



The quality of clinical practice guidelines for chronic respiratory diseases and the reliability of the AGREE II: an observational study

Juliana Souza Uzeloto^{a,*}, Anne M. Moseley^b, Mark R. Elkins^c,
Marcia Rodrigues Franco^a, Rafael Zambelli Pinto^a,
Ana Paula Coelho Figueira Freire^a, Ercy Mara Cipulo Ramos^a

^a Faculdade de Ciências e Tecnologia, UNESP—Univ Estadual Paulista, Presidente Prudente, São Paulo, Brazil

^b The George Institute for Global Health, Sydney Medical School, The University of Sydney, Sydney, New South Wales, Australia

^c Centre for Education & Workforce Development, Sydney Local Health District, Sydney, New South Wales, Australia

Abstract

Objectives To survey the quality of evidence-based clinical practice guidelines for chronic respiratory diseases relevant to physiotherapy practice using the Appraisal of Guidelines for Research and Evaluation version II instrument (AGREE II) and to evaluate the inter-rater reliability of AGREE II.

Design Observational survey.

Procedures Guidelines indexed in the Physiotherapy Evidence Database (PEDro) on chronic respiratory diseases were evaluated by four assessors using AGREE II.

Main outcome measures The six domains and two global items of AGREE II.

Results Thirty-three guidelines were evaluated (58% were published in the last 5 years and 36% were for chronic obstructive pulmonary disease). The domains with the highest scores were scope and purpose (79%, SD 10%) and clarity of presentation (79%, SD 10%). The domain with the lowest score was applicability (37%, SD 23%). Mean overall quality was five out of seven (SD 1). Intraclass correlation coefficients ranged from 0.66 to 0.93 for the six domains and first global item, suggesting good to excellent reliability. The second global item had very poor reliability (Kappa 0.097).

Conclusion The quality of evidence-based clinical practice guidelines for chronic respiratory diseases relevant to physiotherapy could be improved, particularly in consideration with applicability. The number of assessors for AGREE II could be reduced because of the good inter-rater reliability.

© 2016 Published by Elsevier Ltd on behalf of Chartered Society of Physiotherapy.

Keywords: Practice guideline; Respiratory tract diseases; Physiotherapy specialty

Introduction

Chronic respiratory diseases are persistent conditions that affect the respiratory tract, including asthma, allergic rhinitis, bronchiectasis, cystic fibrosis and chronic obstructive pulmonary disease [1]. In 2013, chronic respiratory diseases accounted for more than four million deaths worldwide, an increase of 22% compared to 1990 [2]. The economic burden of chronic respiratory diseases is high. For instance, data from the European Union shows that the economic burden on health services due to chronic obstructive pulmonary disease

* Correspondence: Rua João Lucas Mariotto, 588, Anhumas-São Paulo, Brazil. Fax: +55 18 32295821.

E-mail addresses: juliana.uzeloto@hotmail.com (J.S. Uzeloto), amoseley@georgeinstitute.org.au (A.M. Moseley), mark.elkins@sydney.edu.au (M.R. Elkins), marciarfranco@gmail.com (M.R. Franco), rafaelzambelli@gmail.com (R.Z. Pinto), anapcff@hotmail.com (A.P.C.F. Freire), ercy@bol.com.br (E.M.C. Ramos).

and asthma is about €40 billion [3]. Pulmonary rehabilitation is an important component in the management of chronic respiratory diseases. Physiotherapists are often responsible for implementing pulmonary rehabilitation programs, which include exercise training, behaviour change and collaborative self-management [4]. Recent Cochrane reviews indicate that pulmonary rehabilitation improves quality of life, exercise capacity and dyspnoea for people with chronic obstructive pulmonary disease and interstitial lung disease [5,6]. Physiotherapists also use many other interventions to help manage chronic respiratory disease, such as airway clearance, non-invasive ventilation, and dyspnoea management [7].

Evidence-based clinical practice guidelines are systematically developed statements derived from a rigorous evaluation of high-quality clinical research to guide the overall management of a disease [8]. Many clinical practice guidelines have been developed for the management of chronic respiratory diseases. Unfortunately, the quality of guidelines in healthcare varies widely, particularly in the areas of stakeholder involvement [9–11], rigour of development [9–15], applicability [9–17], and editorial independence [9–11,16]. One explanation for this variability is that the development process is complex, requiring diverse skills from guideline developers including analysis of scientific evidence and synthesis of recommendations [8].

The Appraisal of Guidelines for Research and Evaluation (AGREE) instrument was developed to evaluate the accuracy and methodological transparency of clinical practice guidelines [18]. The original tool has been refined and now is known as AGREE II, which is composed of six domains spanning 23 items plus two global items [19]. In an evaluation of 40 appraisal tools for practice guidelines, the AGREE II tool was considered to be the most comprehensive [20]. While the AGREE II has been validated [18], the AGREE II developers have identified the need to investigate inter-rater reliability [21]. AGREE II has been used to evaluate guidelines in several areas of medicine (including macular degeneration related to age [9], Chagas disease [12], juvenile idiopathic arthritis [16], and respiratory diseases (exclusive to China) [10]), but the quality of English-language guidelines for chronic respiratory diseases relevant to physiotherapy practice does not appear to have been evaluated.

The primary aim of this study is to systematically assess the quality of guidelines for chronic respiratory diseases relevant to physiotherapy practice using the AGREE II instrument. A secondary aim is to evaluate the inter-rater reliability of the AGREE II instrument.

Methods

Design

The present study is an observational survey of published evidence-based clinical practice guidelines for chronic respiratory diseases relevant to physiotherapy practice.

Procedures

The Physiotherapy Evidence Database (PEDro; www.pedro.org.au [22]) was used to identify guidelines because it is likely to provide comprehensive coverage of guidelines relevant to physiotherapy. The guidelines indexed in PEDro are identified through searches of databases of clinical practice guidelines (e.g. National Guideline Clearinghouse in the USA) and bibliographic databases (e.g. Medline, Embase, CINAHL and PsycINFO), receipt of guidelines from guideline development groups (e.g. Dutch clinical practice guideline developers), and notifications from users of PEDro. Furthermore, PEDro indexes all guidelines if they are relevant to physiotherapy practice with no restriction by year or language of publication [22]. Guidelines are included on PEDro if they satisfy the following criteria: (1) produced under the auspices of a health professional association or society, public or private organisation, health care organisation or plan, or government agency; (2) publicly available; (3) a systematic literature search and review of existing scientific evidence published in peer-reviewed journals was performed during the guideline development or the guidelines were based on a systematic review published in the four years preceding publication of the guideline; (4) contains systematically developed statements that include recommendations, strategies, or information to guide decisions about appropriate health care; (5) at least one recommendation concerns at least one intervention that is currently part of physiotherapy practice or that could become part of physiotherapy practice; and, (6) physiotherapy recommendations are based on at least one randomised controlled trial or systematic review [23].

The search strategy used in the PEDro advanced search combined: (i) method: clinical practice guideline; (ii) discipline: cardiothoracics; and (iii) topic: chronic respiratory disease. Only clinical practice guidelines published in English were considered eligible for the present study. Guidelines that were older than the expiry date specified by their developers were excluded. Clinical practice guidelines related to chronic management of respiratory sequelae of other chronic diseases (e.g. chronic neurological diseases) were included.

Full-text copies of each included clinical practice guideline as well as all companion or supplementary documents related to each guideline were retrieved and used for rating. Four assessors who had completed AGREE II training [24] independently evaluated each guideline. The procedures specified in the AGREE II manual were followed [25]. Data acquisition was electronic, with each assessor entering their scores in a pilot-tested Excel spreadsheet.

AGREE II includes 23 specific items and two global items. The specific items are grouped into six domains. The first global item (overall one) rates the overall quality of the guideline, while the second global item (overall two) evaluates recommending the guideline for use [19]. A seven-point Likert scale (range: one *strongly disagree* to seven *strongly*

agree) was used to score each of the 23 specific items and the first global item. A three-point scale was used for the second global item: *yes*; *yes with modification*; or *no*. The specific and global items are listed in Appendix 1 in Supplementary material.

The data from each assessor (*i.e.* scores for each item for each guideline) were merged into a single Excel file. The scaled domain score (expressed as a percentage) for each guideline were calculated using the formula in the AGREE II manual [25]. The scores from all assessors for all items within a domain were summed to produce the “observed score” for the domain. The “minimum possible score” (number of assessors × number of items × 1) and the “maximum possible score” (number of assessors × number of items × 7) were also calculated. The scaled domain score was calculated using the following formula:

$$\frac{\text{Observed score} - \text{Minimum possible score}}{\text{Maximum possible score} - \text{Minimum possible score}} \times 100$$

The average score from the four assessors was calculated for the first global item for each guideline. The AGREE II manual does not provide guidance for producing a summary score from the individual assessor ratings for the second global item. We planned to use the most frequent score (*yes*; *yes with modification*; or *no*) for each guideline.

The following descriptive data were extracted from each guideline: country where the guideline was developed, year of publication, type of disease, type of management, population (adults versus children), number of companion documents, association or society that produced the guideline, and if the Grading of Recommendations Assessment, Development and Evaluation (GRADE) method [26] was used to determine the strength of recommendations. These data were recorded by a single assessor.

Data analysis

The scaled domain scores and the average score for the first global item for each guideline were tabulated. Means and standard deviations were calculated for these seven variables for all guidelines. *Post hoc* analysis was performed for scaled domain scores scoring <60% (*i.e.* evaluation of the mean and standard deviation for the individual items used to calculate the domain). For the second global item, the score from each assessor was tabulated and the most frequent score of each guideline was used. Inter-rater reliability was evaluated for the six domains and the two global items using the scores from the four assessors. Intraclass correlation coefficients (type 2, 1) and 95% confidence intervals were calculated for the six domains and the first global item using SPSS 20.0 software (SPSS, IBM Corporation, NY). A Fleiss Kappa coefficient and 95% confidence interval were calculated for the second global item using a web-based program [27]. The benchmarks used for interpreting the intraclass correlation coefficients were: 0.75 or higher *excellent reliability*;

0.40 to 0.75 *good reliability* and lower than 0.40 *poor reliability* [28]. The benchmarks used for interpreting the Kappa coefficient were: 0.81 or higher *perfect reliability*; 0.61 to 0.80 *substantial reliability*; 0.41 to 0.60 *moderate reliability*; 0.21 to 0.40 *poor reliability* and lower than 0.20 *very poor reliability* [29].

Results

The PEDro search undertaken on 3 November 2015 identified a total of 38 clinical practice guidelines fulfilling the search criteria. Two of these were excluded because they were not in the English language. Three were excluded because they were older than the expiry date specified by their developers. Therefore, 33 guidelines (see Appendix 2 in Supplementary material) were included in the analysis.

Most guidelines were published by countries in North America ($n = 19$, 58%) followed by Europe ($n = 10$, 30%) and Oceania ($n = 2$, 6%), while in two guidelines the country of publication was not specified. The guidelines were published between 2007 and 2015, with 19 (58%) published in the last 5 years. One-third of the guidelines were for chronic obstructive pulmonary disease ($n = 12$, 36%) (Fig. 1). Many were focussed on particular interventions (pulmonary rehabilitation, dyspnoea management, non-invasive ventilation, airway clearance, sleep-disordered breathing management) (Fig. 2). Most guidelines were for the management of adults ($n = 25$, 76% adults only; $n = 5$, 15% adults and children; $n = 3$, 9% children). Among the 23 guidelines that had companion or supplementary documents, the mean number of extra documents was three (SD 3). Twenty-five guidelines were produced by a single association or society. For the remaining guidelines, the number of associations involved in the guideline development ranged from two to four associations. Seven (21%) of the guidelines used GRADE to determine the strength of its recommendations. The method used to determine the strength of recommendations was unclear or not reported for 11 (33%) of the guidelines. All descriptive data are presented in detail in Appendix 2 in Supplementary material.

Means and standard deviations of the scaled domain scores and overall quality of the guideline (first global item) are listed in Table 1. The domains with the highest scores were scope and purpose (79%, SD 10%) and clarity of presentation (79%, SD 10%). In contrast, the domain with the lowest score was applicability (37%, SD 23%). Mean overall quality was five out of seven (SD 1).

Post hoc evaluation of individual items was performed for the three domains scoring <60%: editorial independence (54%), stakeholder involvement (52%) and applicability (37%), see Fig. 3. Lack of reporting of how the funding body influenced the content of the guideline reduced the rating of the editorial independence item. Failing to seek the views and preferences of the target population lowered the rating of stakeholder involvement. Providing tools for putting the

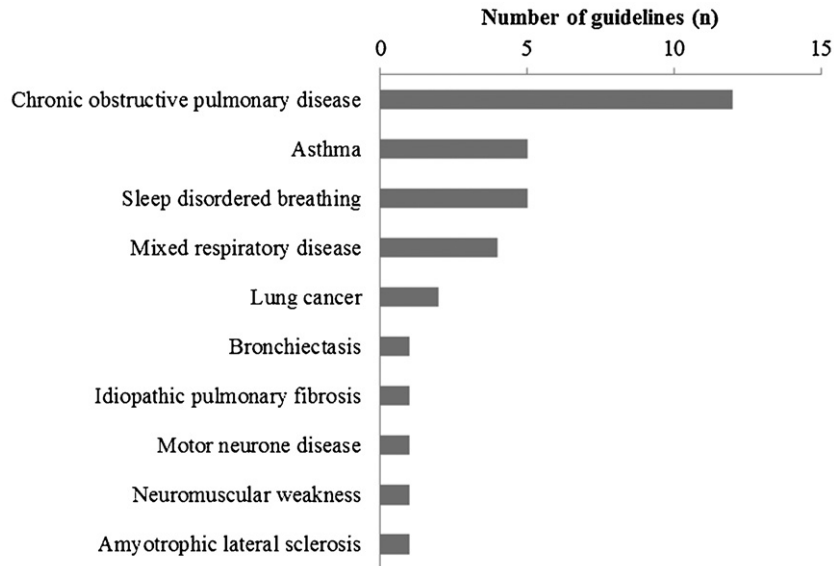


Fig. 1. Disease focus of the guidelines.

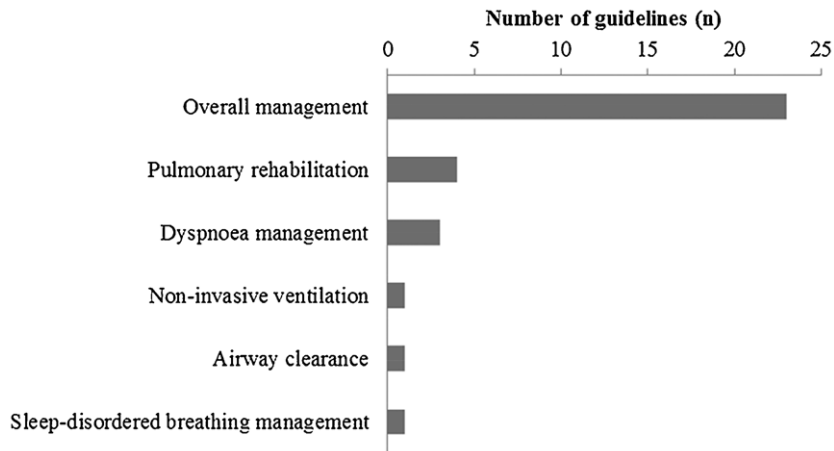


Fig. 2. Management focus of the guidelines.

recommendations into practice, providing monitoring and/or auditing criteria, and considering resource implications had the largest impact on the low scores for the applicability domain.

For the inter-rater reliability analysis, the intraclass correlation coefficients for the six domains and first global item

ranged from 0.66 to 0.93 (see Table 2). All values were in the good or excellent categories [28].

The inter-rater reliability was considered to be very poor (Kappa = 0.097, 95% CI -0.013 to 0.207) for the second global item. It was also not possible to calculate the most common score because of the distribution of the scores from individual assessors (see Appendix 3 in Supplementary material). Instead we calculated the overall frequency of each score (yes; yes with modification; or no) from all assessors and all guidelines. In most cases the assessors recommended the guideline with modifications (i.e. yes with modification; 72%), with 20% recommended without modification (yes), and 8% not recommended (no).

Table 1
Mean and standard deviation scaled score for each domain, ranked from highest to lowest, and the first global item of the AGREE II instrument.

AGREE II domain	Mean	Standard deviation
Scope and purpose	79%	10%
Clarity of presentation	79%	10%
Rigour of development	61%	16%
Editorial independence	54%	24%
Stakeholder involvement	52%	19%
Applicability	37%	23%
Overall quality of the guideline (I7)	5	1

Discussion

Evidence-based clinical practice guidelines relevant to the physiotherapy management of chronic respiratory diseases

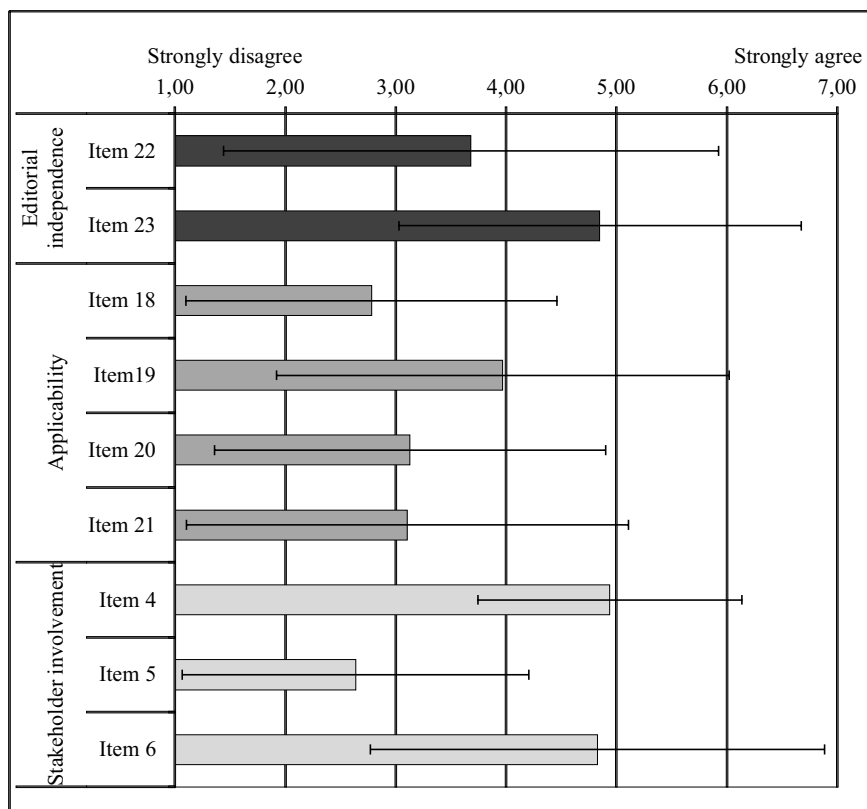


Fig. 3. Post hoc evaluation of individual items for the domains: editorial independence (items 22 and 23), applicability (items 18 to 21) and stakeholder involvement (items 4 to 6).

Table 2
Intraclass correlation coefficients (type 2, 1) and 95% confidence intervals for the six domains, ranked in descending order, and the overall quality of the guideline (first global item) of the AGREE II instrument.

AGREE II domain	Intraclass correlation coefficient	95% confidence interval
Applicability	0.93	0.87 to 0.97
Stakeholder involvement	0.90	0.81 to 0.95
Editorial independence	0.90	0.83 to 0.95
Rigour of development	0.88	0.73 to 0.94
Overall quality of the guideline	0.81	0.61 to 0.91
Clarity of presentation	0.67	0.44 to 0.82
Scope and purpose	0.66	0.42 to 0.81

have a mean overall quality of five out of seven. Although the AGREE II manual explicitly avoids providing benchmarks for interpreting scores, physiotherapists can use existing clinical practice with confidence that they are likely to have clear scope and purpose, acceptable clarity, rigour of development, editorial independence and stakeholder involvement. However, we observed that the applicability domain scored less than 50% (mean 37%), which may indicate that this is an area that requires more attention in future guidelines for chronic respiratory diseases. None of the mean domain scores exceeded 80%, indicating further room for improvement across all domains. The inter-rater reliability was good

to excellent for the six domains and first global item, but poor for the second global item.

A strength of this survey is that it evaluated a relatively large and representative sample of guidelines. One-third of the guidelines were for chronic obstructive pulmonary disease, the most prevalent and cause of most deaths of all chronic respiratory diseases [2]. Most of the guidelines were for the overall management of the disease, but some focussed on one aspect of management (e.g. pulmonary rehabilitation, non-invasive ventilation, airway clearance, or management of a specific symptom). Both overall and focused guidelines provide important information for the management of chronic respiratory diseases.

A limitation of our survey was restriction to guidelines published in English. This may have caused an overestimation of the quality of guidelines for chronic respiratory disease. A recent survey of Chinese-language guidelines for the management of respiratory diseases [10] concluded that the mean scores for the six AGREE II domains for the 16 evidence-based guidelines were all below our values for English-language guidelines, particularly for the editorial independence (5% vs 54%, respectively), rigour of development (26% vs 61%) and stakeholder involvement (25% vs 52%) domains. However, we only excluded two reviews due to language (one French, one Japanese), so any influence on our review's overall estimate of quality was probably very

small. Another limitation of our survey was the identification of guidelines using a single bibliographic database with stringent inclusion criteria (only evidence-based guidelines developed by an organisation are indexed on PEDro). Despite this, our survey evaluated a larger sample ($n = 33$) compared to previous evaluations of English guidelines in other areas of medicine ($n = 5$ to 17) [11–15].

Reporting of the second global item (recommending the guideline for use: *yes*; *yes with modification*; or *no*) was problematic because the AGREE II manual did not provide guidance for producing a summary score and inter-rater reliability was *very poor* ($Kappa = 0.097$) for this item. We were unable to use the most common score for each guideline, instead using a frequency analysis of the scores from all four assessors for all guidelines. In 72% of responses the guidelines were recommended with modification. This difficulty of interpretation was also seen in other studies, with some establishing cut-off values for the AGREE II domains in order to score the second overall item [10–12,14,17], another seeking consensus among assessors [13], and others not reporting this item [9,16]. This variation on the analysis and interpretation of this item makes it difficult to compare between studies and could be the focus of future research.

Our mean domain scores are consistent with several previous studies that used the AGREE II instrument to evaluate guidelines for other areas of healthcare. The domain with the lowest score in our survey, applicability (37%), was at the upper end of the range of values from other studies: 5% to 47% [12–17,30]. The domains with the highest scores, clarity of presentation (79%) and scope and purpose (79%), also compared well to other studies: 51% to 87% [10,13–17,30] and 65% to 92% [9,11,12], respectively. This indicates that guidelines, regardless of the area of healthcare, could be improved (particularly for the applicability domain). This could be addressed in the guideline development process by considering facilitators and barriers to the application of guidelines; providing advice and/or tools for putting the recommendations into practice; developing monitoring and/or auditing criteria; and considering the potential impact of the recommendations on resources within the guideline. The application of clinical practice guidelines to individual patients can also be hampered by features of the trials on which they are based, such as: participants with co-morbidities being excluded; the intervention being delivered in a manner that is not feasible in clinical practice; or poor reporting of details of the intervention. Clinicians may need to adapt guidelines to suit local conditions and obtain additional training in order to implement the recommended interventions. Clinical practice guidelines should be used in combination with the therapist's expertise and practice experience, as well as with the patient's preferences and priorities.

Guidelines synthesise large amounts of clinical research (e.g. one guideline evaluated 655 references [31]) and can improve the outcomes of physiotherapy care [32]. This suggests that a high-quality evidence-based clinical practice

guideline is potentially a more efficient and effective way for clinicians to apply the best evidence, compared to interpreting a large number of individual randomised controlled trials and systematic reviews. Clearly, rigour in the development of guidelines is extremely important for clinical practice. Guideline developers could use the AGREE II tool to guide the construction of guidelines.

Guideline users and researchers could use the AGREE II tool to evaluate guidelines relevant to their area of clinical practice. The inter-rater reliability for individual assessors for the six domains and first global item for AGREE II was *good* to *excellent*, which is consistent with previous reports [10]. This has implications for the scoring recommendations for AGREE II, which encourage the use of four assessors [25]. Perhaps using two assessors to rate each guideline is now justified. However, our reliability evaluation was for individual raters rather than for the summary scores generated by teams of assessors. Future research could evaluate the reproducibility of domain scores produced by teams of raters as well as other measurement properties of the AGREE II tool, such as concurrent validity.

Conclusions

The quality of evidence-based clinical practice guidelines for chronic respiratory diseases relevant to physiotherapy could be improved, especially in consideration with applicability. The AGREE II instrument may be a useful tool for guideline developers and users.

Acknowledgements

We are grateful to Tie Parma Yamato, Patricia do Carmo Silva Parreira and Bruno Saragiotto, who provided practical assistance during the project.

Funding: Grant numbers #2015/13353-4 and #2015/07704-9, São Paulo Research Foundation (FAPESP)—no involvement with study design, collection, analysis and interpretation of data, writing of the report, or in the decision to submit the article for publication.

Conflict of interest: None declared.

Ethics approval: Not required.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.physio.2016.11.002>.

References

- [1] Burney P, Jarvis D, Perez-Padilla R. The global burden of chronic respiratory disease in adults. *Int J Tuberc Lung Dis* 2015;19:10–20.
- [2] Global Burden of Disease 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age–sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015;385:117–71.
- [3] ERS ERS. European lung white book. Available from: <http://www.erswhitebook.org/chapters/the-economic-burden-of-lung-disease/>.
- [4] Spruit MA, Singh SJ, Garvey C, ZuWallack R, Nici L, Rochester C, et al. An official American Thoracic Society/European Respiratory Society statement: key concepts and advances in pulmonary rehabilitation. *Am J Respir Crit Care Med* 2013;188:e13–64.
- [5] McCarthy B, Casey D, Devane D, Murphy K, Murphy E, Lacasse Y. Pulmonary rehabilitation for chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 2015;2:CD003793.
- [6] Dowman L, Hill CJ, Holland AE. Pulmonary rehabilitation for interstitial lung disease. *Cochrane Database Syst Rev* 2014;10:CD006322.
- [7] Opdekamp C, Sergysels R. Respiratory physiotherapy in lung diseases. *Rev Med Brux* 2003;24:A231–5.
- [8] Field MJ, Lohr KN, editors. Clinical practice guidelines: directions for a new program. Washington, DC: 1990.
- [9] Wu AM, Wu CM, Young BK, Wu DJ, Margo CE, Greenberg PB. Critical appraisal of clinical practice guidelines for age-related macular degeneration. *J Ophthalmol* 2015;710324.
- [10] Jiang M, Liao LY, Liu XQ, He WQ, Guan WJ, Chen H, et al. Quality assessment of clinical practice guidelines for respiratory diseases in China: a systematic appraisal. *Chest* 2015;148:759–66.
- [11] Zhang Z, Guo J, Su G, Li J, Wu H, Xie X. Evaluation of the quality of guidelines for myasthenia gravis with the AGREE II instrument. *PLoS One* 2014;9:e111796.
- [12] Olivera MJ, Fory JA, Olivera AJ. Quality assessment of clinical practice guidelines for Chagas disease. *Rev Soc Bras Med Trop* 2015;48:343–6.
- [13] Dersch R, Toews I, Sommer H, Rauer S, Meerpohl JJ. Methodological quality of guidelines for management of Lyme neuroborreliosis. *BMC Neurol* 2015;15:242.
- [14] Kirby JS, Scharnitz T, Seiverling EV, Ahrns HSF. Actinic keratosis clinical practice guidelines: an appraisal of quality. *Dermatol Res Pract* 2015;2015:1–7.
- [15] Parisi P, Vanacore N, Belcastro V, Carotenuto M, Del Giudice E, Mariani R, et al. Clinical guidelines in pediatric headache: evaluation of quality using the AGREE II instrument. *J Headache Pain* 2014;15:57.
- [16] Smith CA, Toupin-April K, Jutai JW, Duffy CM, Rahman P, Cavallo S, et al. A systematic critical appraisal of clinical practice guidelines in juvenile idiopathic arthritis using the Appraisal of Guidelines for Research and Evaluation II (AGREE II) instrument. *PLoS One* 2015;10:e0137180.
- [17] Brosseau L, Rahman P, Toupin-April K, Poitras S, King J, De Angelis G, et al. A systematic critical appraisal for non-pharmacological management of osteoarthritis using the appraisal of guidelines research and evaluation II instrument. *PLoS One* 2014;9:e82986.
- [18] The AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Qual Saf Health Care* 2003;12:18–23.
- [19] Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ* 2010;182:E839–42.
- [20] Siering U, Eikermann M, Hausner E, Hoffmann-Esser W, Neugebauer EA. Appraisal tools for clinical practice guidelines: a systematic review. *PLoS One* 2013;8:e82915.
- [21] Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. Development of the AGREE II, part 1: performance, usefulness and areas for improvement. *CMAJ* 2010;182:1045–52.
- [22] PEDro-Physiotherapy Evidence Database. Available from: <http://www.pedro.org.au/>. [cited 22 December 2015].
- [23] PEDro-Physiotherapy Evidence Database. Available from: <http://www.pedro.org.au/english/downloads/criteria/>. [cited 23 December 2015].
- [24] AGREE II training tools. Available from: <http://www.agreetrust.org/resource-centre/agree-ii-training-tools/>. [cited 17 December 2015].
- [25] Brouwers MC. Consortium Members. Appraisal of guidelines for research & evaluation II instrument, 2009.
- [26] Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64:383–94.
- [27] Computer program to calculate Cohen and Fleiss Kappa for ordinal scales. Available from: <https://www.statstodo.com/CohenKappa.Pgm.php>. [cited 14 December 2015].
- [28] Fleiss JL. The design and analysis of clinical experiments. New York, NY: John Wiley & Sons; 1986.
- [29] Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
- [30] Acuna-Izcaray A, Sanchez-Angarita E, Plaza V, Rodrigo G, Montes de Oca M, Gich I, et al. Quality assessment of asthma clinical practice guidelines: a systematic appraisal. *Chest* 2013;144:390–7.
- [31] Rudolf M, O'Reilly J, Parnham J, Sloan N, Crowe E, O'Mahony R, et al. Chronic obstructive pulmonary disease: management of chronic obstructive pulmonary disease in adults in primary and secondary care 2010. Available from: <http://www.nice.org.uk/guidance/cg101/resources/guidance-chronic-obstructive-pulmonary-disease-pdf>. [cited 3 November 2015].
- [32] Cura Della Redazione A. Clinical practice guidelines: what they are and how are developed. *Assist Inferm Ric* 2014;33:214–8.

Available online at www.sciencedirect.com

ScienceDirect