Psychometric properties of The Maastricht Upper Extremity Questionnaire: systematic review and meta-analysis

Propriedades psicométricas do The Maastricht Upper Extremity Questionnaire: revisão sistemática e meta-análise

Jânio Luiz Correia Júnior[1], Maria Fernanda da Silva Torres[2], Sthefanny Pontes Sampaio Costa[2], Hermê Fellipo Bordoni Caldeira[1], Marina Pereira Gonçalves[3], Ricardo Freitas-Dias[1,2]

DOI 10.5935/2595-0118.20230069-en

ABSTRACT

BACKGROUND AND OBJECTIVES: With the necessity to assess musculoskeletal complaints caused by computer use, The Maastricht Upper Extremity Questionnaire (MUEQ) was created, which aims to assess musculoskeletal complaints of the upper limbs, shoulder complex and cervical spine in computer users. However, there is currently no comprehensive summary in the scientific literature on the psychometric properties of the MUEQ. The objective of this study was to conduct a synthesis of all available scientific evidence that has analyzed the psychometric properties of the MUEQ.
CONTENTS: This study followed the PRISMA recommendations. The bibliographic search was carried out in the following databases: MEDLINE (via VHL), Embase, LILACS (via VHL), Pubmed, PsycINFO, Scielo, Academic Search Premier, CINAHL, Rehabilitation & Sports Medicine Source, MEDLINE Complete, Web of Science CENTRAL, Scopus and SPORTDiscus. Studies that addressed the psychometric properties of the MUEQ were included, as long as they were original articles of research carried out with human beings and indexed in the databases used. The studies were selected in two phases, with two independent reviewers. A total of 6 articles were included in the analysis. The evidence based on internal structure showed acceptable results. The reliability indexes ranged from α=0.52 to α=0.84, and ICC/composite reliability > 0.70 in the analyzed studies, classified as "good" and "excellent," respectively.
CONCLUSION: In general, this research found a lack of detail on the process of content validity and evidence related to external variables and the description of the sample. These problems extended to the evidence based on the internal structure and reliability of the MUEQ, which did not reach levels considered acceptable to ensure its adequacy and accuracy.

Keywords: Musculoskeletal pain, Occupational health, Surveys and questionnaires, Teleworking, Upper extremity.

RESUMO

JUSTIFICATIVA E OBJETIVOS: Com a necessidade de avaliar as queixas musculoesqueléticas ocasionadas pelo uso de computadores, foi criado o *The Maastricht Upper Extremity Questionnaire* (MUEQ), cujo objetivo foi avaliar as queixas musculoesqueléticas relativas aos membros superiores, ao complexo do ombro e à cervical em usuários de computadores. No entanto, atualmente não existe uma sumarização abrangente, na literatura científica, sobre as propriedades psicométricas do MUEQ. O objetivo deste estudo foi realizar uma síntese de evidências científicas disponíveis que analisaram as propriedades psicométricas do MUEQ.
CONTEÚDO: Este estudo seguiu as recomendações do PRISMA. A busca bibliográfica foi realizada nas bases de dados Medline (BVS), Embase, LILACS (via BVS), Pubmed, *PsycINFO,* Scielo, *Academic Search Premier, Cinahl, Rehabilitation* & *Sports Medicine Source,* Medline *Complete, Web of Science CENTRAL, Scopus* e *SPORTDiscus.* Foram incluídos estudos que abordaram as propriedades psicométricas do MUEQ, desde que fossem artigos originais de pesquisas desenvolvidas com seres humanos e indexados nas bases utilizadas. A seleção dos estudos ocorreu em duas fases, com dois revisores independentes. Foram incluídos 6 artigos/publicações na análise. A evidência baseada na estrutura

Jânio Luiz Correia Júnior – https://orcid.org/0000-0001-7122-4923;
Maria Fernanda da Silva Torres – https://orcid.org/0000-0002-7791-9280
Sthefanny Pontes Sampaio Costa – https://orcid.org/0000-0003-2619-7801;
Hermê Fellipo Bordoni Caldeira – https://orcid.org/0000-0002-6979-5134;
Marina Pereira Gonçalves – https://orcid.org/0000-0002-6344-7059;
Ricardo Freitas-Dias – https://orcid.org/0000-0003-4846-9951.

1. University of Pernambuco, Graduate Program in Hebiatrics, Determinants of Health in Adolescence, Pernambuco, PE, Brazil.
2. University of Pernambuco, Physiotherapy Course, Petrolina, PE, Brazil.
3. Federal University of the São Francisco Valley, Psychology Course, Petrolina, PE, Brazil.

Presented on March 22, 2023.
Accepted for publication on September 11, 2023.
Conflict of interests: none - Sponsoring sources: none.

HIGHLIGHTS
- This was the first systematic review to evaluate the psychometric properties of the Maastricht Upper Extremity Questionnaire (MUEQ);
- This systematic review provided comprehensive and up-to-date evidence of the psychometric properties and level of evidence of the MUEQ;
- The evidence based on the internal structure and reliability of the MUEQ did not reach levels considered acceptable in all the studies.

Correspondence to:
Ricardo Freitas-Dias
E-mail: ricardo.freitas@upe.br

interna apresentou resultados aceitáveis. Os índices de fidedignidade variaram de α=0,52 a α=0,84 e ICC/confiabilidade composta foram maiores que 0,70 nos estudos selecionados, classificados como "bom" e "excelente", respectivamente.
CONCLUSÃO: De um modo geral, esta pesquisa constatou a falta de detalhamento sobre o processo de validade de conteúdo e de evidências relacionados a variáveis externas e à descrição da amostra. Esses problemas se estenderam à evidência baseada na estrutura interna e à confiabilidade do MUEQ, que não alcançaram níveis considerados aceitáveis para garantir sua adequação e precisão.
Descritores: Dor musculoesquelética, Extremidade superior, Inquéritos e questionários, Saúde ocupacional, Teletrabalho.

INTRODUCTION

Musculoskeletal complaints in the upper limbs, shoulder complex, and neck have been the subject of much attention in modern society, since disorders in this region constitute a serious problem, negatively impacting workers' health and generating negative effects, including reduced productivity at work, absenteeism, and consequently, loss of employment[1,2]. Musculoskeletal complaints of the arm, neck and/or shoulder (CANS) are defined as complaints not caused by acute trauma or any systemic disease[3].
Worldwide epidemiological data, especially from developed countries, has reported that the annual prevalence of musculoskeletal symptoms related to the hand, arm, shoulder and neck regions in computer users varies between 10% and 51.7%[4,5]. Thus, these symptoms can be considered a work-related global health problem[6,7]. Over the last 20 years, and especially during and after the coronavirus pandemic, there has been a significant increase in the number of computer users, which has also resulted in an increase in CANS[8-10].
It is important to note that the increase in complaints may be related to psychosocial factors associated with the home office, as well as ergonomic issues[9,11]. Many of these users and workers do not have adequate minimum conditions in their homes, which increases the occupational risks in this context[12-15]. In other words, performing tasks at the computer for a long time and in a seated position is considered one of the risk factors for developing musculoskeletal problems in the workplace[16-19]. The outcome of these disorders can be severe and debilitating symptoms such as intense pain, numbness, and tingling in the arms, neck and shoulders[20].
The Maastricht Upper Extremity Questionnaire (MUEQ) was developed to assess upper limb musculoskeletal pain in computer users and its associated physical and psychosocial risk factors, as an instrument designed to assess CANS[21,22]. The original MUEQ version was validated in the Dutch population, specifically in office workers from IT field. The 95 items of the questionnaire were grouped into six different domains: workstation, body posture, work control, work demand, breaks, work environment and social support. Each domain contains between 7 and 10 items, which use a five-point Likert-type response scale, ranging from "always" to "never", or a dichotomous statement, "yes" and "no"[21,23].
To validate the MUEQ, analyses were carried out to assess its psychometric properties, including evidence of validity based on internal structure and reliability. Reliability was assessed using Cronbach's alpha and showed acceptable results (α>0.70), indicating high reliability of the questionnaire. To check the relationship between the items assessed and the questionnaire construct, the item-total correlation was used, in which most of the results found varied between 0.2 and 0.5, indicating that the items assessed were related to the questionnaire construct[21,23]. It is important to emphasize that the validation of an instrument in a specific context does not automatically imply its applicability in other populations or contexts. It is therefore important to carry out new validation studies on different samples in order to use MUEQ in other situations or groups of people[24].
Considering that the most prevalent musculoskeletal symptoms are related to the hand, arm, shoulder and neck regions, it is important to use MUEQ to identify physical and psychosocial risk factors associated with musculoskeletal complaints in computer users[21]. Therefore, MUEQ is a simple, low-cost tool that can facilitate the implementation of public policies on occupational health and the prevention of work-related illnesses, making it an important option for public health[21,23]. These considerations reinforce the need for MUEQ to have acceptable psychometric properties, in accordance with the recommendations of the Standards for Psychological and Educational Testing[25].
Currently, there is no comprehensive summary of the MUEQ psychometric properties in the scientific literature, which makes it difficult to use it in different countries, contexts and clinical studies. Given this gap, the aim of this study was to carry out a synthesis of all the available scientific evidence that has analyzed the psychometric properties of MUEQ. This analysis will enable the compilation of relevant information on the applicability, methodological quality and level of evidence of the available studies.

CONTENTS

This Systematic Review and Meta-analysis study followed the recommendations of PRISMA[26] and PROSPERO[27,28], and was registered under opinion number CRD42022339858.

Eligibility criteria

To ensure inclusion, the studies had to meet the following criteria: 1) studies evaluating the psychometric properties of MUEQ (evidence of content validity, evidence based on the item response process, evidence based on internal structure, evidence based on relationships with external variables and reliability); 2) original research studies involving human subjects; 3) indexed in the electronic databases used. The following were excluded: 1) review studies; 2) editorials; 3) conference publications; 4) theses/dissertations; 5) Course Conclusion Work files.

Search strategies

The literature search included articles/publications published up to June 10, 2022 and listed in the following electronic databases: Medline (via VHL), Embase, LILACS (via VHL), Pubmed, PsycINFO, Scielo, Academic Search Premier, CINAHL, Rehabilitation & Sports Medicine Source, Medline Complete, Web of Science, CENTRAL, Scopus and SPORTDiscus. Grey literature was searched via BVS and Embase databases.

The searches in the electronic databases were carried out without restrictions on language and year of publication. The type of instrument item in the search strategy, recommended by The COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN), was applied in this study[28,29]. The search strategies were developed by an expert researcher (JLCJ) and reviewed following the Peer Review of Electronic Search Strategies (PRESS) guideline[30].

To identify the terms, searches were carried out for the term "The Maastricht Upper Extremity Questionnaire" using keywords and descriptors found in the MeSH and *DeCS*/MeSH term dictionaries and scientific articles related to the topic in question. The synonymous terms identified were combined using the Boolean operator OR (The Maastricht Upper Extremity Questionnaire OR Maastricht Upper Extremity Questionnaire OR MUEQ OR The MUEQ). In the electronic databases, the key [TIAB] was used to limit the display of search terms, related to title and abstract[28].

Study selection

To select the studies, an evaluation form was developed based on the eligibility criteria (inclusion and exclusion). The form was then calibrated before being screened to select the studies. Duplicate studies were identified in the Mendeley software and removed by a trained researcher (JLCJ). The studies were then exported to the Rayyan QCRI web application (http://rayyan.qcri.Org/)[31,32].

The remaining studies underwent analysis by four independent evaluators (BNB; MFST; SCSPS; SPSC) in two stages: 1) screening by title and abstract; and 2) screening by reading the full text. In the first stage, titles and abstracts were analyzed according to eligibility criteria to identify relevant studies. These studies were classified as "yes" (included) or "no" (excluded). Next, the studies selected by title and abstract were read in full and examined by four independent evaluators (BNB; MFST; SCSPS; SPSC), based on the eligibility criteria (inclusion and exclusion), using the assessment form[28,29].

Any inconsistencies among the four evaluators were discussed, and a final decision was reached by consensus. In the absence of consensus, a fifth reviewer was consulted (JLCJ) to determine the inclusion or exclusion of the study. Finally, the studies selected for reading in full were subjected to a search on the list of references to identify relevant articles/publications that were not screened out in the searches on the electronic databases[28,29].

Data extraction

The evaluators were first trained and familiarized with the data extraction spreadsheet, and then calibrated with a study related to the topic in question. The data extracted from the studies that met the eligibility criteria was entered into an Excel spreadsheet. The following data was extracted from the studies: registration data, objective and type of study, characteristics of the instrument and results of the psychometric properties (evidence based on content, evidence based on the item response process, evidence based on internal structure, evidence based on relationships with external variables and reliability).

The four evaluators described previously independently extracted descriptive data and quantitative results from the selected studies. Any unresolved discrepancies between the four reviewers were examined by a fifth reviewer (JLCJ).

Evaluation of the risk of bias

The methodological quality of the studies included in the review was assessed using the Critical Appraisal Tool (CAT) for validity and reliability studies of objective clinical tools[33]. This tool includes 13 items, with items 1, 2, 10, 12 and 13 applied to validity and reliability studies; items 3, 7, 9 and 11 applied to validity studies; and items 4, 5, 6 and 8 applied to reliability studies. Each item was scored using a 3-point evaluation scale (yes - Y, no - N, or not applicable - NA).

Initially, the evaluators were trained and familiarized with the tool to assess the risk of bias and then calibrated with a study related to the topic in question. Next, two independent evaluators (JLCJ and HFBC) assessed the methodological quality of the selected studies. Any inconsistencies between the two evaluators were discussed, and a final decision was reached by consensus. In the absence of consensus, a third rater was consulted (RFD) to determine whether or not to score the item[29,34].

Assessing the certainty of the evidence

The evaluators were first trained and familiarized with the tool to assess the certainty of the evidence, and then calibrated with a study related to the topic in question. Next, two independent, evaluators (HFBS and JLCJ) assess the certainty of the evidence, using the five recommendations of the modified GRADE approach[29,34]. This instrument has 4 evaluation items. The following criteria were considered to reduce the certainty of the evidence: one to two levels - risk of bias, inconsistency (unexplained), imprecision (small sample size) or indirect evidence; and three levels - evidence based on only one inadequate study (high risk of bias)[29,34]. In the end, the level and certainty of the evidence was classified as high, moderate, low or very low.

All inconsistencies between the two evaluators (JLCJ and HFBC) were discussed, and a final decision was reached by consensus. In the absence of consensus, a third rater was consulted (RFD) to determine whether or not to score the item[29,34].

Data analysis

To group the Cronbach's alpha (α) values of each MUEQ factor, meta-analyses of correlation coefficients were calculated using Fisher's Z transformation. The meta-analyses were calculated using random effects (RE) models, calculated using the restricted maximum likelihood method, assuming that the selected studies were sampled from a larger set of studies. This calculation reduces the risk of type I error, as these models take into account the variability between the included studies.

It is important to note that the random effects model was chosen over a fixed effects model due to experimental factors such as the study methodology (for example, sample conditions such as university students and schoolchildren), which can influence the reporting of interpersonal behaviors of the samples and the reliability of the study results[35-38]. In addition to these factors, the random effects model allows for greater external generalization compared to the fixed effects model.

Heterogeneity between studies was assessed using Cochran Q test statistics (adopting a significance level of $p < 0.1$) and inconsistency was assessed using Higgins' I^2 index[39]. The following criteria were adopted: values of ≤ 40% indicate low heterogeneity; 30% to 60% indicate moderate heterogeneity; > 50% to 90% indicate substantial heterogeneity and > 75% to 100% indicate considerable heterogeneity[28]. When $I^2 > 50\%$ and tau squared (τ^2) > 1, were accompanied by statistical significance ($p < 0.05$), significant heterogeneity was considered to have occurred.

For the internal consistency analysis and qualitative interpretation, Cronbach's alpha values were used, adopting the following categories: *excelente* (excellent) ≥ 0.85; *bom* (good) 0.80 – 0.84; *moderado* (moderate) 0.75 – 0.79 e *justo* (fair) 0.70 – 0.74. Cronbach's alpha values were determined considering the factors and sample size[40,41]. For the results of the evidence based on internal structure, the following criteria were considered acceptable: Comparative Fit Index (CFI) or Tucker Lewis Index (TLI) > 0.95; Root-Mean-Square Error of Approximation (RMSEA) < 0.06; or Standardized Root Mean Square Residual (SRMR) < 0.082, Average Variance Extracted (AVE) > 0.5[29.42-44]. For the results of evidence based on relationships with external variables (criterion validity, convergent type), results > 0.70[29] were considered acceptable. For the reliability results, assessed by Intraclass Correlation Coefficients (ICC) and composite reliability, results > 0.70[29] were considered acceptable.

RESULTS

A total of 1,635 articles/publications were found in the 14 electronic databases selected, of which 858 duplicates were excluded. After screening by title/abstract, 756 articles/publications were excluded, leaving 21 articles/publications to read the full text. After reading the full text, 6 articles/publications were eligible for quantitative and qualitative extraction (Figure 1). Among the 6 eligible articles/publications, 3 studies were identified in a single article[20]. So, 8 studies were identified. At the title/abstract selection stage, there was 99.9% agreement between the evaluators; at the full reading stage, there was 71.43% agreement.

Study characteristics

For the narrative synthesis of this study, six articles/publications were included, containing a total of eight studies, with

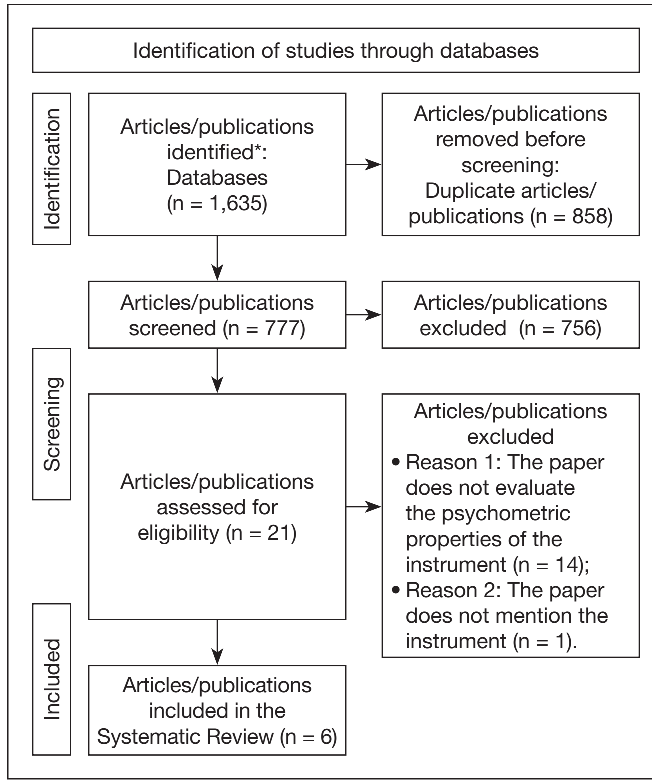

Figure 1. Flowchart of study selection.

*1) Medline (n = 441), 2) CENTRAL (n = 4), 3) Embase (n = 38), 4) Pubmed (n = 570), 5) Rehabilitation & Sports Medicine Source (n = 0), 6) Scopus (n = 29), 7) PsycINFO (n = 3), 8) Academic Search Premier (n = 13), 9) CINAHL (n = 9), 10) Medline Complete (n = 312), 11) SPORTDiscus (n = 0), 12) LILACS (n = 3), 13) Web of Science (n = 213) e 14) Scielo (n = 0).

one article[20] containing three studies[19-23,45]. The studies were published between 2007 and 2021 and were validated in six different languages, including Greek, Dutch, Arabic, Persian, Sinhalese and Portuguese[19-23,45]. The total sample, considering all the studies, was made up of 2,841 individuals, with a variation in sample size in between 50 and 600.

The participants in all the studies were of both genders[19-23,45]. However, females predominated in four studies[20,22,45]. The participants' ages ranged from 20 to 65 in the studies (table 1). In all the studies, the sample was recruited by convenience[19-23,45].

The studies included in this review evaluated the following psychometric properties: content validity, evidence based on internal structure and reliability. However, evidence based on relationships with external variables and the item response process was not evaluated in any of the included studies[19-23,45]. The response rate between the studies ranged from 44% to 97.7%[19-23,45]. For the studies with incomplete and/or missing data, requests were sent via e-mail, but no response was received from any of the authors.

Risk of bias analysis

As for the criteria that assess validity and reliability: only one study[19] described the sample of subjects in detail (item 1); detailed the execution of the test for its replication (item 10); used appropriate statistical methods (item 13); and reported

Table 1. Summary and characteristics of the results of studies using the Maastricht Upper Extremity Questionnaire to assess the nature and occurrence of CANS.

Authors	Studies	Total sample number	Study characteristics	MUEQ application	Types of psychometric properties	Conflict of interests and sources of funding
Bekiari et al.[22]	Study 1	n = 600	Gender: female (60.7%) and male (39.3%). Average age: 37.4. Validation language: Greek. Sample condition: at the workplace. Completion time: 20 minutes. Number of items: 95.	Not reported	Validity of content and evidence based on internal structure	Conflict of interests: declared; Sources of funding: undeclared.
Eltayeb et al.[21]	Study 1	n = 600	Gender: female (49.62%) and male (50.38%). Validation language: Dutch. Sample condition: the study site was the Office of the National Social Security Institute. The questionnaires were distributed through the GAK's internal mail. Participants were asked to fill in the questionnaire and return it using the enclosed envelope. Completion time: 20 minutes. Number of items: 95.	Printed (e-mail)	Evidence based on internal structure	Conflict of interests: declared; Sources of funding: undeclared.
Eltayeb et al.[23]	Study 1	n = 282	Gender: female (35%) and male (65%). Validation language: Arabic. Sample condition: on April 1, 2005, the questionnaires were distributed among the participants, delivered to their workplace. Participants were asked to fill in the questionnaire and return it in specially provided boxes. In mid-April, a reminder note was posted to non-respondents. And the end of April 2005 was set as the last return date. Completion time: 30 minutes. No. of items: 109.	In person	Validity of content and evidence based on internal structure	Conflict of interests: declared; Sources of funding: undeclared.
Ghasemi et al.[19]	Study 1	n = 282	Gender: female (39%) and male (61%). Average age: 35.17. Validation language: Persian. Sample condition: explanations on how to answer the options and the content of the questions were given to the participants. The place of application of the instrument was the Government Information Technology Administration in Tehran, Iran. Completion time: 20 minutes. Number of items: 109 (95 - 107).	In person	Validity of content and evidence based on internal structure	Conflict of interests: declared; Sources of funding: declared.
Ranasinghe et al.[45]	Study 1	n = 450	Gender: female (57.3%) and male (42.7%). Average age: 38.2. Validation language: Sinhalese. Sample condition: Telecommunications company in Colombo - Sri Lanka between January and February 2009. Completion time: 30 minutes. Number of items: 94.	Not reported	Content validity	Conflict of interests: declared; Sources of funding: undeclared.
Turci et al.[20]	Study 1	n = 627	Gender: female (74.5%) and male (25.5%). Average age: 33.56. Validation language: Portuguese. Sample condition: University of São Paulo (USP), Ribeirão Preto Campus. Number of items: 41.	Not reported	Validity of content and evidence based on internal structure	Conflict of interests: declared; Sources of funding: declared.
	Study 2		Gender: female (40%) and male (60%). Validation language: Portuguese. Sample condition: University of São Paulo (USP), Ribeirão Preto Campus. Number of items: 41.			
	Study 3		Gender: female (55.5%) and male (44.5%). Validation language: Portuguese. Sample condition: University of São Paulo (USP), Ribeirão Preto Campus. Completion time: 14.67 min. No. of items: 41.			

clarification of the qualification or competence of the evaluator(s) (item 2 - table 2).

As for the criteria that assess validity, none of the studies[19-23,45]: explained the reference standard test (item 3); reported on the independence of the reference standard test (item 9); detailed the execution of the reference standard test for its replication (item 11); or reported whether the target condition did not change between the application of the two tests (item 7 - table 2).

As for the criteria that assess reliability, no study[19-23,45] reported: inter-rater blinding (item 4); intra-rater blinding (item 5) and varied order of test application (item 6 - table 2).

Looking at the set, it is possible to see that items 3, 4, 5, 6, 7, 9, 11 and 12 were not scored in any of the studies[19-23,45]; which may generate risk of bias in the studies (Figure 2). Finally, the agreement between the two evaluators regarding the risk of bias was 57.7%.

Table 2. Assessment of methodological quality using the CAT tool for validity and reliability studies of objective clinical tools.

Authors	Studies	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	I13
Bekiari et al.[22]	Study 1	N/A	N/A	N	N/A	N/A	N/A	N	N/A	N	N/A	N	N/A	N/A
Eltayeb et al.[21]	Study 1	N/A	N/A	N	N/A	N/A	N/A	N	N/A	N	N/A	N	N/A	N/A
Eltayeb et al.[23]	Study 1	N/A	N/A	N	N/A	N/A	N/A	N	N/A	N	N/A	N	N/A	N/A
Ghasemi et al.[19]	Study 1	Y	Y	N	N	N	N	N	N	N	Y	N	N	Y
Ranasinghe et al.[45]	Study 1	N/A	N/A	N	N/A	N/A	N/A	N	N/A	N	N/A	N	N/A	N/A
Turci et al.[20]	Study 1	N/A	N/A	N	N/A	N/A	N/A	N	N/A	N	N/A	N	N/A	N/A
	Study 2	N/A	N/A	N/A	N	N	N	N/A	Y	N/A	N/A	N/A	N/A	N/A
	Study 3	N/A	N/A	N	N/A	N/A	N/A	N	N/A	N	N/A	N	N/A	N/A

I1 - If human subjects were used, did the authors give a detailed description of the sample of subjects used to perform the test (index)?; I2 - Did the authors clarify the qualification or competence of the rater(s) who performed the test?; I3 - Was the reference standard explained?; I4 - If inter-rater reliability was tested, were the raters blind to the findings of other raters? I5 - If intra-rater reliability was tested, were the raters blind to their own previous findings of the test under evaluation? I6 - Was the order of the test varied? I7 - If human subjects were used, was the period between the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests? I8 - Was the stability (or theoretical stability) of the variable being measured considered when determining the adequacy of the time interval between repeated measurements? I9 - Was the reference standard independent? I10 - Was the execution of the test described in sufficient detail to allow replication of the test? I11 - Was the execution of the reference standard described in sufficient detail to allow its replication? I12 - Were the withdrawals from the study explained? I13 - Were the statistical methods appropriate for the purpose of the study? Y: Yes; N: No; N/A - not applicable.

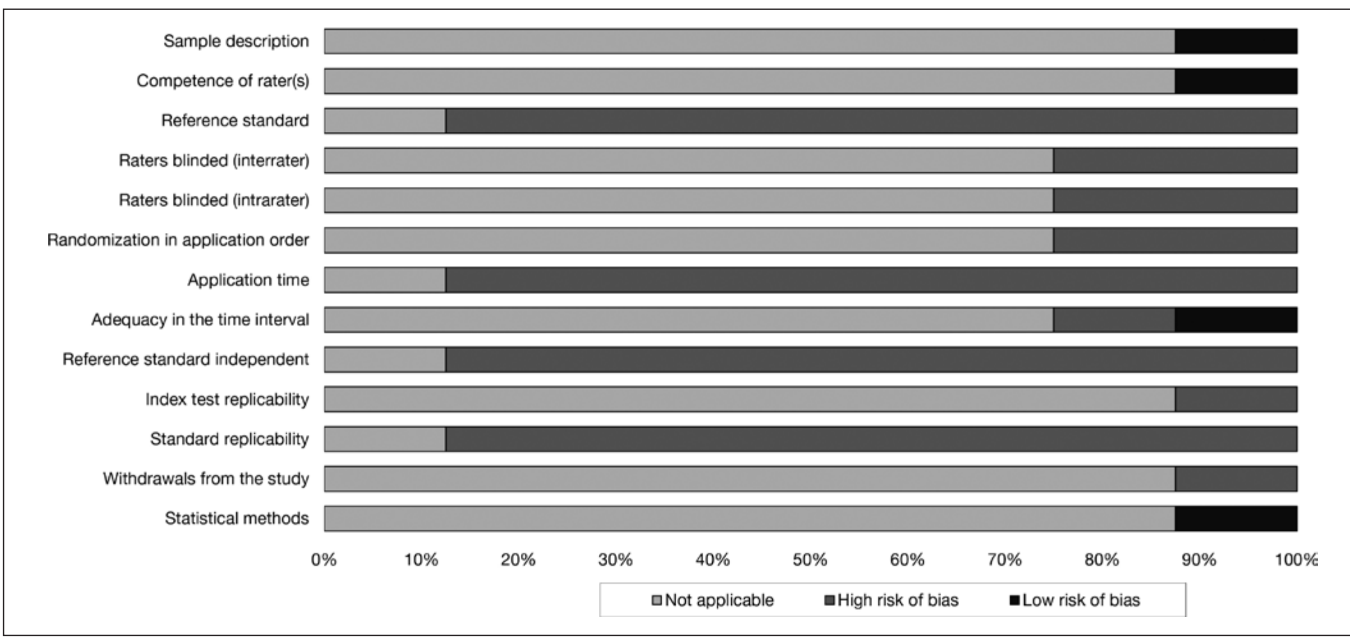

Figure 2. Overall methodological quality using the CAT tool for validity and reliability studies of objective e clinical tools. The data is presented as a percentage (%).

Summary of psychometric properties evidence

Evidence based on internal structure, measured by CFI or RMSEA, showed levels of internal structure validity that were considered acceptable, with excellent adjustments in the studies that reported them[19,20]; except for one study which had CFI results below those recommended by scientific literature (> 0.95 - table 3)[29]. Evidence of reliability, analyzed by ICC, was reported in two studies[19,20], which presented moderate and excellent results, respectively (table 3).

Only two studies included other confirmatory factor analysis indices: a) In one study[19] the following indices were reported: Parsimony comparative fit index (PCFI), with a result of 0.732; Parsimony goodness of fit index (PGFI), with a result of 0.680 and Normed fit index (NFI), with a result of 0.680[19]; and b) In another study[20] the following indices were reported: Consistent Akaike information criterion (CAIC), with a result of 2,230.40; Goodness-of-fit index (GFIÿ), with a result of 0.90; Non-normed fit index (NNFI), with a result of 0.90; and Expected cross-validation index (ECVI), with a result of 3.782[20].

Summary of internal consistency results

The studies[19-23,45] were analyzed using Cronbach's alpha. When all the studies were grouped and meta-analyzed[19-23,45], the results for MUEQ factors were as follows: in the first domain, "body posture", divided into two factors, the average reliabi-

Table 3. Synthesis of quantitative results of the psychometric properties of validity of studies on The Maastricht Upper Extremity Questionnaire instrument for assessing physical, environmental and psychological risks in workplace.

Authors	Studies	Evidence Based on Internal Structure	Reliability
Bekiari et al.[22]	Study 1	Not reported	Not reported
Eltayeb et al.[21]	Study 1	Not reported	Not reported
Eltayeb et al.[23]	Study 1	Not reported	Not reported
Ghasemi et al.[19]	Study 1	PCFI = 0.732; PGFI = 0.680; NFI = 0.680; CFI = 0.800; RMSEA = 0.062.	ICC values: BP = 0.61; WC = 0.77; WD = 0.81; BT = 0.82; WE = 0.67; SS = 0.83; Total = 0.62. Test-retest (alpha): BP = 0.826; WC = 0.977; WD =0.975; BT = 0.789; WE = 0.963; Total: 0.897; SS = 0.986. MUEQ General: 0.897.
Ranasinghe et al.[45]	Study1	Not reported	Not reported
Turci et al.[20]	Study 1	CAIC = 2,230.40; CFI = 0.91; GFIÿ = 0.90; NNFI = 0.90; ECVI = 3.78; RMSEA = 0.04.	ICC values (95%): D: WS: 0.94 (0.90 - 0.96); BP: 0.85 (0.74-0.91); JC: 0.84 (0.71-0.90); WD: 0.95 (0.91-0.97); BT: 0.94 (0.89-0.96); SS: 0.87 (0.77-0.92) Total: 0.95 (0.90-0.97)
	Study 2		Not reported
	Study 3		Not reported

PCFI = Parsimony comparative fit index; PGFI = Parsimony goodness of fit index; NFI = Normed fit index; CFI = Comparative fit index; RMSEA = Root mean squared error approximation; CAIC = Consistent Akaike information criterion; GFIÿ = Goodness-of-fit index; NNFI = Non-normed fit index; ECVI = Expected cross-validation index; ICC = Intraclass correlation coefficients. T = Total; D = Domain; FC = Factor; It: Item; WS = Work Station; BP = Body Posture; WC = Work Control; JC = Job Control; WD = Work Demand; JD = Job Demands; BT = Break Time; WE = Work Environment; SS = Social Support; OE = Office Equipment; CP = Computer Position; WA = Work Area; ABP = Awkward Body Posture; HBP = Head and Body Posture; IBP = Incorrect Body Posture; BH = Bad Habits; DA = Decision Authority; CSD = Creative Skill Development; SD = Skill Discretion; SA = Skills and Abilities; DM = Decision Making; WP = Work Pressure; AU = Autonomous Management; ANC = Alternative, no computer; TC = Task Complexity; TP = Time pressure; A = Autonomy; IWC = Impact of Working Conditions; BQ = Break Quality; WF = Work Flow; TM = Time Management; WO = Work Overload; WB - Work Breaks; VW = Variation in Work.

lity was α = 0.88 [0.85; 0.91 95%CI] for the "body and head posture" factor (Figure 3) and α = 0.65 [0.57; 0.72 95%CI] for the "awkward posture" factor (Figure 3). In the second domain, "breaks", divided into two factors, the average reliability was α = 0.80 [0.65; 0.89 95%CI] for the "autonomy" factor and α = 0.80 [0.77; 0.82 95%CI] for the "quality of breaks" factor (Figure 4). In the third domain, "social support", divided into two factors, the results of the meta-analysis for average reliability were α = 0.84 [0.69; 0.92 95%CI] for the "social support" factor and α = 0.68 [0.58; 0.76 95%CI] for the "workflow" factor (Figure 5).

In the fourth domain, "work control", divided into two factors, the results of the meta-analysis for average reliability were α = 0.73 [0.65; 0.80 95%CI] for the "decision-making authority" factor and α = 0.78 [0.71; 0.84 95%CI] for the "ability criterion" factor (Figure 6). In the fifth domain, "work demands", divided into two factors, the average reliability was α = 0.77 [0.54; 0.90 95%CI] for the "task complexity" factor and α = 0.71 [0.46; 0.86 95%CI] for the "work pressure" factor (Figure 7). In the sixth and final domain presented, "work environment", divided into two factors, the average reliability was α = 0.52 [0.47; 0.57 95%CI] for the "office equipment" factor and α = 0.61 [0.43; 0.74 95%CI] for the "computer position" factor (Figure 8).

The global analysis of MUEQ factors' results in relation to the internal consistency index represented by Cronbach's alpha ranged from 0.52 to 0.84, with a 95% confidence interval of 0.43 to 0.92 (95% CI) and with some factors showing internal consistency values classified as "Good", "Moderate" and "Fair", while others did not reach the minimum acceptable values (Figures 3-8).

Only 25% of the factors achieved Cronbach's alpha values classified as "Good", which indicates satisfactory internal consistency for these specific MUEQ factors (Figures 4-5). A further 25% of the factors achieved Cronbach's alpha values classified as "Moderate" (Figures 3; 6-7), and 16.6% achieved values classified as "Fair" (Figure 6-7), suggesting that the internal consistency of these factors may be questioned. However, the greatest concern lies with the four factors that did not show acceptable minimum values for Cronbach's alpha (Figures 3, 5, 8). With regard to the inconsistency results, assessed by I^2, a substantial variation of 77% to 98% can be observed, with 83.33% of the MUEQ factors showing values higher than 75%, indicating considerable substantial heterogeneity.

Analysis of the certainty of evidence

All the studies presented very low certainty of evidence, showing adequate results for the indirect evidence, imprecision, and inconsistency items, with the exception of two factors ("impact of working conditions" and "working environment"), which presented serious inconsistency. However, the studies did not present adequate results for the risk of bias items (table 4). The studies analyzed by GRADE were the same as those included in the meta-analysis, with the exception of two factors ("impact of working conditions" and "working environment"), which were not meta-analyzed due to the fact that they were present in one study only. The agreement between the two evaluators as to the certainty of evidence was 44.65%[29].

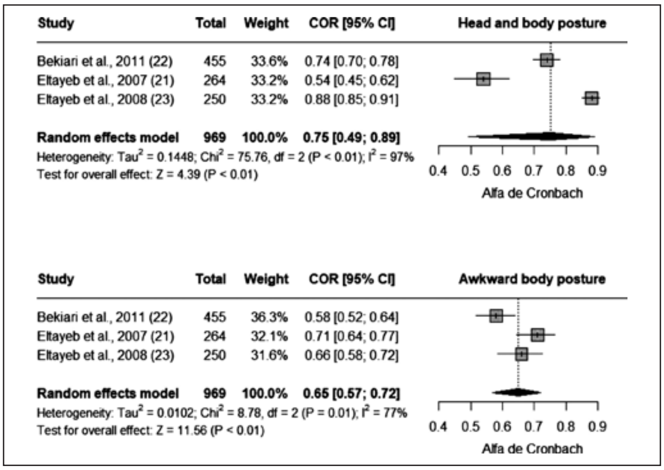

Figure 3. Comparison of reliability evidence for "head and body posture" and "awkward body posture" factors in the "body posture" domain of MUEQ studies.

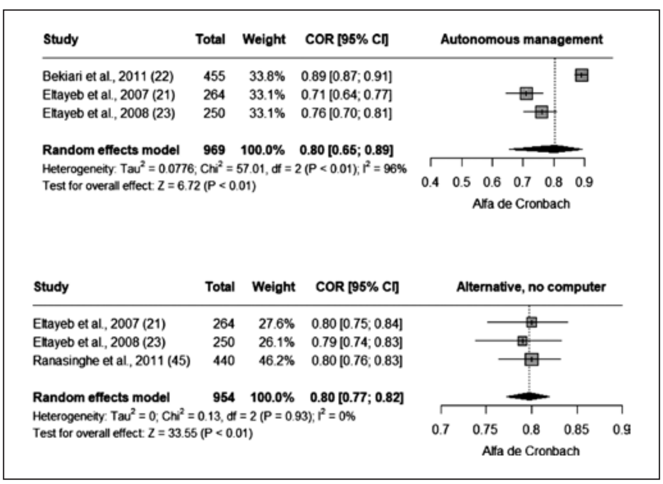

Figure 4. Comparison of reliability evidence for "autonomy" ("autonomous management") and "break quality" ("alternative", "no computer") factors in the "break time" domain of MUEQ studies.

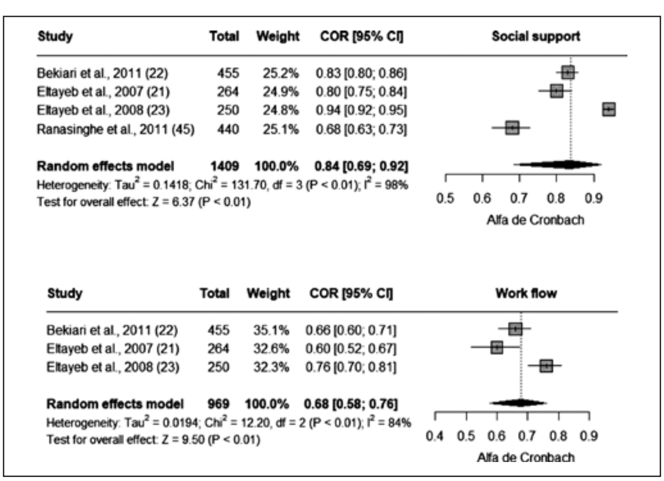

Figure 5. Comparison of reliability evidence for "social support" and "work flow" factors in the "social support" domain of MUEQ studies.

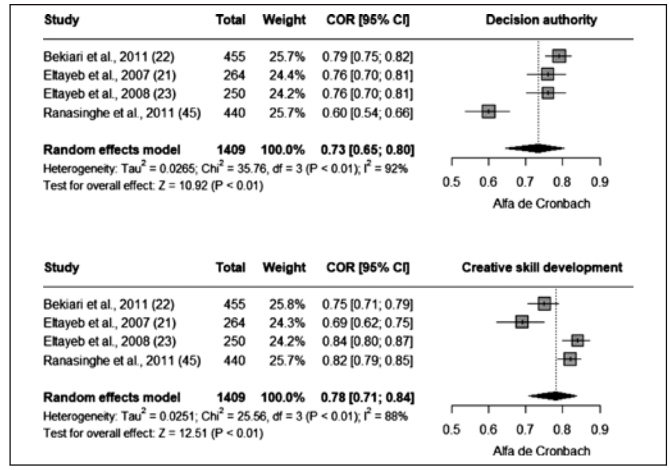

Figure 6. Comparison of reliability evidence for "decision authority" and "creative skill development" factors in the "work control" domain of MUEQ studies.

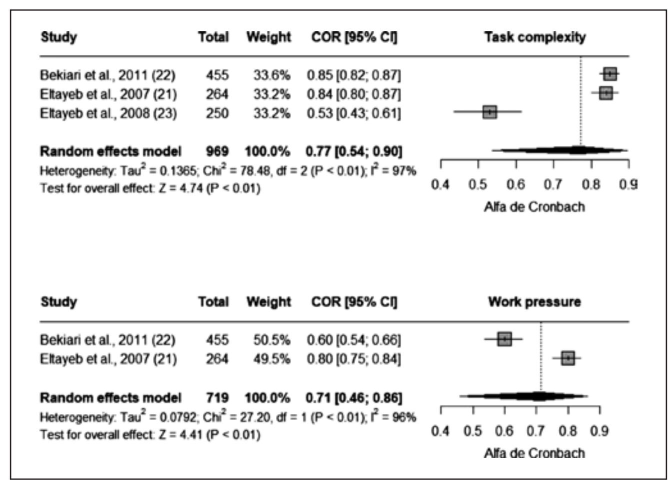

Figure 7. Comparison of reliability evidence for the factors "task complexity" and "work pressure" in the "work demand" domain of MUEQ studies.

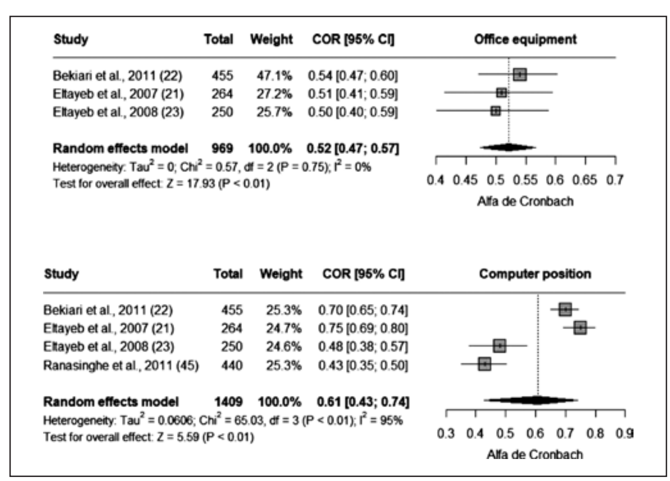

Figure 8. Comparison of the reliability evidence for "office equipment" and "computer position" factors in the "work station" domain of MUEQ studies.

Table 4. Assessment of the certainty of evidence using the Grading of Recommendations Assessment, Development, and Evaluation tool.

Assessing the certainty of the evidence						Effect	Level of certainty of the evidence
Nº of studies	Study design	Risk of bias	Inconsis-tency	Indirect evidence	Imprecision	COR (95% CI)	
Head and body posture							
3	Observational, cross-sectional study	Severe[a]	Not severe	Not severe	Not severe	0.88 [0.85; 0.91]	⊕◯◯◯ Very low
Awkward body posture							
3	Observational, cross-sectional study	Severe[a]	Not severe	Not severe	Not severe	0.65 [0.57; 0.72]	⊕◯◯◯ Very low
Autonomous management							
3	Observational, cross-sectional study	Severe[a]	Not severe	Not severe	Not severe	0.80 [0.65; 0.89]	⊕◯◯◯ Very low
Quality of breaks (alternative, no computer)							
3	Observational, cross-sectional study	Severe[a]	Not severe	Not severe	Not severe	0.80 [0.77; 0.82]	⊕◯◯◯ Very low
Social support							
4	Observational, cross-sectional study	Severe[a]	Not severe	Not severe	Not severe	0.84 [0.69; 0.92]	⊕◯◯◯ Very low
Work flow							
7*	Observational, cross-sectional study	Severe[a]	Not severe	Not severe	Not severe	0.68 [0.58; 0.76]	⊕◯◯◯ Very low
Decision authority							
4	Observational, cross-sectional study	Severe[a]	Not severe	Not severe	Not severe	0.73 [0.65; 0.80]	⊕◯◯◯ Very low
Creative skill development							
4	Observational, cross-sectional study	Severe[a]	Not severe	Not severe	Not severe	0.78 [0.71; 0.84]	⊕◯◯◯ Very low
Task complexity							
3	Observational, cross-sectional study	Severe[a]	Not severe	Not severe	Not severe	0.77 [0.54; 0.90]	⊕◯◯◯ Very low
Work pressure							
2	Observational, cross-sectional study	Severe[a]	Not severe	Not severe	Not severe	0.71 [0.46; 0.86]	⊕◯◯◯ Very low
Office equipment							
2	Observational, cross-sectional study	Severe[a]	Not severe	Not severe	Not severe	0.52 [0.47; 0.57]	⊕◯◯◯ Very low
Computer position							
2	Observational, cross-sectional study	Severe[a]	Not severe	Not severe	Not severe	0.61 [0.43; 0.74]	⊕◯◯◯ Very low
Work environment							
1	Observational, cross-sectional study	Severe[a]	Severe[b]	Not severe	Not severe	Not reported	⊕◯◯◯ Very low
Impact of working conditions							
1	Observational, cross-sectional study	Severe[a]	Severe[b]	Not severe	Not severe	Not reported	⊕◯◯◯ Very low

*The same studies were analyzed, with their respective subdivisions in the forest plots. a. High attrition bias and low reporting of the adequacy of the time interval between repeated measurements. b. Lack of compiled (meta-analyzed) studies.

DISCUSSION

This is the first systematic review to evaluate the MUEQ psychometric properties. The aim of this questionnaire is to assess musculoskeletal pain in computer users, together with the associated physical and psychosocial risk factors[19-23,45]. This review provided evidence of its applicability, methodological quality, evidence of psychometric properties and certainty of evidence based on the studies included. From the results obtained, it can be stated that the questionnaire did not present levels considered acceptable of evidence based on the internal structure (Figures 3-8).

In general, among the results of the internal consistency of the six domains of MUEQ - "body posture"; "breaks"; "social support"; "work control"; "work demand"; "workplace", assessed by Cronbach's alpha and grouped by the meta-analysis, some were considered acceptable (classified as "moderate" to "good") (Figures 3-7). However, other studies obtained values below the acceptable limit (Figures 6 and 7). It is important to note that these estimates may have been inflated by the heterogeneity between the studies, which ranged from 77% ("substantial") to 98% ("considerable"), but were not explained due to the inclusion of few studies.

Regarding to the reliability results of the six domains of MUEQ, only two studies[19,20], evaluated by ICC, were classified as acceptable (ICC > 0.70 - table 3)[25]. Internal consistency measures are used to indicate the amount of measurement error. Thus, the results of this study corroborate the low amount of measurement errors[25]. The evidence based on internal structure, of an incremental nature (TLI/CFI, > 0.95) and of an absolute nature (RMSEA, < 0.06), was analyzed by two studies[19,20], which, however, did not present results considered acceptable by the scientific literature[29,42-44].

No study has evaluated content validity, evidence related to external variables or the item response process[19-23,45]. This evidence is essential to ensure the clarity and coherence of the items in psychometric instruments, and the lack of it can compromise the quality of the items and the understanding of the instruments[25]. Therefore, although MUEQ is a questionnaire widely used to assess musculoskeletal pain in computer users, it is important to carry out other forms of validation, such as content validity and evidence related to external variables and the item response process, to ensure its validity and reliability in clinical practice. These validations are essential for implementing effective preventive and protective measures for workers' health and well-being.

Clinical applicability

This study provided results indicating that MUEQ has psychometric properties in certain countries, such as Greece, the Netherlands, Sudan, Iran, Sri Lanka and Brazil. However, it is crucial to carry out a more thorough analysis before applying it in other countries. In addition to the well-known advantages of questionnaires, MUEQ plays a significant role in clinical applicability, as its results are associated with work environment and musculoskeletal pain.

The findings of this research highlight the importance of carrying out other forms of MUEQ validation, using content validity, evidence based on internal structure, reliability and evidence based on relationships with external variables[25], for its use in clinical contexts, especially with a view to identifying risk factors related to musculoskeletal injuries and psychosocial aspects in workers who spend long hours using computers. This validation is necessary for the implementation of preventive and protective measures aimed at the health and well-being of these workers.

In Brazil, there are some valid instruments that assess aspects of work, such as the Quick Exposure Check[46], the Job Factors Questionnaire[47] and the Nordic Musculoskeletal Questionnaire[48]. However, among the instruments available, MUEQ-Br stands out as the only tool that comprehensively assesses the physical and biopsychosocial aspects related to CANS in Brazilian workers who use computers.

Limitations and strengths of the study

The strength of this study was the systematic approach, using a sensitive and broad search protocol in 14 electronic databases. Additionally, rigorous control measures were implemented at every stages of the process, and eligibility criteria were established which did not restrict inclusion by study type, population, language, age, gender and publication date. This expansive and inclusive approach allowed for a more comprehensive analysis of the available data and contributed to the robustness and validity of the results obtained. While this study presented a synthesis of the MUEQ's internal consistency results, it was not possible to explore publication bias and the factors that may affect the heterogeneity of the results, due to the limited number of studies included in the meta-analysis (≤ 4 studies). Moreover, some studies did not provide detailed information on the internal consistency divided by factors, which compromised the collective analysis of the studies included in this review.

Regarding the results of the evidence based on the MUEQ's internal structure and reliability, which did not demonstrate acceptable psychometric properties, it is recommended that new studies comprehensively evaluate the psychometric psychometric properties, including evidence based on internal structure, reliability, content and the item response process, as well as evidence based on relationships with external variables[25]. These detailed and comprehensive evaluations are essential to better comprehend the validity and reliability of MUEQ in different contexts and populations, ensuring that this questionnaire is an effective tool for assessing musculoskeletal pain in computer users.

CONCLUSION

This study presented evidence of the MUEQ's psychometric properties, but the analysis carried out highlighted the lack of detail in the methodological procedures, especially in relation to content validity, evidence of external variables and sample description. The evidence based on the internal structure and reliability of MUEQ did not reach acceptable levels to guarantee its adequacy and accuracy. For a more complete understanding of MUEQ's psychometric properties, future research with greater methodological rigor, diversified samples, and robust techniques is recommended. This would ensure its reliable application in academic and clinical contexts.

AUTHORS' CONTRIBUTIONS

Jânio Luiz Correia Júnior
Data Collection, Conceptualization, Resource Management, Project Management, Writing - Preparation of the Original, Writing - Review and Editing
Maria Fernanda da Silva Torres
Data Collection, Writing - Preparation of the Original, Writing - Review and Editing
Sthefanny Pontes Sampaio Costa
Data Collection, Writing - Preparation of the Original, Writing - Review and Editing
Hermê Fellipo Bordoni Caldeira
Data Collection, Writing - Preparation of the Original, Writing - Review and Editing
Marina Pereira Gonçalves
Conceptualization, Writing - Preparation of the Original, Writing - Review and Editing
Ricardo Freitas-Dias
Statistical Analysis, Conceptualization, Project Management, Writing - Preparation of the Original, Writing - Review and Editing, Supervision

REFERENCES

1. Bongers PM, Ijmker S, van den Heuvel S, Blatter BM. Epidemiology of work related neck and upper limb problems: Psychosocial and personal risk factors (Part I) and effective interventions from a bio behavioural perspective (Part II). J Occup Rehabil. 2006;16(3):272-95.
2. Hutting N, Staal J, Heerkens YF, Engels JA, Sanden MWN der. A self-management program for employees with complaints of the arm, neck, or shoulder (CANS): study protocol for a randomized controlled trial. Trials. 2013;14(1):258.

3. Huisstede BMA, Miedema HS, Verhagen AP, Koes BW, Verhaar JAN. Multidisciplinary consensus on the terminology and classification of complaints of the arm, neck and/or shoulder. Occup Environ Med. 2007;64(5):313-9.
4. Kashif M, Anwar M, Noor H, Iram H, Hassan HMJ. Prevalence of musculoskeletal complaints of arm, neck and shoulder and associated risk factors in computer office workers. Phys Medizin, Rehabil Kurortmedizin. 2020;30(05):299-305.
5. AlOmar RS, AlShamlan NA, Alawashiz S, Badawood Y, Ghwoidi BA, Abugad H. Musculoskeletal symptoms and their associated risk factors among Saudi office workers: a cross-sectional study. BMC Musculoskelet Disord. 2021;22(1):763.
6. European Commission. Health and safety at work in Europe (1999–2007) - A statistical portrait. In: 2010th ed. Luxembourg: Eurostat; 2010. 41-70p.
7. Punnett L, Wegman DH. Work-related musculoskeletal disorders: the epidemiologic evidence and the debate. J Electromyogr Kinesiol. 2004;14(1):13-23.
8. Sonne M, Villalta DL, Andrews DM. Development and evaluation of an office ergonomic risk checklist: ROSA – Rapid office strain assessment. Appl Ergon. 2012;43(1):98-108.
9. Gerding T, Syck M, Daniel D, Naylor J, Kotowski SE, Gillespie GL, Freeman AM, Huston TR, Davis KG. An assessment of ergonomic issues in the home offices of university employees sent home due to the COVID-19 pandemic. Work. 2021;68(4):981-92.
10. Global Workplace Analytics. Global Work-from-Home Experience Survey [Internet]. 2023 [cited 2023 Aug 7]. Available from: https://globalworkplaceanalytics.com/work-at-home-after-covid-19-our-forecast (2023).
11. Fernandes T, Salgueiro ACF. Dores musculoesqueléticas e ergonomia em tempos de home office. Res Soc Dev. 2022;11(13):e414111335743-e414111335743.
12. Asundi K, Odell D, Luce A, Dennerlein JT. Notebook computer use on a desk, lap and lap support: Effects on posture, performance and comfort. Ergonomics. 2010;53(1):74-82.
13. Davis KG, Kotowski SE, Daniel D, Gerding T, Naylor J, Syck M. The home office: ergonomic lessons from the “new normal”. Ergon Des Q Hum Factors Appl. 2020;28(4):4-10.
14. Werth AJ, Babski-Reeves K. Assessing posture while typing on portable computing devices in traditional work environments and at home. Proc Hum Factors Ergon Soc Annu Meet. 2012;56(1):1258-62.
15. Rodrigues MS, Leite RDV, Lelis CM, Chaves TC. Differences in ergonomic and workstation factors between computer office workers with and without reported musculoskeletal pain. Work. 2017;57(4):563-72.
16. Szeto GPY, Straker LM, O’Sullivan PB. A comparison of symptomatic and asymptomatic office workers performing monotonous keyboard work—2: Neck and shoulder kinematics. Man Ther. 2005;10(4):281-91.
17. Nakatsuka K, Tsuboi Y, Okumura M, Murata S, Isa T, Kawaharada R, Matsuda N, Uchida K, Horibe K, Kogaki M, Ono R. Association between comprehensive workstation and neck and upper-limb pain among office worker. J Occup Health. 2021;63(1):e12194.
18. Marques NR, Hallal CZ, Gonçalves M. Características biomecânicas, ergonômicas e clínicas da postura sentada: uma revisão. Fisioter Pesqui. 2010;17(3):270-6.
19. Ghasemi M, Kamalikhah T, Salesi M, Rahmati F. Evaluation of psychometric properties of the maastricht upper extremity questionnaire (MUEQ) in iranian computer users. J Educ Health Promot. 2021;10:245.
20. Turci AM, Bevilaqua-Grossi D, Pinheiro CF, Bragatto MM, Chaves TC. The Brazilian Portuguese version of the revised Maastricht Upper Extremity Questionnaire (MUEQ-Br revised): translation, cross-cultural adaptation, reliability, and structural validation. BMC Musculoskelet Disord. 2015;16(1):41.
21. Eltayeb S, Staal JB, Kennes J, Lamberts PH, de Bie RA. Prevalence of complaints of arm, neck and shoulder among computer office workers and psychometric evaluation of a risk factor questionnaire. BMC Musculoskelet Disord. 2007;8(1):68.
22. Bekiari EI, Lyrakos GN, Damigos D, Mavreas V, Chanopoulos K, Dimoliatis IDK. A validation study and psychometrical evaluation of the Maastricht Upper Extremity Questionnaire (MUEQ) for the Greek-speaking population. J Musculoskelet Neuronal Interact. 2011;11(1):52-76.
23. Eltayeb SM, Staal JB, Hassan AA, Awad SS, de Bie RA. Complaints of the arm, neck and shoulder among computer office workers in Sudan: a prevalence study with validation of an Arabic risk factors questionnaire. Environ Heal. 2008;7(1):33.
24. Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, Bouter LM, de Vet HCW, Mokkink LB. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res. 2018;27(5):1159-70.
25. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for Educational and Psychological Testing. Lanham, MD: American Educational Research Association; 2014. 1-230p.
26. Chien PF, Khan KS, Siassakos D. Registration of systematic reviews: PROSPERO. BJOG An Int J Obstet Gynaecol. 2012;119(8):903-5.
27. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Moher D. Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. J Clin Epidemiol. 2021;134:103-12.
28. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022) [Internet]. 2nd ed. Chichester (UK): Cochrane; 2022 [cited 2023 Aug 7]. p. 1-728. Available from: www.training.cochrane.org/handbook.
29. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, Terwee CB. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1147-57.
30. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. J Clin Epidemiol. 2016;75:40-6.
31. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210.
32. Hupe M. EndNote X9. J Electron Resour Med Libr. 2019;16(3-4):117-9.
33. Brink Y, Louw QA. Clinical instruments: reliability and validity critical appraisal. J Eval Clin Pract. 2012;18(6):1126-32.
34. GRADEpro GDT: GRADEpro Guideline Development Tool [Software]. McMaster University and Evidence Prime, 2022. Available from gradepro.org
35. Sitoe SA, Codonhato R, Both J, Fiorese L. Educação física e satisfação das necessidades psicológicas básicas em escolares de Maputo-Moçambique. Pensar a Prática. 2019;22:1-11.
36. Souza GS, Duarte MFS. Estágios de mudança de comportamento relacionados à atividade física em adolescentes. Rev Bras Med Esp. 2005;11(2):104-8.
37. Loprinzi PD, Cardinal BJ, Loprinzi KL, Lee H. Benefits and environmental determinants of physical activity in children and adolescents. Obes Facts. 2012;5(4):597-610.
38. De Lara PZM. Fairness, teachers’ non-task behavior and alumni satisfaction: the influence of group commitment. J Educ Adm. 2008;46(4):514-38.
39. Higgins JPT, Thompson SG, Deeks JJ, Altman GD. Measuring inconsistency in meta-analyses Julian. BMJ. 2003;327:132-35.
40. Santos GM, Strathdee SA, El-Bassel N, Patel P, Subramanian D, Horyniak D, Cook RR, McCullagh C, Marotta P, Choksi F, Kang B, Allen I, Shoptaw S. Psychometric properties of measures of substance use: a systematic review and meta-analysis of reliability, validity and diagnostic test accuracy. BMC Med Res Methodol. 2020;20(1):1-22.
41. Ponterotto JG. An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures. Percept Mot Skills. 2007;5(7):997.
42. Terwee CB, Bot SDM, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34-42.
43. Prinsen CAC, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, Williamson PR, Terwee CB. How to select outcome measurement instruments for outcomes included in a ‘Core Outcome Set’ – a practical guideline. Trials. 2016;17(1):1-10.
44. Hair Jr JF, Hult GTM, Ringle CM, Sarstedt M. A primer on partial least squares structural equation modeling (PLS-SEM). 1st ed. Thousand Oaks (CA): Sage publications; 2021. 1-328 p.
45. Ranasinghe P, Perera YS, Lamabadusuriya DA, Kulatunga S, Jayawardana N, Rajapakse S, Katulanda P. Work-related complaints of arm, neck and shoulder among computer office workers in an Asian country: prevalence and validation of a risk-factor questionnaire. BMC Musculoskelet Disord. 2011;12(1):68.
46. Comper MLC, Costa LOP, Padula RS. Quick Exposure Check (QEC): a crosscultural adaptation into Brazilian-Portuguese. Work. 2012;41(Supp1):2056-9.
47. Karasek R, Brisson C, Kawakami N, Houtman I, Bongers P, Amick B. The Job Content Questionnaire (JCQ): An instrument for internationally comparative assessments of psychosocial job characteristics. J Occup Health Psychol. 1998;3(4):322-55.
48. de Barros ENC, Alexandre NMC. Cross-cultural adaptation of the Nordic musculoskeletal questionnaire. Int Nurs Rev. 2003;50(2):101-8.