Nurse Media Journal of Nursing
e-ISSN: 2406-8799, p-ISSN: 2087-7811
https://medianers.undip.ac.id
12(2):269-287, August 2022
https://doi.org/10.14710/nmjn.v12i2.43808

REVIEW

An Evaluation of Psychometric Properties of Needs Assessment Instruments in Patients with Coronary Artery Disease Undergoing Cardiac Rehabilitation Programs: A Systematic Review

Kourosh Zarea[1], Eesa Mohammadi[2], Johanne Alteren[3], Neda Sayadi[1]

[1]*Nursing Department, Nursing Care Research Centre in Chronic Disease, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran*
[2]*Department of Nursing, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran*
[3]*Faculty of Health Sciences and Social Care, Molde University College, Molde, Norway*

Article Info

Article History:
Received: 1 January 2022
Revised: 10 August 2022
Accepted: 18 August 2022
Online: 31 August 2022

Keywords:
Cardiovascular rehabilitation; coronary artery disease; COSMIN checklist; instrument needs assessment

Corresponding Author:
Neda Sayadi
Nursing Department, Nursing Care Research Centre in Chronic Disease, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran.
Email:
sayadi_neda@yahoo.com

Abstract

Background: Coronary Artery Disease (CAD) is the foremost reason of illnesses and death in the world. Assessment of the needs is a priority in these patients. However, there is a paucity of tools for the evaluation of needs, with the controversy surrounding their validity.
Purpose: This review aimed to evaluate the psychometric properties of tools used to assess needs of patients with CAD undergoing cardiac rehabilitation (CR) plans.
Methods: An online literature search combined with manual search was carried out on 11 databases to identify relevant articles. The terms used in the search were: cardiac rehabilitation AND coronary artery disease, cardiac rehabilitation AND acute coronary syndrome, and questionnaires OR need assessment OR tool OR scale. Articles from 1989 to 2021 were selected using some inclusion criteria and no validation studies were excluded. The quality of the questionnaires was evaluated by researchers using consensus-based standards for the selection of health status measurement instruments (COSMIN) list. Data analysis had been done by calculating overall methodological quality scores per study on a measurement property using COSMIN checklist. A methodological quality score per box was obtained by taking the lowest rating of any item in a box ('worse score counts').
Results: Of 653 articles, 15 papers were involved in the study. Six studies reported cross-cultural validity, nine studies for criterion validity, and none reported measurement error, hypothesis testing, and responsiveness. There is no vigorous and valid single scale for the measurement of needs in CAD patients. Overall, the CADE-Q questionnaire was good and a patient self-assessment tool for cardiac rehabilitation was poor based on psychometric properties.
Conclusions: The findings of this study disclosed that even though it has been more than 32 years, from 1989 to 2021, of the development in need assessment instruments, each instrument has as a minimum of one "poor" psychometric property according to the COSMIN checklist. So, it is recommended for the next studies to design and develop instruments with better psychometric validities for clinical environment.

How to cite: Zarea, K., Mohammadi, E., Alteren, J., & Sayadi, N. (2022). An evaluation of psychometric properties of needs assessment instruments in patients with coronary artery disease undergoing cardiac rehabilitation programs: A systematic review. *Nurse Media Journal of Nursing, 12*(2), 269-287. https://doi.org/10.14710/nmjn.v12i2.43808

1. Introduction

Cardiovascular illness is the principal reason for death in the world (Vasan et al., 2016). Patients with Coronary Artery Disease (CAD) are at high risk of recurring cardiovascular events. In spite of the progress in pharmacological and invasive cure approaches, risk factors remain independent forecasters of cardiac death in patients with CAD (De Bacquer et al., 2013).

Consequently, secondary prevention constitutes a crucial part of the current care of patients by cardiovascular illness. The expression "Cardiac Rehabilitation" (CR) refers to coordinated complex interventions planned to enhance a cardiac patient's physical, mental, and social performance, as well as steadying, reducing, or even reversing the development of the causal atherosclerotic progressions, thus decreasing morbidity and mortality (Anderson et al., 2016). As such, CR or secondary prevention plans deliver a critical and cost-effective situation to bring applicable preventive care (Balady et al., 2007). Because of the low quality of life in patients with

CAD particularly in patients undergoing surgery (Fayyazi et al., 2012), a participation in CR and education programs has been established to be related to an enhanced lifestyle and an improved diagnosis after an acute coronary occasion (Perk et al., 2012; Ahyana et al., 2013). These programs are aimed at reducing risk factors and maintaining individuals in ideal physical, psychological, social, and functional conditions (Dibben et al., 2021).

The main activities pursued in a CR program include education, a recommendation for lifestyle modification, risk factor management, psychosocial assistance, secondary prevention, and lasting management plans (Fawcett & Desanto-Madeya, 2012). CR contains four main phases that are phase I (the acute phase), phase II (the sub-acute phase), phase III (the intensive outpatient therapy phase), and phase IV (the self-governing continuing conditioning phase). The first phase of nursing care in all phases of CR is to recognize patients' needs. The result of the research by Mohammadi et al. (2019) showed that the care needs of the patients with CAD in phase I CR including physical, psychological, social, and spiritual care needs. Physical care needs included aspects such as providing patients with information regarding the characteristics of CAD, sexual and physical activities after hospital discharge, post-discharge dietary regimen, medications, physical exercise, smoking cessation, wound care, self-care during physical activity, follow-up medical visits, and CAD signs and symptoms. Psychological care needs were mainly related to stress and anxiety management and depression prevention. Social care needs covered aspects such as social relationships after getting discharged from hospital and returning to work and other social activities (Mohammadi et al., 2019).

In particular, the educational supportive role of nurses by Orem in the process of CR has been expressed and the focus of the nurse in this section is to accurately identify the social, emotional, and physical needs of patients and give appropriate training to them since providing these programs is beyond their ability due to the limited time available to nurses. Having a valid and reliable assessment tool in a clinical setting to assess the needs of the patient and his family can compensate the time limit of nurses and increase the effectiveness of the rehabilitation program (Naghdi et al., 2016). Despite the advantages of CR and the major activities carried out in this regard, there is a paucity of tools for the evaluation of needs, with the controversy surrounding their validity which means that sometimes the author did not implement or did not completely report the psychometric properties of the tools (Fawcett & Desanto-Madeya, 2012). Indeed, one of the critical issues today in studies of this field is the selection of appropriate and relevant assessment tools (Naghdi et al., 2016). In addition, before the selection of a tool, it is vitally important that psychometric properties be evaluated based on appropriate criteria (Mokkink, Terwee, Knol et al., 2010).

The consensus-based standards for the selection of health status measurement instruments (COSMIN) checklist, a tool for evaluating the methodological quality of studies on the measurement properties of health status questionnaires, has been developed as the tool of choice in recent years (Mokkink, Terwee, Knol et al., 2010). Utilizing the COSMIN checklist makes it likely to disapprovingly assess and evaluate the quality of these studies (Menezes Costa et al., 2009). This checklist can be used in systematic review studies to examine the features of the scale with the same purpose. Tool choice must be according to high-quality studies, and the COSMIN checklist can be used as a guide for developing tools and reporting features of instrumental scales in studies (Mokkink, Terwee, Knol et al., 2010).

Nonetheless, although the assessment, planning, and implementation of effective interventions in the CR domain call for precise measurement, and evaluation of CR needs appropriate tools, no systematic review on CR questionnaires to date delivers an explained appraisal on the methodological quality of the studies. Therefore, the objective of present review is to recognize scales that investigate needs in patients with CAD undergoing CR and evaluate the psychometric properties of the instruments.

2. Methods

2.1 Research design

This study is a systematic review of studies that assessed the psychometric properties of needs in patients with CAD undergoing CR programs. Also, this review was done in 2021 and in accordance with the Favoured Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) strategies.

2.2 Search methods

An online literature search was carried out by two research members (NS & KZ) on MagIran, IranMedex (Indexing articles published in Iranian biomedical journals), ISC (Islamic World Science Citation), SID (Scientific Information Database – a Persian database), PubMed, MEDLINE, CINAHL (via EBSCO), Scopus, Wiley, EMBASE (via OVID), and Web of Science to identify relevant articles. The terms used in the search were: cardiac rehabilitation AND coronary artery disease, cardiac rehabilitation AND acute coronary syndrome, and questionnaires OR need assessment OR tool OR scale. The words were applied as keywords or free-text words in all databases except for PubMed, in which Mesh terms were used. The search was supplemented with a separate search for the recognized questionnaires in addition to the authors of these questionnaires. The titles of the related references were searched, and the full texts of the articles meeting the inclusion criteria were studied (Figure 1).

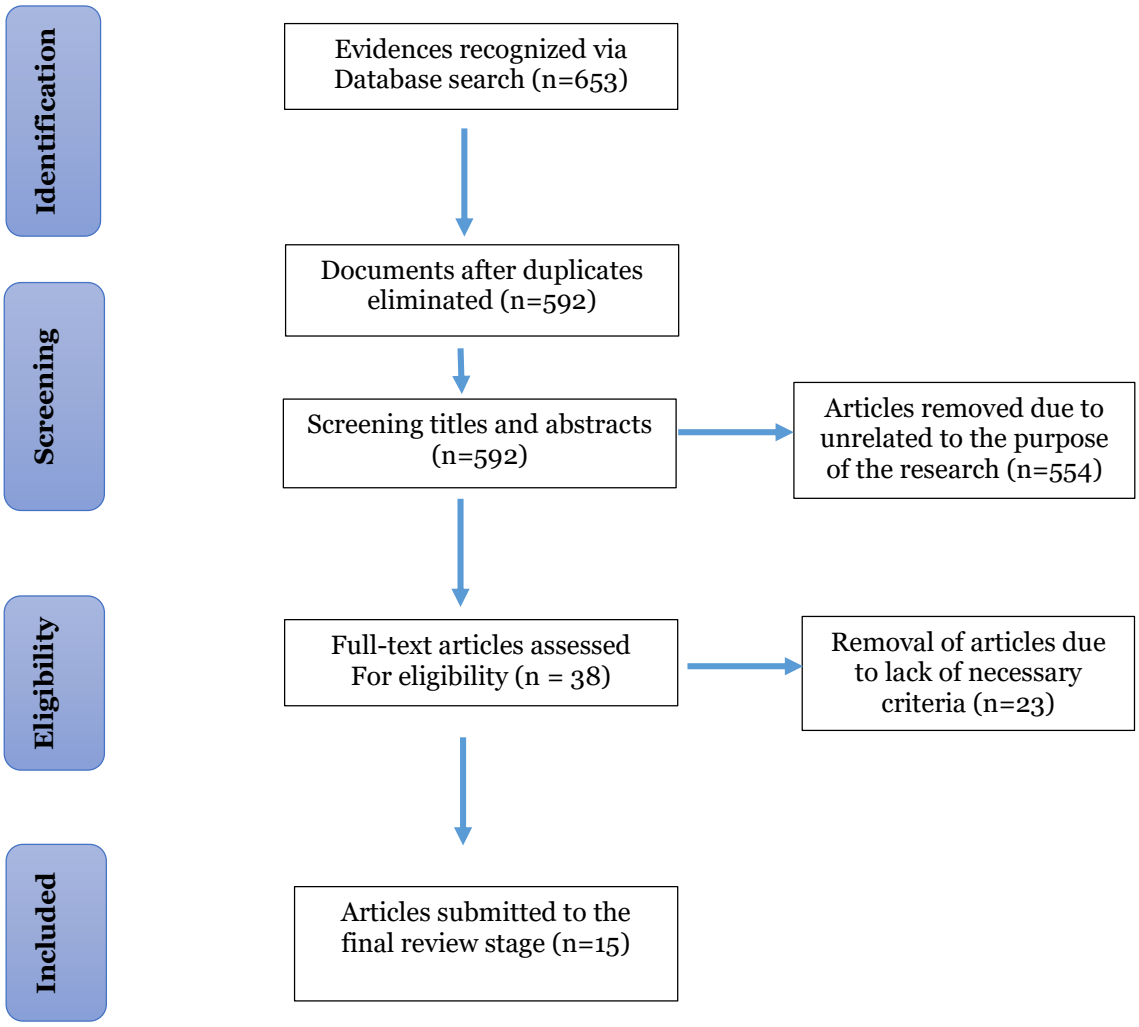


Figure 1. PRISMA flowchart

2.3 Inclusion and exclusion criteria

Articles were included in this study if they were: full-text (original articles), published in English or Persian, published in peer-review journals, and concerned with the development or assessment of the measurement properties of an original form of a need assessment in CR programs questionnaire. Abstracts without full articles, review/systematic review articles, and conference articles were excluded from the study. As per Terwee et al. (2007) standards, publications counted as editorials and case reports were also omitted.

2.4 Screening of articles

Two investigators (NS & KZ) individually assessed titles and abstracts of the recognized records for possible inclusion in the study and evaluated full texts for eligibility by applying the inclusion and exclusion criteria. As can be seen from Figure 1, after screening the titles and abstracts, 554 articles were removed as they were unrelated to the research's purpose. The investigators selected the full-text articles based on the inclusion and exclusion criteria, leaving 15 articles for final review. The investigators determined differences regarding inclusion and exclusion criteria by discussion until they reached an agreement. If agreement could not be reached, the last decision was made by third and fourth investigators (EM & JA).

2.5 Data extraction

Data were extracted from included articles by two research members (NS & KZ) into Table 1 (See Appendix 1). This was carried out in order to summarize the need assessment scales and narrative findings of psychometric characteristics of the scales derived from the included studies.

2.6 Quality appraisal

The methodological quality of articles was assessed using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist by NS & KZ. When agreement could not be reached between two investigators, the last decision was made by the third and fourth investigators (EM & JA).

The COSMIN checklist has newly been improved and available by Mokkink, Terwee, Patrick et al. (2010). The COSMIN list is according to a global Delphi study contributed by 57 experts. Delphi technique considered the most appropriate method to develop a checklist on the methodological quality of studies on measurement properties. Within this Delphi study, the authors have had many interesting discussions, and reached consensus on a number of important issues about the assessment of measurement properties. Therefore, this checklist has been verified to have a suitable inter-rater agreement and reliability (Mokkink, Terwee, Knol et al., 2010). The COSMIN checklist involves twelve boxes. Nine of these boxes denote methodological standards for studies on measurement properties: A) internal consistency, B) reliability, C) measurement error, D) content validity, E) structural validity, F) hypotheses testing, G) cross-cultural validity, H) criterion validity, I) responsiveness. Box J) contains two standards for the interpretability of patient-reported outcomes. Besides, the COSMIN checklist delivers assessment criteria for articles that use the Item-Response-Theory (IRT box) and generalizability of the results (Generalizability box). Each of the boxes A) to I) and the IRT box consists of several items regarding design necessities and statistical analyses.

The scoring system of COSMIN checklist indicated in Table 2-7 (See Appendix 2). The items can be scored on a four-point rating scale representing options for poor, fair, good, or excellent quality. The overall score of the quality of each psychometric property is defined as the lowest score of any item within the box, following the "worst score counts" method. For example, for a reliability study, if one item in the box 'Reliability' is scored poor, the methodological quality of that reliability study is esteemed as poor. At the COSMIN website (www.cosmin.nl), the authors indicate that the checklist mainly emphasizes standards for studies that examine psychometric properties of Health-Related Patient-Reported Outcomes (HR-PROs) (Mokkink, Terwee, Patrick et al., 2010). The quality appraisal of the involved articles is presented in Table 8.

2.7 Data analysis

As this study investigated the psychometric properties of the scales from the articles, the data analysis had also been done by using COSMIN checklist (Table 8). After all studies were assessed their psychometric components using nine boxes (box A-I) in the COSMIN checklist, the authors (NS & KZ) analyzed the data according to overall quality of the scales and each box of the COSMIN checklist.

3. Results

3.1 Study characteristics

As presented in the PRISMA flow chart (Figure 1), 653 articles were found in the first search. Afterward omitting doubled and irrelevant studies, 15 studies remained. The search identified 7

different questionnaires relating to need assessment in patients with CAD undergoing the CR program.

Table 8. COSMIN quality assessment

First author (year)	COSMIN BOXES								
	BOX A INTERNAL consistency	BOX B Reliability	BOX C Measurement error	BOX D Content validity	BOX E Structural validity	BOX F Hypothesis testing	BOX G Cross-cultural validity	BOX H Criterion validity	BOX I Respon-siveness
De Melo Ghisi et al. (2010)	Good	Poor	NR	excellent	Good	NR	-	-	NR
De Melo Ghisi, Oh et al. (2013)	Good	Poor	NR	excellent	Good	NR	Good	Poor	NR
Marofi et al. (2020)	Good	Poor	NR	excellent	Good	NR	Good	-	NR
De Melo Ghisi et al. (2015)	Good	Poor	NR	excellent	Good	NR	-	Poor	NR
Chen et al. (2018)	Good	Poor	NR	excellent	Good	NR	Good	Poor	NR
Santos et al. (2019)	Poor	Poor	NR	excellent	Poor	NR	Good	Poor	NR
De Melo Ghisi et al. (2016)	Poor	Poor	NR	excellent	Poor	NR	-	Poor	NR
De Melo Ghisi et al. (2018)	Good	Poor	NR	excellent	Good	NR	Good	Poor	NR
De Melo Ghisi & Oh (2021)	Good	Poor	NR	excellent	Good	NR	Good	Poor	NR
De Melo Ghisi, Grace et al. (2013)	Poor	Poor	NR	excellent	Poor	NR	-	Poor	NR
De Melo Ghisi et al. (2014)	Poor	Poor	NR	excellent	Poor	NR	Good	Poor	NR
Sayadi et al. (2021)	Good	Poor	NR	excellent	Good	NR	-	-	NR
Van Engen-Verheul et al. (2012)	Poor	Poor	NR	excellent	Poor	NR	-	-	NR
Smith et al. (2015)	Poor	Poor	NR	excellent	Poor	NR	-	-	NR
Phelan et al. (1989)	Poor	Poor	NR	excellent	Poor	NR	-	-	NR

Note. NR: not reported

These questionnaire consisted of: (1) A patient self-assessment tool for cardiac rehabilitation, (2) the cardiac rehabilitation needs assessment tool (CRNAT), (3) the information needs in cardiac rehabilitation (INCR) tool, (4) the coronary artery disease education questionnaire (CADE-Q), (5) the second version of the coronary artery disease education questionnaire (CADE-Q II), (6) short version of the coronary artery disease education questionnaire (CADE-Q SV), (7) care needs questionnaire in phase 1 cardiac rehabilitation for patients with coronary artery disease (CNCR-Q). The rest of the articles (n=7) of included studies showed the cross-cultural validation of these questionnaires in other countries and one study was an algorithm about need assessment for the patient in CR.

Involved studies were available from the year 1989 to 2021. All studies were peer-review original articles. Majority of the studies (n=5) were showed in the Canada (De Melo Ghisi et al., 2016; De Melo Ghisi et al., 2015; De Melo Ghisi, Grace et al., 2013; De Melo Ghisi & Oh, 2021; De Melo Ghisi, Oh et al., 2013) followed by Iran (n=2) (Marofi et al., 2020; Sayadi et al., 2021), Brazil (n=4) (De Melo Ghisi et al., 2010; De Melo Ghisi et al., 2018; De Melo Ghisi et al., 2014 Santos et al., 2019) , China (n=1) (Chen et al. 2018), Australia (n=1) (Smith et al., 2015), Netherlands (n=1) (Van Engen-Verheul et al., 2012) and USA (n=1) (Phelan et al., 1989).

3.2 Quality evaluation of the studies

The results showed that the concept of need was investigated in all studies. Overall the CADEQ questionnaire was good based on psychometric properties. Respecting to the study design, seven articles were cross-cultural assessment studies (Chen et al., 2018; De Melo Ghisi et al., 2014; De Melo Ghisi et al., 2018; De Melo Ghisi & Oh, 2021; De Melo Ghisi, Oh et al., 2013; Marofi et al., 2020; Santos et al., 2019) and others were studied about design and psychometric properties of questionnaires (De Melo Ghisi et al., 2010; De Melo Ghisi et al., 2015; De Melo Ghisi et al., 2016; De Melo Ghisi, Grace et al., 2013; Phelan et al., 1989; Sayadi et al., 2021; Smith et al., 2015; Van Engen-Verheul et al., 2012).

The number of tool items and factors of involved studies were different. The maximum item number was 60 (De Melo Ghisi, Grace et al., 2013), the minimum number of dimensions were three (Marofi et al., 2020) and the maximum number was 11 dimensions in one study (Phelan et al., 1989).

Internal consistency was directed with estimating Cronbach's alpha and Kuder- Richardson-20 test. Intra-Class Corelation (ICC) and Spearman's rank correlation coefficient were used for stability in reliability. Educational level, family income, CR duration, and time of diagnosis were used as a criterion for criterion validity. Majority of studies had construct validity; exploratory factor analysis (n=6) (De Melo Ghisi et al., 2015; De Melo Ghisi et al., 2010; De Melo Ghisi et al., 2018; De Melo Ghisi & Oh, 2021; De Melo Ghisi, Oh et al., 2013; Sayadi et al., 2021), confirmatory factor analysis (n=1) (Chen et al., 2018), and exploratory and confirmatory factor analysis (n=1) (Marofi et al., 2020) while other studies did not report it. Other psychometric characteristics of involved studies are reviewed in Table 1. The results of the COSMIN quality assessment of 15 involved articles are provided in Table 8. None of these articles had "Excellent" quality in all psychometric properties.

3.3 BOX A – Internal consistency

Internal consistency is the extent to which items in a (sub)scale are intercorrelated, so assessing the equal construct (Terwee et al., 2012). In twelve studies, internal consistency calculated based on Cronbach alpha or Kuder-Richardson-20, (Chen et al., 2018; De Melo Ghisi et al., 2014; De Melo Ghisi et al., 2016; De Melo Ghisi et al., 2015; De Melo Ghisi et al., 2010; De Melo Ghisi et al., 2018; De Melo Ghisi, Grace et al., 2013; De Melo Ghisi & Oh, 2021; De Melo Ghisi, Oh et al., 2013; Marofi et al., 2020; Santos et al., 2019; Sayadi et al., 2021). The COSMIN checklist scores for four studies were "poor" because they did not calculate the factor analysis also the author did not reference to other studies (De Melo Ghisi et al., 2014; De Melo Ghisi et al., 2016; De Melo Ghisi, Grace et al., 2013; Santos et al., 2019) and eight studies were evaluated as "Good" because percentage of missing items were not explained. (Chen et al., 2018; De Melo Ghisi et al., 2015; De Melo Ghisi et al., 2010; De Melo Ghisi et al., 2018; De Melo Ghisi & Oh, 2021; De Melo Ghisi, Oh et al., 2013; Marofi et al., 2020; Sayadi et al., 2021). Three studies had not calculated internal consistency so the COSMIN score were "poor" (Phelan et al., 1989; Smith et al., 2015; Van Engen-Verheul et al., 2012).

3.4 BOX B – Reliability

Reliability is the degree to which patients can be discriminated from each other, despite measurement errors (relative measurement error) (Terwee et al., 2012). Ten studies had not calculated reliability so the COSMIN score was "poor" (Chen et al., 2018; De Melo Ghisi et al., 2016; De Melo Ghisi et al., 2015; De Melo Ghisi et al., 2014; De Melo Ghisi & Oh, 2021; De Melo Ghisi, Oh et al., 2013; Phelan et al., 1989; Sayadi et al., 2021; Smith et al., 2015; Van Engen-Verheul et al., 2012). The rest studies described reliability criteria and were estimated as "poor," because all of them used only one measurement (De Melo Ghisi et al., 2010; De Melo Ghisi et al., 2018; De Melo Ghisi, Grace et al., 2013; Santos et al., 2019).

3.5 BOX C – Measurement error

The systematic and random error of a score that is not attributed to true variations in the construct is measured as measurement error. Measurement errors of all studies were not stated.

3.6 BOX D – Content validity

The content validity is described as "the extent to which the domain of interest is comprehensively sampled by the items in the questionnaire". Content validity of all studies were "excellent" because the content of questionnaire was evaluated by expert panels.

3.7 BOX E – Structural validity

According to the COSMIN checklist, structural validity is the degree to which the scores of scales are a sufficient indication of the dimensionality of the construct. In this regard, seven articles did not describe factor analysis and were appraised as "poor."(De Melo Ghisi et al., 2016; De Melo Ghisi, Grace et al., 2013; Santos et al., 2019; De Melo Ghisi et al., 2014; Smith et al., 2015; Van Engen-Verheul et al., 2012; Phelan et al., 1989) and eight studies were evaluated exploratory or confirmatory factor analysis but proportion of missing items NOT defined, so COSMIN score were "good"(Chen et al., 2018; De Melo Ghisi et al., 2015; De Melo Ghisi et al., 2010; De Melo Ghisi et al., 2018; De Melo Ghisi, Oh et al., 2013; De Melo Ghisi, Oh et al., 2013; Marofi et al., 2020; Sayadi et al., 2021).

3.8 BOX F – Hypothesis testing

Hypothesis testing of all studies were not reported.

3.9 BOX G – Cross cultural

In this study, seven questionnaires were used. Cross-cultural adaptation was performed only for four questionnaires (De Melo Ghisi et al., 2010; De Melo Ghisi et al., 2015; De Melo Ghisi et al., 2016; De Melo Ghisi, Grace et al., 2013). Cross-cultural adaptation of INCR was done to the Portuguese language (De Melo Ghisi et al., 2014). Cross-cultural adaptation of CADE-Q were done to Persian and English language (De Melo Ghisi, Oh et al., 2013; Marofi et al., 2020). Cross-cultural adaptation of CADE-Q SV were done to French-Canadian and Brazilian-Portuguese language (De Melo Ghisi et al., 2018; De Melo Ghisi & Oh, 2021). Cross-cultural adaptations of CADE-Q II were done to Chinese and Brazilian language (Chen et al., 2018; Santos et al., 2019). The COSMIN score for all studies were "good" because the percentage of missing items is NOT defined.

3.10BOX H – Criterion validity

Criterion validity is "the degree to which scores on a specific questionnaire relate to a gold standard" (Terwee et al., 2007). Nine studies were performed criterion validity. They used criteria such as duration of CR, monthly family income, and educational level of patients. The COSMIN score was poor because the criterion used could not be respected a sufficient gold standard (Chen et al., 2018; De Melo Ghisi et al., 2016; De Melo Ghisi et al., 2015; De Melo Ghisi et al., 2018; De Melo Ghisi et al., 2014; De Melo Ghisi, Grace et al., 2013; De Melo Ghisi & Oh, 2021; De Melo Ghisi, Oh et al., 2013; Santos et al., 2019).

3.11 BOX I- Responsiveness

Responsiveness is the ability of a questionnaire to distinguish clinically critical changes over time. Responsiveness of all studies was not reported.

4. Discussion

The aim of this study was to evaluate and recognize the best questionairre according to psychometric properties. This review covered fifteen articles that used seven different questionnaires, namely the CRNAT, INCR, CADE-Q, CADE-Q SV, CADE-Q II, CNCR-Q, and the patient self-assessment tool for cardiac rehabilitation. According to the COSMIN checklist, these tools did not score "excellent" quality in all psychometric properties. In other words, there was no vigorous and valid single scale for the measurement of needs in CAD patients.

The findings showed that while the first article published in 1989 (Phelan et al., 1989) yet the psychometric properties of these publications have not significantly increased to the year of 2021. The reason for this problem might be related to the nature of the concept of need which is a subjective concept and each of patients defined this concept based on attitude, social and cultural background. Concerning the country of publication, the majority of studies were performed in

Canada (De Melo Ghisi et al., 2015; De Melo Ghisi et al., 2016; De Melo Ghisi, Grace et al., 2013; De Melo Ghisi & Oh, 2021; De Melo Ghisi, Oh et al., 2013).

The INCR was developed to evaluate information needs in CR. This scale assesses information need and is culturally adapted to Brazil. This tool includes 11 dimensions. After the validation process of Brazilian version of INCR questionnaire, one dimension of questionnaire including barriers/goal setting was deleted. As the validation process was not performed with adequate sample, the author recommended future research was needed to assess whether the scale was sensitive to change, such as following participation in the education components of CR, or to test implementation of new education materials. Finally, whether the INCR is a valuable and valid tool to identify information needs in individual patients should be further explored (De Melo Ghisi et al., 2014).

The CADE-Q is another questionnaire that was developed by Ghisi et al. (2010), to calculate and describe coronary patients' awareness of CR programs. Most of the psychometric properties of this questionnaire had been previously reported, and it had been adapted to English and Persian language (De Melo Ghisi, Oh et al., 2013; Marofi et al., 2020). Due to the limitations of CADE-Q including the lack of attention to all aspects of CR such as the psychosocial dimension, the author designed and validated CADE-Q II (De Melo Ghisi et al., 2015). This questionnaire was culturally adapted to Brazilian and Chinese patients (Chen et al., 2018; Santos et al., 2019). Because CADE-Q II was long, there was little willingness to fill in the questionnaire so the CADE-Q SV was developed (De Melo Ghisi et al., 2016). The CADE-Q SV was cross-culturally adapted to Brazilian-Portuguese and French-Canadian (De Melo Ghisi et al., 2018; De Melo Ghisi & Oh, 2021). When a new version of a CADE-Q questionnaire was developed, it should not only be updated, and have better theoretical basis, but it must also be shown to be at least as good as the original instrument in terms of validity and reliability. In this context, results of the CADE-Q SV were consistent with those presented in previous versions of this instrument, particularly in relation to criterion validity (correlation to educational level) and all areas being considered internally consisted (α>0.70). The overall mean, as well as the means of the areas were high, reinforcing the idea that CR patients are knowledgeable of the information that is important for them. It may also suggest that individuals with low socioeconomic levels or low literacy are not participating in these programs and strategies to eliminate barriers to access CR should be implemented (De Melo Ghisi et al., 2016).

The patient self-assessment tool for cardiac rehabilitation and CRNAT are other tools for the assessment of CR needs. None of the psychometric properties of these questionnaires had been reported except content validity (Phelan et al., 1989; Smith et al., 2015). The newest questionnaire for need assessment is CNCR-Q that is designed by Sayadi et al. (2021). This questionnaire is developed based on a definition of care needs in patients with CAD undergoing CR according to Islamic culture. Sayadi et al. (2021) added spiritual care needs for CAD patients. This feature has not been mentioned in previous studies. The questionnaire is a tool with 40 items. After conducting face validity qualitatively, all tool items were considered important and were retained for the next steps. After completing the steps for determining the content validity ratio (CVR) and content validity index (CVI) of 40 items, all items were preserved for decision making at a later stage. The results of exploratory factor analysis revealed four factors. Moreover, the factor analysis results in the elimination of three items and the final version of the questionnaire with 37 items remained (Sayadi et al., 2021).

The general quality of the articles examining the measurement properties as rated by the COSMIN list was poor to excellent. Several studies did not state adequate information in the article, thus it was difficult to evaluate their quality. All studies which were reviewed in the study had described internal consistency as reliability, but in several studies, there was no information about other crucial properties. Some tools had a lack of face validity and stability evaluation, so future studies should consider these properties once trying to validate tools. Involved studies did not discuss measurement error, hypothesis testing, and responsiveness, which might be related to the nature of the concept of need, and because these criteria were more applicable for concepts which measured objective changes such as blood pressure in patients. The highest methodological quality was the CADE-Q (De Melo Ghisi et al., 2010), CADE-Q II (De Melo Ghisi et al., 2015) and CNCR-Q (Sayadi et al., 2021) that in one box of COSMIN checklist scored as "Excellent," two boxes "Good," and one box "poor."

5. Implications and limitations

The strength of this review could be attributed to the fact that two investigators individually measured all records in full text, and concurrently with the third and fourth reviewer qualities whom were certified by double or triple evaluation of several studies. The implication of this review includes, according to the aim of this study, to assist nurses in choosing valid tools in the field of CR, to assist nurses in selecting credible tools to more quickly identify the care needs of patients undergoing CR, and to reduce the readmission of heart patients by correctly identifying care needs. It appears that the CADE-Q questionnaire can be recommended to nurses for using in need assessment. However, this review had a limitation that required to be addressed. Our search was limited to studies published in Persian or English. So, studies published in other languages were not involved.

6. Conclusion

This systematic review delivers a summary of 7 tools assessing needs in CAD patients undergoing CR. Although, based on the COSMIN checklist, none of the studies was evaluated excellent in all boxes, the results of this study helped researchers to select the best quality questionnaire among existing questionnaires in this field. In other words, there was no vigorous and valid single scale for the measurement of needs in CAD patients; however, the CADE-Q questionnaire was good based on psychometric properties and a patient self-assessment tool for cardiac rehabilitation was poor based on psychometric properties. This study proposes that future assessment studies on psychometric properties concern standards like the COSMIN checklist to increase the quality of the studies and to enhance the assessment of results. Also, we encourage article's readers to explore how an instrument's psychometric properties might be improved and then re-tested with the result of the tool being of even greater use in clinical practice to decrease the mortality and morbidity of cardiovascular disease.

Acknowledgment

As part of the fourth author's Ph.D. thesis (Neda Sayadi), this study was financially supported by the Nursing Care Research Centre in Chronic Diseases, Ahvaz Jundishapur University of Medical Sciences, and Ahvaz, Iran (Number CDCRC-9412).

Author contribution

NS and EM participated in the definition of intellectual content, literature search and data analysis. JA and KZ were being involved in the manuscript preparation, editing and review.

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Ahyana, A., Kritpracha, C., Thaniwattananon, P. (2013). Cardiac rehabilitation enhancing programs in patients with myocardial infarction: A literature review. *Nurse Media Journal of Nursing, 3*(1), 541-556. https://doi.org/10.14710/nmjn.v3i1.4468

Anderson, L., Thompson, D. R., Oldridge, N., Zwisler, A. D., Rees, K., Martin, N., & Taylor, R. S. (2016). Exercise-based cardiac rehabilitation for coronary heart disease. *The Cochrane Database of Systematic Reviews*, *2016*(1). https://doi.org/10.1002/14651858

Balady, G. J., Williams, M. A., Ades, P. A., Bittner, V., Comoss, P., Foody, J. M., Franklin, B., Sanderson, B., Southard, D., American Heart Association Exercise, Cardiac Rehabilitation, and Prevention Committee, the Council on Clinical Cardiology, American Heart Association Council on Cardiovascular Nursing, American Heart Association Council on Epidemiology and Prevention, American Heart Association Council on Nutrition, Physical Activity, and Metabolism, & American Association of Cardiovascular and Pulmonary Rehabilitation (2007). Core components of cardiac rehabilitation/secondary prevention programs: 2007 update: A scientific statement from the American Heart Association Exercise, Cardiac Rehabilitation, and Prevention Committee, the Council on Clinical Cardiology; the Councils on Cardiovascular Nursing, Epidemiology and Prevention, and Nutrition, Physical Activity,

and Metabolism; and the American Association of Cardiovascular and Pulmonary Rehabilitation. *Circulation*, *115*(20), 2675–2682. https://doi.org/10.1161/CIRCULATIONAHA.106.180945

Chen, Z., Geng, J., Wang, M, Hu, L., De Melo Ghisi, G. L., & Yu, H. (2018). The Chinese version of the coronary artery disease education questionnaire-II (CADEQ-II): Translation and validation. *Patient Preference and Adherenc*e, *12*, 1587-1596. https://doi.org/10.2147/PPA.S176639

De Bacquer, D., Dallongeville, J., Kotseva, K., Cooney, M. T., Pajak, A., Deckers, J. W., Mayer, O., Vanuzzo, D., Lehto, S., Fras, Z., Ostor, E., Ambrosio, G. B., De Backer, G., Wood, D., Keil, U., Sans, S., Graham, I., & Pyorala, K. (2013). Residual risk of cardiovascular mortality in patients with coronary heart disease: The EUROASPIRE risk categories. *International Journal Cardiology, 168*(2), 910-914. https://doi.org/10.1016/j.ijcard.2012.10.051

De Melo Ghisi, G. L., Oh, P., Thomas, S., & Benetti, M. (2013). Development and validation of an English version of the Coronary Artery Disease Education Questionnaire (CADE-Q). *European Journal of Preventive Cardiology, 20*(2), 291–300. https://doi.org/10.1177/2047487312437061

De Melo Ghisi, G., Sandison, N., & Oh, P. (2016). Development, pilot testing and psychometric validation of a short version of the coronary artery disease education questionnaire: The CADE-Q SV. *Patient Education and Counseling, 99*(3), 443–447. https://doi.org/10.1016/j.pec.2015.11.002

De Melo Ghisi, G., & Oh, P. (2021). Validation of the French-Canadian version of a short questionnaire to assess knowledge in cardiac patients (CADE-Q SV). *The Canadian Journal of Nursing Research, 54*(1), 51–58. https://doi.org/10.1177/0844562120986001

De Melo Ghisi, G. L., Grace, S. L., Thomas, S., Evans, M. F., & Oh, P. (2013). Development and psychometric validation of a scale to assess information needs in cardiac rehabilitation: The INCR Tool. *Patient Education and Counseling, 91*(3), 337–343. https://doi.org/10.1016/j.pec.2013.01.007

De Melo Ghisi, G. L., Durieux, A., Manfroi, W. C., Herdy, A. H., Carvalho, T. d., Andrade, A., & Benetti, M. (2010). Construction and validation of the CADE-Q for patient education in cardiac rehabilitation programs. *Arquivos Brasileiros de Cardiologia, 94*(6), 813–822. https://doi.org/10.1590/s0066-782x2010005000045

De Melo Ghisi, G., Chaves, G., Loures, J. B., Bonfim, G. M., & Britto, R. (2018). Validation of the Brazilian-Portuguese version of a short questionnaire to assess knowledge in cardiovascular disease patients (CADE-Q SV). *Arquivos Brasileiros de Cardiologia, 111*(6), 841–849. https://doi.org/10.5935/abc.20180169

De Melo Ghisi, G. L., Dos Santos, R. Z., Bonin, C. B., Roussenq, S., Grace, S. L., Oh, P., & Benetti, M. (2014). Validation of a Portuguese version of the Information Needs in Cardiac Rehabilitation (INCR) scale in Brazil. *Heart & Lung: The Journal of Critical Care, 43*(3), 192–197. https://doi.org/10.1016/j.hrtlng.2014.01.009

Dibben, G., Faulkner, J., Oldridge, N., Rees, K., Thompson, D.R., Zwisler, A. D., & Taylor, R.S. (2021). Exercise-based cardiac rehabilitation for coronary heart disease. *Cochrane Database of Systematic Review, 6*(11), CD001800. https://doi.org/ 10.1002/14651858.CD001800.pub4

Fawcett, J., & Desanto-Madeya, S. (2012). *Contemporary nursing knowledge: Analysis and evaluation of nursing models and theories*. FA Davis.

Fayyazi, S., Sayadi, N., & Gheybizadeh, M. (2012). Comparison of quality of life before and after open heart surgery. *Zahedan Journal of Research in Medical Sciences, 14*(9), 98-100.

Ghisi, G. L., Grace, S. L., Thomas, S., Evans, M. F., & Oh, P. (2015). Development and psychometric validation of the second version of the Coronary Artery Disease Education Questionnaire (CADE-Q II). *Patient Education and Counseling, 98*(3), 378–383. https://doi.org/10.1016/j.pec.2014.11.019

Marofi, Z., Bandari, R., Heravi-Karimooi, M., Rejeh, N., & Montazeri, A. (2020). Cultural adoption, and validation of the Persian version of the coronary artery disease education questionnaire (CADE-Q): A second-order confirmatory factor analysis. *BMC Cardiovascular Disorder, 20,* 345. https://doi.org/10.1186/s12872-020-01628-5

Menezes Costa, L., Maher, C. G., McAuley, J. H., & Costa, L. O. (2009). Systematic review of cross-cultural adaptations of McGill pain questionnaire reveals a paucity of clinimetric testing.

Journal of Clinical Epidemiology, 62(9), 934–943. https://doi.org/10.1016/j.jclinepi.2009.03.019

Mohammadi, E., Zarea, K., Alteren, J., & Sayadi, N. (2019). Care needs in the Phase I of cardiac rehabilitation: A hybrid concept analysis. *Nursing and Midwifery Studies, 8*(1), 48-54. http://doi.org/10.4103/nms.nms_76_17

Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., Bouter, L. M., & de Vet, H. C. (2010). The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Medical Research Methodology, 18*(10), 22. https://doi.org/10.1186/1471-2288-10-22

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology, 63*(7), 737–745. https://doi.org/10.1016/j.jclinepi.2010.02.006

Naghdi, S., Ansari, N N., Raji, P., Shamili, A., Amini, M., & Hasson, S. (2016). Cross-cultural validation of the Persian version of the functional independence measure for patients with stroke. *Disability and Rehabilitation, 38*(3), 289–298. https://doi.org/10.3109/09638288.2015.1036173

Perk, J., De Backer, G., Gohlke, H., Graham, I., Reiner, Z., Verschuren, M., Albus, C., Benlian, P., Boysen, G., Cifkova, R., Deaton, C., Ebrahim, S., Fisher, M., Germano, G., Hobbs, R., Hoes, A., Karadeniz, S., Mezzani, A., Prescott, E., … Zannad, F.(2012). European guidelines on cardiovascular disease prevention in clinical practice (version 2012). The fifth joint task force of the European society of cardiology and other societies on cardiovascular disease prevention in clinical practice (constituted by representatives of nine societies and by invited experts). *European Heart Journal*, *33*(13), 1635-1701. https://doi.org/10.1093/eurheartj/ehs092

Phelan, C., Finnell, M. D., & Mottla, K. A. (1989). A patient self-assessment tool for cardiac rehabilitation. *Rehabilitation Nursing, 14*(2), 81–87. https://doi.org/10.1002/j.2048-7940.1989.tb01425.x

Santos, R., Ghisi, G., Bonin, C., Chaves, G., Haase, C. M., Britto, R. R., & Benetti, M. (2019). Validation of the Brazilian version of CADE-Q II to assess knowledge of coronary artery disease patients. *Arquivos Brasileiros de Cardiologia*, *112*(1), 78–84. https://doi.org/10.5935/abc.20180244

Sayadi, N., Alteren, J., Mohammadi, E., & Zarea, K. (2021). Development and psychometric properties evaluation of a care needs questionnaire in phase 1 cardiac rehabilitation for patients with coronary artery disease: CNCR-Q. *Journal of Caring Cciences, 10*(1), 29–36. https://doi.org/10.34172/jcs.2021.006

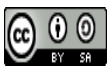
Smith, J., Garton-Smith, J., Briffa, T., & Maiorana, A. (2015). The development of a new cardiac rehabilitation needs assessment tool (CRNAT) for individualised secondary prevention. *Heart, Lung & Circulation, 24*(5), 458–464. https://doi.org/10.1016/j.hlc.2015.01.001

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clininical Epidemiology, 60*(1), 34-42. https://doi.org/10.1016/j.jclinepi.2006.03.012

Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W., Bouter, L. M., & de Vet, H. C. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *21*(4), 651–657. https://doi.org/10.1007/s11136-011-9960-1

Van Engen-Verheul, M. M., Kemps, H. M., de Keizer, N. F., Hellemans, I. M., Goud, R., Kraaijenhagen, RA., & Peek, N. (2012). Revision of the Dutch clinical algorithm for assessing patient needs in cardiac rehabilitation based on identified implementation problems. *European Journal of Preventive Cardiology*, *19*(3), 504-514.

Vasan, R.S., Benjamin, E. J. (2016). The future of cardiovascular epidemiology. *Circulation, 133*(25), 2626-2633. https://doi.org/10.1161/CIRCULATIONAHA.116.023528



Appendix 1

Table 1. Data extraction of included studies

Author (year)	Scale	Country	Target population	Face validity	Content validity	Construct Validity					Reliability	
						Sample size	Factor extraction method (Rotation)	Selection of the number of factors	name of factors	Total variance (%)	Consistency	Stability
De Melo Ghisi et al. (2010)	CADE-Q	Brazil	patient with CAD	patient with CAD(n=30)	19-item4-dimension 4 points Likert expert panel	155	EFA KMO=0.608 Bartlett's Sphericity tests	PCA Eigenvalues>1 screen plot	f1:General factor f2:Causal factor f3:Risk Factors Factor f4:Other Factors	56.1	Total α:0.068	Overall ICC:0.783
De Melo Ghisi, Oh et al. (2013)	English version of CADE-Q	Canada	patient with CAD	patient with CAD(n=50)	19-item4-dimension 4 points Likert expert panel	200	EFA KMO=0.797Bartlett's Sphericity tests	PCA Eigenvalues>1	F1:Exercise F2:Risk Factors F3:Diagnosis,treatment and medicines F4:Pathophysiology,signals and symptoms F5: prevention of risk factors	62.23	Total α:0.809	Overall ICC: 0.846 (2-week)
Marofi et al. (2020)	Persian version of CADE-Q	Iran	patient with CAD	patient with CAD(n=10)	19-item3-dimension 4 points Likert expert pane	500	EFA KMO,Bartlett's Sphericity tests,CFA	prom ax rotation screen plot	F1: lifestyle habits& exercise F2: Risk factors F3: Diagnosis and treatment	48.9	Total α:0.844 F1: α = 0.825 F2: α =0.553 F3: α = 0.507	Overall ICC: 0.886 F1: ICC=0.870 F2: ICC=0.782 F3: ICC=0.825
De Melo Ghisi et al. (2015)	CADE-Q II	Canada	patient with CAD	patient with CAD(n=30)	31-item4-dimension 4 points Likert expert pane	307	EFA KMO=0.859Bartlett's Sphericity tests	PCA Eigenvalues>1	F1: Medical condition F2: Risk factors and exercise F3: nutrition F4: psychosocial risk	62.2	Total α: 0.91 F1: A= 0.71 F2: α = 0.65 F3: α =0.66 F4: α =0.67	NR
Chen et al. (2018)	Chines version of CADE-Q II	China	patient with CAD	patient with CAD(n=40)	28-item 5 dimension 4 points Likert expert panel CVI=0.87	316	CFA	-	F1: medication condition F2: Risk factors F3: Exercise F4: Nutrition F5:Psychological risk	NR	Total α: 0.907 F1: α = 0.692 F2: α = 0.687 F3: α =0.714 F4: α =0.705 F5: α =0.701	NR
Santos et al. (2019)	Brazilian version of CADE-Q II	Brazil	patient with CAD	patient with CAD(n=23)	27-item 5 dimension 4 points Likert expert panel CVI=0.87	307	NR	NR	F1: medical condition F2: Risk factors F3: Exercise F4: Nutrition F5:Psychological risk	NR	Total α: 0.78	Overall ICC:0.77 (2-week)
De Melo Ghisi et al. (2016)	CADE-Q SV	Canada	Patient with CAD	patient with CAD(n=20)	20-item 5 dimension 3 points expert panel	200	NR	NR	F1: medical condition F2: Risk factors F3: Exercise F4: Nutrition F5:Psychological risk	NR	Total α: NRF1: α = 0.94 F2: α = 0.79 F3: α =0.76 F4: α =0.84 F5: α =0.91	NR

Table 1. Continued

Author (year)	Scale	Country	Target population	Face validity	Content validity	Construct Validity					Reliability	
						Sample size	Factor extraction method (Rotation)	Selection of the number of factors	name of factors	Total variance (%)	Consistency	Stability
De Melo Ghisi et al. (2018)	Brazilian-Portuguese version of CADE-Q SV	Brazil	patient with CAD	patient with CAD(n=21)	20-item 5 dimension 3 points expert panel	200	EFA KMO=0.87Bartlett's Sphericity tests	PCA Eigenvalues>1	F1: medical condition F2: Risk factors F3: Exercise F4: Nutrition F5:Psychological risk	59	Total KR-20= 0.7	Overall ICC>0.7
De Melo Ghisi & Oh (2021)	French-Canadian version of CADE-Q SV	Canada	patient with CAD	NR	20-item 5 dimension 3 points expert panel	115	EFA KMO=0.90Bartlett's Sphericity tests	PCA Eigenvalues>1	F1: medical condition F2: Risk factors F3: Exercise F4: Nutrition F5:Psychological risk	69.9	Total KR-20= 0.72	NR
De Melo Ghisi, Grace et al. (2013)	INCR	Canada	patient with CAD	patient with CAD(n=34)	55-item 10 dimension 5 points Likert expert panel	203	NR	NR	F1: The heart(physiology,symptoms, surgical treatments),F2:Nutrition F3:Exercise/physical activity,F4:Medication F5:Work/vocational/social F6:Stress/psychological factors F7:General/social concerns F8:Emergency/safety F9:Diagnosis and treatment F10:Risk factors	NR	Total α: NRF1: α = 0.88 F2: A= 0.87 F3: α =0.88 F4: α =0.92 F5: α =0.85 F6: α =0.87 F7: α =0.84 F8: α =0.90 F9: α =0.86 F10: α =0.87	NR
De Melo Ghisi et al. (2014)	Portuguese version of INCR	Brazil	patient with CAD	NR	55-item 10 dimension 5 points Likert expert panel	300	NR	NR	F1: The heart(physiology,symptoms, surgical treatments),F2:Nutrition F3:Exercise/physical activity,F4:Medication F5:Work/vocational/social F6:Stress/psychological factors F7:General/social concerns F8:Emergency/safety F9:Diagnosis and treatment F10:Risk factors	NR	Total α: NRF1: α = 0.83 F2: α = 0.84 F3: α =0.91 F4: α =0.83 F5: α =0.71 F6: α =0.81 F7: A=NR F8: α =0.80 F9: α =0.72 F10: α =0.93	NR
Sayadi et al. (2021)	CNCR-Q	Iran	patient with CAD	patient with CAD(n==10)	37-item 4 dimension 5 points Likert and 4 points expert panel CVI CVR	200	EFA KMO=0.76Bartlett'sSphericity tests	PCA Eigenvalues>1 screen plot	F1:physical care needs F2:Spiritual psychological care needs F3:Social family care needs f4: requierment for discharge plane	40.467	Total α: 0.78 F1: α =0.85 F2: α =0.83 F3: α =0.73 F4: α =0.39	NR

Table 1. Continued

Author (year)	Scale	Country	Target population	Face validity	Content validity	Construct Validity				Reliability		
						Sample size	Factor extraction method (Rotation)	Selection of the number of factors	name of factors	Total variance (%)	Consistency	Stability
Enguen et al. (2012)	Dutch clinical algorithm for assessing patient needs in Cardiac Rehabilitation	Nether-land	patient with CAD	NR	5 dimension expert panel	NR	NR	NR	F1: Physical functioning F2:Psychological functioning F3:Distruption or treat to social functioning F4: Cardiovascular risk profile F5: Lifestyle	NR	NR	NR
Smith et al. (2015)	CRNAT	Australia	patient with CAD	NR	expert panel	NR	NR	NR	NR	NR	NR	NR
Phelan et al. (1989)	A Patient Self-Assessment Tool for Cardiac Rehabilitation	USA	patient with CAD	NR	38-item 11 dimension	31	NR	NR	F1: Health perception F2:Nutrition F3:Elimination F4:Activity/exercise F5:Perceptual F6:Sleep/rest patients F7:Self-perception/concept F8:Roles/relationship F9:Sextuality F10:Coping/stress F11:Value/belief F12:Discharge planning F13:Teaching needs	NR	NR	NR

Notes: CADE-Q: Coronary Artery Disease Education-Questionnaire, CAD: Coronary Artery Disease, EFA: Exploratory Factor Analysis, KMO: Kaiser-Meyer-Olkin, PCA: Principal Component Analysis, ICC: Intra- Class Correlation, CFA: Confirmatory Factor Analysis, NR: Not Reported, CADE-Q SV Coronary Artery Disease Education-Questionnaire Short Version, INCR: Information Needs in Cardiac Rehabilitation, CNCR-Q: Care Needs in Cardiac Rehabilitation Questionnaire, CRNAT: Cardiac Rehabilitation Needs Assessment Tool, α: alpha, KR; Kuder Richardson

Appendix 2

Table 2. Scoring system of internal consistency (Box A)

Box A. Internal consistency	Excellent	Good	Fair	Poor
1 Does the scale consist of effect indicators, i.e. is it based on a reflective model?				
Design requirements				
2 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
3 Was there a description of how missing items were handled	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
4 Was the sample size included in the internal consistency analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
5 Was the unidimensionality of the scale checked? i.e. was factor analysis or IRT model applied?	Factor analysis performed in the study population	Authors refer to another study in which factor analysis was performed in a similar study population	Authors refer to another study in which factor analysis was performed, but not in a similar study population	Factor analysis NOT performed and no reference to another stud
6 Was the sample size included in the unidimensionality analysis adequate?	7* #items and ≥100	5* #items and ≥100 OR 6-7* #items but <100	5* #items but <100	<5#items
7 Was an internal consistency statistic calculated for each (unidimensional) (sub)scale separately?	Internal consistency statistic calculated for each subscale separately			Internal consistency statistic NOT calculated for each subscale separately
8 Were there any important flaws in the design or methods of the study	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods 9 for Classical Test Theory (CTT), continuous scores: Was Cronbach's alpha calculated?	Cronbach's alpha calculated		Only item-total correlations calculated	No Cronbach's alpha and no item-total correlations calculated
10 for CTT, dichotomous scores: Was Cronbach's alpha or KR-20 calculated?	Cronbach's alpha or KR-20 calculated		Only item-total correlations calculated	No Cronbach's alpha or KR-20 and no itemtotal correlations calculated
11 for IRT: Was a goodness of fit statistic at a global level calculated? E.g. $\chi 2$, reliability coefficient of estimated latent trait value (index of (subject or item) separation)	Goodness of fit statistic at a global level calculated			Goodness of fit statistic at a global level NOT calculated

Table 3. Scoring system of reliability (Box B)

Box B. Reliability: relative measures (including test-retest reliability, inter-rater reliability and intra-rater reliability)	Excellent	Good	Fair	Poor
Design requirements 1 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3 Was the sample size included in the analysis adequate?	Adequate sample size (≥100)	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4 Were at least two measurements available?	At least two measurements			Only one measurement
5 Were the administrations independent?	Independent measurements	Assumable that the measurements were independent	Doubtful whether the measurements were independent	measurements NOT independent
6 Was the time interval stated?	Time interval stated		Time interval NOT stated	
7 Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable
8 Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate	Time interval NOT appropriate
9 Were the test conditions similar for both measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar
10 Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods 11 for continuous scores: Was an intraclass correlation coefficient (ICC) calculated?)	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated
12 for dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			Only percentage agreement calculated
13 for ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated	Only percentage agreement calculated
14 for ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described		

Table 4. Scoring system of content validity (Box D)

Box D. Content validity (including face validity)	Excellent	Good	Fair	Poor
General requirements 1 Was there an assessment of whether all items refer to relevant aspects of the construct to be measured?	Assessed if all items refer to relevant aspects of the construct to be measured		Aspects of the construct to be measured poorly described AND this was not taken into consideration	NOT assessed if all items refer to relevant aspects of the construct to be measured
2 Was there an assessment of whether all items are relevant for the study population? (e.g. age, gender, disease characteristics, country, setting)	Assessed if all items are relevant for the study population in adequate sample size (≥10)	Assessed if all items are relevant for the study population in moderate sample size (5-9)	Assessed if all items are relevant for the study population in small sample size (	NOT assessed if all items are relevant for the study population OR target population not involved
3 Was there an assessment of whether all items are relevant for the purpose of the measurement instrument? (discriminative, evaluative, and/or predictive)	Assessed if all items are relevant for the purpose of the application	Purpose of the instrument was not described but assumed	NOT assessed if all items are relevant for the purpose of the application	
4 Was there an assessment of whether all items together comprehensively reflect the construct to be measured?	Assessed if all items together comprehensively reflect the construct to be measured		No theoretical foundation of the construct and this was not taken into consideration	NOT assessed if all items together comprehensively reflect the construct to be measured
5 Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study

Table 5. Scoring system of structural validity (Box E)

Box E. Structural validity	Excellent	Good	Fair	Poor
1 Does the scale consist of effect indicators, i.e. is it based on a reflective model?				
Design requirements 2 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
3 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
4 Was the sample size included in the analysis adequate?	7* #items and ≥100	5* #items and ≥100 OR 5-7* #items but <100	5* #items but<100	5* #items
5 Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study (e.g. rotation method not described)	Other important methodological flaws in the design or execution of the study (e.g. inappropriate rotation method)
Statistical methods 6 for CTT: Was exploratory or confirmatory factor analysis performed?	Exploratory or confirmatory factor analysis performed and type of factor analysis appropriate in view of existing information	Exploratory factor analysis performed while confirmatory would have been more appropriate		No exploratory or confirmatory factor analysis performed
7 for IRT: Were IRT tests for determining the (uni-) dimensionality of the items performed?	RT test for determining (uni)dimensionality performed			IRT test for determining (uni)dimensionality NOT performed

Table 6. Scoring system of cross-cultural validity (Box G)

Box G. Cross-cultural validity	Excellent	Good	Fair	Poor
Design requirements	Percentage of missing items described	Percentage of missing items NOT described		
1 Was the percentage of missing items given?				
2 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled.	Not clear how missing items were handled.	
3 Was the sample size included in the analysis adequate?	CTT: 7* #items and ≥100 IRT: ≥200 per group	CTT: 5* #items and ≥100 OR 5-7* #items but<100 IRT:>200 IN 1 group and 100-199 in 1 group	CTT: 5* #items but<100 IRT:100-199 per group	CTT:< 5* #items IRT:(<100 in 1 or both group)
4 Were both the original language in which the HR-PRO instrument was developed, and the language in which the HR-PRO instrument was translated described?	Both source language and target language described			Source language NOT known
5 Was the expertise of the people involved in the translation process adequately described? e.g. expertise in the disease(s) involved, expertise in the construct to be measured, expertise in both languages	Expertise of the translators described with respect to disease, construct, and language	Expertise of the translators with respect to disease or construct poor or not described	Expertise of the translators with respect to language not described	
6 Did the translators work independently from each other?	Translators worked independent	Assumable that the translators worked independent	Unclear whether translators worked independent	Translators worked NOT independent
7 Were items translated forward and backward?	Multiple forward and multiple backward translations	Multiple forward translations but one backward translation	One forward and one backward translation	Only a forward translation
8 Was there an adequate description of how differences between the original and translated versions were resolved?	Adequate description of how differences between translators were resolved.	Poorly or NOT described how differences between translators were resolved.		
9 Was the translation reviewed by a committee (e.g. original developers)?	Translation reviewed by a committee (involving other people than the translators, e.g. the original developers)	Translation NOT reviewed by (such) a committee		
10 Was the HR-PRO instrument pre-tested (e.g. cognitive interviews) to check interpretation, cultural relevance of the translation, and ease of comprehension	Translated instrument pretested in the target population	Translated instrument pretested, but unclear if this was done in the target population	Translated instrument pretested, but NOT in the target population	Translated instrument NOT pre-tested
11 Was the sample used in the pre-test adequately described?	Sample used in the pre-test adequately described		Sample used in the pre-test NOT (adequately) described	
12 Were the samples similar for all characteristics except language and/or cultural background?	Shown that samples were similar for all characteristics except language /culture	Stated (but not shown) that samples were similar for all characteristics except language /culture	Unclear whether samples were similar for all characteristics except language /culture	Samples were NOT similar for all characteristics except language /culture

Table 6. Continued

Box G. Cross-cultural validity	Excellent	Good	Fair	Poor
13 Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods 14 for CTT: Was confirmatory factor analysis performed?	Multiple-group confirmatory factor analysis performed			Multiple-group confirmatory factor analysis NOT performed
15 for IRT: Was differential item function (DIF) between language groups assessed?	DIF between language groups assessed.			DIF between language groups NOT assessed

Table 7. Scoring System of Criterion Validity (Box H)

Box H. Criterion validity	Excellent	Good	Fair	Poor
Design requirements 1 Was the percentage of missing items given?	Percentage of missing items described	Percentage of missing items NOT described		
2 Was there a description of how missing items were handled?	Described how missing items were handled	Not described but it can be deduced how missing items were handled	Not clear how missing items were handled	
3 Was the sample size included in the analysis adequate?	Adequate sample size (≥100	Good sample size (50-99)	Moderate sample size (30-49)	Small sample size (<30)
4 Can the criterion used or employed be considered as a reasonable 'gold standard'?	Criterion used can be considered an adequate 'gold standard' (evidence provided	No evidence provided, but assumable that the criterion used can be considered an adequate 'gold standard'	Unclear whether the criterion used can be considered an adequate 'gold standard'	Criterion used can NOT be considered an adequate 'gold standard'
5 Were there any important flaws in the design or methods of the study?	No other important methodological flaws in the design or execution of the study		Other minor methodological flaws in the design or execution of the study	Other important methodological flaws in the design or execution of the study
Statistical methods 6 for continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated
7 for dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated