Ageing Research Reviews 76 (2022) 101588

Contents lists available at ScienceDirect

Ageing Research Reviews

journal homepage: www.elsevier.com/locate/arr

Review

The Tilburg Frailty Indicator: A psychometric systematic review

Juan-José Zamora-Sánchez [a,b,1], Ana-María Urpí-Fernández [a,2], Meritxell Sastre-Rus [c,3], Iris Lumillo-Gutiérrez [b,d,4], Vicente Gea-Caballero [e,f,g,5], Lina Jodar-Fernández [h,6], Iván Julián-Rochina [i,j,7], Edurne Zabaleta-del-Olmo [a,k,l,*,8]

[a] Gerència Territorial de Barcelona, Institut Català de la Salut, Barcelona, Spain
[b] School of Nursing, Universitat de Barcelona, Barcelona, Spain
[c] Escola Universitària d'Infermeria Gimbernat, Universitat Autònoma de Barcelona, Sant Cugat del Vallès, Barcelona, Spain
[d] Chronic Disease Management Team, Baix Llobregat Centre Primary Care Service, Costa de Ponent Primary Care Directorate, Gerència Territorial Metropolitana Sud, Institut Català de la Salut, Cornellà de Llobregat, Spain
[e] Deanery, Faculty of Health Sciences, Valencian International University, Valencia, Spain
[f] Nursing School "La Fe", University of Valencia, Valencia, Spain
[g] Instituto de Investigación Sanitaria del Hospital Universitario La Paz (IdiPAZ), Madrid, Spain
[h] Montbaig Primary Care Centre, Delta Primary Care Service, Costa de Ponent Primary Care Directorate, Gerència Territorial Metropolitana Sud, Institut Català de la Salut, Viladecans, Spain
[i] Nursing Department, Universitat de València, Valencia, Spain
[j] Frailty and Cognitive Impairment Group (FROG), Universitat de València, Valencia, Spain
[k] Nursing Department, Faculty of Nursing, Universitat de Girona, Girona, Spain
[l] Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain

ARTICLE INFO

Keywords:
Aged
Frailty
Patient-reported outcome measures
Psychometrics
Systematic Review

ABSTRACT

Background: The Tilburg Frailty Indicator (TFI) is one of the most prominent multidimensional frailty assessment instruments. This review aimed to critically appraise and summarise its measurement properties.
Methods: Reports were eligible if they included results of studies aimed at developing the TFI or evaluating its measurement properties. We performed a literature search in MEDLINE, CINAHL, and PsycINFO databases from their inception until December 8, 2021. We also searched grey literature databases. We assessed the methodological quality of the included studies using the "COSMIN Risk of Bias". The measurement properties were evaluated using specific criteria. We graded the quality of the evidence using a GRADE approach.
Results: Sixty-three studies were included. We found moderate sufficient evidence for TFI content validity, although it is still insufficient for the comprehensiveness of its items. TFI construct validity was based on sufficient evidence from two studies of its structural validity as well as multiple hypothesis-testing for construct validity studies with inconsistent results. We did not find any studies that assessed cross-cultural validity. Only one of TFI's three dimensions showed sufficient evidence for the internal consistency of its scores, and results in test-retest reliability were inconsistent. The TFI showed high sufficient concurrent validity with the comprehensive geriatric assessment. We identified several studies assessing its predictive validity for adverse frailty-related outcomes, although most of the evidence from these studies was insufficient. We did not find any studies that assessed the responsiveness of TFI scores.

* Correspondence to: Gran Via Corts Catalanes, 587 àtic, 08007 Barcelona, Spain.
E-mail addresses: jjzamora.bcn.ics@gencat.cat (J.-J. Zamora-Sánchez), amurpics@gmail.com (A.-M. Urpí-Fernández), meritxell.sastre@eug.es (M. Sastre-Rus), irislumillo.bcn.ics@gencat.cat (I. Lumillo-Gutiérrez), vagea@universidadviu.com (V. Gea-Caballero), ljodar@ambitcp.catsalut.net (L. Jodar-Fernández), ivan.julian@uv.es (I. Julián-Rochina), ezabaleta@idiapjgol.org, nieves.zabaleta@udg.edu (E. Zabaleta-del-Olmo).
[1] ORCID iD: https://orcid.org/0000–0002-3487–7789
[2] ORCID iD: https://orcid.org/0000–0002-2446–1077
[3] ORCID iD: https://orcid.org/0000–0002-7043–0706
[4] ORCID iD: https://orcid.org/0000–0002-4787–3939
[5] ORCID iD: https://orcid.org/0000–0001-8607–3195
[6] ORCID iD: https://orcid.org/0000–0003-4316–6724
[7] ORCID iD: https://orcid.org/0000–0002-1782–0019
[8] ORCID iD: https://orcid.org/0000–0002-5072–8548

https://doi.org/10.1016/j.arr.2022.101588
Received 13 November 2021; Received in revised form 11 January 2022; Accepted 7 February 2022
Available online 10 February 2022

Conclusions: The TFI had evidence gaps in several relevant measurement properties. Further research is needed to strengthen its usefulness as a clinical decision-making tool.

1. Introduction

Identifying older people who are frail or at risk of becoming frail has become the centrepiece of geriatric care in recent years (Hoogendijk et al., 2019). This identification can be approached from three perspectives: predominantly clinical, multidimensional, and holistic functional (Junius-Walker et al., 2018). Regarding the clinical perspective, frailty is understood as a clinical state determined by a series of signs and symptoms (Xue, 2011); however, this perspective may be insufficient to capture the full range of potential criteria determining frailty (Junius--Walker et al., 2018; Sezgin et al., 2019). Moreover, other researchers suggest that frailty should not be limited to the physical domain and they advocate carrying out a multidimensional assessment that includes other factors such as psychological, cognitive, emotional, social or spiritual (Junius-Walker et al., 2018; Sezgin et al., 2019). On the other hand, World Health Organisation recommends holistic functional perspective, which includes a multidimensional approach to frailty. Nevertheless, its focus is on "total person functioning" rather than deficits and diseases (Junius-Walker et al., 2018). Because of this, it seems that the multidimensional perspective would be the most appropriate to assess all the possible factors involved in frailty (Sezgin et al., 2019).

Among the instruments available to perform a multidimensional assessment of frailty, the Tilburg Frailty Indicator (TFI) (Gobbens et al., 2010b) is prominent. It is based on an integral conceptual model of frailty which authors define as a "dynamic state affecting an individual who experiences losses in one or more domains of human functioning (physical, psychological, social), which is caused by the influence of a range of variables, and which increases the risk of adverse outcomes" (Gobbens et al., 2010a). Therefore, the TFI assesses physical, psychological, and social dimensions of frailty. Existing systematic reviews show that it is one of the most robust instruments, especially to use in primary care (Pialoux et al., 2012; Sutton et al., 2016). However, these systematic reviews conclude that further in-depth evaluation of the measurement properties of TFI is needed. Moreover, these previous reviews do not cover all available evidence due to their completion date and focus on evaluating several measurement instruments rather than on a single instrument. Therefore, a specific systematic review on TFI may be more efficient in identifying all available evidence.

Developing and improving an effective individual care plan that meets the person's life goals involve a regular and multidimensional assessment of the person's needs (Dent et al., 2019). Multidimensional measures of frailty can provide more insight into these needs and enable tailored care management. The TFI is a multidimensional frailty assessment tool frequently used in both clinical practice and geriatric research, so there is a need to update the available evidence to determine its suitability for this purpose. This systematic and psychometric review aimed to critically appraise and summarise the measurement properties of TFI to support evidence-based recommendations on its use and identify gaps in knowledge on its measurement properties, which can be used to design new studies.

2. Methods

We conducted this psychometric systematic review following the COSMIN initiative (Mokkink et al., 2018; Prinsen et al., 2018; Terwee et al., 2018). The review is reporting according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) 2020 (Page et al., 2021) statement and the PRISMA literature search extension (Rethlefsen et al., 2021). The review protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO) on January 4, 2021 (registration number: CRD42021224427); no changes have been made to the protocol.

2.1. Search strategy and eligibility criteria

We performed a comprehensive literature search in MEDLINE (PubMed), CINAHL (EBSCOhost), and PsycINFO (EBSCOhost) databases from their inception until December 8, 2021. We also searched the grey literature on OpenGrey and Grey literature Report databases. Reports were eligible if they included results of studies aimed at developing the TFI or evaluating one or more of its measurement properties. Studies that only use the TFI as an outcome measure (e.g., clinical trials) or those used to validate another measurement instrument were excluded. Only full-text reports were included because the minimum information about a study is often found in the abstracts. Likewise, for our MEDLINE (PubMed) search, we added a highly sensitive filter to identify studies on measurement properties(Terwee et al., 2009). We imposed no language restrictions on any of the searches. The reproducible searches for all databases are available at https://doi.org/10.5281/zenodo.5513482. We manually screened reference lists of included studies to identify additional studies.

2.2. Selection process

We imported the retrieved references into the Rayyan QCRI web application program (Ouzzani et al., 2016). Two reviewers removed duplicates using the program's duplicate identification strategy and then manually. These two reviewers independently assessed the titles and abstracts of the retrieved records against the eligibility criteria. If a record seemed relevant to at least one of the reviewers, the full text of the report was independently reviewed by these two reviewers. Reviewers discussed conflicts over inclusion, and a third reviewer was consulted in case of not reaching a consensus.

2.3. Data collection process and study risk of bias assessment

We assessed the methodological quality of the included studies using the "COSMIN Risk of Bias" (available at: https://www.cosmin.nl/wp-content/uploads/COSMIN-RoB-checklist-V2–0-v17_rev3.pdf). The evaluation was carried out by each member of the review team independently. All assessments were compared, discussed, and agreed at regular meetings. Evaluation data were collected using forms designed by COSMIN (available at https://cosmin.nl/wp-content/uploads/Scoring-form-COSMIN-boxes_april_final.xlsx).

We considered that the TFI was based on a reflective model, and consequently, we assessed its structural validity and the internal consistency of its scores (Mokkink et al., 2018). Regarding criterion validity, we agreed, based on the available literature on frailty and following the COSMIN guidelines, that only comprehensive geriatric assessment (CGA) could be considered a gold standard for TFI (Hoogendijk et al., 2019; Mokkink et al., 2018; Parker et al., 2018; Prinsen et al., 2018). Predictive validity was assessed, as most of the identified studies aimed to assess this property; however, only longitudinal studies were assessed. We did not assess cross-sectional studies for this property; even though the authors described them as predictive validity studies, we considered them for construct validity.

2.4. Synthesis methods

The measurement properties were evaluated using specific criteria

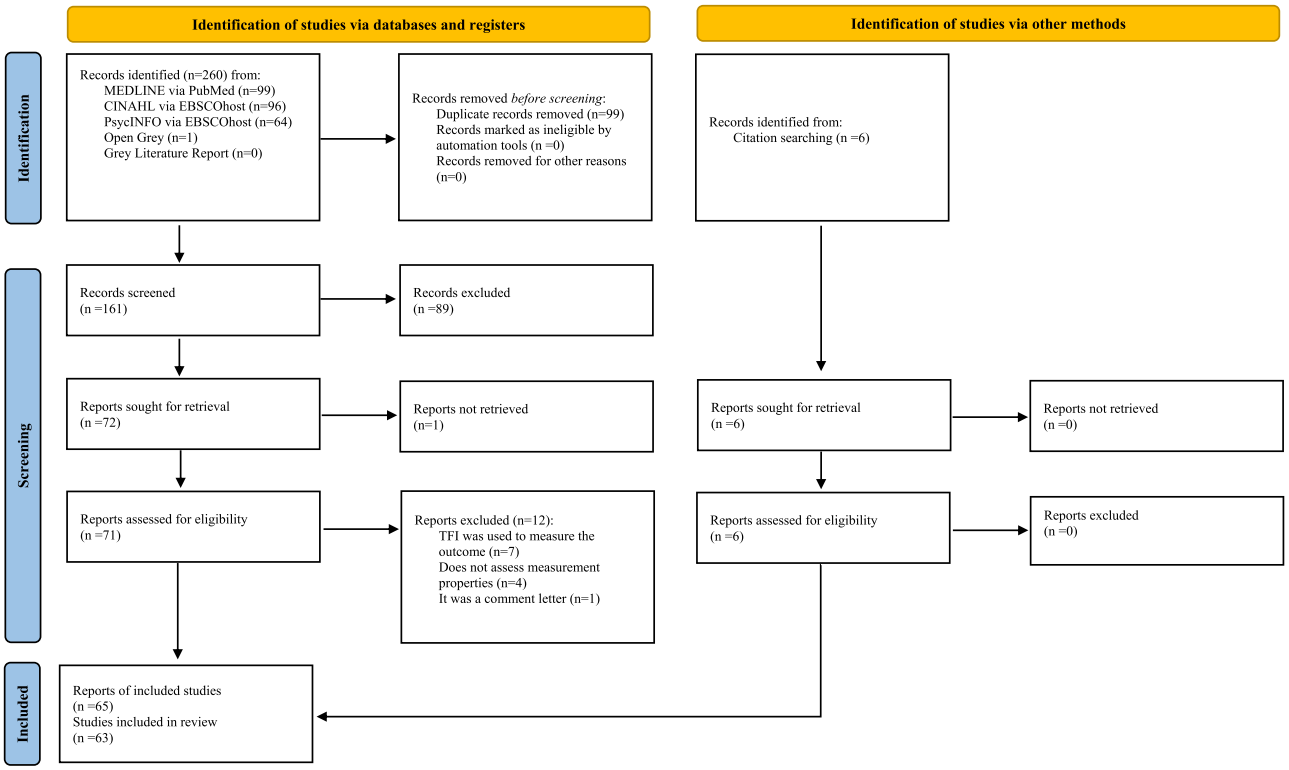

Fig. 1. PRISMA 2020 Flow-Chart of Study Selection (Page et al., 2021).

developed and agreed by experts (Prinsen et al., 2018). This assessment was carried out based on the number of studies available and the consistency of their results. We pooled the internal consistency results of the TFI scores by the Meta-Essentials tool (Suurmond et al., 2017) for correlational data version 1.5. We chose random-effects models based on the diversity of populations studied and in the expectation that internal consistency coefficients would differ. The meta-analyses were run with the Fisher-transformed values, which are transformed back into normal correlation coefficients for presentation (van Rhee et al., 2015). The extent and impact of between study heterogeneity were assessed by the tau^2 and the I^2 statistics, respectively. We explored possible causes of variation of results across studies by subgroup analyses based on the mode of administration. We summarised the rest of the measurement properties qualitatively.

The quality of the evidence was graded using a "Grading of Recommendations Assessment, Development and Evaluation (GRADE)" approach modified by COSMIN (Prinsen et al., 2018). As with the assessment of studies' quality, all these assessments were made independently by each review team member and subsequently compared, discussed, and agreed at regular meetings.

3. Results

The literature search and study selection process is detailed in Fig. 1. A total of 63 studies were included; a study had three reports (Hayajneh, 2019, 2016; Hayajneh et al., 2021). The summary characteristics of the first study (TFI development) and all included studies are shown in Table 1 and Supplementary file Table 1. Twenty of the 63 studies were from The Netherlands, 18 from other European countries, 16 from Asian countries, five from Brazil, two from Turkey, and two were multi-country studies. All studies assessed the measurement properties of part B of the TFI (described in Table 1). In addition, the 13 reports that were excluded and the reasons for exclusion are listed in Supplementary file Text 1.

The results of the methodological quality assessment of each study on a measurement property using the COSMIN Risk of Bias checklist (Mokkink et al., 2018) are shown in Supplementary file Table 2 and Table 2. Supplementary file Table 3 provides the ratings of each study against the criteria for good measurement properties (Prinsen et al., 2018). Finally, Table 3 summarises the evidence and its grade of quality by using the GRADE approach (Prinsen et al., 2018).

Table 1
Main characteristics of the Tilburg Frailty Indicator (Gobbens et al., 2010b).

Target population	Mode of administration	Recall period	(Sub)scale (s) (number of items)	Response options	Range of scores/scoring	Original language	Available translations
Community-dwelling older people	Self-administered	Now	It consists of 2 parts:- Part A contains ten questions on determinants of frailty and diseases (multimorbidity) - Part B contains 15 questions divided into three domains of frailty: physical (8 questions), psychological (4 questions), and social (3 questins)	- Yes/No (11 questions)- Yes/ Sometimes /No (4 questions)	- The score for frailty and the three domains of frailty are determined by adding items' responses belonging to each scale. The response options "sometimes" and "yes" are grouped into a single score.- Score ranges from 0 to 15.- Frailty cut off point≥ 5 - Higher score=Higher degree of frailty	Dutch	-Arabic (Jordan & Saudi Arabia)-Chinese (China)-Croatian -Danish -English (UK)-French (France)-German -Greek -Italian -Persian -Polish-Portuguese (Brazil & Portugal)-Spanish (Spain)-Turkish

Table 2
Quality of studies on measurement properties of Tilburg Frailty Indicator.

	Instrument development	Content validity					Structural validity	Internal consistency	Cross-cultural validity	Reliability	Measurement error	Criterion validity	Construct validity	
		Asking patients			Asking experts									
Study		Relevance	Comprehensiveness	Comprehensibility	Relevance	Comprehensiveness							Convergent validity	Known groups validity
Gobbens et al. (2010b)	I							V		D		D	V	
Metzelthin et al. (2010)								I					V	
Daniels et al. (2012)												D		
Gobbens and van Assen (2012)												D		
Gobbens et al. (2012a)												D	D	
Gobbens et al. (2012b)								V						
Gobbens et al. (2012c)													D	
Santiago et al. (2012)				D										
De Witte et al. (2013)													V	D
Gobbens et al. (2013)								V					V	
Santiago et al. (2013)								V		A			D	
Andreasen et al. (2014)				A										
Gobbens and van Assen (2014)								V				D		
Gobbens et al. (2014)								V				D		
Uchmanowicz et al. (2014)				D				I						
Andreasen et al. (2015)		D	D											
Coelho et al. (2015)				D				V		D			V	
Gobbens et al. (2015)													D	
Mulasso et al. (2015)				D										
Roppolo et al. (2015)													V	
Uchmanowicz and Gobbens (2015)													V	D
Uchmanowicz et al. (2015)													V	D
Freitag et al. (2016)				D				V		I			V	
Mulasso et al. (2016)								V					V	
Uchmanowicz et al. (2016)								V		D				
Chong et al. (2017)												I		
Dong et al. (2017)				D				V		D			V	
Gobbens (2017)													V	D
Gobbens and van Assen (2017)													V	
Gonzalez-Colaço Harmand et al. (2017)												D	V	
Mulasso et al. (2017)												I		V
Andreasen et al. (2018)												D		
Chong et al. (2018)												D		
Renne and Gobbens (2018)								V					V	D
Santiago et al. (2018)												I		
Uchmanowicz et al. (2018)												V		
van der Vorst et al. (2018)												I		
Vrotsou et al. (2018)				I			V	V		A	A		V	D
Hayajneh, (2016)Hayajneh, (2019) Hayajneh et al. (2021)				D				V					V	D
Kendhapedi and Devasenapathy (2019)													V	
Op het Veld et al. (2019a)												V		
Op het Veld et al. (2019b)												V		
Santiago et al. (2019)													V	
Topcu et al. (2019)				I				I		D				
Zhang et al. (2019)														V
Alqahtani et al. (2020)				D				V		D			V	
Giacomini et al. (2020)													V	D
Gobbens and Andreasen (2020)												I		
Gobbens et al. (2020)												D		
Mazoochi et al. (2020)												V		
Si et al. (2020)												V		
Xie et al. (2020)													V	
Yang et al. (2020)													V	

(*continued on next page*)

Table 2 (*continued*)

Study	Instrument development	Content validity				Structural validity	Internal consistency	Cross-cultural validity	Reliability	Measurement error	Criterion validity	Construct validity	
		Asking patients		Asking experts								Convergent validity	Known groups validity
		Relevance	Comprehensiveness	Comprehensibility	Relevance	Comprehensiveness							
Zhang et al. (2020)							V				V	V	
Gobbens et al. (2021)											V		
Ding et al. (2021)											V	V	
Lin et al. (2021a)				D		V	V		D			V	D
Lin et al. (2021b)												V	V
Ozsoy et al. (2021),												V	V
Papathanasiou et al. (2021)							V					V	
Qin et al. (2021)													V
Si et al. (2021a)											D		
Si et al. (2021b)											D		

V, Very good; A, Adequate; D, Doubtful; I, Inadequate. Empty cells indicate that this measurement property had not been assessed.

3.1. Content validity

Description and origin of the construct to be measured are clear, as well as the target population and the context of use. The TFI was developed based on previous research on frailty, but only experts were involved (Gobbens et al., 2010a). Therefore, it appears that its development was not performed on a sample representing the target population. Comprehensibility and comprehensiveness are tested in a prefinal form of the instrument by representatives of professional disciplines (n = 10) and by people aged 75 years and older (n = 33), but these samples were not clearly described. No adjustment was necessary as this pilot test indicated that this version of the instrument was clear and comprehensive. However, the method and approach used to analyse data were not reported. For these reasons, the total quality of instrument design and the pilot test performance were rated as inadequate and doubtful, respectively. This meant that the total rating of the instrument development was inadequate (see Supplementary file Table 2).

We identified 13 studies that analysed content validity aspects (see Table 2). Twelve of these studies assessed the comprehensibility of TFI versions translated into other languages. Among these studies, the study of the Danish version was the most salient (Andreasen et al., 2014). The quality of ten studies was assessed as doubtful mainly because the methodology's crucial aspects were not sufficiently clarified. One study on the TFI's Turkish version (Topcu et al., 2019) was assessed as inadequate because it did not describe that a cognitive interview or pilot test had been conducted in the target population. Moreover, in this study, the comprehensibility of the instrument was assessed by ten geriatricians. Only one study assessed aspects linked to the relevance and comprehensiveness of the instrument (Andreasen et al., 2015). In this study conducted with the TFI's Danish version, participants confirmed that the instrument covered most aspects of frailty. However, they identified aspects that they believed were not covered. The methodological quality of this study was rated as doubtful because it did not describe the experience with qualitative methods of the interviewers.

In summary (see Table 3), the content validity of the TFI was considered sufficient with moderate quality evidence, as the methodological quality of the studies was mostly doubtful. Evidence on comprehensiveness was considered insufficient as one content validity study indicated that the instrument did not cover important aspects related to frailty; in contrast, the evidence on the relevance and comprehensibility of the instrument was considered sufficient, although, concerning the latter aspect of content validity, the quality of evidence was considered high as the methodological quality of one of the content validity studies was adequate.

3.2. Construct validity

Two studies assessed the structural validity of the TFI (Lin et al., 2021a; Vrotsou et al., 2018), and 36 studies performed hypotheses testing for its construct validity. We did not identify any study that assessed its cross-cultural validity. Structural validity was analysed in the Spanish and Taiwanese TFI versions, and the results indicated a good fit of the three-factor model. The methodological quality of these studies was rated very good, so the evidence on this measurement property was graded as high. We identified a total of 298 hypotheses for construct validity of which we considered 220 (74%) to be confirmed according to the criteria proposed by COSMIN (Prinsen et al., 2018). Nevertheless, the results were inconsistent, so we did not grade the evidence.

3.3. Criterion validity

Concurrent validity with CGA was assessed in three studies (Mazoochi et al., 2020; Si et al., 2020, 2021a), for which the evidence was rated as sufficient with a high-quality grade. Several studies assessed the predictive validity of TFI for different adverse frailty-related outcomes such as higher health care use, disability,

Table 3
Summary of findings.

Measurement property	Summary or pooled result	Overall rating	Quality of evidence
Content validity	Not applicable	Sufficient	Moderate: TFI development study inadequate quality, and there are multiple content validity studies of doubtful quality available
Relevance	Not applicable	Sufficient	Moderate: TFI development study inadequate quality, and there is one content validity study of doubtful quality available
Comprehensiveness	Not applicable	Insufficient	Moderate: TFI development study inadequate quality, and there is one content validity study of doubtful quality available
Comprehensibility	Not applicable	Sufficient	High: TFI development study inadequate quality, and there is at least one content validity study of adequate quality available
Structural validity	Three-factor model: CFI range 0.91–0.96 and RMSEA range 0.035–0.005	Sufficient	High: There are two studies of very good quality available
Internal consistency			
Physical domain	Global: coefficient 0.70 (95% CI, 0.67–0.73), n = 8203, $I^2 = 67.1\%$	Sufficient	Moderate: There are multiple studies of very good quality available, and there is moderate inconsistency
	Self-administered mode: coefficient 0.68 (95% CI, 0.65–0.71), n = 7218, $I^2 = 51.8\%$	Sufficient	Moderate: There is multiple studies of very good quality available, and there is moderate inconsistency
	Interviewer-administered mode: coefficient 0.75 (95% CI, 0.71–0.79), n = 985, $I^2 = 12.7\%$	Sufficient	High: There is multiple studies of very good quality available, and there is low inconsistency
Psychological domain	Global: coefficient 0.56 (95% CI, 0.50–0.62), n = 7774, $I^2 = 84.6\%$	Insufficient	Low: There is multiple studies of very good quality available, and there is high inconsistency
	Self-administered mode: coefficient 0.57 (95% CI, 0.51–0.62), n = 6789, $I^2 = 77.3\%$	Insufficient	Low: There is multiple studies of very good quality available, and there is high inconsistency
	Interviewer-administered mode: coefficient 0.55 (95% CI, 0.34–0.71), n = 985, $I^2 = 91.3\%$	Sufficient	Low: There is multiple studies of very good quality available, and there is high inconsistency
Social domain	Global: coefficient 0.44 (95% CI, 0.34–0.52), n = 7970, $I^2 = 93.9\%$	Insufficient	Low: There is multiple studies of very good quality available, and there is high inconsistency
	Self-administered mode: coefficient 0.38 (95% CI, 0.28–0.46), n = 6985, $I^2 = 91.4\%$	Insufficient	Low: There is multiple studies of very good quality available, and there is high inconsistency
	Interviewer-administered mode: coefficient 0.55 (95% CI, 0.31–0.72), n = 985, $I^2 = 92.5\%$	Sufficient	Low: There is multiple studies of very good quality available, and there is high inconsistency
Cross-cultural validity \Measurement invariance	No info available		
Reliability	ICC range 0.23–0.99	Inconsistent	
Measurement error	LoA were calculated but MIC was not defined	Indeterminate	
Criterion validity			
- *Concurrent validity*	3 out of 3 hypotheses confirmed (100%)	Sufficient	High: There are studies of very good quality available
- *Predictive validity*	Health care use: 18 out of 48 hypotheses confirmed (38%)	Insufficient	High: There are studies of very good quality available
	Disability (all studies): 6 out of 14 hypotheses confirmed (43%)	Insufficient	High: There are studies of very good quality available
	Disability (one year follow-up): 5 out of 9 hypotheses confirmed (56%)	Insufficient	Moderate: There are studies of doubtful quality available
	Disability (two-year follow-up): 1 out of 2 hypotheses confirmed (50%)	Insufficient	High: There are studies of very good quality available
	Disability (12-year follow-up): 1 out of 1 hypothesis confirmed (100%)	Sufficient	Low: There is one study of doubtful quality available
	Mortality (all studies): 6 out of 9 hypotheses confirmed (67%)	Insufficient	High: There are studies of very good quality available
	Mortality (one-year follow-up): 2 out of 2 hypotheses confirmed (100%)	Sufficient	Moderate: There are studies of doubtful quality available
	Mortality (two-year follow-up): 1 out of 1 hypothesis unconfirmed (0%)	Insufficient	High: There are studies of very good quality available
	Mortality (≥four-year follow-up): 3 out of 5 hypotheses confirmed (60%)	Insufficient	High: There is one study of very good quality available
	QoL: 1 out of 4 hypotheses confirmed (25%)	Insufficient	Moderate: There are studies of doubtful quality available
	Falls: 1 out of 5 hypotheses confirmed (20%)	Insufficient	Moderate: There are studies of doubtful quality available
Hypotheses testing for construct validity	220 out of 298 hypotheses confirmed (74%)	Inconsistent	
Responsiveness	No info available		

CFI, Comparative Fit Index; CI, Confidence Interval; ICC, Intraclass Correlation Coefficient; LoA, Limits of Agreement; MIC, Minimal Important Change; QoL, Quality of life; RMSEA, Root Mean Square Error of Approximation; TFI, Tilburg Frailty Indicator.

mortality, lower quality of life or falls. We also identified one study that used a composite outcome variable with unequal importance of events such as readmission or death (Andreasen et al., 2018). The most frequent follow-up period was one year and two years. In some studies, it was six months, four years, seven years, and 12 years. One study analysed the TFI prediction for in-hospital mortality (Chong et al., 2017). The percentage lost to follow-up, in studies providing this information, was less than 20% in ten studies and higher in seven. Health care use was assessed using a wide range of variables: hospital admission, hospitalisation, unplanned readmission, receiving personal care or nursing care or informal care, general practitioner visits, contacts with health care professionals, residential care facilities, or institutionalisation. In most of these studies, the participants self-reported these variables, which contributed to the study's methodological quality being rated as doubtful. Concerning disability, all the identified studies, except one that used the Katz index (Gonzalez-Colaço Harmand et al., 2017), used the Groningen Activity Restriction Scale. Quality of life was measured in identified studies with the brief version of the World Health Organisation Quality of Life (WHOQOL-BREF) questionnaire. The evidence for TFI predictive validity was rated as sufficient for 12-year disability

(Gonzalez-Colaço Harmand et al., 2017) and one-year mortality prediction (Chong et al., 2018; Daniels et al., 2012), the quality of this evidence was graded as low and moderate, respectively (see Table 3). It was rated as insufficient for all other adverse frailty-related outcomes studied.

3.4. Reliability

Twenty-one studies assessed the internal consistency of TFI scores. The methodological quality of these studies was mostly rated as very good, except for three studies that only reported overall internal consistency coefficient instead of each of three TFI domains (Metzelthin et al., 2010; Topcu et al., 2019; Uchmanowicz et al., 2014). We, therefore, rated the methodological quality of these three studies as inadequate. The pooled results showed that only the physical domain had sufficient evidence in this measurement property with a moderate quality due to its inconsistency (see Table 3). Likewise, the psychological and social domains only showed sufficient evidence when the instrument was interview-administered, although the quality of the evidence was low due to the high inconsistency observed. Test-retest reliability was assessed in ten studies. Statistical analyses were approached by treating the response options as continuous or dichotomous, whereby the studies provide information on intraclass correlation coefficient as well as Cohen's kappa index. However, two studies assessed this property using Pearson's correlation coefficient (Gobbens et al., 2010b; Lin et al., 2021a). The methodological quality of seven studies was rated as doubtful mainly because they did not provide evidence that participants were stable, it was unclear whether the test conditions were similar, or the time interval was not appropriate. The time interval was 20 weeks in one study, so its methodological quality was considered inadequate (Freitag et al., 2016). Two studies were of adequate methodological quality (Santiago et al., 2013; Vrotsou et al., 2018). We rated the overall evidence as inconsistent, so we did not grade its quality. Only one study assessed the measurement error of TFI scores (Vrotsou et al., 2018). In this study, the limits of agreement of scores were calculated, but there was no information about the minimal important change, so this evidence was rated indeterminate.

3.5. Responsiveness

We did not identify any study that assessed the responsiveness of a TFI change score.

4. Discussion

This psychometric review was conducted to assess the quality of TFI measurement properties. We found moderate sufficient evidence for TFI content validity, although it is still insufficient for aspects such as the comprehensiveness of its items. TFI construct validity was based on sufficient evidence from two studies of its structural validity, as well as multiple hypothesis-testing for construct validity studies with inconsistent results. However, we did not find any studies that assessed cross-cultural validity. Regarding criterion validity, the TFI showed high sufficient concurrent validity with the CGA. We also identified a substantial number of studies assessing its predictive validity for adverse frailty-related outcomes, although most of the evidence from these studies was insufficient. Internal consistency of scores was the most assessed measurement property; however, only the physical domain scale showed sufficient moderate evidence. We did not find any studies that assessed the responsiveness of TFI scores.

This review has included a much larger number of studies on the TFI compared to other reviews (Pialoux et al., 2012; Sutton et al., 2016). Focusing on a single measurement instrument has allowed for a more in-depth analysis. Furthermore, the fact of not using language limits may also have contributed to the identification of a more significant number of studies. On the other hand, scoring and grading the quality of methods, the interpretation of results and the grading of evidence is a subjective process. However, all included articles were independently reviewed by the reviewers and agreed by consensus amongst the reviewer team. This process helped to resolve discrepancies and reduce variability in interpretation.

In patient-reported outcome measures such as TFI, the target population is the most appropriate assessor of the content validity of a measurement instrument. The target population's comprehensibility of the TFI items has been evaluated in different studies due to the existence of multiple language versions. However, the assessment of the relevance and comprehensiveness of its items by the target population has been scarcely evaluated. We found only two studies that evaluated these content validity aspects (Andreasen et al., 2015; Gobbens et al., 2010b). Andreasen et al. (2015) find that most of the TFI items are relevant to older people; however, some items in the TFI physical domain do not seem relevant to them, such as those related to unintentional weight loss or hand strength. They also observe that the TFI does not consider some important aspects for this population, such as pain, sleep disturbances, spirituality, or meaningful activities.

We identified that structural validity was assessed in two studies (Lin et al., 2021a; Vrotsou et al., 2018) which confirmed a three-factor model for three TFI domains. However, it is surprising that no further studies have been found that analyse this measurement property since such a widely studied measurement instrument. In contrast, many studies have been carried out to analyse the convergent and discriminative validity of TFI scores. Multiple measurement instruments and variables were used for this purpose. It also highlights the large number of hypotheses tested in some studies. Hypotheses for construct validity were confirmed at a high rate (74%), but inconsistent results were observed regarding the association between TFI scores and variables measuring similar or related constructs. Moreover, most studies used *P*-values rather than to assess whether the magnitude of correlations or observed differences were similar or greater than expected (Prinsen et al., 2018). We did not find any studies that analysed cross-cultural validity of the TFI despite the numerous adaptations and translations that have been carried out. This type of validity is essential to determine the equivalence of scores between the original population and the new target population (Prinsen et al., 2018).

We found three studies showing good concurrent validity of TFI with CGA (Mazoochi et al., 2020; Si et al., 2020, 2021a). This finding has important implications, as CGA is a time-consuming and high-resource intensive process (Hoogendijk et al., 2019; Parker et al., 2018) that can be problematic in some healthcare settings such as primary care. In these settings, the TFI may be a simpler and more feasible tool to capture similar aspects of frailty. We identified a significant number of TFI validation studies for adverse frailty-related outcomes prediction, but their results are inconclusive. Measuring the predictive validity of a frailty measurement instrument is no easy matter. Although most people remain in their baseline frailty state at a follow-up of 1–5 years, a substantial proportion (up to 37%) experience at least one transition, including both worsening and improvement of the frailty state (Hoogendijk et al., 2019). Therefore, there is a need to conduct validation studies that consider frailty as a dynamic process.

Only the TFI physical domain subscale showed sufficient internal consistency of its scores. One possible explanation for this might be that the internal consistency coefficients are highly dependent on the number of items and the two remaining subscales have a tiny number of items. An alternative explanation is that the psychological and social domains are poorly comprehensive, as Andreasen et al. (2015) reported. We identified a substantial number of studies that administered the TFI by interview even though it was designed to be self-administered. One interesting finding is that the consistency of the results was different according to the mode of administration of the TFI, except in the social domain. This indicates a possible effect of the mode of administration of the TFI on their data. This influence has been described in the literature, and it may be necessary to conduct experimental studies to determine

the origin, magnitude, and direction of this influence (Bowling, 2005). Results in test-retest reliability were inconsistent, probably due to the variability of different methodological aspects of the studies, such as the lack of evidence of whether respondents were stable, the different time intervals used, and above all the fact that the instrument was not administered under similar conditions on both occasions. Test-retest reliability is an essential requirement of all measurement instruments in clinical practice and research, so this is a measurement property that needs further investigation. On the other hand, in order to evaluate the measurement error, it is necessary to have information on the minimal important change, defined as the smallest change in score that people consider important (Prinsen et al., 2018). We found no information on this issue, so there is also a need to generate evidence on it.

Responsiveness was not explored in any of the studies identified. The lack of assessment of this measurement property is common in frailty measurement instruments (Hoogendijk et al., 2019). However, it is a very relevant measurement property, as clinicians and researchers need measurement instruments that can be used to monitor changes in frailty over time.

5. Conclusions

This review found that the TFI had evidence gaps in several relevant measurement properties. There are important issues for future research, and more studies are needed to strengthen its usefulness as a clinical decision-making tool. A patient-reported outcome measure must be valid for a wide range of uses in different populations, and each use may require new evidence. However, most studies on the TFI focus on analysing the same measurement properties and other properties such as structural validity, cross-cultural validity, reliability, and responsiveness have received much less attention. Moreover, it would be interesting to establish the methodologies and evaluation criteria for testing hypotheses for construct validity. In this way, more homogeneous results would be obtained, which would give more strength to available evidence. On the other hand, it would also be desirable to conduct studies that focus on confirming fewer but strongly supported hypotheses. Finally, it is essential to highlight the importance of generating more evidence regarding the content validity of the TFI, especially in aspects related to the comprehensiveness of its items from the perspective of people who are frail.

Funding

This work was supported by the Health Department grant number from the Generalitat of Catalunya (Spain) (SLT008/18/00011); and by a grant from the Territorial Management of Barcelona of the Institut Català de la Salut (Catalan Institute of Health) to the first author (JJZS) in the 2018 and 2021 editions. The funders had no role in review design, decision to publish, or preparation of this manuscript.

Authors' contribution

JJZS and EZO conceived and designed the review. JJZS and EZO selected the reports. All authors extracted and analysed de data. All authors accessed and verified the data. EZO wrote the first draft of the manuscript. All authors interpreted the data and contributed to the writing of the final version of the manuscript. All authors agreed with the results and conclusions of this manuscript. All authors had full access to all the data in the review and had final responsibility for the decision to submit for publication.

Declarations of interest

None.

Data Availability

Data were extracted from published research papers. All data generated or analysed during this review are included in this manuscript and its supplementary information files.

Acknowledgements

The authors appreciate the review of the English text by Patryk Bialoskorski, MA.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.arr.2022.101588.

References

Alqahtani, B.A., Abdelbasset, W.K., Alenazi, A.M., 2020. Psychometric analysis of the Arabic (Saudi) tilburg frailty indicator among Saudi community-dwelling older adults. Arch. Gerontol. Geriatr. 90, 104128 https://doi.org/10.1016/j.archger.2020.104128.

Andreasen, J., Aadahl, M., Sørensen, E.E., Eriksen, H.H., Lund, H., Overvad, K., 2018. Associations and predictions of readmission or death in acutely admitted older medical patients using self-reported frailty and functional measures. A Danish cohort study. Arch. Gerontol. Geriatr. 76, 65–72. https://doi.org/10.1016/j.archger.2018.01.013.

Andreasen, J., Lund, H., Aadahl, M., Gobbens, R.J.J., Sorensen, E.E., 2015. Content validation of the Tilburg Frailty Indicator from the perspective of frail elderly. A qualitative explorative study. Arch. Gerontol. Geriatr. 61, 392–399. https://doi.org/10.1016/j.archger.2015.08.017.

Andreasen, J., Sørensen, E.E., Gobbens, R.J.J., Lund, H., Aadahl, M., 2014. Danish version of the tilburg frailty indicator – translation, cross-cultural adaption and validity pretest by cognitive interviewing. Arch. Gerontol. Geriatr. 59, 32–38. https://doi.org/10.1016/j.archger.2014.02.007.

Bowling, A., 2005. Mode of questionnaire administration can have serious effects on data quality. J. Public Health 27, 281–291. https://doi.org/10.1093/pubmed/fdi031.

Chong, E., Ho, E., Baldevarona-Llego, J., Chan, M., Wu, L., Tay, L., 2017. Frailty and risk of adverse outcomes in hospitalized older adults: a comparison of different frailty measures. J. Am. Med. Dir. Assoc. 18, 638.e7–638.e11. https://doi.org/10.1016/j.jamda.2017.04.011.

Chong, E., Ho, E., Baldevarona-Llego, J., Chan, M., Wu, L., Tay, L., Ding, Y.Y., Lim, W.S., 2018. Frailty in hospitalized older adults: comparing different frailty measures in predicting short- and long-term patient outcomes. J. Am. Med. Dir. Assoc. 19 (450–457), e3 https://doi.org/10.1016/j.jamda.2017.10.006.

Coelho, T., Santos, R., Paúl, C., Gobbens, R.J., Fernandes, L., 2015. Portuguese version of the tilburg frailty indicator: transcultural adaptation and psychometric validation. Geriatr. Gerontol. Int. 15, 951–960. https://doi.org/10.1111/ggi.12373.

Daniels, R., van Rossum, E., Beurskens, A., van den Heuvel, W., de Witte, L., 2012. The predictive validity of three self-report screening instruments for identifying frail older people in the community. BMC Public Health 12, 69. https://doi.org/10.1186/1471-2458-12-69.

De Witte, N., Gobbens, R., De Donder, L., Dury, S., Buffel, T., Verté, D., Schols, J.M.G.A., 2013. Validation of the comprehensive frailty assessment instrument against the tilburg frailty indicator. Eur. Geriatr. Med. 4, 248–254. https://doi.org/10.1016/j.eurger.2013.03.001.

Dent, E., Martin, F.C., Bergman, H., Woo, J., Romero-Ortuno, R., Walston, J.D., 2019. Management of frailty: opportunities, challenges, and future directions. Lancet 394, 1376–1386. https://doi.org/10.1016/S0140-6736(19)31785-4.

Ding, L., Miao, X., Lu, J., Hu, J., Xu, X., Zhu, H., Xu, Q., Zhu, S., 2021. Comparing the performance of different instruments for diagnosing frailty and predicting adverse outcomes among elderly patients with gastric cancer. J. Nutr. Health Aging 25, 1241–1247. https://doi.org/10.1007/s12603-021-1701-8.

Dong, L., Liu, N., Tian, X., Qiao, X., Gobbens, R.J.J., Kane, R.L., Wang, C., 2017. Reliability and validity of the tilburg frailty indicator (TFI) among Chinese community-dwelling older people. Arch. Gerontol. Geriatr. 73, 21–28. https://doi.org/10.1016/j.archger.2017.07.001.

Freitag, S., Schmidt, S., Gobbens, R.J.J., 2016. Tilburg frailty indicator. Z. Für Gerontol. Geriatr. 49, 86–93. https://doi.org/10.1007/s00391-015-0889-9.

Giacomini, S.B.L., Fhon, J.R., Rodrigues, R.A.P., 2020. Frailty and risk of falling in the older adult living at home. Acta Paul. Enferm. 33. https://doi.org/10.37689/acta-ape/2020AO0124.

Gobbens, R.J., van Assen, M.A., Luijkx, K.G., Schols, J.M., 2012c. Testing an integral conceptual model of frailty. J. Adv. Nurs. 68, 2047–2060. https://doi.org/10.1111/j.1365-2648.2011.05896.x.

Gobbens, R.J.J., 2017. Physical and mental dimensions of quality of life of frail older people] (Dutch). Tijdschr. Gerontol. Geriatr. 48, 160–168. https://doi.org/10.1007/s12439-017-0221-9.

Gobbens, R.J.J., Andreasen, J., 2020. The prediction of readmission and mortality by the domains and components of the tilburg frailty indicator (TFI). Arch. Gerontol. Geriatr. 89, 104077 https://doi.org/10.1016/j.archger.2020.104077.

Gobbens, R.J.J., Boersma, P., Uchmanowicz, I., Santiago, L.M., 2020. The tilburg frailty indicator (TFI): new evidence for its validity. Clin. Interv. Aging 15, 265–274. https://doi.org/10.2147/CIA.S243233.

Gobbens, R.J.J., Krans, A., van Assen, M.A.L.M., 2015. Validation of an integral conceptual model of frailty in older residents of assisted living facilities. Arch. Gerontol. Geriatr. 61, 400–410. https://doi.org/10.1016/j.archger.2015.06.001.

Gobbens, R.J.J., Luijkx, K.G., van Assen, M.A.L.M., 2013. Explaining quality of life of older people in the Netherlands using a multidimensional assessment of frailty. Qual. Life Res. 22, 2051–2061. https://doi.org/10.1007/s11136-012-0341-1.

Gobbens, R.J.J., Luijkx, K.G., Wijnen-Sponselee, M.T., Schols, J.M.G.A., 2010a. Towards an integral conceptual model of frailty. J. Nutr. Health Aging 14, 175–181. https://doi.org/10.1007/s12603-010-0045-6.

Gobbens, R.J.J., van Assen, M.A.L.M., 2017. Associations between multidimensional frailty and quality of life among Dutch older people. Arch. Gerontol. Geriatr. 73, 69–76. https://doi.org/10.1016/j.archger.2017.07.007.

Gobbens, R.J.J., van Assen, M.A.L.M., 2014. The prediction of quality of life by physical, psychological and social components of frailty in community-dwelling older people. Qual. Life Res. 23, 2289–2300. https://doi.org/10.1007/s11136-014-0672-1.

Gobbens, R.J.J., van Assen, M.A.L.M., 2012. Frailty and its prediction of disability and health care utilization: the added value of interviews and physical measures following a self-report questionnaire. Arch. Gerontol. Geriatr. 55, 369–379. https://doi.org/10.1016/j.archger.2012.04.008.

Gobbens, R.J.J., van Assen, M.A.L.M., Augustijn, H., Goumans, M., van der Ploeg, T., 2021. Prediction of mortality by the tilburg frailty indicator (TFI). J. Am. Med. Dir. Assoc. 22, 607.e1–607.e6. https://doi.org/10.1016/j.jamda.2020.07.033.

Gobbens, R.J.J., van Assen, M.A.L.M., Luijkx, K.G., Schols, J.M.G.A., 2012a. The predictive validity of the tilburg frailty indicator: disability, health care utilization, and quality of life in a population at risk. Gerontologist 52, 619–631. https://doi.org/10.1093/geront/gnr135.

Gobbens, R.J.J., van Assen, M.A.L.M., Luijkx, K.G., Wijnen-Sponselee, M.T., Schols, J.M.G.A., 2012b. Young frail elderly: assessed using the tilburg frailty indicator (Dutch). Tijdschr. Gerontol. Geriatr. 43, 296–307. https://doi.org/10.1007/s12439-012-0043-8.

Gobbens, Robbert J.J., van Assen, M.A.L.M., Luijkx, K.G., Wijnen-Sponselee, M.Th, Schols, J.M.G.A., 2010b. The tilburg frailty indicator: psychometric properties. J. Am. Med. Dir. Assoc. 11, 344–355. https://doi.org/10.1016/j.jamda.2009.11.003.

Gobbens, R.J.J., van Assen, M.A.L.M., Schalk, M.J.D., 2014. The prediction of disability by self-reported physical frailty components of the tilburg frailty indicator (TFI). Arch. Gerontol. Geriatr. 59, 280–287. https://doi.org/10.1016/j.archger.2014.06.008.

Gonzalez-Colaço Harmand, M., Meillon, C., Bergua, V., Tabue Teguo, M., Dartigues, J.-F., Avila-Funes, J.A., Amieva, H., 2017. Comparing the predictive value of three definitions of frailty: results from the three-city study. Arch. Gerontol. Geriatr. 72, 153–163. https://doi.org/10.1016/j.archger.2017.06.005.

Hayajneh, A.A., 2019. The psychometric properties of the arabic version of the tilburg frailty indicator. Glob. J. Health Sci. 11 https://doi.org/10.5539/gjhs.v11n9p123, 123.

Hayajneh, A.A., Hammouri, H., Rababa, M., Al-Rawashedeh, S., Wallace, D.C., Alsatari, E.S., 2021. Frailty and its correlates in cognitively intact community-dwelling older adults. Dement. Geriatr. Cogn. Disord. 50, 357–363. https://doi.org/10.1159/000519054.

Hayajneh, A.A.R., 2016. The Arabic (Jordan) Version of the Tilburg Frailty Indicator. University of North Carolina, Greensboro.

Hoogendijk, E.O., Afilalo, J., Ensrud, K.E., Kowal, P., Onder, G., Fried, L.P., 2019. Frailty: implications for clinical practice and public health. Lancet 394, 1365–1375. https://doi.org/10.1016/S0140-6736(19)31786-6.

Junius-Walker, U., Onder, G., Soleymani, D., Wiese, B., Albaina, O., Bernabei, R., Marzetti, E., 2018. The essence of frailty: a systematic review and qualitative synthesis on frailty concepts and definitions. Eur. J. Intern. Med. 56, 3–10. https://doi.org/10.1016/j.ejim.2018.04.023.

Kendhapedi, K.K., Devasenapathy, N., 2019. Prevalence and factors associated with frailty among community-dwelling older people in rural Thanjavur district of South India: a cross-sectional study. e032904–e032904 BMJ Open 9. https://doi.org/10.1136/bmjopen-2019-032904.

Lin, C., Yu, N., Wu, H., Liu, Y., 2021b. Risk factors associated with frailty in older adults with type 2 diabetes: a cross-sectional study. J. Clin. Nurs. https://doi.org/10.1111/jocn.15953.

Lin, C.-H., Liu, C.-Y., Rong, J.-R., 2021a. Psychometric properties of the Taiwanese version of the tilburg frailty indicator for community-dwelling older adults. Healthc. Basel Switz. 9. https://doi.org/10.3390/healthcare9091193.

Mazoochi, F., Gobbens, R.J.J., Lotfi, M., Fadayevatan, R., 2020. Diagnostic accuracy of the tilburg frailty indicator (TFI) for early frailty detection in elderly people in Iran. Arch. Gerontol. Geriatr. 91, 104187 https://doi.org/10.1016/j.archger.2020.104187.

Metzelthin, S.F., Daniëls, R., van Rossum, E., de Witte, L., van den Heuvel, W.J., Kempen, G.I., 2010. The psychometric properties of three self-report screening instruments for identifying frail older people in the community. BMC Public Health 10, 176. https://doi.org/10.1186/1471-2458-10-176.

Mokkink, L.B., de Vet, H.C.W., Prinsen, C.A.C., Patrick, D.L., Alonso, J., Bouter, L.M., Terwee, C.B., 2018. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. Qual. Life Res. 27, 1171–1179. https://doi.org/10.1007/s11136-017-1765-4.

Mulasso, A., Roppolo, M., Gobbens, R.J., Rabaglietti, E., 2017. Mobility, balance and frailty in community-dwelling older adults: what is the best 1-year predictor of falls? Geriatr. Gerontol. Int. 17, 1463–1469. https://doi.org/10.1111/ggi.12893.

Mulasso, A., Roppolo, M., Gobbens, R.J.J., Rabaglietti, E., 2016. The Italian version of the tilburg frailty indicator. Res. Aging 38, 842–863. https://doi.org/10.1177/0164027515606192.

Mulasso, A., Roppolo, M., Gobbens, R.J.J., Rabaglietti, E., 2015. The evaluation of frailty in older adults: translation and cross-cultural adaptation of the tilburg frailty indicator to the Italian context (Italian). G. Ital. Psicol. 42, 593–615. https://doi.org/10.1421/81168.

Op het Veld, Linda, P.M., Beurskens, A.J.H.M., de Vet, H.C.W., van Kuijk, S.M.J., Hajema, K., Kempen, G.I.J.M., van Rossum, E., 2019b. The ability of four frailty screening instruments to predict mortality, hospitalization and dependency in (instrumental) activities of daily living. Eur. J. Ageing 16, 387–394. https://doi.org/10.1007/s10433-019-00502-4.

Op het Veld, van Rossum, L.P.M., Kempen, E., Beurskens, G.I.J.M., Hajema, A.J.H.M., de Vet, K.J., H.C.W, 2019a. Can the combined use of two screening instruments improve the predictive power of dependency in (instrumental) activities of daily living, mortality and hospitalization in old age? J. Frailty Aging 1–6. https://doi.org/10.14283/jfa.2019.17.

Ouzzani, M., Hammady, H., Fedorowicz, Z., Elmagarmid, A., 2016. Rayyan-a web and mobile app for systematic reviews. Syst. Rev. 5, 210. https://doi.org/10.1186/s13643-016-0384-4.

Ozsoy, G., Ates Bulut, E., Gurpinar, B., Ilcin, N., Isik, A.T., 2021. Determination of an optimal frailty cut-off score of tilburg frailty indicator and frailty associated factors in community-dwelling turkish older adults. Ann. Geriatr. Med. Res. https://doi.org/10.4235/agmr.21.0086.

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372, n71. https://doi.org/10.1136/bmj.n71.

Papathanasiou, I.V., Rammogianni, A., Papagiannis, D., Malli, F., Mantzaris, D.C., Tsaras, K., Kontopoulou, L., Kaba, E., Kelesi, M., Fradelos, E.C., 2021. Frailty and quality of life among community-dwelling older adults. Cureus 13, e13049. https://doi.org/10.7759/cureus.13049.

Parker, S.G., McCue, P., Phelps, K., McCleod, A., Arora, S., Nockels, K., Kennedy, S., Roberts, H., Conroy, S., 2018. What is Comprehensive geriatric assessment (CGA)? An umbrella review. Age Ageing 47, 149–155. https://doi.org/10.1093/ageing/afx166.

Pialoux, T., Goyard, J., Lesourd, B., 2012. Screening tools for frailty in primary health care: a systematic review: screening tools for frailty. Geriatr. Gerontol. Int. 12, 189–197. https://doi.org/10.1111/j.1447-0594.2011.00797.x.

Prinsen, C.A.C., Mokkink, L.B., Bouter, L.M., Alonso, J., Patrick, D.L., de Vet, H.C.W., Terwee, C.B., 2018. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual. Life Res. 27, 1147–1157. https://doi.org/10.1007/s11136-018-1798-3.

Qin, Y., Li, J., McPhillips, M., Lukkahatai, N., Yu, F., Li, K., 2021. Association of fear of falling with frailty in community-dwelling older adults: a cross-sectional study. Nurs. Health Sci. 23, 516–524. https://doi.org/10.1111/nhs.12840.

Renne, I., Gobbens, R.J., 2018. Effects of frailty and chronic diseases on quality of life in Dutch community-dwelling older adults: a cross-sectional study. Clin. Interv. Aging 13, 325–334. https://doi.org/10.2147/CIA.S156116.

Rethlefsen, M.L., Kirtley, S., Waffenschmidt, S., Ayala, A.P., Moher, D., Page, M.J., Koffel, J.B., 2021. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. Syst. Rev. 10 https://doi.org/10.1186/s13643-020-01542-z, 39.

Roppolo, M., Mulasso, A., Gobbens, R.J., Mosso, C.O., Rabaglietti, E., 2015. A comparison between uni- and multidimensional frailty measures: prevalence, functional status, and relationships with disability. Clin. Interv. Aging 10, 1669–1678. https://doi.org/10.2147/CIA.S92328.

Santiago, L.M., Gobbens, R.J.J., Mattos, I.E., Ferreira, D.B., 2019. A comparison between physical and biopsychosocial measures of frailty: prevalence and associated factors in Brazilian older adults. Arch. Gerontol. Geriatr. 81, 111–118. https://doi.org/10.1016/j.archger.2018.12.003.

Santiago, L.M., Gobbens, R.J.J., van Assen, M.A.L.M., Carmo, C.N., Ferreira, D.B., Mattos, I.E., 2018. Predictive validity of the Brazilian version of the Tilburg Frailty Indicator for adverse health outcomes in older adults. Arch. Gerontol. Geriatr. 76, 114–119. https://doi.org/10.1016/j.archger.2018.02.013.

Santiago, L.M., Luz, L.L., Mattos, I.E., Gobbens, R.J.J., 2012. [Cross-cultural adaptation of the tilburg frailty indicator (TFI) for use in the Brazilian population] (Portuguese). Cad. Saúde Pública 28, 1795–1801. https://doi.org/10.1590/S0102-311X2012000900018.

Santiago, L.M., Luz, L.L., Mattos, I.E., Gobbens, R.J.J., van Assen, M.A.L.M., 2013. Psychometric properties of the Brazilian version of the Tilburg frailty indicator (TFI). Arch. Gerontol. Geriatr. 57, 39–45. https://doi.org/10.1016/j.archger.2013.03.001.

Sezgin, D., O'Donovan, M., Cornally, N., Liew, A., O'Caoimh, R., 2019. Defining frailty for healthcare practice and research: a qualitative systematic review with thematic analysis. Int. J. Nurs. Stud. 92, 16–26. https://doi.org/10.1016/j.ijnurstu.2018.12.014.

Si, H., Jin, Y., Qiao, X., Tian, X., Liu, X., Wang, C., 2020. Comparing diagnostic properties of the FRAIL-NH scale and 4 frailty screening instruments among Chinese institutionalized older adults. J. Nutr. Health Aging 24, 188–193. https://doi.org/10.1007/s12603-019-1301-z.

Si, H., Jin, Y., Qiao, X., Tian, X., Liu, X., Wang, C., 2021a. Comparison of 6 frailty screening tools in diagnostic properties among Chinese community-dwelling older people. Geriatr. Nurs. 42, 276–282. https://doi.org/10.1016/j.gerinurse.2020.08.017.

Si, H., Jin, Y., Qiao, X., Tian, X., Liu, X., Wang, C., 2021b. Predictive performance of 7 frailty instruments for short-term disability, falls and hospitalization among Chinese community-dwelling older adults: a prospective cohort study. Int. J. Nurs. Stud. 117, 103875 https://doi.org/10.1016/j.ijnurstu.2021.103875.

Sutton, J.L., Gould, R.L., Daley, S., Coulson, M.C., Ward, E.V., Butler, A.M., Nunn, S.P., Howard, R.J., 2016. Psychometric properties of multicomponent tools designed to assess frailty in older adults: a systematic review. BMC Geriatr. 16, 55. https://doi.org/10.1186/s12877-016-0225-2.

Suurmond, R., van Rhee, H., Hak, T., 2017. Introduction, comparison, and validation of meta-essentials: a free and simple tool for meta-analysis. Res. Synth. Methods 8, 537–553. https://doi.org/10.1002/jrsm.1260.

Terwee, C.B., Jansma, E.P., Riphagen, I.I., de Vet, H.C.W., 2009. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual. Life Res. 18, 1115–1123. https://doi.org/10.1007/s11136-009-9528-5.

Terwee, C.B., Prinsen, C.A.C., Chiarotto, A., Westerman, M.J., Patrick, D.L., Alonso, J., Bouter, L.M., de Vet, H.C.W., Mokkink, L.B., 2018. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual. Life Res. 27, 1159–1170. https://doi.org/10.1007/s11136-018-1829-0.

Topcu, Y., Tufan, F., Kılıç, C., 2019. Turkish version of the tilburg frailty indicator. Clin. Interv. Aging 14, 615–620. https://doi.org/10.2147/CIA.S197512.

Uchmanowicz, I., Gobbens, R., Jankowska-Polanska, B., Loboz-Rudnicka, M., Manulik, S., Loboz-Grudzien, K., 2014. Cross-cultural adaptation and reliability testing of the tilburg frailty indicator for optimizing care of Polish patients with frailty syndrome. Clin. Interv. Aging 997. https://doi.org/10.2147/CIA.S64853.

Uchmanowicz, I., Gobbens, R.J., 2015. The relationship between frailty, anxiety and depression, and health-related quality of life in elderly patients with heart failure. Clin. Interv. Aging 1595. https://doi.org/10.2147/CIA.S90077.

Uchmanowicz, I., Jankowska-Polańska, B., Uchmanowicz, B., Kowalczuk, K., Gobbens, R. J.J., 2016. Validity and Reliability of the Polish Version of the Tilburg Frailty Indicator (TFI). J. Frailty Aging 5, 27–32. https://doi.org/10.14283/jfa.2015.66.

Uchmanowicz, I., Kuśnierz, M., Wleklik, M., Jankowska-Polańska, B., Jaroch, J., Łoboz-Grudzień, K., 2018. Frailty syndrome and rehospitalizations in elderly heart failure patients. Aging Clin. Exp. Res. 30, 617–623. https://doi.org/10.1007/s40520-017-0824-6.

Uchmanowicz, I., Lisiak, M., Wontor, R., Łoboz-Grudzień, K., 2015. Frailty in patients with acute coronary syndrome: comparison between tools for comprehensive geriatric assessment and the Tilburg Frailty Indicator. Clin. Interv. Aging 521. https://doi.org/10.2147/CIA.S78365.

van der Vorst, A., Op het Veld, L.P.M., De Witte, N., Schols, J.M.G.A., Kempen, G.I.J.M., Zijlstra, G.A.R., 2018. The impact of multidimensional frailty on dependency in activities of daily living and the moderating effects of protective factors. Arch. Gerontol. Geriatr. 78, 255–260. https://doi.org/10.1016/j.archger.2018.06.017.

van Rhee, H., Suurmond, R., Hak, T., 2015. User manual for Meta-Essentials: Workbooks for meta-analysis (Version 1.4).

Vrotsou, K., Machón, M., Rivas-Ruíz, F., Carrasco, E., Contreras-Fernández, E., Mateo-Abad, M., Güell, C., Vergara, I., 2018. Psychometric properties of the tilburg frailty indicator in older Spanish people. Arch. Gerontol. Geriatr. 78, 203–212. https://doi.org/10.1016/j.archger.2018.05.024.

Xie, J., Yu, Q., Yang, L., Sun, Y., 2020. A preliminary study of the effects of the tilburg frailty indicator, frailty phenotype, and silver nanoparticle-silver needle therapy in senile inpatients with frailty. J. Nanosci. Nanotechnol. 20, 6518–6524. https://doi.org/10.1166/jnn.2020.18602.

Xue, Q.-L., 2011. The frailty syndrome: definition and natural history. Clin. Geriatr. Med. 27, 1–15. https://doi.org/10.1016/j.cger.2010.08.009.

Yang, Y., Liao, L., Deng, Q., 2020. Comparison of FRAIL scale and tilburg frailty indicator scale in assessment of frailty status of elderly patients with diabetes in the community (Chinese). Nurs. Integr. Tradit. Chin. West. Med. 6, 1–4. https://doi.org/10.11997/nitcwm.2020001001.

Zhang, X., Tan, S.S., Bilajac, L., Alhambra-Borrás, T., Garcés-Ferrer, J., Verma, A., Koppelaar, E., Markaki, A., Mattace-Raso, F., Franse, C.B., Raat, H., 2020. Reliability and validity of the tilburg frailty indicator in 5 European countries. J. Am. Med. Dir. Assoc. 21 (772–779), e6 https://doi.org/10.1016/j.jamda.2020.03.019.

Zhang, X., Tan, S.S., Franse, C.B., Alhambra-Borrás, T., Durá-Ferrandis, E., Bilajac, L., Markaki, A., Verma, A., Mattace-Raso, F., Voorham, A.J.J., Raat, H., 2019. Association between physical, psychological and social frailty and health-related quality of life among older people. Eur. J. Public Health 29, 936–942. https://doi.org/10.1093/eurpub/ckz099.