



Revista Brasileira de Fisioterapia

ISSN: 1413-3555

rbfisio@ufscar.br

Associação Brasileira de Pesquisa e Pós-
Graduação em Fisioterapia
Brasil

Silva, Anderson M.; Costa, Lucíola C. M.; Comper, Maria L.; Padula, Rosimeire S.
Cross-cultural adaptation and reproducibility of the Brazilian-Portuguese version of the
modified FRESNO Test to evaluate the competence in evidence based practice by
physical therapists

Revista Brasileira de Fisioterapia, vol. 20, núm. 1, enero-febrero, 2016, pp. 26-47
Associação Brasileira de Pesquisa e Pós-Graduação em Fisioterapia
São Carlos, Brasil

Available in: <http://www.redalyc.org/articulo.oa?id=235045322004>

- How to cite
- Complete issue
- More information about this article
- Journal's homepage in redalyc.org

redalyc.org

Scientific Information System

Network of Scientific Journals from Latin America, the Caribbean, Spain and Portugal

Non-profit academic project, developed under the open access initiative

Cross-cultural adaptation and reproducibility of the Brazilian-Portuguese version of the modified FRESNO Test to evaluate the competence in evidence based practice by physical therapists

Anderson M. Silva^{1,2}, Lucíola C. M. Costa¹, Maria L. Comper^{1,3}, Rosimeire S. Padula¹

ABSTRACT | Background: The *Modified Fresno Test* was developed to assess knowledge and skills of both physical therapy (PT) professionals and students to use evidence-based practice (EBP). **Objectives:** To translate the Modified Fresno Test into Brazilian-Portuguese and to evaluate the test's reproducibility. **Method:** The first step consisted of adapting the instrument into the Brazilian-Portuguese language. Then, a total of 57 participants, including PT students, PT professors and PT practitioners, completed the translated instrument. The responses from the participants were used to evaluate reproducibility of the translated instrument. Internal consistency was calculated using the Cronbach's alpha. Reliability was calculated using the intraclass correlation coefficient (ICC) for continuous variables, and the Kappa coefficient (K) for categorical variables. The agreement was assessed using the standard error of the measurement (SEM). **Results:** The cross-cultural adaptation process was appropriate, providing an adequate Brazilian-Portuguese version of the instrument. The internal consistency was good ($\alpha=0.769$). The reliability for inter- and intra-rater assessment were ICC=0.89 (95% CI 0.82 to 0.93); for evaluator 1 was ICC=0.85 (95% CI 0.57 to 0.93); and for evaluator 2 was ICC=0.98 (95% CI 0.97 to 0.99). The SEM was 13.04 points for inter-rater assessment, 12.57 points for rater 1 and 4.59 points for rater 2. **Conclusion:** The Brazilian-Portuguese language version of the *Modified Fresno Test* showed satisfactory results in terms of reproducibility. The *Modified Fresno Test* will allow physical therapy professionals and students to be evaluated on the use of understanding EBP.

Keywords: evidence-based practice; modified Fresno scale; physical therapy; professional practice; professional education.

BULLET POINTS

- Evidence-based practice (EBP) is critical to clinical decision-making.
- Physical therapists should be evaluated for their knowledge and skills to use EBP.
- The Modified Fresno Test is a self-explanatory instrument designed to test the knowledge and skills of PTs.
- The results of the cross-cultural adaptation and instrument reproducibility for the Brazilian Portuguese version of the modified Fresno Test were satisfactory.

HOW TO CITE THIS ARTICLE

Silva AM, Costa LCM, Comper ML, Padula RS. Cross-cultural adaptation and reproducibility of the Brazilian-Portuguese version of the modified FRESNO Test to evaluate the competence in evidence based practice by physical therapists. *Braz J Phys Ther*. 2016 Jan-Feb; 20(1):26-47. <http://dx.doi.org/10.1590/bjpt-rbf.2014.0140>

● Introduction

Evidence-Based Practice (EBP) can be defined as the use of relevant scientific evidence to guide clinical decision making and to optimize health outcomes of patients¹. Therefore, EBP incorporates knowledge related to clinical research of high quality, professional knowledge, and patient's preferences¹⁻⁴. This approach has been increasingly used by physical

therapists. In order to be effective and applicable in clinical practice, EBP should be able to^{3,5}: formulate a clinical question; drive effective searches of databases; critically assess the methodological quality of study findings; and provide the evidence for therapeutic decisions^{2,6-10}. Any difficulty in performing these steps could be a barrier to the adoption of EBP⁵. To that end,

¹ Programa de Mestrado e Doutorado em Fisioterapia, Universidade Cidade de São Paulo (UNICID), São Paulo, SP, Brazil

² Curso de Fisioterapia, Faculdades Integradas do Vale do Ribeira, Registro, SP, Brazil

³ Curso de Fisioterapia, União Metropolitana de Ensino e Cultura, Itabuna, BA, Brazil

Received: Oct. 23, 2014 Revised: Apr. 27, 2015 Accepted: Sept. 11, 2015

adequate training¹¹⁻¹³ and the constant evaluation of knowledge and skills acquired by the professionals during their training using reliable instruments are fundamental^{14,15}.

The concept of competence is broad¹⁶, and although there is no single definition, the consensus is that integration of personal attributes to the actions carried out in specific contexts, aimed at achieving results, are key aspects¹⁷. In the physical therapy context, competence indicates the capacity to manipulate cognitive resources that add knowledge, information, skills, and, above all, intelligence, to effectively and with relevance, solve a number of situations in clinical practice^{18,19}. The term could also incorporate the knowledge domain², ability to work in teams, and to be effective in verbal and written communication^{2,17,20}.

Although there are a number of instruments¹² that are used to assess the EBP in health professionals have been identified, few have had their validity and reliability tested or have failed to evaluate EBP in its entirety²¹. The Fresno Test^{1,2,14}, originally developed to assess the ability of medical professionals to evaluate medical literature, is the only instrument that evaluates all stages of EBP^{2,12}. It was later adapted to evaluate other health professionals such as nurses¹⁸, occupational therapists²² and physical therapists^{7,18}. This Modified Fresno Test⁷ was developed to assess the knowledge and skills of physical therapy professionals and students to use EBP principles in their daily professional practices & classes. The instrument identifies the whether the individual has the academic and professional knowledge necessary to be able to use high methodological quality evidence for clinical decision-making.

In the absence of instruments that assess the knowledge and skills of Brazilian PT students and professionals to use EBP, the need for a cross-cultural adaptation of a suitable test instrument to be used in the Brazilian-Portuguese language becomes evident. Thus, this study aimed to translate and cross-culturally adapt the Modified Fresno Test for physical therapists, and to evaluate the reproducibility of the instrument in the Brazilian-Portuguese version.

● Method

The Modified Fresno Test instrument

The Modified Fresno Test is a self-explanatory tool that evaluates competencies and skills of professionals and students to use EBP⁷. The instrument is composed of an initial text with instructions for completion and

two clinical scenarios. Each participant must choose one of the scenarios provided and answer 13 questions pertaining to the chosen scenario. Questions 9 and 10 require mathematical calculations. The questionnaire must be answered within the maximum permitted time of sixty minutes.

The total score is calculated from the partial score of each question which has 4 possible answers (a to d) to choose from. For example, question 2 asks the participant to name the research databases selected to answer the clinical question, which has the following items listed: (a) different resources; (b) convenient resources; (c) clinical relevant resources; and (d) valid resources. The response or answer for each item is then scored using one of five classification categories as follows: (1) not clear; (2) limited; (3) minimum; (4) strong; and (5) excellent⁷. The sum of scores for each criterion results in a score per question that ranges from 0 to 24 points: Q1 to Q7 allow a maximum score of 24 points; Q8 and Q10 a maximum of 16 points; Q9 allows 12 points; and Q11 to Q13 a maximum of 4 points. The total test score is the sum of the points from all questions, ranging from 0 to 224 points⁷ (Table 1). After the participant responded to all of the questions, two trained raters scored the responses.

Cross-cultural adaptation of the Modified Fresno Test

The first step was to cross-culturally adapt the Modified Fresno Test⁷ to the Brazilian-Portuguese language. To translate and adapt the instrument, the guidelines for translation and adaptation of health care questionnaires, proposed by Beaton et al.¹⁹, were adopted¹⁹. Thus, the following steps were conducted: translation, synthesis of the translation, back-translation into English, summary of the back-translation, evaluation of the new document by an expert committee followed by a pre-test.

The resulting version of the translation process was pre-tested using 30 physical therapists who had 1 to 10 years (mean=5.7; SD=3.1) experience who may or may not have had experience using EBP principles. The participants were 22 females and 8 males, between the ages of 25 and 36 years, who had graduated from public or private universities that were not involved in the first part of the study. Of those, 21 participants received the questionnaire in digital format and 9 in print format. Each participant was asked to read and answer the questionnaire and at the end of each question, there was a blank field with the text "*here, describe your suggestions and/or difficulty*

Table 1. Modified Fresno Test for physical therapists – Brazilian-Portuguese Edition: Percentage and scoring of the questions according to the steps of Evidence Based Practice.

Step/Action	Questions - Q	Scores	% and total score
Step 1: Elaboration of the question	Q1 - Formulate a clinical question	Q1 – (0 to 24)	11% (24)
Step 2: Search for the best available evidence	Q2 - Information Sources Q4 - Search (search strategy)	Q2 – (0 to 24) Q4 – (0 to 24)	21% (48)
Step 3a: Critical evaluation (qualitative evidence)	Q3 - Study design Q5 - Relevance Q6 - Internal Validity Q7 - Magnitude and significance Q12 - Best study design (diagnosis) Q13 - Best study design (prognosis)	Q3 – (0 to 24) Q5 – (0 to 24) Q6 – (0 to 24) Q7 – (0 to 24) Q12 – (0 or 4) Q13 – (0 or 4)	46% (104)
Step 3b: Critical evaluation (quantitative evidence)	Q9 - Sensitivity, positive predictive value and positive likelihood ratio Q10 - Absolute risk reduction, relative risk, NNT, and p-value Q11 - Confidence Interval	Q9 – (0 to 12) Q10 – (0 to 16) Q11 – (0 or 4)	14% (32)
Step 4: Implementation of evidence into clinical practice	Q8 - Questioning the patient/family	Q8 – (0 to 16)	7% (16)
		Total (0 to 224)	Total 100% (224)

Q: Question; NNT: Number needed to treat.

in answering this question”, so that they could report any difficulties encountered when understanding the questions. Once completed, the questionnaires were scored by two investigators with experience in EBP practice, working in the areas of musculoskeletal and cardiorespiratory physical therapy for 7 years.

Measurement properties

This phase tested the measurement properties of the Modified Fresno Test for the Brazilian-Portuguese version, following the guidelines proposed by Mokkink et al.²³ for testing internal consistency and intra- and inter-rater reliability, and the guidelines of Terwee et al.²⁴ to test the compliance of the instrument.

Reliability measures the accuracy of an instrument to provide the same answer in repeated measurements, representing the relative error of the measurement. It can be measured when the instrument is used more than once by the same examiner (intra-rater reliability) and when the instrument is used to evaluate the same condition by different raters (inter-rater reliability)^{23,25,26}. The correlation represents the degree to which the repeated analysis of the same individual provided similar responses^{24,25}.

The statistical tests adopted for the reliability analysis were the same ones used with the original

instrument⁷. The Cronbach’s alpha (α) was used to assess the internal consistency of the instrument (ranging between 0.70 and 0.95), the intraclass correlation coefficient type_{2,1} (ICC_{2,1}) was used to test the reliability of the numeric or quantitative variables in continuous scales²⁷, and the Kappa coefficient (K) was used for the categorical variables. The ICC was classified according to Fleiss²⁷ as poor (<0.4), good (0.4 to 0.75) and excellent reliability (>0.75). The Kappa was classified according to Landis and Koch²⁸ as reliability being almost perfect (>0.81), substantial (0.61 to 0.80), moderate (0.41 to 0.60), low (0.21 to 0.40) and very low (0.00 to 0.20). The reliability tests were performed with a 95% confidence interval.

The agreement between the assessors was assessed using the standard measurement error (SME), which reflected the error of the instrument, expressed by the standard deviation of the differences of the test and re-test, divided by the square root of 2²⁴. The relationship between the SME and the total score of the questionnaire was interpreted as proposed by Ostelo et al.²⁹ who considered values $\leq 5\%$ as very good, $>5\%$ and $\leq 10\%$ as good, $>10\%$ and $\leq 20\%$ as doubtful and $>20\%$ as negative. All data collected were transferred to the software Statistical Package for Social Sciences (SPSS) version 17.0 for analysis.

Procedures

This study was approved by the Ethics Committee of the Universidade Cidade de São Paulo (UNICID), São Paulo, SP, Brazil, under the Research Protocol number 13696713/2012 and authorized by the authors who developed the original instrument. All participants signed an informed consent form.

The Modified Fresno Test was translated into Brazilian-Portuguese by two bilingual (Portuguese and English) translators, both with over 15 years of linguistics training, whose first language was Brazilian-Portuguese. Translator one had 11 years experience in higher education, was knowledgeable in health terminology and was aware of the instrument goals, while translator two had 9 years experience in higher education, but no knowledge of how EPB was used, or experience in the health area. Next, the translated versions were compared and analyzed and a consensual approach by the 2 translators to resolve any differences of interpretation and translation was conducted. Thus, the two independent translations (T_1 and T_2) were created and then synthesized into a single consensual translation, termed T_{12} .

The T_{12} version was then back translated into English by two different independent translators (back-translator 1 and back-translator 2). These translators had backgrounds in physical therapy and were knowledgeable in the EBP methodology and were fluent in Portuguese and English, with the latter being their mother tongue. In addition, these back-translators, although aware of EBP methodology, were unaware of the Modified Fresno Test. The results of the two back-translations (RT_1 and RT_2) were synthesized in a consensual manner similar to the original translations resulting in a single consensual translation called RT_{12} .

Finally, a Committee of Experts, consisting of methodologists, health professionals and professional translators (2 translators, 2 back-translators, 6 physical therapists with teaching experience in EBP), evaluated and consolidated all versions ($R12$ & $T12$). The committee compared the translations and discussed the semantic, structural, idiomatic and conceptual properties from the Brazilian-Portuguese version and developed a pre-final version. After conducting the pre-test, barely any changes were made to the final version.

To test the final version of the instrument, a convenience sample of 159 physical therapists were invited to participate. The sample size was estimated as proposed by the guidelines for reliability tests^{24,30,31}. The sample consisted of: (1) physical therapy professors of public and private institutions involved in the clinical

practice, (2) students enrolled on the third year of their PT program and (3) physical therapists not connected to higher education institutions, independent of the their familiarity with EBP.

The volunteers were instructed to use a notebook and a calculator to assist in answering all of the questions. They were asked to complete the task within the stipulated time (60 minutes) even though the original study had not justified the time parameters. Data collection was conducted from April to September 2013.

The evaluators received a single training, divided into three steps of one hour each. The first hour was an orientation about the criteria and scoring of the questions. The second hour was a pilot test in which each evaluator scored one sample test. The third hour was for analysis and discussion of the score results. The evaluators were blinded to avoid identification of the participants and the score of the instrument during the test and re-test.

After the training, the evaluators received the questionnaires answered by 57 participants and scored the answers to obtain the partial and total test scores of the Modified Fresno Test. Inter- and intra-rater reliability tests and inter-rater agreement of the partial and full scores of the instrument were carried out using a test/retest design with an interval of seven days between scoring sessions. The test/retest was performed by two independent evaluators who did not participate in the previous stages.

Results

Of the 159 individuals invited to participate in the study, only 32% ($n=57$) returned the questionnaires. Of those, 36 were PT professionals (13 were also PT teachers in the clinic) with an average working experience time of 6.6 years, ($SD=3.8$) and 21 were PT academics (last year program). Thirty-seven answered the questionnaire in print and 20 in the digital version. The 95 participants who did not return the questionnaires were asked about their reasons for not returning the test information. Some claimed lack of knowledge of the topic and seven expressed lack of interest in the research. None of them reported difficulty in understanding the questions, but reported a lack of knowledge for answering the topics covered. The percentage of missing items per question for the respondents of the questionnaire ($n=57$) is shown in Table 2. The evaluators who scored the answers of the test/re-test did not report any problems.

Table 2. Measurement properties of the Modified Fresno Test – Brazilian-Portuguese version per item and the total sum of the questions.

Questions	Inter-rater		Intra-rater 1		Intra-rater 2	
	(n=57)	(n=57)	(n=57)	(n=57)	(n=57)	(n=57)
	ICC _{2,1} (95% CI)	Agreement – (SEM)	ICC _{2,1} (95% CI)	Agreement – (SEM)	ICC _{2,1} (95% CI)	Agreement – (SEM)
Q1	0.75 (0.60-0.84)	3.40	0.82 (0.69-0.90)	2.80	0.98 (0.97-0.99)	0.76
Q2	0.52 (0.30-0.68)	3.81	0.47 (0.20-0.67)	4.06	0.96 (0.94-0.97)	1.07
Q3	0.89 (0.82-0.93)	2.51	0.86 (0.78-0.91)	3.01	0.97 (0.95-0.98)	1.22
Q4	0.75 (0.61-0.84)	3.11	0.80 (0.68-0.88)	2.64	0.97 (0.95-0.98)	1.05
Q5	0.66 (0.48-0.78)	2.97	0.49 (0.22-0.68)	3.70	0.91 (0.85-0.94)	4.06
Q6	0.68 (0.51-0.79)	4.44	0.79 (0.67-0.87)	3.60	0.94 (0.90-0.96)	1.84
Q7	0.58 (0.38-0.73)	4.59	0.67 (0.50-0.79)	4.10	0.94 (0.90-0.96)	4.85
Q8	0.65 (0.47-0.78)	2.53	0.64 (0.31-0.81)	2.52	0.96 (0.93-0.97)	0.83
Q9	0.90 (0.84-0.94)	1.11	0.97 (0.95-0.98)	0.65	0.99 (0.98-0.99)	0.30
Q10	0.94 (0.90-0.96)	0.94	0.94 (0.90-0.96)	0.95	0.99 (0.99-0.99)	1.96
Q11	0.92 (0.87-0.95)	0.37	0.92 (0.87-0.95)	NC	1.00 (1.00-1.00)	0.37
Q12	0.75 (0.60-0.84)	0.97	0.89 (0.83-0.93)	0.65	1.00 (1.00-1.00)	0.97
Q13	0.93 (0.88-0.95)	0.53	0.93 (0.88-0.95)	0.53	1.00 (1.00-1.00)	NC
Total	0.89 (0.82-0.93)	13.04	0.85 (0.57-0.93)	12.57	0.98 (0.97-0.99)	4.59

Missing data (n=57) – Q9=7 (12.3%); Q10=5 (8.7%); Q11=4 (7.0%); Q13=1(1.8%). NC: The statistical program did not allow the calculation of these items.

Cross-cultural adaptation of the Modified Fresno Test for Brazilian-Portuguese

When the original version was compared with the versions resulting from the translations (T₁ and T₂) and back-translations (RT₁ and RT₂), no significant differences were found for the content and meaning. Grammatical adjustments were conducted for different words but the same meaning was maintained between the two versions.

In the pre-test, participants reported no problems understanding the terms and questions. However, six of the 30 physical therapists who answered the instrument (two answered in print version and four in digital) reported difficulties understanding the instrument’s instructions. Through these results, adjustments were made to facilitate understanding and the text was forwarded to the Committee of Experts, resulting in the final Modified Fresno Test – Brazilian-Portuguese version (Appendix 1 and 2).

Evaluation of reliability

The Brazilian-Portuguese version of the test showed good internal consistency (α=0.769). For evaluator 1, the ICC ranged from 0.47 to 0.97. The agreement of the SEM varied from 0.53 to 4.10 points. Evaluator 1 showed excellent reliability, except for questions 2 and 5,

which showed an ICC of 0.47 and 0.49, respectively, indicating moderate reliability. In relation to questions 11, 12 and 13, the Kappa coefficient obtained was excellent and ranged from 0.89 to 0.93.

For evaluator 2, reliability was found to be higher compared to values reported for evaluator 1. The ICC ranged from 0.98 to 0.99 and the SEM agreement ranged from 0.30 to 4.85 points. These results represented excellent reliability for all questions tested in the instrument. Questions 11, 12 and 13 showed excellent reliability with a Kappa index of 1.0. The inter-rater reliability presented an ICC that ranged from 0.52 to 0.94 and an agreement by the SEM that ranged from 0.37 to 4.59 points.

The inter-rater reliability between the 2 evaluators proved to be excellent except for question 2, which resulted in an ICC of 0.52 indicating moderate reliability. Questions 11, 12 and 13 showed good to excellent reliability, with a Kappa index ranging from 0.75 to 0.93.

For the total score of the instrument, the intra-rater reliability was excellent for both evaluators. Evaluator 1 had an ICC of 0.85 (0.57 to 0.93) and evaluator 2, an ICC of 0.98 (0.97 to 0.99). The inter-rater reliability was excellent with an ICC of 0.89 (from 0.82 to 0.93). The agreement by the SEM was 13.04 points for the

inter-rater assessment, 12.57 points for evaluator 1 and 4.59 points for evaluator 2.

● Discussion

The purpose of this study was to adapt and test the reliability properties and agreement of a Brazilian-Portuguese version of the Modified Fresno Test, an instrument, that to date, is considered to be the most complete and appropriate test for evaluating PT professionals and students knowledge and skills on the use of EBP¹².

The adaptation process to the Brazilian-Portuguese version of the Modified Fresno Test followed the steps indicated by the available literature¹⁹. The steps were completed without difficulty and adjustments were made to the content and meaning of the text. Although most of the participants failed to complete the final version of the questionnaire, the reasons were unrelated to the difficulties with the topic¹⁰ or in understanding the questions. Therefore, it was not considered a barrier in the instrument adaptation process, since different levels of skills and competence in relation to EBP was expected, as mentioned by the authors who developed the instrument⁷.

The question with the highest amount of missing data was Q9, which required statistical knowledge, followed by Q10, Q11 and Q13. The omission of EBP learning during professional training, the difficulty in dealing with statistics and conducting the databases search were the reasons for participants dropping out of this study³². It is possible that those reports were the major obstacles facing physical therapists, independent of their professional time in the process of adopting EBP. Difficulty in the searching process, interpretation and in translating the evidence into clinical practice is related to the competencies and skills of a professional; however, younger professionals have shown a more positive attitude to the adoption of EBP³³. Nevertheless, other obstacles have been pointed out by other studies in the adoption of EBP, such as limited access to databases and complete manuscripts, language issues, and the time available to learn the topic^{5,6,33-35}.

The original version of the instrument lacked classification categories for the total score of the questions, but the higher the values, the better the participants were in answering each question or a group of questions. The results of the reliability tests for the total score of the modified instrument, obtained by using the Brazilian-Portuguese version, were considered excellent for the intra- and inter-rater

assessment. The study, which validated the modified instrument for physical therapists, presented good internal consistency ($\alpha=0.769$), similar to the values obtained with the original unmodified instrument (0.780). In addition, the modified English version of the instrument presented an excellent reliability inter- and intra-observer of 0.92 (95% CI 0.88 to 0.94), for evaluator 1 an ICC of 0.96 (95% CI .91-.98), and for evaluator 2 of 0.96 (95% CI from 0.91 to 0.98)⁷.

The scores of the partial questions of the adapted instrument showed moderate to excellent results for intra- and inter-rater reliability of the evaluators. These results were similar to the original English version of the modified instrument which presented moderate to excellent intra-observer reliability for the evaluators for all questions (ICC ranging from 0.62-1.0) and moderate to excellent inter-rater reliability (ICC ranging from 0.61 to 0.99) for all questions of the instrument; however, differences between the two versions were observed for questions 2, 5 and 8.

Question 8 evaluated the ability of the participants to obtain information about the patient's perspective, which presented unsatisfactory reliability ($ICC_{2,1}=0.47$) in the original version and good reliability ($ICC_{2,1}=0.65$) in the Brazilian-Portuguese version for the inter-rater analysis. Question 2 evaluated the participant's knowledge and ability to search the database, and presented a moderate reliability for evaluator 1 and for the inter-rater assessment with the Brazilian-Portuguese version. In the original English version, reliability values were excellent for this question ($ICC_{2,1}=0.83$), as well as for evaluator 1 and for the inter-rater assessment ($ICC_{2,1}=0.90$). Question 5 referred to the participant's ability to determine the clinical relevance of the identified studies, and showed lower reliability values in the Brazilian-Portuguese version than in the original version for evaluator 1 ($ICC_{2,1}=0.49$). The score differences between the two versions for this question might be related to the level of knowledge acquired during their professional training¹⁵ of EBP between participants.

The values of the inter-rater reliability and intra-rater SEM for the Brazilian-Portuguese version showed little variability in the intra- and inter-rater scores, considering that the total instrument score ranged from 0 to 224 points. These results could not be compared with the original modified version because this property was not assessed in the original version⁷. Upon completion of the translation phases and testing of the measurement properties, the Brazilian-Portuguese version of the Modified Fresno Test for physical therapists proved to be similar in terms of the language

and clearly understood by the Brazilian population (Appendix 1*). In addition, it could be a useful tool for researchers and educational institutions in assessing the ability of Brazilian professionals to adopt EBP. According to the strategy of using the Fresno Test for medical doctors and the Modified Fresno Test for physical therapists, cited in the respective studies^{7,14}, the authors believe that this instrument is best used for continuous evaluation, which involves a constant process of assessments and reassessments to ensure people understand EBP principles and if they do not, have taken the time or are given the opportunity to develop skills in EBP.

The difficulty in obtaining a greater number of participants to answer the questionnaire during the test and after 24 hours, ensuring that there was no time to supplement the information from websites or books, was a limitation of this study because only 11 individuals returned the questionnaire on time. Therefore, the present study only tested the reliability and agreement of the evaluators of the Modified Fresno Test version adapted to Brazilian-Portuguese. Therefore, the authors recommend the development of new studies in order to test other properties of the instrument, such as construct validity and responsiveness of the Brazilian-Portuguese version. Further studies are needed to continually assess the new learning methods and the use of scientific evidence in therapeutic decisions during students professional training, because these methods are fundamental prerequisite for the adoption of EBP principles.

● Conclusion

The Modified Fresno Test for physical therapists revealed satisfactory results in the adaptation process. The Brazilian-Portuguese version of the instrument showed good internal consistency, excellent reliability and low variability in the intra-agreement test and for inter-evaluators. There are other properties that should be tested. However, the steps taken so far may contribute to the development of studies aimed at assessing comprehensively the knowledge and skills of physical therapy professionals and students in the use of EBP in their clinical practice.

● References

- Herbert R, Jamtvedt G, Mead J, Hagen KB. Practical evidence-based physiotherapy. 2nd ed. London: Churchill Livingstone; 2011.
- Ilic D. Assessing competency in evidence based practice: strengths and limitations of current tools in practice. BMC Med Educ. 2009;9(1):53. <http://dx.doi.org/10.1186/1472-6920-9-53>. PMID:19656412.
- Herbert R, Jamtvedt G, Mead J, Hagen KB. Practical evidence-based physiotherapy. Edinburgh: Elsevier Butterworth Heinemann; 2005. 234 p.
- Shiwa SR, Costa LOP, Moser ADL, Aguiar IC, Oliveira LVF. PEDro: a base de dados de evidências em fisioterapia. Fisioter Mov. 2011;24(3):523-33. <http://dx.doi.org/10.1590/S0103-51502011000300017>.
- Jette DU, Bacon K, Batty C, Carlson M, Ferland A, Hemingway RD, et al. Evidence-based practice: beliefs, attitudes, knowledge, and behaviors of physical therapists. Phys Ther. 2003;83(9):786-805. PMID:12940766.
- Connolly BH, Lupinacci NS, Bush AJ. Changes in attitudes and perceptions about research in physical therapy among professional physical therapist students and new graduates. Phys Ther. 2001;81(5):1127-34. PMID:11319938.
- Tilson JK. Validation of the modified Fresno Test: assessing physical therapists' evidence based practice knowledge and skills. BMC Med Educ. 2010;10(38):1-9.
- Glasziou P, Del Mar C, Salisbury J. Prática clínica baseada em evidências: livro de exercícios. 2nd ed. Porto Alegre: Art Med; 2010.
- Claridge JA, Fabian TC. History and development of evidence-based medicine. World J Surg. 2005;29(5):547-53. <http://dx.doi.org/10.1007/s00268-005-7910-1>. PMID:15827845.
- Silva TM, Costa LCM, Costa LOP. Evidence-based practice: a survey regarding behavior, knowledge, skills, resources, opinions and perceived barriers of Brazilian physical therapists from São Paulo state. Braz J Phys Ther. 2015;19(4):294-303. <http://dx.doi.org/10.1590/bjpt-rbf.2014.0102>. PMID:26443977.
- Marques AP, Peccin MS. Pesquisa em fisioterapia: a prática baseada em evidências e modelos de estudos. Fisioter Pesqui. 2005;11(1):43-8.
- Shaneyfelt T, Baum KD, Bell D, Feldstein D, Houston TK, Kaatz S, et al. Instruments for evaluating education in evidence-based practice: a systematic review. JAMA. 2006;296(9):1116-27. <http://dx.doi.org/10.1001/jama.296.9.1116>. PMID:16954491.
- Coury H, Vilella I. Profile of the Brazilian physical therapy researcher. Rev Bras Fisioter. 2009;13(4):356-63. <http://dx.doi.org/10.1590/S1413-355520090005000048>.
- Ramos KD, Schafer S, Tracz SM. Validation of the Fresno test of competence in evidence based medicine. BMJ. 2003;326(7384):319-21. <http://dx.doi.org/10.1136/bmj.326.7384.319>. PMID:12574047.
- Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312(7023):71-2. <http://dx.doi.org/10.1136/bmj.312.7023.71>. PMID:8555924.
- Domingues RCL, Amaral E, Bicudo-Zeferino AM. Global rating: a method for assessing clinical competence. Rev Bras Educ Med. 2009;33(1):148-51. <http://dx.doi.org/10.1590/S0100-55022009000100019>.
- Lima VV. Competência: distintas abordagens e implicações na formação de profissionais de saúde. Rev Interface. 2005;9:369-79.

18. Morris J, Maynard V. The feasibility of introducing an evidence based practice cycle into a clinical area: an evaluation of process and outcome. *Nurse Educ Pract.* 2009;9(3):190-8. <http://dx.doi.org/10.1016/j.nepr.2008.06.002>. PMID:18707920.
19. Beaton D, 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.
20. Brasil. Conselho Nacional de Educação – CNE. Resolução CNE/CES nº 4, de 19 de fevereiro de 2002. Institui Diretrizes Curriculares Nacionais do Curso de Graduação em Fisioterapia. Diário Oficial da União; Brasília; 4 março 2002. Seção 1.
21. Kogan JR, Holmboe ES, Hauer KE. Tools for direct observation and assessment of clinical skills of medical trainees: a systematic review. *JAMA.* 2009;302(12):1316-26. <http://dx.doi.org/10.1001/jama.2009.1365>. PMID:19773567.
22. McCluskey A, Bishop B. The adapted Fresno test of competence in evidence-based practice. *J Contin Educ Health Prof.* 2009;29(2):119-26. <http://dx.doi.org/10.1002/chp.20021>. PMID:19530195.
23. 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.
24. 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.
25. Maher CG, Latimer J, Costa LOP. The relevance of cross-cultural adaptation and clinimetrics for physical therapy instruments. *Rev Bras Fisioter.* 2007;11(4):245-52. <http://dx.doi.org/10.1590/S1413-35552007000400002>.
26. Puga VOO, Lopes AD, Costa LOP. Assessment of cross-cultural adaptations and measurement properties of self-report outcome measures relevant to shoulder disability in portuguese: a systematic review. *Rev Bras Fisioter.* 2012;16(2):85-93. <http://dx.doi.org/10.1590/S1413-35552012005000012>. PMID:22588123.
27. Fleiss JL. The design and analysis of clinical experiments. New York: John Wiley & Sons; 1986.
28. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-74. <http://dx.doi.org/10.2307/2529310>. PMID:843571.
29. Ostelo RWJG, de Vet HCW, Knol DL, van den Brandt PA. 24-item Roland-Morris Disability Questionnaire was preferred out of six functional status questionnaires for post-lumbar disc surgery. *J Clin Epidemiol.* 2004;57(3):268-76. <http://dx.doi.org/10.1016/j.jclinepi.2003.09.005>. PMID:15066687.
30. Moriguchi CS, Alem MER, Veldhoven M, Coury HJCG. Cultural adaptation and psychometric properties of Brazilian need for recovery scale. *Rev Saude Publica.* 2010;44(1):131-9. <http://dx.doi.org/10.1590/S0034-89102010000100014>. PMID:20140337.
31. Argimon-Pallas JM, Flores-Mateo G, Jimenez-Villa J, Pujol-Ribera E. Psychometric properties of a test in evidence based practice: the spanish version of the Fresno test. *BMC Med Educ.* 2010;10(1):45. <http://dx.doi.org/10.1186/1472-6920-10-45>. PMID:20553577.
32. Silva TM, Costa LCM, Garcia AN, Costa LOP. What do physical therapists think about evidence-based practice? A systematic review. *Man Ther.* 2015;20(3):388-401. <http://dx.doi.org/10.1016/j.math.2014.10.009>. PMID:25458142.
33. Queiroz PS, Santos MJD. Facilidades e habilidades do fisioterapeuta na procura, interpretação e aplicação do conhecimento científico na prática clínica: um estudo piloto. *Fisioter Mov.* 2013;26(1):13-23. <http://dx.doi.org/10.1590/S0103-51502013000100002>.
34. Akinbo S, Odebiyi D, Okunola T, Aderoba O. Evidence-based practice: knowledge, attitudes and beliefs of physiotherapists in Nigeria. *Int J Med Inform.* 2008;4(2).
35. Schreiber J, Stern P. A review of the literature on evidence-based practice in physical therapy. *Internet J Allied Health Sci Pract.* 2005;3(4).

Correspondence

Rosimeire Simprini Padula

Universidade Cidade de São Paulo (UNICID)

Rua Cesário Galeno, 448/475, Tatuapé

CEP 03071-000, São Paulo, SP, Brazil

e-mail: rosimeire.padula@unicid.edu.br

Appendix 1. Versão adaptada do Teste Modificado de Fresno para fisioterapeutas.

● **Instruções:**

Prática Baseada em Evidências (PBE) envolve competências e habilidades relacionadas com a identificação e avaliação de evidências para informar a prática. Esta ferramenta, o Teste Modificado de Fresno é projetado para avaliar suas habilidades para a PBE.

Há oito perguntas de resposta curta, 2 perguntas que exigem uma série de cálculos matemáticos, e 3 questões de preenchimento de lacunas. Uma calculadora e papel para anotação lhe foram fornecidos. Recursos adicionais (sites da internet, livros, etc.) não são permitidos.

Por favor, preencha todo o teste de uma só vez e permita-se até 60 minutos para completar o teste.

Responda às perguntas 1 a 4 e 8 com base nos seguintes cenários clínicos:

Cenário 1: Você acabou de avaliar Maria, uma secretária, que recentemente sofreu uma lesão lombar relacionada ao trabalho, movimentando caixas de arquivos de 4,54 Kg há 3 dias. Suas radiografias são negativas e seu único sintoma é a dor na coluna lombar com intensidade 2/10, durante movimentos de flexão e ficar sentada por tempo prolongado. Ela ficou fora do trabalho por dois dias e está ansiosa para voltar, mas está também preocupada com uma nova lesão. Você está pensando em um programa de exercícios de estabilização, mas pergunta a si mesmo se a terapia manual deve ser incluída no programa de fisioterapia do paciente.

Cenário 2: Marvin é um garoto de 10 anos com hemiparesia secundária, devido a um acidente vascular cerebral associado a uma má formação artério-venosa. Ele apresenta-se à terapia ambulatorial e seus pais expressam preocupação particular sobre o braço de Marvin e fraqueza nas pernas. Você está pensando em implementar um programa intensivo de tarefa específica de fortalecimento, mas um colega adverte que tal programa é capaz de aumentar o tônus flexor e espasticidade moderada do paciente e sugere baixa intensidade de alongamento e um programa de posicionamento passivo.

Pergunta # 1: Escolha um dos cenários clínicos acima. Escreva uma pergunta clínica, focada neste cenário, que irá ajudá-lo a organizar uma busca na literatura clínica.

Resposta:

Pergunta # 2: Onde você poderia encontrar respostas para essa e outras perguntas clínicas semelhantes? Nomeie quantas fontes de informação e/ou pesquisa puder e não apenas as que você acha que são “boas” fontes. Descrever as vantagens e desvantagens de cada tipo de fonte de informação que você colocou na lista.

Resposta:

Pergunta # 3: Que tipo de estudo (projeto de estudo) melhor responderia a sua pergunta clínica descrita na pergunta 01 e por quê?

Resposta:

Pergunta # 4: Se você estivesse em uma busca no Medline, CINAHL ou qualquer outro banco de dados para a pesquisa original, para responder a sua pergunta clínica relacionada ao cenário selecionado para a pergunta 1, descreva a estratégia de busca que você pode usar. Seja o mais específico possível sobre os termos de pesquisa e campos de pesquisa que você usaria. Explique sua justificativa para esta abordagem. Descreva como você pode limitar sua pesquisa, se necessário, e explique o seu raciocínio.

Resposta:

Pergunta # 5: Quando você encontrar um artigo sobre a sua pergunta clínica ou quaisquer outras, que características do estudo você vai considerar para determinar se ele é relevante? Inclua exemplos. As questões 6 e 7 irão perguntar a você como determinar se o estudo é válido, e quão importante são os resultados. Para esta pergunta, por favor, concentre-se em como determinar se ele é realmente relevante para a sua prática.

Resposta:

Pergunta # 6: Quando você encontrar um artigo relacionado à sua pergunta clínica ou quaisquer outras, que características do estudo você vai considerar para determinar se as suas conclusões são válidas? (Você já foi abordado sobre relevância e a questão 7 vai perguntar como determinar a importância dos resultados. Para esta pergunta, por favor, concentre-se na validade do estudo).

Resposta:

Pergunta # 7: Quando você encontrar um artigo, que se relacione com a sua pergunta clínica ou quaisquer outras, que características dos resultados você vai considerar para determinar sua magnitude e significância (clínica e estatística)?

Resposta:

Pergunta # 8: Para o cenário clínico que você escolheu, liste duas questões que você gostaria de fazer ao paciente / família para obter uma melhor compreensão de suas preferências pessoais e / ou circunstâncias em relação a sua questão clínica.

Resposta:

Pergunta # 09: Um estudo sobre a precisão do diagnóstico do teste ergométrico (TE) para o diagnóstico de doença arterial coronariana (DAC) incluiu 96 mulheres com suspeita de DAC, 29 das quais foram posteriormente determinadas tendo DAC (estenose > 50% em um ou mais vasos coronarianos). Das mulheres com DAC, 9 teve um TE anormal. Dos 67 pacientes determinados não tendo DAC, 32 tinham um TE anormal.

a: Com base nesses resultados, a sensibilidade do TE para DAC é?

b: Com base nesses resultados, o valor preditivo positivo de TE para DAC é?

c: Com base nesses resultados, a razão de verossimilhança positiva para um TE anormal para DAC é?

*** Caso não tenha calculadora, para facilitar os valores de frequência utilize o texto alternativo abaixo:**

Um estudo sobre a precisão do diagnóstico do teste ergométrico (TE) para o diagnóstico de doença arterial coronariana (DAC) incluiu 120 mulheres com suspeita de DAC, 30 dos quais foram posteriormente determinadas tendo DAC (estenose > 50% em um ou mais vasos coronarianos). Daquelas com DAC, 10 tinham um TE anormal. Dos 90 pacientes determinados não tendo DAC, 30 tiveram um TE anormal.

Resposta:

Arredondamento é aceitável (por exemplo 21,9-22 é aceitável)

Pergunta # 10: Um estudo aleatorizado recente de mulheres grávidas com incontinência descobriu que, após o treinamento do assoalho pélvico 20% tinham incontinência, comparado com 32% no grupo controle três meses após o parto. O Nível alfa para o estudo foi estabelecido ao nível de significância de 0,05.

a: A redução do risco absoluto para eventos recorrentes é?

b: A redução do risco relativo de eventos recorrentes é?

c: O número necessário para tratar (NNT) para prevenir um evento recorrente é?

d: O valor-p indicando diferença estatisticamente significativa entre os grupos seria?

*** Caso não tenha calculadora, para facilitar os valores de frequência utilize o texto alternativo abaixo:**

Um recente estudo randomizado de gestantes com incontinência descobriu que, após o treinamento do assoalho pélvico 20% tinham incontinência em comparação com 30% no grupo controle em três meses após o parto. O Nível alfa para o estudo foi estabelecida ao nível de significância de 0,05.

Resposta:

Pergunta # 11: O mesmo estudo descrito na questão 10 revelou um risco relativo de incontinência de 0,61 para as mulheres que receberam treinamento do assoalho pélvico. Isto sugere que o tratamento com treinamento do assoalho pélvico reduz o risco para a incontinência. Gostaríamos de saber se esta diferença é estatisticamente significativa, por isso, olhe para o intervalo de confiança. Dê um exemplo de um intervalo de confiança que suporte a conclusão de que a taxa de incontinência foi realmente (estatisticamente) diferente para estes dois grupos de tratamento.

*** Caso não tenha calculadora, para facilitar os valores de frequência utilize o texto alternativo abaixo:**

O mesmo estudo descrito na questão 11 revelou um risco relativo de incontinência de 0,66 para as mulheres que receberam treinamento do assoalho pélvico. Isto sugere que o tratamento com treinamento do assoalho pélvico reduz o risco para a incontinência. Gostaríamos de saber se esta diferença é estatisticamente significativa, por isso, olhe para o intervalo de confiança. Dê um exemplo de um intervalo de confiança que suporte a conclusão de que a taxa de incontinência foi realmente (estatisticamente) diferente para estes dois grupos de tratamento.

Resposta:

Pergunta # 12: Qual desenho do estudo é o melhor para um estudo sobre o diagnóstico?

Resposta:

Pergunta # 13: Qual desenho do estudo é o melhor para um estudo sobre o prognóstico?

Resposta:

Tempo gasto para completar o teste: _____

Appendix 2. Escore de pontuação da versão adaptada do Teste Modificado de Fresno para fisioterapeutas.

Pergunta # 1:Escolha um dos cenários clínicos acima. Escreva uma pergunta clínica, focada neste cenário, que irá ajudá-lo a organizar uma busca na literatura clínica.

	Excelente	Forte	Limitado	Mínimo	Não evidente
a: População	6: Vários descritores relevantes; por exemplo, “lesão relacionada com o trabalho”, “mulher” ou “grave” ou “dor lombar”, por exemplo, “menino com hemiparesia” grupo etário específico, sexo, diagnóstico, apresentação motora.	4: Um descritor apropriado como os exemplos acima, “Mulheres” ou “trabalhador” ou “dor lombar”, por exemplo, “hemiparesia” “menino” “10 anos” “pós-AVC”	2: Um único descritor geral capaz de contribuir para pesquisa; por exemplo, “Paciente”		0: Nenhum dos presentes acima.
b: Intervenção	6: Inclui intervenção de interesse específico; (intervenção poderia ser uma técnica de diagnóstico); <ul style="list-style-type: none">• terapia manual;• componentes específicos individuais de terapia manual;• combinação de exercício e terapia manual;• tarefas específicas de fortalecimento.		2: Menciona intervenção, mas é capaz de contribuir para pesquisa; por exemplo, “Métodos” “opções” “tratamentos”.		0: Nenhum dos presentes acima.
c: Comparação	6: Identifica alternativa de interesse específico; exemplo “Nenhuma terapia manual”, “alongamento de baixa intensidade”.		2: Menciona comparação, mas é capaz de contribuir para pesquisa; exemplo “Métodos alternativos”		0: Nenhum dos presentes acima.
d: Resultado	6: Resultado que é objetivo e significativo para o paciente ou o caso do paciente (se a pergunta é de diagnóstico, deve estar relacionado com o que o diagnóstico está tentando detectar); exemplo, retorno ao trabalho, redução da dor, prevenção de lesões; por exemplo, controle motor seletivo ou uso funcional de extremidades paralisadas, velocidade no andar.	4: Sem resultado específico: <ul style="list-style-type: none">• Recuperação• Espasticidade• Tônus• Força	2: Referência ao resultado, mas de modo geral a ser capaz de contribuir para pesquisar: <ul style="list-style-type: none">• Efeitos• Alteração do resultado• Eficaz• Melhoria• Sucesso		0: Nenhum dos presentes acima.

Pergunta # 2: Onde você poderia encontrar respostas para essa e outras perguntas clínicas semelhantes? Nomeie quantas fontes de informação puder - e não apenas as que você acha que são “boas” fontes. Descrever as vantagens e desvantagens de cada tipo de fonte de informação que você colocou na lista.

	Excelente	Forte	Limitado	Min.	Não evidente
a: Variedade de fontes	<p>6: Pelo menos quatro tipos de fontes listados. Os tipos incluem:</p> <ul style="list-style-type: none"> • Bases de dados eletrônicas da literatura original [Medline (PubMed / Ovid), CINAHL]; • Bases de dados de disciplina específica (baseados em evidências, PEDro); • Periódicos (JAMA, NEJM, incluindo o acesso através da biblioteca); • Livro de texto (Merck, Harrison's, monografias); • Revisões Sistemáticas (Cochrane); • EBM publicações ou bases de dados de informação pré-avaliados (Best Evidence, InfoRetriever, DynaMed, EBM, ACPJC, EBP, Clinical Evidence); • Sites médicos (MDConsult, PraxisMD, SumSearch); • Pesquisa da Internet em Geral (google, yahoo); • Diretrizes Clínicas (Guideline Clearinghouse); • Organizações Profissionais (AAFP, La Leche League, site do NIH); • Pessoas (colega, consultor, atendendo bibliotecário,). 	4: Três tipos de fontes listados.	2: Dois tipos de fontes listados.		0: Nenhuma variedade. Apenas uma fonte listada, ou todas as fontes do mesmo tipo.
b: Conveniência	<p>6: A discussão inclui pelo menos duas questões específicas relacionadas com a conveniência, ou menciona a mesma questão ao discutir duas fontes diferentes. Problemas podem incluir:</p> <ul style="list-style-type: none"> • Custo (por exemplo, “gratuito”, “apenas por assinatura”); • Velocidade (por exemplo, “rápida”, “leva tempo”); • Facilidade de busca (por exemplo, “deve saber como estreitar busca”, “fácil de navegar”); • Facilidade de uso (por exemplo, “concisa” e “NNTs já calculado”); • Disponibilidade (por exemplo, “disponíveis online”). 	4: Inclui uma questão específica / explicações relativas à conveniência.	2: Menciona conveniência envolvida na utilização de uma ou mais fontes, mas sem explicação, por exemplo, “Conveniente” ou “fácil” ou “difícil”.		0: Nenhuma menção de conveniência.

Pergunta # 2: Continuação...

	Excelente	Forte	Limitado	Min.	Não evidente
c: Relevância Clínica	<p>6: A discussão inclui pelo menos duas questões específicas relacionadas com a relevância, ou menciona a mesma questão ao discutir duas fontes diferentes. Problemas podem incluir:</p> <ul style="list-style-type: none">• Resultados clinicamente relevantes;• Escrito para aplicação clínica (por exemplo, “pertinente” “informações sobre os efeitos adversos” ou “há formulários de informação do paciente”);• Foco em especialidade adequada (por exemplo, “dirigido à fisioterapeutas”);• Informação aplicável ao paciente em questão (por exemplo, “pode examinar mais detalhes deste paciente em particular” ou “a maioria dos estudos são da Europa”);• Inclui intervenções específicas em questão;• Especificidade (visão geral x alvo) (por exemplo, “pode obter informações básicas” ou “mais especializada”);• Abrangência da fonte (probabilidade de encontrar uma resposta nesta fonte) (por exemplo, “ela pode encontrar qualquer coisa” ou “contém referências úteis” ou “não susceptíveis de ter resposta para esta pergunta”).	<p>4: Inclui uma questão específica / explicação relacionada com a relevância.</p>	<p>2: Menciona relevância do uso de uma ou mais fontes mas sem explicação, por exemplo, “Relevante”.</p>		<p>0: Nenhuma menção de relevância.</p>
d: Validade	<p>6: A discussão inclui pelo menos duas questões específicas relacionadas com a validade, ou menciona a mesma questão ao discutir duas fontes diferentes. Problemas podem incluir:</p> <ul style="list-style-type: none">• Certeza de validade (por exemplo, “qualidade é incerto” ou “não foi exibido” ou “precisa ser avaliado criticamente”);• Abordagem baseada em evidências (por exemplo, “baseada em evidências” ou “Evidencia de grau 1” ou “não há referências fornecidas”);• Viés de especialista (por exemplo, “geralmente apenas a opinião de alguém”);• Abordagem sistemática;• A revisão por pares;• Capacidade para verificar;• Padrão de cuidados (por exemplo, “aceito na comunidade médica”);• Informação fornecida suficiente para uma validade crítica (por exemplo, “apenas abstrato” ou “texto completo não disponível”);• Atualizado/desatualizado (por exemplo, “a pesquisa mais recente”);• Confiabilidade - no contexto do grau de confiabilidade que pode ser colocado na fonte.	<p>4: Inclui uma questão específica / explicação relacionada à validade.</p>	<p>2: Menciona validade do uso de uma ou mais fontes, mas sem explicação, por exemplo, “Bom”, “lixo”.</p>		<p>0: Nenhuma menção de validade.</p>

Pergunta # 3: Que tipo de estudo (projeto de estudo) melhor responderia a sua pergunta clínica descrita no Q1 e por quê?

	Excelente	Forte	Limitado	Mínimo	Não evidente
a: Projeto de estudo	<p>12: Nomeia uma das melhores fontes:</p> <ul style="list-style-type: none"> • Experiência Controlada Aleatorizada; • Experiência Aleatorizada; • Revisão Sistemática; • Meta-Análise; • Experiência Aleatorizada, duplamente cego. 	<p>9: Descreve, mas não chama pelo nome, uma das melhores fontes como acima:</p> <ul style="list-style-type: none"> • Comparando dois grupos, um obtém-se o tratamento, o outro placebo; • Estudo duplamente cego. 	<p>6: Descreve ou nomeia um projeto de estudo menos desejável:</p> <ul style="list-style-type: none"> • Estudo de coorte; • Ensaio clínico prospectivo; • Meta-análise de tais estudos; • Longitudinal ou prospectivo. 	<p>3: Descreve ou nomeia um projeto de estudo pobre para responder a uma pergunta de tratamento:</p> <ul style="list-style-type: none"> • Controle de casos, estudo transversal, estudo de casos, "retrospectivo"; • Ou descreve um estudo com detalhe insuficiente para identificar um modelo: por exemplo, um estudo comparativo. 	<p>0: Nenhum dos presentes acima.</p>
b: Justificativa	<p>12: Inclui justificativa bem fundamentada, que reflete a compreensão da importância da aleatorização e / ou cegueira. Conecta explicitamente aleatorização para redução de confusão e / ou cegueira de observador ou viés de aferição. Por exemplo: "Um estudo controlado aleatorizado tentará evitar qualquer viés que possa influenciar o resultado do estudo através de aleatorização" OU "mais adequada para questões de terapia, pois reduz preconceitos e controles para fatores de confusão."</p>	<p>9: A justificativa está presente, e toca em questões relacionadas com a aleatorização e / ou cegueira, mas menos claramente articulada, por exemplo, "Os grupos devem ser semelhantes" ou "tentar eliminar fatores de confusão" ou "evitar viés de seleção" ou "ser objetivo" ou "eliminar o preconceito".</p>	<p>6: A justificativa está presente, e levanta questões legítimas não relacionados à aleatorização ou cegueira, tal como a relação custo-benefício, recolhimento de produto.</p> <p>Talvez mencionar aleatorização ou cegueira, mas sem explicação. (Por exemplo: "o melhor em um cenário aleatório e cego"),</p> <p>Por exemplo: "Opiniões de gráfico fornecem muitos dados sem muito custo"</p>	<p>3: Tentativa de justificativa, mas os argumentos não são específicos e não demonstram compreensão da relação entre o projeto e as várias ameaças à validade</p> <p>Pode mencionar aleatorização ou cegueira, mas sem explicação. (Por exemplo: "o melhor em um cenário aleatório e cego"), por exemplo, "Para garantir a qualidade" ou "para reduzir potenciais conflitos" ou "para comparar".</p>	<p>0: Nenhum dos presentes acima</p>

Pergunta # 4: Se você estivesse em uma busca no Medline, CINAHL ou qualquer outro banco de dados para a pesquisa original, para responder a sua pergunta clínica relacionada ao cenário selecionado para a pergunta 1, descreva a estratégia de busca que você pode usar. Seja o mais específico possível sobre os termos de pesquisa e campos de pesquisa que você usaria. Explique sua justificativa para esta abordagem. Descreva como você pode limitar sua pesquisa, se necessário, e explique o seu raciocínio.

	Excelente	Forte	Limitado	M	Não evidente
a: Termos de busca	8: 3 ou mais termos que refletem paciente, intervenção, comparação, e resultado (PICR) sendo considerados.	6: 2 termos do PICR.	3: 1 termo do PICR.		0: Não presente
b: Instruções/ Estratégia	8: Descrição da estratégia de pesquisa reflete o entendimento de que os artigos no banco de dados são indexados por mais de um campo. Discute um ou mais campo/índice/rótulo pelo nome (MESH, Palavra do título, nome da publicação, idioma, palavras-chave, autor, título da revista, utilização de operadores booleanos, etc.) E fornece justificativa plausível para a estratégia de busca usando um ou mais destes índices, por exemplo, “Palavra-chave é menos específico do que a estrutura”.	6: Nomeia um ou mais campo ou índice de categoria, mas não fornece defesa plausível de estratégia de pesquisa com base neste conhecimento, por exemplo, “Eu faria uma pesquisa por palavra... seguido por...” “Eu usaria termos... desta forma”.	3: Fraca descrição da estratégia, nenhum nome dado aos rótulos, ou estratégia abertamente equivocada. Por exemplo: “Gostaria de usar termos...” [nenhuma descrição da estratégia].		0: Nenhuma compreensão evidente para os artigos “marcados” por diversos campos ou índices.
c: Delimitadores	8: Descreve mais de uma abordagem para a busca de limitação (por exemplo, “limite humano” ou “Inglês” ou “adulto”), nomeia um tipo de publicação específica, ou descrição de consultas clínicas no PubMed, ou o uso de operadores booleanos ou pesquisa combinações ou inclui termos relacionados a um projeto de estudo ideal (por exemplo, ao acaso) ou sugere uso de subposições. * NOTA: Se o assunto inclui o nome do índice ao descrever um delimitador (por exemplo, “verificar a linguagem como Inglês”), então nós damos crédito para um rótulo, assim como a um método de delimitação.	6: Descreve apenas um método comum de limitar a pesquisa. Por exemplo, descreve as formas de estreitar a pesquisa usando palavras-chave, mas nenhuma outra das estratégias listadas.	3: Fornece uma explicação fraca ou descrição do uso de limitadores / estreitamento de busca.		0: Nenhuma técnicas válidas para limitar uma pesquisa listada.

Pergunta # 5: Quando você encontrar um relatório de pesquisa original sobre esta questão, ou quaisquer outras, que características do estudo você vai considerar para determinar se ele é relevante? Inclua exemplos. Questões 6 e 7 irão perguntar a você como determinar se o estudo é válido, e quão importante são os resultados. Para esta pergunta, por favor, concentre-se em como determinar se ele é realmente relevante para a sua prática.

	Excelente	Forte	Limitado	Não evidente
a: A questão	<p>12: Discussão bem-fundamentada e refletida da relevância das variáveis dependentes e independentes utilizadas no estudo, incluindo exemplos / razões específicas. Pode discutir (bem-fundamentada e refletida):</p> <ul style="list-style-type: none"> • A viabilidade do teste ou intervenção; • “O teste pode funcionar, mas se a minha prática não pode se dar ao luxo de comprar a máquina, não importa”; • O paciente ou doença orientada natureza do resultado; • “Elas medem a capacidade das crianças para usar a função melhorada em atividades lúdicas?” • A congruência entre a definição operacional e o a questão de pesquisa, por exemplo, “Se o seu método de aferição do resultado é uma representação realista do resultado que nos interessa”. 	<p>9: Menos discussão reflexiva da relevância das variáveis dependentes e independentes utilizadas no estudo. Pode incluir conceitos ou exemplos específicos sem razão clara. Pode referir-se aos mesmos itens listados na “excelente”, mas sem demonstrar profundidade de compreensão.</p>	<p>5: Resposta implica na consideração de quão bem o estudo aborda a questão em mãos, mas oferece pouca discussão sobre por que isso pode ser importante,</p> <ul style="list-style-type: none"> • Por exemplo: “Quais são as variáveis?”; • “Ela respondeu à minha questão?”; • “O desfecho”; • “O propósito do estudo”; • “Impactará minha prática?”; • “Comprimento de acompanhamento”. 	<p>0: Nenhuma discussão sobre a questão de pesquisa e variáveis utilizadas para respondê-la.</p>
b: Descrição dos assuntos.	<p>12: Inclui ambos:</p> <p>Uma expressão clara da importância da relação entre os sujeitos do estudo e da população alvo. E pelo menos um exemplo de uma doença ou característica demográfica pertinente.</p> <ul style="list-style-type: none"> • Por exemplo: “Os pacientes eram semelhantes ao meu em termos de idade e raça?” ou • “era um amostra hospitalar ou clínica como os meus pacientes?” ou • “Os pacientes tinham mesmo nível de gravidade da doença como meu paciente?” ou • “fez seleção ou inclusão inadequado de critérios dos resultados em uma população que difere da minha em raça, idade, etc.”. 	<p>9: Inclui um, mas não ambos:</p> <p>Uma expressão clara da importância da relação entre os sujeitos do estudo e da população alvo, ou pelo menos um exemplo de uma doença relevante ou característica demográfica</p> <p>Por exemplo, “O paciente é como o meu?” ou “o nível de educação da população”.</p>	<p>5: Resposta indica a consideração dos sujeitos do estudo, mas não oferece nenhuma discussão sobre a conexão entre sujeitos do estudo e a população alvo ou características específicas da amostra</p> <ul style="list-style-type: none"> • Por exemplo: “É uma amostra adequada?” ou • “qual foi a resposta ou taxa de participação?” ou • “quais foram os critérios de exclusão?” ou • “viés de seleção” ou • “ajuste” ou • “onde o estudo foi conduzido”. 	<p>0: Nenhuma discussão sobre as características dos sujeitos de pesquisa.</p>

NOTA: AS RESPOSTAS ÀS PERGUNTAS 5, 6 E 7, PODEM SER APLICADAS A QUALQUER PARTE DA RUBRICA DE CLASSIFICAÇÃO PARA ESSES ITENS

Pergunta # 6: Quando você encontrar um relatório de pesquisa original relacionado à sua pergunta clínica ou quaisquer outras, que características do estudo você vai considerar para determinar se as suas conclusões são válidas? (Você já foi abordado sobre relevância e a questão 7 vai perguntar como determinar a importância dos resultados. Para esta pergunta, por favor, concentre-se na validade do estudo).

	Excelente	Forte	Limitado	Mínimo	Não evidente
a: Validade Interna	24: Lista ou descreve pelo menos 5 questões importantes para a validade interna, tais como: <ul style="list-style-type: none">• Adequação do projeto de estudo;• Adequação da cegueira;• Cancelamento de Alocação;• Aleatorização de trabalho de grupo;• Medição inválida ou tendenciosas (“seguido próprio protocolo?”);• Importância do grupo de comparação ou controle;• Intenção de tratar a análise;• Consideração de co-variáveis apropriadas (“foram outros fatores relevantes considerados?”);• Conclusões consistentes com a evidência (“Os resultados fazem sentido?”);• Importância do acompanhamento de todos os participantes do estudo;• Análise estatística apropriada;• Tamanho da amostra /Poder;• Patrocínio;• Quando o estudo foi realizado;• Confirmação com outros estudos;• Medidas de resultado válidas.	18: Identifica 3-4 questões específicas como acima.	10: Identifica duas questões específicas como acima.	5: Menciona validade interna ou um conceito específico da lista de exemplos acima.	0: Nenhum dos presentes acima.

NOTA: AS RESPOSTAS ÀS PERGUNTAS 5, 6 E 7, PODEM SER APLICADAS A QUALQUER PARTE DA RUBRICA DE CLASSIFICAÇÃO PARA ESSES ITENS

Pergunta # 7: Quando você encontrar um relatório de pesquisa original, que se relacione com a sua questão clínica ou quaisquer outras, que características dos resultados você vai considerar para determinar sua magnitude e significância (clínica e estatística)?

	Excelente	Forte	Limitado	Min	Não relevante
a: Magnitude	<p>12: Resposta deve discutir claramente ambos:</p> <ul style="list-style-type: none"> Significado clínico (“qual é o significado clínico?” Ou “quão ampla foi a diferença encontrada”, a alteração excede MCID) <p>E</p> <ul style="list-style-type: none"> Exemplo (s) de tamanho de efeito de medições (por exemplo, especificidade, sensibilidade, razão de probabilidade de um teste, número necessário para tratar, o risco relativo, redução do risco absoluto, diferença média de resultados contínuos, positivo ou negativo, o valor preditivo). 	<p>9: Resposta discute um, mas não ambos:</p> <ul style="list-style-type: none"> Significado clínico (“qual é o significado clínico?” Ou “quão ampla foi a diferença encontrada?”) <p>OU</p> <ul style="list-style-type: none"> Exemplo (s) tamanho de efeito de medições (por exemplo, especificidade, sensibilidade, razão de probabilidade de um teste, número necessário para tratar, o risco relativo, redução do risco absoluto, diferença média de resultados contínuos, positivo ou negativo, o valor preditivo). 	<p>5: Resposta só sugere uma consideração de significado clínico ou o tamanho do efeito.</p> <ul style="list-style-type: none"> Por exemplo: “Será que isso importa?” “será que vai impactar minha prática” ou Por exemplo: (menciona “mínima diferença clinicamente importante”, mas não explica como esse valor seria usado para determinar a significância clínica). 		0: Nenhum dos presentes acima.
b: Significância Estatística	<p>12: Discussão refletida e bem-fundamentada dos índices de significância estatística, incluindo pelo menos dois exemplos específicos de importantes conceitos relacionados, tais como:</p> <ul style="list-style-type: none"> P- valores; Intervalos de confiança; Poder; Precisão das estimativas; Erro Tipo 1 ou tipo 2. 	<p>9: Lista mais do que um conceito (como acima) com a discussão insuficiente ou ausente (por exemplo, “p-valor e intervalos de confiança”)</p> <p>OU</p> <p>Lista e discute apenas um conceito (por exemplo, “p-valor menor que <.05”).</p>	<p>5: Menciona a necessidade de avaliar a significância estatística ou nomeia somente um dos conceitos acima sem mais discussão (por exemplo, “p-valores”, “estatisticamente significativos”).</p>		0: Nenhum dos presentes acima

NOTA: AS RESPOSTAS ÀS PERGUNTAS 5, 6 E 7, PODEM SER APLICADAS A QUALQUER PARTE DA RUBRICA DE CLASSIFICAÇÃO PARA ESSES ITENS

Pergunta # 8: Para o cenário clínico que você escolheu, liste duas questões que você gostaria de fazer ao paciente / família para obter uma melhor compreensão de suas preferências pessoais e / ou circunstâncias em relação a sua questão clínica.

	Excelente	Forte	Limitado	Mínimo	Não evidente
a: Questão1	8: A pergunta é apropriada para obter informações importantes sobre as preferências do paciente, os valores, as circunstâncias, as expectativas, e / ou motivações que irão impactar diretamente os cuidados clínicos.		4: A pergunta é geral, mas trata de questões relevantes para a compreensão da perspectiva do paciente.	2: A pergunta é geral e não trata de questões específicas para as perspectivas do paciente. Por exemplo: Pergunta padrão de avaliação subjetiva não específica para perspectivas do paciente. Por exemplo: Perguntas sim / não ou factual que não são susceptíveis de trazer à tona detalhes sobre perspectiva do paciente.	0: Nenhuma pergunta ou não uma questão atual • “histórico médico antigo” ou • “preferências”
b: Questão2	8: Mesmo que acima, mas traz à tona informações diferentes do que a primeira pergunta (caso contrário 0).		4: Mesmo que acima, mas traz à tona informações diferentes do que a primeira pergunta (caso contrário 0).	2: O mesmo que acima, mas traz à tona informações diferentes do que a primeira pergunta (caso contrário 0).	0: Nenhuma pergunta ou não uma questão real.

Pergunta # 09: Um estudo sobre a precisão do diagnóstico do teste ergométrico (TE) para o diagnóstico de doença arterial coronariana (DAC) incluiu 96 mulheres com suspeita de DAC, 29 dos quais foram posteriormente determinadas tendo DAC (estenose> 50% em um ou mais vasos coronarianos). Daqueles com DAC, 9 teve um TE anormal. Dos 67 pacientes determinados não tendo CAD, 32 tinham um TE anormal.

*** Texto alternativo utilizando valores de frequência naturais:**

Um estudo sobre a precisão do diagnóstico do teste ergométrico (TE) para o diagnóstico de doença arterial coronariana (DAC) incluiu 120 mulheres com suspeita de DAC, 30 dos quais foram posteriormente determinada tendo DAC (estenose> 50% em um ou mais vasos coronarianos). Daqueles com DAC, 10 tinham uma ET anormal. Dos 90 pacientes determinados não tendo DAC, 30 tiveram um TE anormal.

	Excelente	Forte	Limitado	Mínimo	Não evidente
a: Com base nesses resultados, a sensibilidade do TE para DAC é...	4: 0.31; 31%; 9/29 *4: 0.33; 33%; 10/30	3: Dentro 5%: 26-36%			0: Sem resposta ou resposta errada.
b: Com base nesses resultados, o valor preditivo positivo de TE para DAC é...	4: 0.22; 22%; 9/41 *4: 0.25; 25%; 10/40	3: Dentro 5%: 17-27%			0: Sem resposta ou resposta errada.
c: Com base nesses resultados, a razão de verossimilhança positiva para um TE anormal para DAC é...	4: 0.65; 31/47 *4: 1.0; 0.333/1-0.666	3: Dentro 5%: 0.60-0.70			0: Sem resposta ou resposta errada.

Arredondamento é aceitável (por exemplo 21,9-22 é aceitável)

Pergunta # 10: Um estudo aleatorizado recente de mulheres grávidas com incontinência descobriu que, após o treinamento do assoalho pélvico 20% tinham incontinência, comparado com 32% no grupo controle em três meses após o parto. Nível alfa para o estudo foi estabelecida ao nível de significância de 0,05.

*** Texto alternativo utilizando valores de frequência naturais:**

Um recente estudo randomizado de gestantes com incontinência descobriu que, após o treinamento do assoalho pélvico 20% tinham incontinência em comparação com 30% no grupo controle em três meses após o parto. Nível alfa para o estudo foi estabelecida ao nível de significância de 0,05.

	Excelente	Forte	Limitado	Mínimo	Não Evidente
a: A redução do risco absoluto para eventos recorrentes é...	4: 0.12; 12% *4: 0.10; 10%	3: Dentro 2%: 10-14%			0: Sem resposta ou resposta errada.
b: A redução do risco relativo de eventos recorrentes é...	4: 0.38; 38%; 12/32 *4: 0.33; 33%; 10/30	3: Dentro 2%: 36-40%			0: Sem resposta ou resposta errada.
c: O número necessário para tratar (NNT) para prevenir um evento recorrente é...	4: 9; 1/0.12 *4: 10; 1/0.10	3: Dentro 1: 8-10			0: Sem resposta ou resposta errada.
d: O valor-p indicando diferença estatisticamente significativa entre os grupos seria...	4: <0.05	3: 0.05			0: Sem resposta ou resposta errada.

Pergunta # 11: O mesmo estudo descrito na questão 11 revelou um risco relativo de incontinência de 0,61 para as mulheres que receberam treinamento do assoalho pélvico. Isto sugere que o tratamento treinamento do assoalho pélvico reduz o risco para a incontinência. Gostaríamos de saber se esta diferença é estatisticamente significativa, por isso, olhe para o intervalo de confiança. Dê um exemplo de um intervalo de confiança que suporte a conclusão de que a taxa de incontinência foi realmente (estatisticamente) diferente para estes dois grupos de tratamento.

*** Texto alternativo utilizando valores de frequência naturais:**

O mesmo estudo descrito na questão 11 revelou um risco relativo de incontinência de 0,66 para as mulheres que receberam treinamento do assoalho pélvico. Isto sugere que o tratamento treinamento do assoalho pélvico reduz o risco para a incontinência. Gostaríamos de saber se esta diferença é estatisticamente significativa, por isso, olhe para o intervalo de confiança. Dê um exemplo de um intervalo de confiança que suporte a conclusão de que a taxa de incontinência foi realmente (estatisticamente) diferente para estes dois grupos de tratamento.

	Excelente	Forte	Limitado	Mínimo	Não evidente
A (resposta)	4: Indicação de que qualquer IC que não incluía 1 indicaria significância estatística.				0: Outros

Pergunta # 12: Qual desenho do estudo é o melhor para um estudo sobre o diagnóstico?

A (resposta)	4: Estudo de Coorte; estudo transversal; comparação com o padrão ouro; revisão sistemática.				0: Outros
--------------	---	--	--	--	-----------

Pergunta # 13: Qual desenho do estudo é o melhor para um estudo sobre o prognóstico?

A (resposta)	4: Coorte; prospectivo, longitudinal; Revisão Sistemática.				0: Outros
--------------	--	--	--	--	-----------