscript must sign a copyright transfer agreement form valid from the date of the acceptance of the manuscript.

Manuscript form and presentation

Original manuscripts

The official language of the BJPT is English. The BJPT accepts the submission of manuscripts with up to 3,500 words (excluding title page, abstract, references, tables, figures, and legends). The information contained in appendices will be included in the total number of words allowed.

The manuscript must include a title and identification page, abstract, and keywords before the body of the manuscript. References, tables, figures, and appendices should be inserted at the end of the manuscript.

Title and identification page

The title of the manuscript must not exceed 25 words and must include as much information about the study as possible. Ideally, the terms used in the title should not appear in the list of keywords. The identification page must also contain the following details:

Full title and short title of up to 45 characters to be used as a legend on the printed pages;

Authors: author's first and last name in capital letters without title followed by a superscript number (exponent) identifying the institutional affiliation (department, institution, city, state, country). For more than one author, separate using commas;

Corresponding author: name, full address, email, and telephone number of the corresponding author who is authorized to approve editorial revisions and provide additional information if needed.

Keywords: up to six indexing terms or keywords.

Abstract

The abstract must be concise, not exceeding 250 words in a single paragraph in English, and must be inserted immediately after the title page. Do not include references, footnotes or undefined abbreviations in the abstract. It must be written in a structured format.

Bullet points

On a separate page, the manuscript must identify three to five phrases that capture the essence of the topic under investigation and the main conclusions of the paper. Each bullet point must be written in a summarized fashion and provide the main contributions of the study to the current literature, as well as the clinical implications (i.e., how the results can influence clinical practice or scientific research in the area of physical therapy and rehabilitation). These points must be presented in a text box in the beginning of the article, after the abstract. Each bullet point must have no more than 80 characters (with spaces).

Introduction

This part of the manuscript should describe and define the topic under investigation, explain the relationships with to other studies in the same field, justify the need for the study, and specify the objective(s) of the study and hypotheses, if applicable.

Methods

This section consists in describing the methodological design of the study and presenting a clear and detailed report of the study participants and data collection procedures, transformation/reduction, and analysis in order to allow reproducibility of the study. For clinical trials, the participant selection and allocation process must be organized in a flowchart containing the number of participants in each phase as well as their main characteristics (see model of CONSORT flow diagram).

Whenever relevant to the type of study, the author should include the calculation that adequately justifies the sample size for investigation of the intervention effects. All of the information needed to estimate and justify the sample size used in the study must be clearly stated.

The authors must describe the dependent and independent variables; whether the parametric assumptions were met; specify the software used in

the data analysis and the level of significance; and specify the statistical tests and their purpose.

Results

The results should be presented briefly and concisely. Pertinent results must be reported with the use of text and/or tables and/or figures. Data included in tables and figures must not be duplicated in the text.

The results must be summarized into self-explanatory graphs or tables using measures of central tendency and variability (e.g. mean (SD) instead of mean±SD); must include measures of magnitude of effect (e.g. effect size) and/or indicators of the precision of the estimates (e.g. confidence intervals); must report the power of the non-significant statistical tests.

Discussion

The purpose of the discussion is to interpret the results and to relate them to existing and available knowledge, especially the knowledge already presented in the Introduction. Be cautious when emphasizing recent findings. The data presented in the Methods and/or in the Results sections should not be repeated. Study limitations, implications, and clinical application to the areas of physical therapy and rehabilitation sciences must be described.

References

The recommended number of references is 30, except for systematic reviews of the literature. Avoid references that are not available internationally, such as theses and dissertations, unpublished results and articles, and personal communication. References should be organized in numerical order of first appearance in the text, following the Uniform Requirements for Manuscripts Submitted to Biomedical Journals prepared by the ICMJE.

Journal titles should be written in abbreviated form, according to the List of Journals of Index Medicus. Citations should be included in the text as superscript (exponent) numbers without dates. The accuracy of the references appearing in the manuscript and their correct citation in the text are the responsibility of the author(s).

Examples: http://www.nlm.nih.gov/bsd/uniform_requirements.html.

Tables, Figures, and Appendices

An overall total of five (5) tables and figures is allowed. Appendices must be included in the number of words allowed in the manuscript. In the case of previously published tables, figures, and appendices, the authors must provide a signed permission from the author or editor at the time of submission.

- **Tables:** these must include only indispensable data and must not be excessively long (maximum allowed: one A4 page with double spacing). They should be numbered consecutively using Arabic numerals and should be inserted at the end of the text. Small tables that can be described in the text are not recommended. Simple results are best presented in a phrase rather than a table.
- **Figures:** these must be cited and numbered consecutively using Arabic numerals in the order in which they appear in the text. The information in the figures must not repeat data described in tables or in the text. The title and legend(s) should explain the tables and figures without the need to refer to the text. All legends must be double-spaced, and all symbols and abbreviations must be defined. Use uppercase letters (A, B, C, etc.) to identify the individual parts of multiple figures.

Whenever possible, all symbols should be placed in the legends. However, symbols identifying curves in a graph can be included in the body of the figure, provided this does not hinder the analysis of the data. Figures in color will only be published in the online version. With regard to the final artwork, all figures must be in high resolution or in its original version. Low-quality figures will not be accepted and may result in delays in the process of review and publication.

- **Acknowledgements:** these must include statements of important contributions specifying their nature. The authors are responsible for obtaining the authorization of individuals/institutions named in the acknowledgements.

Authors are strongly encouraged to use EQUATOR network checklists that are specific for their research design (for example, CONSORT statement for clinical trials, PRISMA statement for systematic reviews or STROBE statement for observational studies).

All statements from the EQUATOR network can be found on the following website: http://www.equator-network.org

Electronic submission

Manuscripts must be submitted, in English, via the website http://www.scielo.br/rbfis.

It is the authors' responsibility to remove all information (except on the title and identification page) that may identify the article's source or authorship.

When submitting a manuscript for publication, the authors must include, in addition to the files described above, the following supplementary documents: Cover letter; 2) Conflict of interest statement; and 3) Copyright transfer statement signed by all authors.

The review process

The submissions that meet the journal's standards and are in accordance with the BJPT editorial policies will be forwarded to the area editors, who will perform an initial assessment and recommend them or not to the chief editor for peer-review. The criteria used for the initial analysis of the area editor include: originality, pertinence, clinical relevance, and methodology. The manuscripts that do not have merit or do not conform to the editorial policies will be rejected in the pre-analysis phase, regardless of the adequacy of the text and methodological quality. Therefore, the manuscript may be rejected based solely on the recommendation of the area editor without the need for further review, in which case, the decision is not subject to appeal. The manuscripts selected for pre-analysis will be submitted to review by specialists, who will work independently. The reviewers will remain anonymous to the authors, and the authors will not be identified to the reviewers. The editors will coordinate the exchange between authors and reviewers and will make the final decision on which articles will be published based on the recommendations of the reviewers and area editors. If accepted for publication, the articles may be subject to minor changes that will not affect the author's style. If an article is rejected, the authors will receive a justification letter from the editor. After publication or at the end of the review process, all documentation regarding the review process will be destroyed.

Areas of expertise

1. Physiology, Kinesiology, and Biomechanics; 2. Kinesiotherapy/therapeutic resources; 3. Motor development, acquisition, control, and behavior; 4. Education, Ethics, Deontology, and Physical Therapy History; 5. Assessment, prevention, and treatment of cardiovascular and respiratory disorders; 6. Assessment, prevention, and treatment of aging disorders; 7. Assessment, prevention, and treatment of musculoskeletal disorders; 8. Assessment, prevention, and treatment of neurological disorders; 9. Assessment, prevention, and treatment of gynecological disorders; 10. Assessment and measurement in Physical Therapy; 11. Ergonomics/ Occupational Health.

APPENDIX:

Checklist for reviewers/authors of studies on the translation and cross-cultural adaptation of questionnaires/assessment instruments

Instructions to reviewers/authors:

- Stage I: Translation into Portuguese:
  - ✓ Did the authors mention the presence of at least two translators?
  - ✓ Were the bilingual translators native speakers of Brazilian Portuguese?
  - ✓ Did the translators have different professional backgrounds and profiles (i.e. one translator has knowledge of the concepts assessed by the instrument and the other is not related to the health area)?
  - ✓ Did the translators work on the translation independently?
  - ✓ Did the authors describe the translator's questions or changes and the rationale behind the translation?

- Stage II: Synthesis of translation
  - ✓ Were the translators involved in the reporting of the process?
  - ✓ Were the translated versions compared to the original questionnaire to extract a synthesis of the first Portuguese version of the questionnaire (version I)?
  - ✓ Did the authors present a report of the synthesis process containing the questions that required changes and how they were resolved?
  - ✓ Was there mention of the process of consensus between the translators?
- Stage III: Back translation
  - ✓ Was version I of the translated questionnaire translated back into the original instrument's language?
  - ✓ Were at least two translators involved?
  - ✓ Were the bilingual translators native speakers of the original instrument's language?
  - ✓ Did the authors ensure that the translators were not familiar with the original version of the questionnaire?
  - ✓ Did the authors ensure that these translators did not have a background in the area of health or information about the concepts explored by the questionnaire or instrument?
- Stage IV: Expert committee
  - ✓ Did the Committee include methodologists, health professionals, language professionals, and translators (Stage I and II translators and Stage III back translators)?
  - ✓ Were the authors of the original questionnaire contacted and did they grant approval for the cross-cultural adaptation? (Required)
  - ✓ Is there mention of the participation of the authors of the original questionnaire during this stage? (Not required)
  - ✓ Did the consolidation of a pre-final version consider all reports, translations, and back translations?
  - ✓ Is there mention of the aspects that required changes at this stage and of how they were resolved?
  - ✓ Were the Committee's decisions aimed at ensuring semantic, idiomatic, experimental, and conceptual equivalence between the versions?
- Stage V: Pre-test of pre-final version
  - ✓ Was the pre-final version tested on at least 30 subjects?
  - ✓ Were these subjects part of the target population of the assessed questionnaire or instrument?
  - ✓ Did every subject answer the questionnaire or instrument and was each one interviewed to explore their comprehension of each item and answer of the questionnaire? Guillemin et al. (1993) suggest posing the question: "What did you mean?" to assess their understanding of the item.
  - ✓ Did the authors report the percentage of uncertainties during this part of the process (pre-final version)? Uncertainties reported by 15 or 20% or more of the sample indicate the need for revision of the questionnaire (Ciconelli et al., 1999; Nusbaum et al., 2001). If the percentage is greater than 15% or if more subjects were included, the translated and adapted version of the questionnaire or instrument must be changed and a new pre-test must be conducted and reported.
  - ✓ For original instruments already established in the literature and whose construct has been assessed, the authors should briefly describe the results of this assessment. Otherwise, the authors of the current version must assess the construct using the data from the translation.
  - ✓ We recommend that the original instrument be submitted with the manuscript as a separate file.

Checklist – Submission of studies on translations and cross-cultural adaptations and validation	Author		Reviewer
Items of translation and cross-cultural adaptation	**Mark with an X**	**Reported on page no.**	**Mark with an X**
1 - Title mentions that it is a translation and cross-cultural adaptation.			
2 - Reference to original instrument was included in Methods.			
3 - Reference to original instrument was included in References.			
4 - Translated instrument was included in full at the time of submission.			
5 - Original instrument was submitted in full.			
6 - Authorization was given by the authors of the original instrument.			
7 - Guidelines by Beaton et al. (2000) were followed in the translation and adaptation stages and the authors clearly mention the use of this guideline.			
Translation - 2 translators (1 lay translator and 1 specialized in the area).			
Meeting of translation committee (synthesis of translation).			
Back translation - 2 lay translators.			
Meeting of expert committee.			
Test of pre-final version (n>30).			
Rate of comprehension was described in the test of the pre-final version – uncertainties reported by 15 or 20% or more of the sample indicate the need for revision of the questionnaire (Nusbaum et al., 2001).			
8 - All items of the questionnaire were translated and cross-culturally adapted, including alternative answers and instructions.			
9 - A clear description was given of the cultural adaptations made during the study.			
10 - A clear description of the sample characteristics was included in the stages of the study.			
Measurement properties			
Required			
11 - Was the translated instrument's reproducibility (test-retest) assessed?			
12 - Was the sample size adequate for assessment of the reproducibility? (Terwee et al., 2007)			
13 - Was the translated instrument's internal consistency assessed?			
14 - Was the sample size adequate for assessment of the internal consistency? (Terwee et al., 2007)			
Recommended			
15 - Was confirmatory factor analysis of the translated instrument conducted?			
16 - Was the sample size adequate (Mokkink et al., 2010) for confirmatory factor analysis of the translated instrument?			
OR			
17 - If exploratory factor analysis was not conducted in the original study, was it conducted in the translation study?			
18 - Was the sample size adequate (Mokkink et al., 2010) for exploratory factor analysis of the translated instrument?			

References

- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976). 2000;25(24):3186-3191.
- Ciconelli RM, Ferraz MB, Santos W, Meinão I, Quaresma MR. Tradução para Língua Portuguesa e validação do questionário genérico de avaliação de qualidade de vida SF-36. Revista Brasileira de Reumatologia. 1999;39:143-150.
- Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol. 1993;46(12):1417-1432.
- Nusbaum L, Natour J, Ferraz MB, Goldenberg J. Translation, adaptation and validation of the Roland-Morris questionnaire. Braz J Med Biol Res. 2001;34(2):203-210.
- Chen WH, Lenderking W, Jin Y, Wyrwich KW, Gelhorn H, Revicki DA. Is Rasch model analysis applicable in small sample size pilot studies for assessing item characteristics? An example using PROMIS pain behavior item bank data. Qual Life Res. 2014 Mar;23(2):485-93.
- Terwee CB1, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007 Jan;60(1):34-42.

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA REABILITAÇÃO MESTRADO E DOUTORADO

Recomendado pela CAPES – Conceito 5

O Programa de Pós-graduação em Ciências da Reabilitação tem como base a perspectiva apresentada no modelo proposto pela Organização Mundial de Saúde e propõe que as dissertações e trabalhos científicos desenvolvidos estejam relacionados com o desempenho funcional humano. Com a utilização de um modelo internacional, espera-se estimular o desenvolvimento de pesquisas que possam contribuir para uma melhor compreensão do processo de função e disfunção humana, contribuir para a organização da informação e estimular a produção científica numa estrutura conceitual mundialmente reconhecida. O Programa de Pós-graduação em Ciências da Reabilitação tem como objetivo tanto formar como aprofundar o conhecimento profissional e acadêmico, possibilitando ao aluno desenvolver habilidades para a condução de pesquisas na área de desempenho funcional humano.

O programa conta com parcerias nacionais e internacionais sedimentadas, e os seus laboratórios de pesquisa contam com equipamentos de ponta para o desenvolvimento de estudos na área de Ciências da Reabilitação.

Mais informações

Fone/Fax: (31) 3409-4781
www.eef.ufmg.br/mreab

Universidade Federal de São Carlos

Programa de Pós-Graduação em Fisioterapia

O Programa de Pós-Graduação em Fisioterapia tem como área de concentração: "Processos de Avaliação e Intervenção em Fisioterapia". Nosso objetivo é oferecer condições acadêmicas necessárias para que o aluno adquira um repertório teórico e metodológico, tornando-se apto a exercer as atividades de docente de nível universitário e iniciá-lo na carreira de pesquisador.

Os cursos de mestrado e doutorado (stricto sensu) foram os primeiros criados na área de fisioterapia do país.

Linhas de pesquisa do programa são:

- ☐ Instrumentação e Análise Cinesiológica e Biomecânica do Movimento
- ☐ Processos de Avaliação e Intervenção em Fisioterapia do Sistema Músculo-Esquelético
- ☐ Processos Básicos, Desenvolvimento e Recuperação Funcional do Sistema Nervoso Central
- ☐ Processos de Avaliação e Intervenção em Fisioterapia Cardiovascular e Respiratória

Recomendado pela CAPES – Conceito 6

Mais informações
Fone: (16) 3351-8448
www.ppgft.ufscar.br
e-mail ppg-cr@ufscar.br

O que é PEDro?

PEDro, Physiotherapy Evidence Database, é uma base de dados eletrônica gratuita de evidências relevantes em fisioterapia. PEDro permite acesso rápido a mais de 17.000 estudos clínicos aleatorizados, revisões sistemáticas e diretrizes de prática clínica em fisioterapia, fornecendo, quando possível, resumo e link para o texto completo de cada documento indexado.

Todos os estudos clínicos na PEDro são avaliados para fins de classificação de qualidade. Esses critérios de qualidade permitem aos usuários identificar de forma rápida os estudos que mais possivelmente contenham informações válidas e úteis para guiar a prática clínica.

PEDro é gratuita. E PEDro está disponível em português:

www.pedro.org.au/portuguese

PEDro

PHYSIOTHERAPY EVIDENCE DATABASE

CONHEÇA VISITE CONSULTE CONTRIBUA

BJPT Brazilian Journal of Physical Therapy

ISSN 1413-3555

2016 Mar/Apr; 20(2)