Accepted Manuscript

Submission Date: 2021-10-13
Accepted Date: 2022-07-05
Publication Date: 2022-07-11

Journal of the American Academy of Audiology

Hyperacusis and Misophonia: A Systematic Review of Psychometric Measures

Fatma B Kula, Mark Cropley, Hashir Aazh.

Affiliations below.

DOI: 10.1055/a-1896-5032

Please cite this article as: Kula F B, Cropley M, Aazh H. Hyperacusis and Misophonia: A Systematic Review of Psychometric Measures. Journal of the American Academy of Audiology 2022. doi: 10.1055/a-1896-5032

Conflict of Interest: The authors declare that they have no conflict of interest.

Abstract:
Abstract
Background: Hyperacusis can be defined as intolerance of certain everyday sounds, which are perceived as too loud or uncomfortable and which cause significant distress and impairment in the individual‘s day-to-day activities. Misophonia is defined as a high magnitude of emotional and behavioural reaction to certain sounds produced by human beings, such as eating sounds and breathing sounds. Several psychometric instruments have been developed to assess symptoms and the impact of hyperacusis and misophonia; however, to the author‘s knowledge, no study has evaluated and compared the methodological quality of the studies on psychometric properties of the existing instruments.
Purpose: To systematically review the research studies assessing the psychometric properties of the instruments used for hyperacusis and misophonia and assess the quality and appropriateness of the methodologies used.
Research Design: Systematic review.
Data Collection and Analysis: A systematic literature search was performed using five electronic literature databases (PubMed, Scopus, PsycINFO, Google Scholar and Web of Science). Studies were included if they were written in English and reported information about the psychometric properties of instruments measuring hyperacusis or misophonia symptoms or their impact. The quality of the studies and that of the psychometric instruments were evaluated using the consensus-based standards for the selection of health measurement instruments (COSMIN) tool.
Results: The title and abstracts of 916 articles were screened and 39 articles were selected for full-text evaluation, with 14 articles meeting the inclusion criteria. From these 14 articles, eight different instruments (5 for hyperacusis and 3 for misophonia) were identified and reviewed comprising: (1) Hyperacusis Questionnaire (HQ), (2) Inventory of Hyperacusis Symptoms (IHS), (3) questionnaire on hypersensitivity to sound (GUF), (4) Hyperacusis Handicap Questionnaire (HHQ), (5) Short Hyperacusis Questionnaire, (6) Amsterdam Misophonia Scale (A-MISO-S), (7) MisoQuest, and (8) the Misophonia Questionnaire (MQ).
Conclusion: None of the papers reviewed reported all the information required to meet the COSMIN standards. The studies‘ methodological quality varied between ‚very good‘ and ‚inadequate‘ depending on their grade on the COSMIN tool. There is a need for further research on the psychometric properties of the instruments included in this review.

Corresponding Author:

Affiliations:
Mark Cropley, University of Surrey, University of Surrey, Guildford, United Kingdom of Great Britain and Northern Ireland
Hashir Aazh, The Royal Surrey County Hospital NHS Foundation Trust, Audiology, Surrey, United Kingdom of Great Britain and Northern Ireland

Appendix-1 Definitions and criteria for good measurement properties by COSMIN guidance.

(The table is reproduced from the COSMIN guidance, using their definitions and criteria.)

Measurement Property	**Definition**	**Rating**	**Criteria**
Structural Validity	The degree to which the scores of a PROM are an adequate reflection of the dimensionality of the construct to be measured	+	**CTT:** CFA: CFI or TLI or comparable measure >0.95 OR RMSEA <0.06 OR SRMR <0.082 **IRT/Rasch**: No violation of unidimensionality3: CFI or TLI or comparable measure >0.95 OR RMSEA <0.06 OR SRMR <0.08 *AND* no violation of local independence: residual correlations among the items after controlling for the dominant factor < 0.20 OR Q3's < 0.37 *AND* no violation of monotonicity: adequate looking graphs OR item scalability >0.30

			AND adequate model fit: IRT: χ2 >0.01 Rasch: infit and outfit mean squares ≥ 0.5 and ≤ 1.5 OR standardized values > -2 and <2
		?	CTT: Not all information for '+' reported IRT/Rasch: Model fit not reported
		-	Criteria for '+' not met
Internal Consistency	The degree of the interrelatedness among the items	+	At least low evidence for sufficient structural validity AND Cronbach's alpha(s) ≥ 0.70 for each unidimensional scale or Subscale
		?	Criteria for "At least low evidence4 for sufficient structural validity" not met
		-	At least low evidence4 for sufficient structural validity AND Cronbach's alpha(s) < 0.70 for each unidimensional scale or subscale
Reliability	The proportion of	+	ICC or weighted Kappa ≥ 0.70

	the total variance in the measurements which is due to 'true'† differences between patients	**?**	ICC or weighted Kappa not reported
		-	ICC or weighted Kappa < 0.70
Measurement Error	The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured	+	SDC or LoA < MIC5
		?	MIC not defined
		-	SDC or LoA > MIC5
Hypotheses testing for construct validity	The degree to which the scores of a PROM are consistent with hypotheses *(for instance with regard to internal relationships, relationships to*	+	The result is in accordance with the hypothesis
		?	No hypothesis defined (by the review team)
		-	The result is not in accordance with the hypothesis

	scores of other instruments, or differences between relevant groups) based on the assumption		
Cross-cultural validity\measurement invariance	The degree to which the performance of the items on a translated or culturally adapted PROM are an adequate reflection of the performance of the items of the original version of the PROM	+	No important differences found between group factors (such as age, gender, language) in multiple group factor analysis OR no important DIF for group factors (McFadden's R2 < 0.02)
		?	o multiple group factor analysis OR DIF analysis performed
		-	Important differences between group factors OR DIF was found
Criterion validity	The degree to	+	Correlation with gold standard ≥

	which the scores of a PROM are an adequate reflection of a 'gold standard'		0.70 OR AUC ≥ 0.70
		?	Not all information for '+' reported
		-	Correlation with gold standard < 0.70 OR AUC < 0.70
Responsiveness	The ability of a PROM to detect change over time in the construct to be measured	**+**	The result is in accordance with the hypothesis7 OR AUC ≥ 0.70
		?	No hypothesis defined (by the review team)
		-	The result is not in accordance with the hypothesis7 OR AUC < 0.70

Hyperacusis and Misophonia: A Systematic Review of Psychometric Measures

Fatma B. Kula,[1] Mark Cropley,[1] Hashir Aazh[2,3,4]

[1]Department of Psychology, The University of Surrey, Guildford, United Kingdom

[2]Audiology Department, Royal Surrey NHS Foundation Trust, United Kingdom

[3]Department of Communication Sciences & Disorders, Florida Atlantic University, United States

[4]Faculty of Engineering and Physical Sciences, University of Surrey, United Kingdom

Address for correspondence: Fatma B. Kula, Department of Psychology, Faculty of Health and Medical Sciences, The University of Surrey, Guildford, GU2 7XH, United Kingdom (e-mail: f.kula@surrey.ac.uk).

Abstract

Background Hyperacusis can be defined as an intolerance of certain everyday sounds, which are perceived as too loud or uncomfortable and which cause significant distress and impairment in the individual's day-to-day activities. Misophonia is defined as a high magnitude of emotional and behavioural reaction to certain sounds produced by human beings, such as eating sounds and breathing sounds. Several psychometric instruments have been developed to assess symptoms and the impact of hyperacusis and misophonia; however, to the author's knowledge, no study has evaluated and compared the methodological quality of the studies on psychometric properties of the existing instruments.

Purpose To systematically review the research studies assessing the psychometric properties of the instruments used for hyperacusis and misophonia and assess the quality and appropriateness of the methodologies used.

Research Design Systematic review.

Data Collection and Analysis A systematic literature search was performed using five electronic literature databases (PubMed, Scopus, PsycINFO, Google Scholar and Web of Science). Studies were included if they were written in English and reported information about the psychometric properties of instruments measuring hyperacusis or misophonia symptoms or their impact. The quality of the studies and that of the psychometric instruments were evaluated using the consensus-based standards for the selection of health measurement instruments (COSMIN) tool.

Results The title and abstracts of 916 articles were screened and 39 articles were selected for full-text evaluation, with 14 articles meeting the inclusion criteria. From these 14 articles, eight different instruments (5 for hyperacusis and 3 for misophonia) were identified and reviewed comprising: (1) Hyperacusis Questionnaire (HQ), (2) Inventory of Hyperacusis Symptoms (IHS), (3) questionnaire on hypersensitivity to sound (GUF), (4) Hyperacusis Handicap Questionnaire (HHQ), (5) Short Hyperacusis Questionnaire, (6) Amsterdam Misophonia Scale (A-MISO-S), (7) MisoQuest, and (8) the Misophonia Questionnaire (MQ).

Conclusion None of the papers reviewed reported all the information required to meet the COSMIN standards. The studies' methodological quality varied between 'very good' and 'inadequate' depending on their grade on the COSMIN tool. There is a need for further research on the psychometric properties of the instruments included in this review.

Keywords

hyperacusis

misophonia

psychometric instruments and properties

Hyperacusis can be defined as an intolerance of certain everyday sounds, which are perceived as too loud or uncomfortable and which cause significant distress and impairment in the individual's day-to-day activities (1). Other definitions of hyperacusis are largely in agreement with this definition, with some differences in details (2, 3). Tyler, Pienkowski (4) described four categories of hyperacusis comprising (1) loudness hyperacusis, (2) fear hyperacusis, (3) pain hyperacusis, and (4) annoyance hyperacusis. There is some overlap between annoyance hyperacusis and misophonia. A recent consensus study described that misophonia is characterized by the experience of unpleasant or distressing emotions when exposed to certain sounds generated by another individual, especially (but not exclusively) those produced by the human body (5). In misophonia, the specific pattern or meaning of the sound to the individual as opposed to its loudness seem to be the key contributing factor to the individual's reaction. Individuals with misophonia often experience suffering, distress or cannot tolerate sounds associated with oral functions (e.g., chewing, eating), nasal sounds (e.g., breathing and sniffing), as well as non-oral/nasal sounds (e.g., pen clicking, keyboard typing, clock ticking) (5).

Prevalence estimates range from 2% to 15.2% for hyperacusis (6, 7) and 6% to 49.1% for misophonia (8-10). It is likely that the discrepancy in prevalence reports is related the differences in study populations and the way that hyperacusis and misophonia were assessed and diagnosed.

Several psychometric instruments have been developed and applied in research and clinical practice to evaluate hyperacusis and/or misophonia. The methodologies used to design and evaluate the psychometric properties of these instruments (e.g., validity, reliability, sensitivity to change) are very diverse. The two important psychometric properties are reliability and validity which are essential for choosing suitable instruments for research or clinical purposes (11). Reliability comprises measures of internal consistency (degree of interrelatedness

among the items), test-retest reliability (consistency of scores obtained at different times), inter-rater reliability (consistency of scores obtained by different raters), and measurement error (the systematic and random error of a patient`s score that is not attributed to true changes in the construct to be measured) (12). Validity is defined as the extent to which an instrument measures what it claims to measure (13) and comprises (1) content validity (the degree to which the questions on the instrument represent the construct that it seeks to measure (14)), (2) construct validity (the extent to which the instrument validly measures the construct it purports to measure), (3) structural validity (the degree to which the scores of the instrument is an adequate reflection of the dimensionality of the construct to be measured), (4) hypotheses testing (the degree to which scores on the instrument are consistent with hypothesized relationship with other instruments), (5) cross-cultural validity (the degree to which items on a translated or adapted measure correspond to the performance of the original items), and (6) criterion validity (the degree to which scores correspond with a gold standard measure).

Studies assessing the psychometric properties of hyperacusis and misophonia instruments have used inconsistent methods. For example, participants in some studies were recruited from hospital patient referrals (15-17) while others from the general population or university students (9, 18). Some of these instruments are validated in languages other than English and the English versions although published, have not been validated (19). In addition, among the published papers there are some discrepancies regarding reporting of the important psychometric properties of the instruments they evaluated or developed. As the result of these discrepancies, it may not be clear to many audiologists whether the psychometric properties of the existing hyperacusis and misophonia questionnaires meet the standards required for them to be used effectively in research and/or clinical practice.

To develop a greater understanding of the reliability and validity of the existing hyperacusis and misophonia instruments, a systematic review of the literature can be extremely informative. Systematic reviews provide a summary of the strengths and weaknesses of the existing questionnaires, appraise the methodological quality of published studies, and discuss the differences between them (20); the results of which, can guide clinical practice and research.

Consensus-based standards for the selection of health measurement instruments (COSMIN) were developed to provide a comprehensive methodological tool for assessing the methodological quality of patient-reported outcome measures (21). COSMIN is an initiative of an international multidisciplinary team of researchers with a background in epidemiology, psychometric, medicine, qualitative research, and health care, who have expertise in the development and evolution of outcome measurement instruments. They developed the COSMIN risk of bias checklist that can be used in systematic reviews to assess the methodological quality of the studies included to the review (22, 23).

The present study aimed to systematically review the psychometric properties of the existing hyperacusis and misophonia questionnaires, summarise their strengths and weaknesses, and appraise the methodological quality of published studies against the criteria set by COSMIN tool (21, 23).

Methods

This systematic review was conducted in compliance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guideline (24) and it was registered with the PROSPERO database (https://www.crd.york.ac.uk/prospero; registration number: CRD42021235539).

Inclusion and Exclusion Criteria

The following inclusion criteria for articles were applied: i) published in English, ii) published in a peer-reviewed journal, iii) detailed the development or evaluated the measurement properties of instruments measuring hyperacusis or misophonia symptoms or their impact.

Articles were excluded if they: i) were not indexed in a recognised database, ii) did not report at least one psychometric property as defined by the COSMIN checklist (information relating to the psychometric properties are presented below), iii) were a review, personal/expert opinions and manuals, guidelines, or reported animal studies and any unpublished and incomplete studies.

Search Strategy

An initial search was conducted on 29th January 2021. A systematic search was presented in the following electronic databases: PubMed, Scopus, Web of Science, PsycINFO and Google Scholar. We entered a specific search term strategy in each search engine (see Table 1), searching in articles topics, titles, abstracts, and keywords. The database search was conducted without setting any limits in terms of the publication date of the studies. If possible, filters were applied to find related articles in the English language only and with humans only. The reference lists of any relevant articles were checked throughout the process to ensure that any related studies were not missed. Original searches were last updated on 29th April 2021. Prior to submitting the final revision of this paper on 17th June 2022, a quick search was conducted to double check if any new studies have emerged with regard to the questionnaires reviewed in this paper which did not show any new relevant studies.

[Table-1 about here.]

Selection Criteria

After the removal of duplicates articles, one reviewer (FK) screened titles and abstracts to identify eligible articles. The reference list of any relevant articles was also reviewed by the

first author. Then, two reviewers (FK and HA) screened the full text of the articles independently. The decision regarding the inclusion/exclusion of studies was made as a result of two reviewers' judgment about the selection of the articles and to verify inter-rater reliability of the full text screening, we calculated the Kohen's kappa value which was 0.65 indicating substantial agreement between the two reviewers (25). Any disagreements were resolved by the third reviewer (MC).

Data Extraction

Psychometric properties including content validity, structural validity, internal consistency, reliability, hypothesis testing for construct validity, cross-cultural validity, measurement error, criterion validity and responsiveness were extracted from studies in line with recommendations specified in the COSMIN guidelines (22). Other extracted information was country of origin, number of samples, gender, study population, and instrument-related factors such as construct measured, number of items, range of total score, and response options. All data were extracted by the first author in May 2021.

Evaluation of methodological quality

Two reviewers independently applied the COSMIN checklist for all included studies according to the recommended guidelines. Discrepancies of opinions were resolved by consensus between the two reviewers or, if the agreement was not achieved, disagreements were discussed and resolved through consultation with the third reviewer. No one graded any of their own papers.

The methodological quality of studies and their psychometric properties were assessed using the COSMIN checklist (21) as shown in Appendix-1. Based on this assessment we reported whether the above mentioned nine domains were assessed or adequately reported by various studies on psychometric properties of the hyperacusis and misophonia questionnaires.

Evaluation of Psychometric Properties of the Included Instruments

Each measurement property was rated by applying a four-point COSMIN risk of bias scale (4= 'very good', 3= 'adequate', 2= 'doubtful', 1= 'inadequate'). Consistent with COSMIN instructions, the overall quality rating for each measurement property was determined by taking the 'worst score counts' method (i.e., the lowest rating of any of the items in a given category) (23). For the next step, the result of individual studies on measurement properties was also evaluated against COSMIN 2018 updated criteria for good measurement properties (Appendix 1). The assessment resulted in rating for each questionnaire: sufficient (+), insufficient (–), or indeterminate (?). We used this information to create a table that demonstrates whether the key nine psychometric properties were reported for each questionnaire and if they meet the COSMIN criteria.

Inter-rater between the two reviewers was 82.0% (Kappa: 0.73) for the risk of bias ratings, and 84.5% (Kappa= 0.82) for the measurement properties, indicating substantial agreement between the two reviewers (25).

Results

Study Selection

A total of 1040 articles were identified through the initial search (Figure 1), and ten additional articles were identified through a review of citations. After removing duplicates, 916 articles were screened based on their title and abstract, and 39 articles were selected for full-text assessment. As a result of this full-text evaluation, 25 of the 39 articles were removed because they focused on the different constructs or did not report any psychometric property defined by the COSMIN checklist. In addition, one of the articles was not included this study because it was published in a predatory journal (26). Fourteen articles were included in this review, and from these 14 articles, eight different hyperacusis and

misophonia instruments were evaluated (9, 10, 15-19, 27-33). See the PRISMA flow diagram in Figure 1.

[Figure-1 about here.]

Study and Participant's Characteristics

Table 2 summarizes the characteristics of the included studies. The eligible studies were published from 2002 to 2020. Approximately 20% of the studies were conducted in the UK (9, 15, 28), 13% in the USA (10, 29), and Italy (17, 33). The rest of the studies were conducted in India (32), Belgium (30), Germany (16), Japan (31), Turkey (27), Poland (19) and France (18). The most used questionnaire reported was the Hyperacusis Questionnaire developed by Khalfa in 2002 (34). All questionnaires were developed to assess or diagnose hyperacusis or misophonia.

Table 2 also shows the participants' characteristics of the studies included to this review. Sample sizes for these studies ranged between 46 and 705 individuals from the general population and/or clinical population. Most studies included clinical populations (n=9) and two studies reported student populations, with the remainder utilising individuals from the general population (n=3).

Psychometric Instruments for Hyperacusis and Misophonia

Table 3 provides a summary of the description of the questionnaires including five hyperacusis and three misophonia instruments. All measures utilised the Likert type scales using 3 to 5 points scale. In addition, the structure of the included instruments varies greatly; five measures have three factor-structure (10, 16, 18, 32, 33), two measures have one factor-structure (9, 19), and the other measure has a five factor-structure (29).

[Table-2 about here.]

[Table-3 about here.]

The Methodological Quality of the Included Studies

Table 4 summarises methodological quality ratings for the 14 studies included to the review. All the studies reported more than one psychometric property. In addition, all studies reported internal consistency. Most studies described hypotheses testing for construct validity (11/14) and structural validity (8/14). Only a small number of studies included psychometric data on cross-cultural validity (2 studies), reliability (1 study), and measurement error (1 study). No information was retrieved on responsiveness and criterion validity in any study.

Psychometric Properties of the Included Instruments

Table 5 presents ratings for each psychometric instrument. The psychometric properties extracted from the studies were evaluated against the criteria for good psychometric properties on the COSMIN. None of the instruments could be fully evaluated over all nine psychometric properties as the necessary data was not always reported.

[Table 4 about here.]

[Table 5 about here.]

Discussion

The purpose of this systematic review was to evaluate the quality of psychometric properties of the current hyperacusis and misophonia instruments (until April 2021) using the COSMIN guidelines. The COSMIN checklist is a well-known tool and has been developed in conjunction with other existing guidelines for systematic reviews, such as the Cochrane Handbook for systematic reviews of intervention (35), the PRISMA statement (36) and the

Grading of Recommendations Assessment, Development and Emulation (GRADE) principles.

To our knowledge, this is the first systematic review to evaluate the measurement properties of instruments designed to measure hyperacusis or misophonia across a range of healthcare contexts and settings. This review identified eight measures (five for hyperacusis and three for misophonia) and 14 studies on the psychometric properties of these instruments. In general, the methodological quality of the included studies in this review varied between 'very good' and 'inadequate' across all psychometric properties based on the COSMIN tool. None of the identified instruments has reported all nine psychometric properties recommended by COSMIN.

The Methodological Quality of the Included Studies and Psychometric Properties of the Instruments

According to the COSMIN guideline (2018), content validity is considered an important measurement property of an instrument. However, none of the included articles reported using adequate methods to assess content validity. One explanation is that the constructs of hyperacusis and misophonia are not fully understood. Therefore, it was not possible to rate this following the COSMIN recommendation. However, all the questions within the various questionnaires appeared to have good content validity, as the questionnaires appeared to have included all the relevant items measuring the constructs in question. In addition, the questionnaires have been designed by clinicians and/or researchers working with patients who experience hyperacusis and/or misophonia so they were in a good position to create relevant items.

In terms of structural validity, six studies did not report any psychometric data. The rest of the studies methodological quality for structural validity varies between “very good” and “inadequate” according to COSMIN risk of bias checklist assessment. This mainly was due to

studies only reporting exploratory factor analysis (EFA) without confirmatory factor analysis (CFA). To test the factor structure, CFA or item response theory (IRT) analysis are preferred according to the COSMIN checklist (37).

None of the instruments reported on all three psychometric properties within the domain of reliability (reliability, internal consistency, and measurement error). Only one measurement instrument (MisoQuest) reported reliability with measuring interclass correlation coefficient (ICC), while all instruments reported internal consistency with receiving a very good score for study quality. Although measurement error is clinically important because as more error is introduced into the score, the lower reliability will be, only one article that tested MisoQuest (18), reported it.

None of the studies reported information on criterion validity. As there is no universally accepted gold standard to measure hyperacusis and misophonia, this feature of criterion validity could not be reported in this review. In addition, cross-cultural validity was reported in two studies (17, 27) with doubtful ratings. However, five studies (10, 16, 18, 19, 29) included in this review did not conduct cross-cultural validity because the measures were developed and validated in the original language.

Hypotheses testing for construct validity was reported in 11 studies (78.6%) with ratings of either very good, adequate, or doubtful. Only four studies (15, 16, 29, 30), reported both convergent and discriminant validity according to COSMIN risk of bias assessment. Except for these four studies, the remaining studies had limited evidence for construct validity.

Table 5 gives information about the results of each study on the different measurement properties, and it was rated as sufficient (+), insufficient (-), or indeterminate (?) following COSMIN criteria for good measurement properties. There is insufficient evidence within the included papers to making a judgment on their overall quality. Therefore, we chose not to

summarize the results and thus not to grade the total level of evidence per psychometric instruments.

There are some other hyperacusis questionnaires used in clinics and research, but these were not reviewed as their relevant publications did not provide the psychometric properties required by COSMIN. One questionnaire, for example, is the Multiple-Activity-Scale for Hyperacusis (MASH), by Dauman and Bouscau-Faure (38). The development procedure and metrics were not reported in this paper, so it was not possible to review its psychometric properties.

Several newly developed hyperacusis and misophonia questionnaires were not included in this review as the results of their psychometric properties were not published in a peer-review journal at the time our original literature search (1, 39, 40). Therefore, it was not possible to evaluate them with the COSMIN checklist in this review. Future reviews should assess the questionnaires which were published from April 2021.

In this systematic review, the populations within included studies varied, with both clinical and non-clinical samples. Clinicians desiring to select measures for clinical use should consider how generalizable the results are to the intended population, taking into account the populations from which the data in these studies were generated. For example, IHS (29) appears to be internally consistent in both clinical and non-clinical populations. The MisoQuest (19) is internally consistent for the clinical population. In terms of Hyperacusis Questionnaire originally developed by Khalfa (18) was internally consistent for just general population and Fackrell, Fearnley, Hoare and Sereda (28) investigated the validity and reliability of the HQ in a population who had tinnitus. They found the HQ to have high internal consistency (Cronbach`s alpha= 0.88) but confirmatory factor analysis revealed that the proposed three-factor, and an alternative one-factor structure were poor. Therefore, HQ does not seem to work well within a tinnitus population. Future studies should endeavour to

use clinical population of patients with hyperacusis or misophonia when developing questionnaires.

Implications for Future Research

Given the recent measures being adapted for use in other countries and languages, we believe that there is a need for appropriate and more testing for cross-cultural validity. Studies with different cultural groups should perform factor analyses for multiple groups and complete measurement invariance or DIF (differential item functioning) to give information on whether the measures are equivalent when used in different cultures/languages. For example, MisoQuest was developed in Polish, and validation has only been performed in a Polish population. Therefore, for future directions, validation and cross-cultural evaluation of MisoQuest are needed for other countries and different languages.

Regarding structural validity, future studies should perform factor analyses using CFA (confirmatory factor analysis) or IRT (item response theory) for seven instruments (HQ, IHS, HHQ, SHQ, GUF, MQ, A-MISO-S).

To gain a comprehensive picture of reliability, all elements of reliability should be assessed. Internal consistency has been assessed for all instruments, but future studies should assess test-retest, interrater, and intra-rater reliability for HQ, IHS, SHQ, GUF, MQ, MisoQuest, A-MISO-S and HHQ. Measurement error also need to be assessed for all eight instruments.

We also believe that future studies measuring content validity should state more explicitly how they evaluated content validity and follow COSMIN criteria when developing and reporting a new measure. This may include exploring the relevance, comprehensiveness, and comprehensibility of the measure among a sufficient sample of participants and professionals, which could lead to more credible evidence of its content validity.

All the available questionnaires regarding hyperacusis and misophonia are designed for adults and therefore may not be appropriate for use in children and adolescents. Therefore, future studies are needed for the development of new questionnaires in these specific groups.

Responsiveness is defined as the ability of the psychometric instrument to detect change over time in the construct measured (37). This review showed that responsiveness to change has not formally been tested for hyperacusis and misophonia questionnaires. However, HQ and A-MISO-S have been used in several interventional studies and appear to be sensitive to change (41-49) (scores have changed following treatment). This provides some evidence for responsiveness. More systematic studies are needed to further explore responsiveness to change and the cut off for meaningful or clinically significant change in hyperacusis and misophonia questionnaires.

Conclusion

This study systematically reviewed publications that evaluated the psychometric properties of eight hyperacusis and misophonia instruments using COSMIN guidelines (i.e., HQ, IHS, HHQ, SHQ, GUF, MQ, A-MISO-S and MisoQuest). Evidence concerning psychometric properties was limited and no single measure of hyperacusis and/or misophonia was found to meet all nine methodological quality standards according to the COSMIN guideline. There is a need for further research on the psychometric properties of the instruments included in this review.

Conflict of Interest

No conflict of interest has been declared by the authors.

Acknowledgements

The first author would like to thank the Ministry of National Education of the Republic of Turkey for supporting her graduate education.

REFERENCES

1. Aazh H, Hayes C, Moore BCJ, Danesh AA, Vitoratou S. Psychometric Evaluation of the Hyperacusis Impact Questionnaire (HIQ) and Sound Sensitivity Symptoms Questionnaire (SSSQ) Using a Clinical Population of Adult Patients with Tinnitus Alone or Combined with Hyperacusis. Journal of American Academy of Audiology. 2022;[Online ahead of print](Feb 23).

2. Adams B, Sereda M, Casey A, Byrom P, Stockdale D, Hoare DJ. A Delphi survey to determine a definition and description of hyperacusis by clinician consensus. International journal of audiology. 2020:1-7.

3. Aazh H, Moore BCJ, Lammaing K, Cropley M. Tinnitus and hyperacusis therapy in a UK National Health Service audiology department: Patients' evaluations of the effectiveness of treatments. International journal of audiology. 2016;55(9):514-22.

4. Tyler RS, Pienkowski M, Rojas Roncancio E, Jun HJ, Brozoski T, Dauman N, et al. A review of hyperacusis and future directions: part I. definitions and manifestations. American journal of audiology. 2014;23(4):402-19.

5. Swedo S, Baguley DM, Denys D, Dixon LJ, Erfanian M, Fioretti A, et al. A Consensus Definition of Misophonia: Using a Delphi Process to Reach Expert Agreement. Front Neurosci 2022;16(March 17):2021.04.05.21254951.

6. Andersson G, Lindvall N, Hursti T, Carlbring P, Andersson G. Hypersensitivity to sound (hyperacusis): a prevalence study conducted via the internet and post: Hipersensibilidad al sonido (hiperacusia): un estudio de prevalencia realizado por internet y por correo. International journal of audiology. 2002;41(8):545-54.

7. Smit AL, Stegeman I, Eikelboom RH, Baguley DM, Bennett RJ, Tegg-Quinn S, et al. Prevalence of Hyperacusis and Its Relation to Health: The Busselton Healthy Ageing Study. The Laryngoscope. 2021.

8. Zhou X, Wu MS, Storch EA. Misophonia symptoms among Chinese university students: Incidence, associated impairment, and clinical correlates. Journal of Obsessive-Compulsive and Related Disorders. 2017;14:7-12.

9. Naylor J, Caimino C, Scutt P, Hoare DJ, Baguley DM. The Prevalence and Severity of Misophonia in a UK Undergraduate Medical Student Population and Validation of the Amsterdam Misophonia Scale. The Psychiatric quarterly. 2021;92(2):609-19.

10. Wu M, Lewin AB, Murphy TK, Storch EA. Misophonia: incidence, phenomenology, and clinical correlates in an undergraduate student sample. J Clin Psychol. 2014;70(10):994-1007.

11. SÜRÜCÜ L, MaslakÇI A. VALIDITY AND RELIABILITY IN QUANTITATIVE RESEARCH. Business & Management Studies: An International Journal. 2020;8(3):2694-726.

12. Gillespie DFPBEOUP. Key concepts in measures. 2015.

13. Altheide DL, Johnson JM. Criteria for assessing interpretive validity in qualitative research. Handbook of qualitative research. Thousand Oaks, CA, US: Sage Publications, Inc; 1994. p. 485-99.

14. Creswell JW. Educational research: Planning, conducting, and evaluating quantitative: Prentice Hall Upper Saddle River, NJ; 2002.

15. Aazh H, Danesh AA, Moore BCJ. Internal Consistency and Convergent Validity of the Inventory of Hyperacusis Symptoms. Ear Hear. 2021;42(4):917-26.

16. Blasing L, Goebel G, Flotzinger U, Berthold A, Kroner-Herwig B. Hypersensitivity to sound in tinnitus patients: an analysis of a construct based on questionnaire and audiological data. International journal of audiology. 2010;49(7):518-26.

17. Fioretti A, Tortorella F, Masedu F, Valenti M, Fusetti M, Pavaci S. Validity of the Italian version of Khalfa's questionnaire on hyperacusis. Acta otorhinolaryngologica Italica : organo ufficiale della Societa italiana di otorinolaringologia e chirurgia cervico-facciale. 2015;35(2):110-5.

18. Khalfa S, Dubal S, Veuillet E, Perez-Diaz F, Jouvent R, Collet L. Psychometric normalization of a hyperacusis questionnaire. ORL J Otorhinolaryngol Relat Spec. 2002;64(6):436-42.

19. Siepsiak M, Śliwerski A, Łukasz Dragan W. Development and Psychometric Properties of MisoQuest-A New Self-Report Questionnaire for Misophonia. Int J Environ Res Public Health. 2020;17(5).

20. Green S. Systematic reviews and meta-analysis. Singapore medical journal. 2005;46(6):270-3; quiz 4.

21. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2018;27(5):1147-57.

22. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2010;19(4):539-49.

23. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res. 2012;21(4):651-7.

24. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg. 2010;8(5):336-41.

25. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med. 2005;37(5):360-3.

26. Rice DB, Skidmore B, Cobey KD. Dealing with predatory journal articles captured in systematic reviews. Systematic reviews. 2021;10(1):175.

27. Erinc M, Derinsu U. Turkish Adaptation of Khalfa Hyperacusis Questionnaire. Medeniyet medical journal. 2020;35(2):142-50.

28. Fackrell K, Fearnley C, Hoare DJ, Sereda M. Hyperacusis Questionnaire as a tool for measuring hypersensitivity to sound in a tinnitus research population. BioMed research international. 2015;2015:290425.

29. Greenberg B, Carlos M. Psychometric Properties and Factor Structure of a New Scale to Measure Hyperacusis: Introducing the Inventory of Hyperacusis Symptoms. Ear and hearing. 2018;39(5):1025-34.

30. Meeus OM, Spaepen M, Ridder DD, Heyning PH. Correlation between hyperacusis measurements in daily ENT practice. International journal of audiology. 2010;49(1):7-13.

31. Oishi N, Yamada H, Kanzaki S, Kurita A, Takiguchi Y, Yuge I, et al. Assessment of hyperacusis with a newly produced Japanese version of the Khalfa hyperacusis questionnaire. Acta Otolaryngol. 2017;137(9):957-61.

32. Prabhu P, Nagaraj MK. Development and validation of Hyperacusis Handicap Questionnaire in individuals with tinnitus associated with hyperacusis. J Otol. 2020;15(4):124-8.

33. Tortorella F, Pavaci S, Fioretti AB, Masedu F, Lauriello M, Eibenstein A. The short hyperacusis questionnaire: A tool for the identification and measurement of hyperacusis in the Italian tinnitus population. Audiol Res. 2017;7(2):182.

34. Khalfa S, Dubal S, Veuillet E, Perez-Diaz F, Jouvent R, Collet L. Psychometric normalization of a hyperacusis questionnaire. Journal for Oto-rhino-laryngology and its Related Specialties. 2002;64(6):436-42.

35. Johnston BC, Patrick DL, Devji T, Maxwell LJ, Bingham III CO, Beaton DE, et al. Patient-reported outcomes. Cochrane Handbook for Systematic Reviews of Interventions2019. p. 479-92.

36. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ : British Medical Journal. 2015;349:g7647.

37. Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2018;27(5):1171-9.

38. Dauman R, Bouscau-Faure F. Assessment and amelioration of hyperacusis in tinnitus patients. Acta oto-laryngologica. 2005;125(5):503-9.

39. Rosenthal MZ, Anand D, Cassiello-Robbins C, Williams ZJ, Guetta RE, Trumbull J, et al. Development and Initial Validation of the Duke Misophonia Questionnaire. Frontiers in psychology. 2021;12(4197).

40. Dibb B, Golding SE, Dozier TH. The development and validation of the Misophonia response scale. Journal of psychosomatic research. 2021;149:110587.

41. Nolan DR, Gupta R, Huber CG, Schneeberger AR. An Effective Treatment for Tinnitus and Hyperacusis Based on Cognitive Behavioral Therapy in an Inpatient Setting: A 10-Year Retrospective Outcome Analysis. Frontiers in psychiatry. 2020;11:25.

42. Aazh H, Bryant C, Moore BCJ. Patients' perspectives about the acceptability and effectiveness of audiologist-delivered cognitive behavioral therapy for tinnitus and/or hyperacusis rehabilitation. American journal of audiology. 2020;28(4):973-85.

43. Aazh H, Landgrebe M, Danesh A, Moore BCJ. Cognitive behavioral therapy for alleviating the distress caused by tinnitus, hyperacusis and misophonia: Current perspectives. Psychology Research and Behavior Management 2019;23(12):991-1002.

44. Aazh H, Moore BCJ. Proportion and characteristics of patients who were offered, enrolled in and completed audiologist-delivered cognitive behavioural therapy for tinnitus and hyperacusis rehabilitation in a specialist UK clinic. International journal of audiology. 2018;57(6):415-25.

45. Aazh H, Moore BCJ. Effectiveness of audiologist-delivered cognitive behavioral therapy for tinnitus and hyperacusis rehabilitation: outcomes for patients treated in routine practice. American journal of audiology. 2018;27(4):547-58.

46. Schroder AE, Vulink NC, van Loon AJ, Denys DA. Cognitive behavioral therapy is effective in misophonia: An open trial. Journal of affective disorders. 2017;217:289-94.

47. Juris L, Andersson G, Larsen HC, Ekselius L. Cognitive behaviour therapy for hyperacusis: A randomized controlled trial. Behaviour research and therapy. 2014;54c:30-7.

48. Beukes EW, Andersson G, Allen PM, Manchaiah V, Baguley DM. Effectiveness of Guided Internet-Based Cognitive Behavioral Therapy vs Face-to-Face Clinical Care for Treatment of Tinnitus: A Randomized Clinical Trial. JAMA Otolaryngology Head & Neck Surgery 2018;144(12):1126-33.

49. Beukes EW, Baguley DM, Allen PM, Manchaiah V, Andersson G. Audiologist-Guided Internet-Based Cognitive Behavior Therapy for Adults With Tinnitus in the United Kingdom: A Randomized Controlled Trial. Ear and hearing. 2018;39(3):423-33.

Table 1. Search term strategies applied in databases

	Construct	**Instrument**	**Psychometric Properties**
Search Terms	Hyperacus* OR Misophon* OR "Reduc* sound intolerance" OR "Noise Sensitivity" OR "Sound Intolerance" OR "Sensory intolerance" OR "Sound Sensitivity" OR "Selective Sound Sensitivity Syndrome" OR "Soft Sound Sensitivity Syndrome" OR "aversive sounds" OR "trigger sounds" OR "decreased sound tolerance"	Assess* OR measur* OR Questionnaire OR instrument* OR self-report OR inventory OR instrument OR Checklist	Psychometr* OR Valid* OR "Reliab* OR Sensitiv* "internal consistency" OR "Factor Analysis"

Table 2. Summary of study characteristics

#	Author and Year	Sample (N)	Study Population	Sample Age Range (years) or overall rates	Gender	Country	The measure of Hyperacusis or Misophonia
1	Aazh, 2021	100	Clinical population-patients attended a tinnitus and hyperacusis clinic	21 to 81	48 Female 52 Male	UK	Inventory of Hyperacusis Symptoms
2	Blasing, 2010	91	Clinical population-patients suffered from tinnitus	15 to 76	36 Females 55 Males	Germany	GÜF: hypersensitivity to sound

3	Erinc, 2020	529	General population	18 to 73	320 Female 209 Male	Turkey	Hyperacusis Questionnaire
4	Fackrell, 2015	264	Clinical population-data collected from tinnitus studies.	24 to 85	158 Male 106 Female	UK	Hyperacusis Questionnaire
5	Fioretti, 2015	117	Clinical Population-Patients with tinnitus complaints	14 to 88	53 Female 64 Male	Italy	Hyperacusis Questionnaire

6	Greenberg, 2018	469	Patients attending an online support group or social media sites	34.8	40% Male 58% Female 2% not disclosed	USA	Inventory of Hyperacusis Symptoms
7	Khalfa, 2002	201	General population	17 to 72	132 Female 69 Male	France	Hyperacusis Questionnaire
8	Meeus, 2010	46	Clinical Population	21 to 81	14 Female 32 Male	Belgium	Hyperacusis Questionnaire
9	Oishi, 2017	215	Clinical population	Group A: 59.2	GroupA:52.7% Female	Japan	Hyperacusis Questionnaire

				Group B: 63.4	GroupB:46.6% Female		
10	Naylor, 2020	336	University medical students	18 to24	73%Female	UK	The Amsterdam Misophonia Scale
11	Prabhu, 2020	77	Clinical Population (Participants with tinnitus complaints)	20 to55	36 Female 41 Male	India	Hyperacusis Handicap Questionnaire

12	Siepsiak, 2020	705	Clinical Population (Mixed group for misophonia and other health conditions patients)	18 to 68	86.2% and 80% Female for each phase	Poland	MisoQuest
13	Tortorella, 2017	117	Clinical Population (Participants with a primary complaint of tinnitus)	23 to 82	49 Female 68 Male	Italy	The Short Hyperacusis Questionnaire
14	Wu, 2014	483	Undergraduate university students	18 to 54	404 Female 79 Male	USA	Misophonia Questionnaire

Table 3. Description of questionnaires

#	Measure	Construct Measured		Structure	Number of items	Response Options	Total Range
1	Hyperacusis Questionnaire (HQ)	Hyperacusis		Three factors	14	4-Point Likert Scale (0= no, 3= Yes, a lot)	0-42
2	Inventory of Hyperacusis Symptoms	Hyperacusis		Five-factor structures	25	4- Point Likert scale	25-100
3	GÜF: (questionnaire on	Hyperacusis		Three factors	15	4- Point Likert Scale	0-45

	hypersensitivity to sound)						
4	Amsterdam Misophonia Scale	Misophonia		One Factor	6	5-Point Likert scale	0-24
5	Hyperacusis Handicap Questionnaire	Hyperacusis		Three factors	21	3-Point Likert Scale	0-84
6	MisoQuest	Misophonia		One Factor	14	5- Point Likert Scale	14-70

7	Misophonia Questionnaire	Misophonia		Three factors	17	4- Point Likert Scale	0-68(for the first two sections)
8	Short Hyperacusis Questionnaire	Hyperacusis		Three factors	6	4- Point Likert Scale	0-24

Table 4. Methodological quality ratings of each study based on COSMIN.

#	Instrument	Study	Structural Validity	Internal consistency	Cross-cultural validity\ measurement invariance	Reliability	Measurement error	Criterion validity	Hypotheses testing for construct validity	Responsiveness
1	**HQ**	Khalfa et	V	V	N	N	N	N	A	N

		al., 2002								
2	**HQ**	Erinc and Derinsu, 2020	V	V	D	N	N	N	A	N
3	**HQ**	Oishi et al., 2017	N	V	N	N	N	N	D	N
4	**HQ**	Fioretti et al., 2015	N	V	D	N	N	N	N	N
5	**HQ**	Meeus et	I	V	N	N	N	N	D	N

		al., 2010								
6	**HQ**	Fackrell et al., 2015	V	V	N	N	N	N	V	N
7	**IHS**	Greenberg and Carlos, 2018	N	V	N	N	N	N	V	N
8	**IHS**	Aazh et al., 2021	N	V	N	N	N	N	A	N

9	**HHQ**	Prabhu and Nagaraj, 2020	N	V	N	N	N	N	D	N
10	**SHQ**	Tortorella et al., 2017	N	V	N	N	N	N	D	N
11	**GUF**	Blasing et al., 2010	D	V	N	N	N	N	A	N

12	**MQ**	Wu et al., 2014	D	V	N	N	N	N	N	N
13	**MisoQuest**	Siepsiak et al., 2020	A	V	N	D	D	N	A	N
14	**A-MISO-S**	Naylor et al., 2020	A	V	N	N	N	N	N	N

COSMIN rating: V: Very Good; A: Adequate; D: Doubtful; I: Inadequate; N: Not reported by the study authors

Table 5. Ratings for each psychometric property quality per instrument based on COSMIN.

	Structural validity	**Internal consistency**	**Cross-cultural validity**	**Measurement invariance**	**Reliability**	**Measurement error**	**Criterion validity**	**Hypothesis testing for construct validity**	**Responsiveness**
HQ									
Khalfa et al., 2002	**-**	**-**	**NR**	**NR**	**?**	**NR**	**NR**	**NR**	**NR**
Erinc and Derinsu, 2020	**?**	+	+	**NR**	**NR**	**NR**	**NR**	+	**NR**
Oishi et al.,	**NR**	+	**?**	**NR**	**NR**	**NR**	**NR**	**NR**	**NR**

2017									
Fioretti et al., 2015	**NR**	+	**?**	**NR**	**NR**	**NR**	**NR**	**NR**	**NR**
Meeus et al., 2010	**-**	**?**	**?**	**NR**	**NR**	**NR**	**NR**	+	**NR**
Fackrell et al., 2015	+	+	**NR**	**NR**	**NR**	**NR**	**NR**	**?**	**NR**

IHS									
Greenberg and Carlos, 2018	**NR**	+	**NR**	**NR**	**NR**	**NR**	**NR**	+	**NR**
Aazh et al., 2021	**NR**	+	**NR**	**NR**	**NR**	**NR**	**NR**	+	**NR**
MQ									
Wu et al., 2014	**?**	**?**	**NR**	**NR**	**NR**	**NR**	**NR**	**?**	**NR**

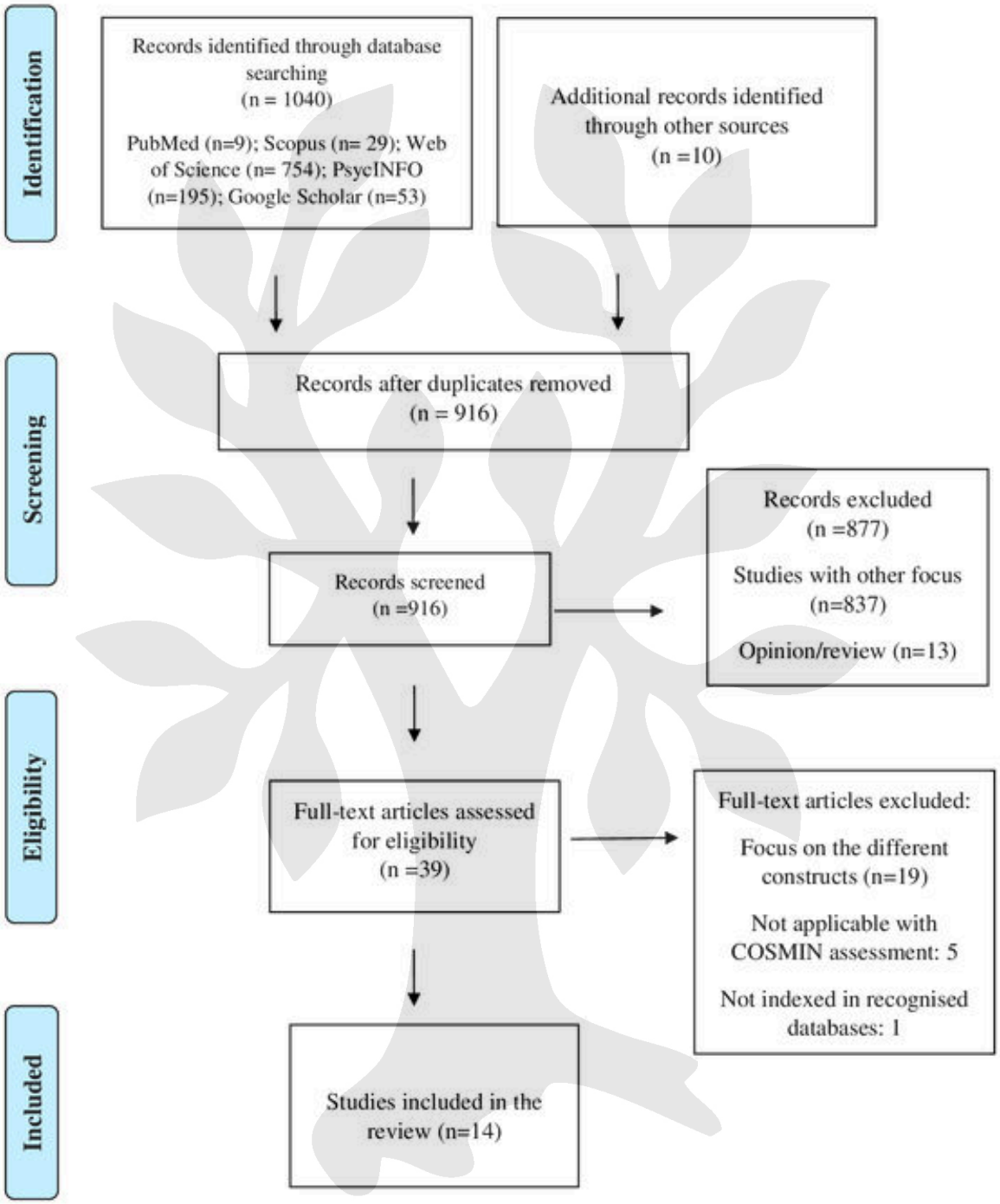
MisoQuest									
Siepsiak et al., 2020	+	+	**NR**	**NR**	+	**?**	**NR**	+	**NR**
HHQ									
Prabhu and	**NR**	+	**NR**	**NR**	**NR**	**NR**	**NR**	**NR**	**NR**

Nagaraj, 2020									
A-MISO-S									
Naylor et al., 2020	**-**	+	**NR**	**NR**	**NR**	**NR**	**NR**	**NR**	**NR**
GUF									
Blasing et al., 2010	**?**	+	**NR**	**NR**	**NR**	**NR**	**NR**	+	**NR**

Short-HQ									
Tortorella et al., 2017	**NR**	**-**	**NR**	**NR**	**NR**	**NR**	**NR**	**NR**	**NR**

COSMIN Rating: (+) 'sufficient', (-) 'insufficient', (?) 'indeterminate', NR= not reported by the study author

Fig. 1. Flowchart of paper selection based on PRISMA guidance.



Kommentare

System user am 11.07.2022 09:21:12

Exifcleaner ausgeführt. 30 Merkmale entfernt. vorher: 49 nachher: 19;

verbleibende Merkmale: ExifTool:ExifToolVersion=12.30, System:FileName=tmp12682009483402643035, System:Directory=/tmp, System:FileSize=332 KiB, System:FileModifyDate=2022:07:11 09:21:12+02:00, System:FileAccessDate=2022:07:11 09:21:12+02:00, System:FileInodeChangeDate=2022:07:11 09:21:12+02:00, System:FilePermissions=-rw-r--r--, File:FileType=JPEG, File:FileTypeExtension=jpg, File:MIMEType=image/jpeg, File:ImageWidth=1654, File:ImageHeight=2339, File:EncodingProcess=Baseline DCT, Huffman coding, File:BitsPerSample=8, File:ColorComponents=3, File:YCbCrSubSampling=YCbCr4:4:4 (1 1), Composite:ImageSize=1654x2339, Composite:Megapixels=3.9