National Institute for Health and Care Excellence

Final

Stroke rehabilitation in adults (update)

[B] Evidence reviews for the optimal tool for the assessment of fatigue

NICE guideline NG236

Evidence reviews underpinning recommendations 1.7.1 and 1.7.2 and recommendations for research in the NICE guideline

October 2023

Final

These evidence reviews were developed by NICE

Update information

April 2024: We added text to section 1.1.14.4 to explain why the Neurological Fatigue Index – Stroke was not recommended.

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Local commissioners and/or providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

ISBN: 978-1-4731-5448-3

Contents

1 Tools to assess fatigue

1.1 Review question

In people after stroke, what is the optimal tool for assessment of fatigue?

1.1.1 Introduction

Fatigue is a common problem after stroke or Transient Ischaemic Attack (TIA), estimated to affect around 50% of stroke survivors. In a large survey by the Stroke Association, 86% of stroke survivors reported having experienced post-stroke fatigue (PSF). It can affect any individual soon after stroke, or months later and may persist years after stroke. It manifests as extreme tiredness that does not improve with rest and is different from normal tiredness. Affected individuals feel lacking in physical and mental energy most of the time. Tiredness post stroke can be caused by several reasons such as side effects of medications, lack of sleep, low mood, and other underlying medical conditions. Risk factors associated with PSF include being female, thalamic stroke, leucoaraiosis (white matter changes), depression, sleeping disturbances, diabetes mellitus, anxiety, and multi-morbidities.

Post-stroke fatigue affects rehabilitation after stroke and subsequent care. It has a huge impact on health-related quality of life, ability to carry out everyday activities, and relationships. It is associated with negative outcomes of stroke, such as changes in cognition, mobility, depression, severity of disability, difficulty returning to paid work, institutionalisation, and mortality.

To support stroke survivors and carers, it is important to detect and assess post-stroke fatigue using a patient reported outcome measure (PROM) that is sensitive, reliable, and valid. There has been a lack of consensus about which measure to use. Many measures have been used to detect and assess post-stroke fatigue. These have not always been developed specifically for stroke and in some cases do not capture the full nature of poststroke fatigue. There has been uncertainty therefore, about which measure to use, which limits the ability of stroke services to routinely screen for post-stroke fatigue to guide stroke rehabilitation.

1.1.2 Summary of the protocol

Table 1: PICO characteristics of review question

Population	Inclusion: • Adults (age ≥16 years) who have had a first or recurrent stroke (including people after subarachnoid haemorrhage) Exclusion: • Children (age <16 years) • People who have had a transient ischaemic attack

Interventions/Tests	Assess-and-treat review
	• Tools for assessment of fatigue after a stroke:
	  ○ Fatigue Assessment Scale
	  – Cut off: 23
	  ○ Fatigue Severity Scale
	  – Cut off: 36
	  ○ Brief Fatigue Inventory
	  – Cut offs:
	  • 1-3 (mild)

	• 4-7 (moderate) • 8-10 (severe) ○ Combinations of the above Where studies include a mixture of the above categories studies will be included if at least 80% satisfy the criteria for one category. If <10% of participants are in a different category (for example: 9% have a Fatigue Severity Scale assessment, 91% have a Modified Fatigue Impact Scale assessment this study will be included in the majority category without downgrading for indirectness. If 10-20% are in a different category, this study will be included in the majority category and downgraded for intervention indirectness. Validity and Reliability review • Any tools for assessment of fatigue after a stroke (either designed for a stroke survivor population, or later validated for stroke survivors)
Comparisons	Assess-and-treat review • Each other Confounding factors (for non-randomised studies only): • Presence of comorbidities • Stroke severity • Time period since stroke • Medication usage • Age • Presence of communication difficulties • Baseline psychological distress scores Confounding factors to be considered in the inclusion criteria (studies will not be excluded if they do not adjust for this in a multivariate/univariate analysis or with matched groups): • Time of day of test administration (to consider in the inclusion criteria) Validity and reliability review Compared to the validity and reliability in a healthy population (or other non-stroke survivor population) or to itself (for reliability)
Outcomes	Clinical effectiveness (assess-and-treat) outcomes: At time period • <1 year • ≥1 year • Person/participant generic health-related quality of life (continuous outcomes will be prioritised) • Carer generic health-related quality of life (continuous outcomes will be prioritised) • Activities of daily living (continuous outcomes will be prioritised) • Psychological distress (continuous outcomes will be prioritised) ○ Depression ○ Anxiety ○ Distress • Stroke-specific Patient-Reported Outcome Measures (continuous outcomes will be prioritised) • Participation in leisure activities/social groups scores (continuous outcomes will be prioritised)

	• Withdrawal due to adverse events (dichotomous outcome)
	
	If not mentioned above, other validated scores will be considered and discussed with the committee to deliberate on their inclusion.
	
	Validity and reliability outcomes:
	Validity:
	• Face/content/construct validity
	• Criterion/Concurrent validity
	• Discriminant/convergent validity
	Reliability:
	• Test-retest reliability
	• Internal consistency (including Cronbach's alpha and intraclass correlation coefficient):
	  ○ Intertest reliability
	  ○ Intratest reliability
	• Inter-rater reliability
	Responsiveness to change
	Dimensions of fatigue considered (for example: physical, emotional, cognitive)

Study design	Assess-and-treat review
	• Systematic reviews of RCTs
	• Parallel RCTs (assess-and-treat)
	• Non-randomised studies (if insufficient evidence from parallel RCTs)
	  ○ Prospective cohort study
	  ○ Retrospective cohort study
	
	Validity and reliability review
	Cohort studies and cross sectional studies (investigating tool validation)

For full details see the review protocol in Appendix A.

1.1.3 Methods and process

This evidence review was developed using the methods and process described in Developing NICE guidelines: the manual. Methods specific to this review question are described in the review protocol in Appendix A and the methods document.

Declarations of interest were recorded according to NICE's conflicts of interest policy.

1.1.4 Effectiveness (assess-to-treat) and validity and reliability evidence

1.1.4.1 Included studies

No relevant assess-to-treat studies were identified. Twenty four reliability and validity studies (included one prospective cohort study and twenty three cross-sectional studies) studies were included in the review;[1, 2, 4-8, 13-15, 17-21, 23-25, 29 , 31, 32, 34-36] these are summarised in Table 2 below. Evidence from these studies is summarised in the clinical evidence summaries Table 3, Table 4 and Table 5.

The studies included evidence discussing fourteen tools (including translations of tools). These were:

- A case definition for fatigue[18]
- Detection List Fatigue[15]
- Dutch Multifactor Fatigue Scale (DMFS)[35]
- Fatigue Assessment Scale (FAS) – including the English, Chinese and Swedish versions[5, 7, 14, 19, 29]
- Fatigue Impact Scale (FIS) and the Modified Fatigue Impact Scale (MFIS) – including the Chinese, Persian and Turkish versions[2, 23, 25, 36]
- Fatigue Severity Scale (FSS) – including the English, Arabic, German, Norwegian and Turkish versions[1, 17, 21, 24, 34]
- Functional Assessment of Chronic Illness Therapy – Fatigue (FACIT-F)[6]
- Lee Fatigue Scale (LFS)[4]
- Multidimensional Fatigue Symptom Inventory-general subscale (MFSI-general)[19]
- Neurological Fatigue Index-Stroke (NFI-S) – including the English, Chinese and Norwegian version[13, 20, 31]
- Numeric Rating Scale – Faces Rating Scale (NRS-FRS)[8]
- Profile of Mood States – Fatigue subscale (POMS-fatigue)[7, 19]
- SF-36v2 vitality subscale (SF-36v2 vitality)[7, 19]
- Visual Analogue Fatigue Scale (VAFS)[32]

See also the study selection flow chart in Appendix C, study evidence tables in Appendix D, forest plots in Appendix E and GRADE tables in Appendix F.

1.1.4.2 Excluded studies

See the excluded studies list in Appendix J.

1.1.5 Summary of studies included in the effectiveness (assess-to-treat) evidence

No assess-to-treat evidence was identified.

1.1.6 Summary of studies included in the validity and reliability evidence

Table 2: Summary of studies included in the evidence review

Study	Tool and comparison tools	Population(s)	Outcomes	Comments
Abdulla 2019[1]	**Fatigue severity scale – Arabic (FSS-A)**	**People after stroke** (n=147)	Face/content/construct validity	Cross-sectional study

Study	Tool and comparison tools	Population(s)	Outcomes	Comments
	Comparison tools: Fatigue Visual Analogue Scale (VAS-F) Stroke Specific quality of Life (SSQOL-A) including an energy domain (SSQOL-A-E) SF-36, including a vitality domain (SF-36v) Beck Depression Inventory-II (BDI-II) National Institutes of Health Stroke Scale (NIHSS)	Mean age (SD): 59.63 (10.97) years Male/female: 69:78 Mean time after stroke (SD) = 12.0 (15.6) months Haemorrhagic:Ischaemic = 23:122 **Healthy participants** (n=70) Mean age (SD): 57.830 (11.52) years Male/female: 28:42	Test-retest reliability (7 days) Internal consistency (Cronbach's alpha) Responsiveness to change	Setting: Saudi Arabia, three general hospitals in the area Funding: No additional information
Batur 2022[2]	**Fatigue impact scale (FIS) – Turkish version** Comparison tools: SF-36 vitality subscale (SF-36v) Fatigue severity scale (FSS) Hospital Anxiety and Depression Scale (HADS)	**People after stroke** (n=41) Mean age (SD): 56.4 (12.9) years Male/female: 19:22 **People with musculoskeletal conditions** (n=41) Mean age (SD) = 55.2 (14.6) years Male/female: 19:22 N=82	Face/content/construct validity Criterion/concurrent validity Discriminant/convergent validity Test-retest reliability Internal consistency Dimensions of fatigue considered	Cross-sectional study Setting: Turkey, physical medicine and rehabilitation department. Funding: No funding was received in support of this work.
Bragstad 2020[4]	**Lee Fatigue Scale (LFS)** Five item and Three item score (latter considered the final score)	**People after stroke** (n=322) Mean age (SD): 66.4 (12.8) years Male/female: 190:132 **People with osteoarthritis** (n=203) Mean age (SD): 68.2 (9.2) years Male/female: 64:139 N=525	Face/content/construct validity Discriminant/convergent validity Internal consistency	Cross-sectional study Setting: Norway, acute stroke or rehabilitation units (people after stroke), inpatient or surgical clinical attendees (people with osteoarthritis). Funding: Academic/government funding from various sources

Study	Tool and comparison tools	Population(s)	Outcomes	Comments
Brandal 2016[5]	**Functional Assessment Scale – Swedish version (FAS-S)** Comparison tools: SF-36 vitality subscale (SF-36v) Geriatric Depression Scale (GDS-15) Rivermead Mobility Index (RMI)	**People after stroke** (n=90) Mean age (range): 68 (39 to 87) years Male/female: 49:22 Diagnosis of stroke: Ischaemic stroke (anterior circulation) = 60 Ischaemic stroke (posterior circulation) = 6 Intracerebral haemorrhage (anterior circulation) = 5 Intracerebral haemorrhage (posterior circulation) = 1 Severity (mean NIHSS [SD]) = 2 (2.8) Median time between stroke onset and assessment (IQR): 132 (115-148) days	Face/criterion/construct validity Test-retest reliability (mean 9.6 days) Internal consistency (Cronbach's alpha)	Cross-sectional study Setting: Sweden, stroke unit Funding: Academic/government funding from various sources
Butt 2013[6]	**Functional Assessment of Chronic Illness Therapy – Fatigue (FACIT-F)** Comparison tool: ECOG performance status rating	**People after stroke** (n=51) Mean age (SD): 62.6 (13.9) years Female = 51.0% Type of stroke = Infarct 70.0% (12.1% subarachnoid haemorrhage) Location: Superficial/cortical = 27.3%) Subcortical = 56.8% Combination or other = 15.9% **People with cancer** (n=297) Mean age (SD): 58.1 (13.5) years	Discriminant/convergent validity Internal consistency (Cronbach's alpha)	Cross-sectional study Setting: United States of America, outpatients (recruited from rehabilitation institutes or other outpatient departments). Funding: NIH grant #K23 MH 084551.

Study	Tool and comparison tools	Population(s)	Outcomes	Comments
		Female = 64.3% **People with HIV/AIDS** (n=51) Mean age (SD): 40.2 (6.9) years Female = 11.8% N = 399		
Cheraghifard 2022[7]	**Fatigue Assessment Scale (FAS)** **Profile of Mood States – Fatigue subscale (POMS-F)** **SF-36 Vitality subscale (SF-36 VT)**	**People after stroke** (n=124) Mean age (SD): 59.48 (11.78) years Male:female: 74:51 Mean time after stroke (SD): 25.51 (12.24) months	Test-retest reliability Internal consistency Responsiveness to change	Cross-sectional studies Setting: Iran, community dwelling people who attended four rehabilitation centers Funding: This work was supported by the Iran University of Medical Sciences.
Chuang 2015[8]	**Numeric Rating Scale – Faces Rating Scale (NRS-FRS)** Comparison tool: Numeric Rating Scale	**People after stroke** (n=106) Mean age (SD): 53.63 (11.25) years Male:Female = 77:29 Mean time after stroke (SD): 24.40 (24.11) months	Criterion/concurrent validity (Spearman Rank Correlation Coefficient) Test-retest reliability (1 week) Responsiveness to change	Cross-sectional study Setting: Taiwan, outpatient rehabilitation program Funding: Academic/government funding from various sources
Ho 2020[14]	**Fatigue Assessment Scale – Chinese version (FAS-C)** Comparison tools: Fatigue Severity Scale – Chinese version (FSS-C) Mental Fatigue Scale – Chinese version (MFS-C) Epworth Sleepiness Scale – Chinese version (ESS-C) Five times Sit-To-Stand Test (FTSTS)	**People after stroke** (n=112) Mean age (SD): 64.15 (5.79) years Female:Male = 38:74 Mean time since stroke (SD): 6.13 (4.79) years	Face/content/construct validity Test-retest reliability Internal consistency (Cronbach's alpha) Responsiveness to change Dimensions of fatigue considered	Cross-sectional study Setting: Hong Kong, local self-help groups Funding: This work was supported by the Hong Kong Polytechnic University, and the Departmental Research Grant (ref: 90013897) from the Department of Rehabilitation Sciences, The Hong Kong Polytechnic University.

Study	Tool and comparison tools	Population(s)	Outcomes	Comments
	Fugl-Meyer Assessment (FMA) 15-Item Geriatric Depression Scale – Chinese Version (GDS-15-C)			
Ho 2021[13]	**Chinese version of the Neurological Fatigue Index-Stroke (C-NFI-Stroke)** Comparison tools: Fatigue Severity Scale (FSS) Mental Fatigue Scale (MFS) General Self-Efficacy Scale (GSES) Geriatric Depression Scale (GDS)	**People after stroke** (n=112) Mean time after stroke (SD): 73.60 (57.43) months **Healthy older people** (n=65) N = 177 Mean age (SD): 64.15 (5.79) years Male = 66.1%	Face/content/construct validity Criterion/Concurrent validity (Spearman's correlation coefficients) Discriminant/convergent validity Test-retest reliability (1 week) Internal consistency (Cronbach's alpha, intra-class correlation coefficient) Inter-rater reliability Responsiveness to change Dimensions of fatigue considered	Cross-sectional study Setting: Hong Kong, self-help groups and a community center. Funding: Academic/government funding from various sources
Kruithof 2016[15]	**Detection List Fatigue (tested in Dutch) (DLF)** Comparison tools: Hospital Anxiety and Depression Scale – Anxiety and Depression subscales (HADS-A/HADS-D) – T1 only Fatigue rating scale (FRS) – T1 and T2 Checklist Individual Strength subscale fatigue (CIS-f) – T2 only Fatigue Severity Scale – 7 item version (FSS) – T2 only	**People after stroke** (n=107) Mean age (SD): 60 (10) years Male = 72 (67.3%) Type of lesion: Infarct = 74 Haemorrhage = 24 Subarachnoid haemorrhage = 6 Unknown = 3 Lesion location: Left = 41 Right = 57 Diffuse = 9	Face/content/construct validity Discriminant/convergent validity Dimensions of fatigue considered	Prospective cohort study Setting: The Netherlands, inpatient and outpatient Funding: No additional information
Lerdal 2011[17]	**Fatigue Severity Scale –**	**People after stroke** (n=119)	Face/content/construct validity	Cross-sectional study

Study	Tool and comparison tools	Population(s)	Outcomes	Comments
	Norwegian version (FSS-N) Comparison tools: SF-36 vitality subscale (SF-36v) Energy-VAS (E-VAS)	Mean age (SD): 68.3 (13.1) years Women:Men = 47:72	Criterion/concurrent validity Internal consistency (Cronbach's alpha)	Setting: Norway, participants in the post-stroke fatigue study (PS-Study) Funding: This project is funded by the Research Council of Norway and Buskerud University College for 2006 to 1010 (Grant: 176503).
Lynch 2007[18]	**A case definition for fatigue** Comparison tools: SF-36 (raw total) Fatigue Assessment Scale (FAS) Profile of Mood States (POMS) Modified Fatigue Severity Index (MFSI)	**People after stroke** (n=55) Median age (IQR): 74 (66-81) years Male = 31 (56%) Haemorrhagic stroke = 3 (6%) Location of stroke (TACS = 11, PACS = 22, LACS = 16, POCS = 6)	Face/content/construct validity Concurrent validity Test-retest reliability Inter-rater reliability Dimensions of fatigue considered	Cross-sectional study Setting: United Kingdom. Inpatient and community settings. Funding: No additional information.
Mead 2007[19]	**SF-36v2 vitality subscale (SF-36v2 vitality)** **Fatigue subscale of the Profile of Mood States (POMS-fatigue)** **Fatigue Assessment Scale (FAS)** **General subscale of Multidimensional Fatigue Symptom Inventory (MFSI-general)** Comparison tool: Brief Fatigue Inventory (BFI)	**People after stroke** (n=55) Median age (IQR): 73 (66-81) years Male = 31 Inpatient/community = 40:15 Type of stroke: Right hemisphere = 27 Haemorrhagic = 3 Total anterior circulation syndromes = 11 Partial anterior circulation syndromes = 22 Lacunar syndromes = 16 Posterior circulation syndromes = 6 Median time (IQR) between	Face/content/construct validity Discriminant/convergent validity Test-retest reliability Internal consistency (intratest reliability) Inter-rater reliability	Cross-sectional study Setting: United Kingdom. Inpatient and community settings. Funding: The study received funding from the Chief Scientist Office of the Scottish Executive Health Department (reference CZG/2/161).

Study	Tool and comparison tools	Population(s)	Outcomes	Comments
		stroke and first assessment: Inpatient = 23 (10-53) days Community = 137 (93 to 217) days		
Mills 2012[20]	**Neurological Fatigue Index-Stroke (NFI-Stroke)** Comparison tools: Fatigue severity scale (FSS) Visual analogue scale (VAS) Stroke Impact Scale (SIS)	**People after stroke** (N=282) Mean age (SD): 67.3 (13.4) years Male = 61.3% Mean time after stroke (SD): 17.2 (11.4) months Ischaemic stroke = 78.7%	Criterion/Concurrent validity Test-retest reliability (2-4 weeks) Internal consistency (person-item separation index) Dimensions of fatigue considered	Cross-sectional study Setting: United Kingdom, outpatient departments Funding: No additional information
Nadarajah 2017[21]	**Fatigue Severity Scale** Comparison tools: Visual analogue scale – Fatigue (VAS-F) SF-36v2 Vitality	**Stroke survivors** (n=50) Mean age (SD): 63.6 (10.3) years Female:Male = 16:34 Mean time after stroke (SD): 35.1 (50.0) months **Healthy participants** (n=50) Mean age (SD): 61.1 (7.4) years Female:male = 28:22 N = 100	Criterion/Concurrent validity (Spearman correlation coefficient) Discriminant/convergent validity Test-retest reliability (intraclass correlation coefficient) Internal consistency (Cronbach's alpha)	Cross-sectional study Setting: United States of America, teaching hospital outpatient setting Funding: This study was supported by a research fund from the University of Malaya (PG107-2014A).
Ng 2022[23]	**Modified Fatigue Impact Scale (MFIS) – Cantonese version** Comparison tools: Fugl Meyer Assessment – Lower Extremity Timed Up and Go completion time Community Integration Measure	**Stroke survivors** (n=101) Mean age (SD): 63.82 (6.40) years Female:Male = 43:58 Mean time after stroke (SD): 6.74 (4.42) years **Healthy participants** (n=50) Mean age (SD): 61.78 (7.41) years	Face/content/construct validity Discriminant/convergent validity Test-retest reliability Internal consistency Responsiveness to change Dimensions of fatigue considered	Cross-sectional study Setting: China, outpatient basis. Funding: Supported by Departmental Research Grant (ref P0013897) from Department of Rehabilitation Sciences, the Hong Kong Polytechnic University.

Study	Tool and comparison tools	Population(s)	Outcomes	Comments
	12-Item Short Form Health Survey version 2 Physical and Mental Component scores	Female:Male = 35:15 N = 151		
Ozyemisci-Taskiran 2019[24]	**Fatigue Severity Scale – Turkish versions (FSS-T)** Comparison tools: Hospital Anxiety and Depression Scale – Turkish version (HADS-T) Visual Analogue Scale – Fatigue (VAS-F) SF-36 vitality subscale (SF-36v)	**People after stroke** (n=46) Mean age (SD): 57.9 (13.3) years Male = 21 Mean duration after stroke (SD): 10.4 (16.7) months **Orthopaedic control group** (n=52) Mean age (SD): 53.0 (15.2) years Male = 20 Diagnoses: Low back pain = 16 Knee osteoarthritis = 14 Meniscal degeneration = 6 Strain = 9 Sprain = 7 N=98	Face/content/construct validity Criterion/Concurrent validity Discriminant/convergent validity Test-retest reliability (7 days) Internal consistency (Cronbach's alpha)	Cross-sectional study Setting: Turkey, inpatients Funding: No additional information.
Saneii 2020[25]	**Fatigue Impact Scale – Persian version (FIS-P)** Comparison tools: Fatigue severity scale – Persian version (FSS-P) SF-36 questionnaire – Persian version (SF-36-P)	**People after stroke** (n=140) Mean age (SD): 58.85 (7.88) years Male:Female = 71:69 Mean time after stroke (SD): 20.27 (15.82) months **Healthy adults** (n=140) Mean age (SD): 58.16 (10.51) years Male:Female = 77:63 N = 280	Face/content/construct validity Discriminant/convergent validity Test-retest reliability (1 week) Internal consistency (Cronbach's alpha) Interrater reliability Responsiveness to change Dimensions of fatigue considered	Cross-sectional study Setting: Iran, occupational therapy clinics Funding: No additional information
Smith, 2008 [29]	**Fatigue Assessment Scale (FAS)**	**People after stroke** (n=80)	Face/content/construct validity	Cross-sectional study

Study	Tool and comparison tools	Population(s)	Outcomes	Comments
	Comparison tools: Beck Depression Inventory (BDI) Stroke-Adapted Sickness Impact Profile (SA-SIP30)	Mean age (SD): 74.1 (6.6) years Male = 44 Right hemisphere stroke = 41 Brainstem lesion = 19 Subcortical lesion = 26 Cortical lesion = 22 No information on stroke location = 13 Mean time since stroke (SD): 7.6 (5.4) months **People with congestive heart failure** (n=137) Mean age (SD): 67.6 (8.8) years Male = 98 **Healthy adults** (n=160) Mean age (SD): 69.3 (6.0) years Male = 74	Convergent/discriminant validity Test-retest reliability Internal consistency (Cronbach's alpha)	Setting: The Netherlands, stroke rehabilitation unit of a nursing home Funding: This research was supported by a VICI grant (453-04-004) from the Netherlands Organisation for Scientific Research, The Hague, the Netherlands, and by a grant from the Dutch Heart Foundation (2003B038) to Johan Denollet.
Taasen 2020[31]	**Neurological Fatigue Index-Stroke Norwegian version (N-NFI-Stroke)**	**People after stroke** (n=63) Mean age (SD): 60.25 (14.69) years Women = 36 >12 months post stroke = 42 <3 months post stroke = 24 More than one stroke = 7	Face/content/construct validity Test-retest reliability (within 2 days up to a week apart) Internal consistency (Cronbach's alpha) Dimensions of fatigue considered	Cross-sectional study Setting: Norway, mixture of community and inpatient Funding: No additional information
Tseng 2010[32]	**Visual Analogue Fatigue Scale (VAFS)** Comparison tools: Heart rate Systolic blood pressure increase	**People after stroke** (n=21) Mean age (SD): 59.5 (10.3) years Men:Women = 12:9 Stroke lesion side (right:left:brain stem) = 15:4:2	Criterion/Concurrent validity Test-retest reliability Intratest reliability (14 days) Responsiveness to change	Cross-sectional study Setting: United States of America, people recruited from local stroke support groups and the ASTRA (Advancing Stroke Treatment

Study	Tool and comparison tools	Population(s)	Outcomes	Comments
	Rate of perceived exertion	Stroke subtype (ischaemic:haemorrhagic) = 18:3 Mean time after stroke (SD): 4.1 (3.5) years		through Research Alliances) participant database. Funding: Academic/government funding from various sources
Valko 2008[34]	**Fatigue Severity Scale-German (FSS-G)** Comparison tools: Visual analogue scale-fatigue (VAS-F) Epworth Sleepiness Scale-German (ESS-G)	**People after stroke** (n=234) Mean age (SD): 63 (14) years Female = 31% Mean time after stroke (SD): 1.21 (0.62) years **Healthy subjects** (n=454) Mean age (SD): 47 (18) years Female = 60% **Multiple sclerosis** (n=188) Mean age (SD): 45 (13) years Female = 67% **Sleep-wake disorders** (n=429) Mean age (SD): 52 (15) years Female = 35% N=1306	Face/content/construct validity Discriminant/convergent validity Test-retest reliability (21 days) Internal consistency (Cronbach's alpha)	Cross-sectional study Setting: Switzerland, Neurology and Pulmonary Departments of the University Hospital of Zurich, Switzerland. Funding: This was not an industry supported study. Dr. Blochas received research support from Respironics, ResMed and Weinmann AG and has had the free use of monitoring equipment from VivoMetrics.
Visser-Keizer 2015[35]	**Dutch Multifactor Fatigue Scale (DMFS)** Checklist Individual Strength (CIS) Hospital Anxiety and Depression Scale – Anxiety and Depression subscales (HADS-A and HADS-D) Dutch Personality Questionnaire – Self-esteem (PDQ-Self-esteem)	**Ischaemic stroke** (n=55) Mean age (SD): 55.4 (9.7) years Women/Men: 25:30 Mean time since injury (SD): 33.7 (34.6) months Lesion location (left:right:bilateral: diffuse:no lesion visible on CT or MRI): 22:25:6:0:2 **Haemorrhagic stroke** (n=22)	Face/content/construct validity Convergent/discriminant validity Internal consistency (Cronbach's alpha) Domains of fatigue considered	Cross-sectional study Setting: The Netherlands, academic rehabilitation center Funding: Supported by the Foundation Beatrixoord North Netherlands, Center for Rehabilitation, University Medical Center Groningen, Groningen, The

Study	Tool and comparison tools	Population(s)	Outcomes	Comments
		Mean age (SD): 53.3 (9.4) years Women/Men: 15:7 Mean time since injury (SD): 29.1 (27.9) months Lesion location (left:right:bilateral: diffuse:no lesion visible on CT or MRI): 6:14:2:0:0 **Traumatic brain injury** (n=35) Mean age (SD): 43.4 (13.3) years Women/Men: 20:15 Mean time since injury (SD): 43.1 (59.2) months Lesion location (left:right:bilateral: diffuse:no lesion visible on CT or MRI): 2:3:8:12:10 **Other acute brain injuries** (n=22) Mean age (SD): 49.8 (12.0) years Women/Men: 9:13 Mean time since injury (SD): 36.9 (43.1) months Lesion location (left:right:bilateral: diffuse:no lesion visible on CT or MRI): 5:3:2:12:0		Netherlands (grant no. 210.101). All outcomes apart from convergent/discriminant validity includes data including participants with other brain injury types.
Wu 2008[36]	**Fatigue Impact Scale – Chinese version (FIS-C)**	**People after stroke** (n=330) Age ranging between 35 and >80 years, median range = 51-65 years. Male:female = 139:75 Duration from survey to cerebral infarction ranged between <1	Face/content/construct validity Internal consistency (Cronbach's alpha) Domains of fatigue considered	Cross-sectional study Setting: China. Department of rehabilitation at a hospital. Funding: No additional information.

Study	Tool and comparison tools	Population(s)	Outcomes	Comments
		month to 7-13 months, median time 1-6 months.		

See Appendix D for full evidence tables.

1.1.7 Summary of the effectiveness (assess-to-treat) evidence

No studies investigating assess-to-treat evidence.

1.1.8 Summary of the validity and reliability evidence

Table 3: Summary of the psychometric parameters (validity, responsiveness to change and dimensions of fatigue considered) of tools to assess fatigue for people after a stroke

Tool name	Face/content/construct validity	Criterion/Concurrent validity	Discriminant/convergent validity	Responsiveness to change	Dimensions of fatigue considered
A case definition for fatigue[18]	Feasibility: At both interviews, all participating patients provided satisfactory answers to all case definition probe questions. Clinical characteristics of cases: 20 (36%) patients fulfilled the case definition at the first interview. The study reported that people fulfilling the case definition were more likely to be of female gender or to have higher levels of emotional distress.	Concurrent validity: Assessed in graphs. The paper stated that patients fulfilling the case definition generally had substantially higher fatigue scores on all four fatigue scales, suggesting good concurrent validity ($p<0.001$, Mann-Whitney U test).	Not available	Not available	Holistic
Detection List Fatigue (DLF)[15]	According to the assessors, the DLF was generally easy to understand for patients and quick to administer. At T1, 1 person was not able to answer item 3 after it was repeated to them 3 times. Four people had no idea to what degree item 3 was true or not, and 1 person had no idea to what degree items 6 and 9 were true or not.	Not available	Compared to: • HADS-A = 0.45 (p: <0.001) • HADS-D = 0.31 (p: <0.010) • FRS (T1) = 0.58 (p: <0.001) • FRS (T2) = 0.63 (p: <0.001) • CIS-f = 0.85 (p: <0.001) • FSS = 0.79 (p: <0.001)	Not available	Physical and mental

Tool name	Face/content/construct validity	Criterion/Concurrent validity	Discriminant/convergent validity	Responsiveness to change	Dimensions of fatigue considered
	At T2, 1 person had no idea to what degree item 5 was true ot not, and 2 people had no idea to what degree item 9 was true or not. In a small number of cases, it was difficult to know whether a patient comprehended an item properly because cognitive deficits, aphasia or a combination of both resulted in communication difficulties.				
Dutch Multifactor Fatigue Scale (DMFS)[35]	5 factors identified: Impact of fatigue, Mental fatigue, Signs and Direct consequences of fatigue, Physical fatigue, Coping with fatigue. Factor loading ranged with the lowest range being in the coping with fatigue factor (between 0.53-0.82) and highest for impact of fatigue (between 0.46 to 0.90)	Not available	Scores may be slightly higher in people with traumatic brain injury and other acute brain injuries (in particular for mental fatigue and signs and direct consequences of fatigue). Therefore, other factors may not be comparable.	Not availabe	Physical and mental
Fatigue Assessment Scale – English version (FAS)[19, 29]	Tools included in the Mead study were deemed to have adequate face validity by consensus agreement. Smith study: Principle component analysis with oblimin rotation revealed two factors. Two of the FAS items had higher values in the BDI, while the BDI item on fatigue had a higher score on the	Not available	Compared to: • MFSI-general = 0.71 (p: <0.001) • POMS-fatigue = 0.59 (p: <0.001) • SF-36v2 vitality = -0.41 (p: 0.03) Mean scores compared between populations:	Not available	Not reported in study Appears to consider physical and cognitive

Tool name	Face/content/construct validity	Criterion/Concurrent validity	Discriminant/convergent validity	Responsiveness to change	Dimensions of fatigue considered
	FAS. Overall, this indicates that the two scales measure different entities. However, the FAS may not be a unidimensional construct in people with stroke, as four of the 10 items had low component loadings (FAS-3 and FAS-7) or loaded on the BDI component (FAS-8 and FAS-6). Therefore, more research on the content validity is required. The association between the FAS and BDI was 0.44 (p <0.001)		• People with stroke = 15.3 (7.6) • People with congestive heart failure = 16.5 (7.9) (when compared to people with stroke, p 0.44) • Healthy controls = 9.2 (5.6) (when compared to people with stroke, p <0.001).		
Fatigue Assessment Scale – Chinese version (FAS-C) [14]	Content validity index: The item level-content validity index (assessed by five expert panel members) ranged from 0.80 to 1.00. Ceiling and floor effects: None of the participants received the higher and lowest sum score. Construct validity: Kaiser-Meyer-Olkin measure = 0.84, indicating sufficient items for each factor. The FAS-C was correlated with:	Not available	Not available	Minimal detectable change: • Summary score = 4.69 • Physical score = 2.44 • Mental score = 4.10	Physical and mental

Tool name	Face/content/construct validity	Criterion/Concurrent validity	Discriminant/convergent validity	Responsiveness to change	Dimensions of fatigue considered
	• MFS-C (rs = 0.68, p <0.001) • FSS-C (rs = 0.57, p <0.001) • ESS-C (rs = 0.36, p <0.001) • FMA upper extremities (rs = 0.24, p 0.011) • FMA lower extremities (rs = 0.24, p 0.012) It did not correlate with FTSTS time (r_s = 0.13, p 0.170)				
Fatigue Assessment Scale – Language unclear (FAS-I)[7]	Not available	Not available	Not available	MCID of FAS = 3.16 to 8.76 (6.3-17.5% of the total score)	See Fatigue Assessment Scale (English version)
Fatigue Assessment Scale – Swedish version (FAS-S)[5]	Floor/ceiling effects: None of the participants received the highest or lowest scores, which implied that there were no floor or ceiling effects. Construct validity: The FAS correlated with the SF-36 subscale for vitality (r_s = -0.73) and with the GDS-15 (r_s = 0.62), indicating that the Swedish FAS had convergent construct validity, but not divergent construct validity.	Not available	Not available	Not available	See Fatigue Assessment Scale (English version)

Tool name	Face/content/construct validity	Criterion/Concurrent validity	Discriminant/convergent validity	Responsiveness to change	Dimensions of fatigue considered
Fatigue Impact Scale – Chinese version (FIS-C)[36]	Content validity: Correlation coefficient was calculated between each item and total scores, and the three subscale were >0.5, which was significantly different. Structural validity: Kaiser-Meyer-Olkin measurement of sampling adequacy = 0.934. Suitable for factor analysis. Factor 1 – Physiological states Factor 2 – Cognitive states was captured Factor 3 – The effects of emotion and fatigue on living Factor 4 – Social communication Factor 5 – Effect on work Factor 6 – Economic situation and sex life	Not available	Not available	Not available	Cognitive, physiological and social
Fatigue Impact Scale – Turkish version (FIS-T)[2]	• Content validity: 82% of people with stroke found the questions understandable, and 65% reported the items represented the impact of their fatigue on their health. • Internal construct validity: The Kaiser-Meyer Olkin	In the ROC analysis, the area under the curve was 0.730 (standard error = 0.055) (p <0.001) for the total DIS score. The sensitivity and specificity values were 76% and 68% respectively for an optimal cutoff value of 43.5.	A negative significant correlation was detected between the FIS scores and the SF-36 mental and vitality survey results. There was a positive but weak correlation between the total FIS score, the FIS cognitive and psychosocial scores and the FSS score. HAD depression was not correlated with the FIS scores, except for a moderate	Not available.	Cognitive impact, physical impact and psychosocial impact domains

Tool name	Face/content/construct validity	Criterion/Concurrent validity	Discriminant/convergent validity	Responsiveness to change	Dimensions of fatigue considered
	Measures of sampling adequacy were 0.82, 0.73 and 0.63, which confirm the appropriateness of factor analysis for the cognitive, physical and psychosocial dimensions respectively. Each component explained the 58%, 45% and 38% variances of the cognitive, psychosocial and physical dimensions respectively.		correlation with the FIS cognitive scores. The FIS psychosocial and FIS total scores were weakly correlated with the HAD anxiety score.		
Fatigue Impact Scale – Persian version (FIS-P)[25]	• Face validity – "The relevance, suitability, clarity and simplicity of all questions were acceptable." • Content validity: Ranged from 0.6-1, average of 0.85, universal agreement was 0.48. • Floor and ceiling effects – 2.1%	Not available	Convergent validity: • Significant negative correlations between FIS-P and all SF-36 subscales, with the only exception found for the cognitive subscale of FIS-P and the emotional domain of SF-36. The FIS-P and FSS scales had a significant positive correlation. Discriminant validity: • Significantly different results between stroke patients and healthy adults.	Minimum detectable change FIS-P Total = 8.26 FIS-P Physical = 3.87 FIS-P Cognitive = 2.76 FIS-P Social = 5.27	Cognitive impact, physical impact and social impact domains
Fatigue Severity Scale – English version (FSS) [21]	Not available in the study	Spearman correlation coefficients	• Stroke survivors, total score: 29.2 (11.3) • Healthy participants, total score: 16.9 (9.5)	Not available	Not reported in study Appears to be physical, social, mental

Tool name	Face/content/construct validity	Criterion/Concurrent validity	Discriminant/convergent validity	Responsiveness to change	Dimensions of fatigue considered
		Stroke survivors. Compared to: • VAS-F: 0.68 (0.50 to 0.81) (p=<0.01) • SF36-v2 vitality: -0.32 (-0.55 to -0.05) (p=<0.02) Healthy participants. Compared to: • VAS-F: 0.66 (0.48-0.79) (p=<0.01) • SF36-v2 vitality: -0.01 (-0.29 to 0.27) (p=<0.02)	• P-value comparing between groups: <0.01		
Fatigue Severity Scale – Arabic version [1]	91.6% of the patients and 93.3% of the healthy participants returned acceptable questionnaires. The Shapiro-Wilk test and visual inspection of frequency distribution histograms were relatively symmetrical with a skewness of -0.59. The scale item skewness fell between -0.19 to -0.76. 7.5% or fewer of participants selected the minimum score while 10.9% or fewer respondents selected the maximum score. The total score showed a strong positive correlation with VAS-F, strong negative	Not available	Not available	Minimal detectable change at 95% confidence interval: Total score = 1.02	See Fatigue Severity Scale – English version

Tool name	Face/content/construct validity	Criterion/Concurrent validity	Discriminant/convergent validity	Responsiveness to change	Dimensions of fatigue considered
	correlations with SSQOL-A-E, moderate negative correlations with SSQOL-A, SF-36 and SF-36v and a moderate positive correlation to BDI-II.				
Fatigue Severity Scale – German version (FSS-G) [34]	In people after stroke, no correlation was found between FSS-G scores and age, duration from disease onset, gender or educational status. The results of a linear regression analysis showed a significantly higher FSS-G score for each of the 3 patient groups than healthy controls.	Not available	• Previous ischaemic stroke, total score: 3.90 (1.85) • Healthy subjects, total score: 3.00 (1.08) • Multiple sclerosis, total score: 4.66 (1.64) • Sleep-wave disorders, total score: 4.34 (1.64)	Not available	See Fatigue Severity Scale – English version
Fatigue Severity Scale – Norwegian version (FSS-N) [17]	Psychometric properties (assessed in 428 measurements): The 9-item version has 1 item not meeting the three criteria for rating scale and 1 item misfit. The 8-item version had 2 items that had misfit. The 7-item version (with items 1 and 2 removed) had no misfit and all items met the criteria for rating scale. First latent variable, % = 83.1% 2nd dimension, % = 3.6% Eigenvalue = 1.5	SF-36v: • Baseline = -0.56 • 6 months = -0.62 • 12 months = -0.72 • 18 months = -0.67 E-VAS: • Baseline = -0.40 • 6 months = -0.56 • 12 months = -0.61 • 18 months = -0.53	Not available	Not available	See Fatigue Severity Scale – English version

Tool name	Face/content/construct validity	Criterion/Concurrent validity	Discriminant/convergent validity	Responsiveness to change	Dimensions of fatigue considered
	Person misfit, n (%) = 30 (7.0%) (maximum score achieved in 7, minimum score achieved in 23) Person-separation index (without extremes) = 2.40 Person-separation reliability = 0.85				
Fatigue Severity Scale – Turkish version (FSS-T) [24]	Feasibility – Among 46 subjects, 93.5% responded to all items of the FSS. The remaining two subjects had left the 6th item incomplete, and one subject did not respond to the 8th item. Among the 52 control subjects, all completed each item, apart from one who missed the 7th item. Content validity – Among people with stroke, 89% found the questions understandable, and 72% viewed the scale as representative of fatigue.	There is no gold standard to diagnose fatigue, hence criterion validity was not assessed with another measure. When FSS-T scores of stroke and control groups were evaluated, it was observed that FSS did not distinguish between two groups. Mean FSS scores and subjects with scores greater than 4 were similar in both groups.	FSS test 1: • Stroke survivors = 4.2 (1.7) • Control = 4.1 (1.4) • P value = 0.717 FSS test 2: • Stroke survivors = 4.2 (1.7) • Control = 3.7 (1.6) • P value = 0.233 For people after stroke: FSS-T correlated moderately with SF-36v: r = -0.531, p = 0.002. No correlation between FSS-T and VAS-F: r = 0.197, p = 0.281. Weak correlation between FSS-T and HADS-anxiety subscale: r = 0.310, p = 0.041.	Not available	See Fatigue Severity Scale – English version

Tool name	Face/content/construct validity	Criterion/Concurrent validity	Discriminant/convergent validity	Responsiveness to change	Dimensions of fatigue considered
			Weak correlation between FSS-T and HADS-depression subscale: r = 0.334, p = 0.027.		
Functional Assessment of Chronic Illness Therapy – Fatigue (FACIT-F)[6]	Not available in this study	Not available	• Stroke, total score: 38.1 (9.6) • Cancer, total score: 36.0 (12.1) • HIV, total score: 34.0 (12.6)	Not available	Not reported in this study Appears physical, social
Lee Fatigue Scale (LFS)[4]	Meets criteria for rating scale functioning. Final LFS short form scale has adequate item goodness-of-fit statistics (between 0.7 and 1.3). Variance could be explained by the first latent variable 81.6% of the time. Person goodness-of-fit statistics were above the criterion of <5% (5.6%). Ceiling effects were less than 10% (0.6%) while floor effects were more than 10% (16.9%). Person separation reliability was above 2.0 (2.49) and internal consistency was above 0.80 for all variable (Rasch 0.89, correlation between sum scores and Rasch measures 0.98).	Not available	• Stroke, LFS-3 total score: 3.88 (2.10) • Osteoarthritis, LFS-3 total score: 3.23 (2.13) • P value = 0.001	Not available	Not reported in this study
Modified Fatigue Impact Scale – Cantonese Version (MFIS-C)[23]	Content validity: All the I-CVI, S-CVI/Ave and S-CVI/UA values are 1, suggesting that the validity of each individual	Not available.	All MFIS-C subscale scores have no significant correlations with the FMA-LE score and TUG completion time, but significant weak to moderate negative	Minimally detectable change$_{95}$ (standard error) and minimally detectable change$_{95\%}$	Cognitive, physical and psychosocial domains.

Tool name	Face/content/construct validity	Criterion/Concurrent validity	Discriminant/convergent validity	Responsiveness to change	Dimensions of fatigue considered
	item and the overall MFIS-C are satisfactory.		correlations with the CIM-C, and SF-12 PCS and MCS scores.	overall score = 14.86 (5.38) and 38.3% Minimally detectable change$_{95}$ (standard error) and minimally detectable change$_{95\%}$ cognitive subscale = 7.49 (2.71) and 44.8% Minimally detectable change$_{95}$ (standard error) and minimally detectable change$_{95\%}$ physical and psychosocial subscale = 9.70 (3.51) and 44.0%	
Multidimensional Fatigue Symptom Inventory-general subscale (MFSI-general)[19]	Tools included in the study were deemed to have adequate face validity by consensus agreement. One MFSI-general question ("I feel pooped") was poorly understood and was changed to "I feel exhausted".	Not available	Compared to: • FAS = 0.71 (p: <0.001) • POMS-fatigue = 0.75 (p: <0.001) • SF-36v2 vitality = -0.71 (p: <0.001)	Not available	Not reported in the study. Appears to be general.
Neurological Fatigue Index-Stroke English version (NFI-Stroke)[20]	Not available in this study	NFI-Stroke Summary when compared to: • FSS = 0.622 • VAS = 0.534 • SIS = 0.628	Not available	Not available	Physical and cognitive subscales.

Tool name	Face/content/construct validity	Criterion/Concurrent validity	Discriminant/convergent validity	Responsiveness to change	Dimensions of fatigue considered
Neurological Fatigue Index-Stroke Chinese version (C-NFI-Stroke)[13]	Item-level content validity index scores ranged from 0.6-1, with the scale-level content validity index score of 0.95 (a score of 0.78 was considered good). The Kaiser-Meyer-Olkin measure was 0.91, implying sufficient items for a factor analysis (a score of 0.90 was considered good).	C-NFI-Stroke summary scale when compared to: • FSS = 0.62 • Mental fatigue scale: 0.63 • General Self-Efficacy scale = -0.35 • Geriatric Depression Scale = 0.60	Summary score: • Stroke survivors (median [IQR): 16.00 (6.50) • Healthy older people (median [IQR]): 12.00 (10.50) • P value = <0.001	Summary scale = 2.92 Physical subscale = 2.68 Cognitive subscale = 1.57	Physical and cognitive subscales.
Neurological Fatigue Index – Stroke Norwegian version (N-NFI-Stroke) 31	Floor and ceiling effects – There was a normal distribution of results with the average score being 17.88. Answered varied between totally disagree (1-17%), disagree (19-49%), agree (34-55%), totally agree (7-31%).	Not available	Not available	Not available	Physical and cognitive subscales.
Numeric Rating Scale – Faces Rating Scale (NRS-FRS)[8]	Not available in this study	• Test NRS-FRS = 0.85 (0.75-0.91) • Retest NRS-FRS = 0.84 (0.73-0.91)	Not available	Minimal detectable change at the 95% confidence interval level: 1.39	Not reported in the study. Appears to be general.
Profile of Mood States – Fatigue subscale (POMS-fatigue)[7, 19]	Tools included in the study were deemed to have adequate face validity by consensus agreement.	Not available	Compared to: • FAS = 0.59 (p: <0.001) • MFSI-general = 0.75 (p: <0.001) • SF-36v2 vitality = -0.58 (p: <0.001)	MCID of POMS-F = 1.49 to 5.63 (5.3-20.1% of the total score)	Not reported in the study. Appears to be general.

Tool name	Face/content/construct validity	Criterion/Concurrent validity	Discriminant/convergent validity	Responsiveness to change	Dimensions of fatigue considered
SF-36v2 vitality subscale (SF-36v2 vitality)[7, 19]	Tools included in the study were deemed to have adequate face validity by consensus agreement.	Not available	Compared to: • FAS = -0.41 (p: 0.03) • MFSI-general = -0.47 (p: <0.001) • POMS-fatigue = -0.58 (p: <0.001)	MCID of SF-36-VT = -5.58 to -15.43	Not reported in the study. Appears to be general.
Visual Analogue Fatigue Scale (VAFS)[32]	Not available in this study	A significant positive relationship was found using Pearson's correlation coefficient for exertion fatigue and systolic blood pressure increase (r=0.630, P=0.02) and exertion fatigue and rate of perceived exertion (r=0.802, P=0.00). Exertion fatigue and heart rate (r=0.738, P=<0.01).	Not available	Not available	Not reported in the study. Appears to be general.

Table 4: Summary of the psychometric parameters (reliability) of tools to assess fatigue for people after a stroke

Tool name	Test-retest reliability	Internal consistency	Inter-rater reliability	Other factors
A case definition for fatigue[18]	Kappa (95% confidence interval) = 0.78 (0.60-0.96), n=51	Not available	Kappa (95% confidence interval) = 0.82 (0.64-0.99), n=43	No additional information. See evidence table for a full list of questions and interpretation of answers.
Detection List Fatigue (DLF)[15]	Not available	Not available	Not available	Tested in Dutch. The mean time to administer the test including the instruction time was 6 minutes (+/- 2, range 3-14).

Tool name	Test-retest reliability	Internal consistency	Inter-rater reliability	Other factors
Dutch Multifactor Fatigue Scale (DMFS)[35]	Not available	Cronbach's alpha (note: this includes people with acute brain injury types that are not stroke, and so the accuracy of these results for a stroke population may be questionable) • Impact of fatigue = 0.91 • Mental fatigue = 0.86 • Signs and Direct consequences of fatigue = 0.83 • Physical fatigue = 0.77 • Coping with fatigue = 0.69	Not available	Tested in Dutch. Limited information available about feasibility.
Fatigue Assessment Scale – English version (FAS)[19, 29]	Between questions: • Percentage agreement (range): 37-71% • Kappa (95% CI) (range): 0.33 (0.05 to 0.61) - 0.77 (0.61 to 0.92) • Intra-class correlation coefficient (Mead) = 0.77 (95% CI: 0.62 to 0.86) • Intra-class correlation coefficient (Smith) = 0.81	Cronbach's alpha • 0.77	Between questions: • Percentage agreement (range): 84-94% • Kappa (95% CI) (range): 0.67 (0.37 to 0.97) – 0.95 (0.90-1.00) • Intra-class correlation coefficient = 0.88 (0.77 to 0.94)	10 items, each item ranked 1-5 (never, sometimes, regularly, often, always) Time to administer: 2 minutes Languages: Arabic, Chinese, Croatian, Danish, Dutch, English, Finnish, French, German, Greek, Italian, Japanese, Norwegian, Polish, Portuguese, Russian, Serbian, Spanish, Swedish, Turkish, Ukrainian Usage in clinical settings is permitted without any condition[26]
Fatigue Assessment Scale – Chinese version (FAS-C) [14]	Intra-class correlation coefficient: • Summary score = 0.92 ($p < 0.001$)	Cronbach's alpha: • Summary score = 0.82 • Physical score = 0.78 • Mental score = 0.71	Not available	See Fatigue Assessment Scale (English version)

Tool name	Test-retest reliability	Internal consistency	Inter-rater reliability	Other factors
	• Physical score = 0.95 (p <0.001) • Mental score = 0.77 (p <0.001) Weighted Kappa value between questions ranging from: 0.38 (0.06-0.69) to 0.83 (0.64-1.01)			
Fatigue Assessment Scale – Language unclear (FAS-I)[7]	Intra-class correlation coefficient of FAS (95% CI) = 0.87 (0.78-0.92)	Cronbach's alpha for FAS = 0.86	Not available	See Fatigue Assessment Scale (English version)
Fatigue Assessment Scale – Swedish version (FAS-S)[5]	Interclass correlation coefficient = 0.73 Weighted Kappa value between questions ranging from: 0.22 (-0.02-0.45) to 0.74 (0.61-0.87)	Cronbach's alpha = 0.82	Not available	See Fatigue Assessment Scale (English version)
Fatigue Impact Scale – Chinese version (FIS-C)[36]	Not available	Cronbach's alpha: • Cognitive subscale = 0.937 • Physiological subscale = 0.918 • Social subscale = 0.940	Not available	10 items, scored 0-4 (0 = no problem, 4 = extreme problem) Time to administer: Around 3 minutes for a non-fatigued person Languages: Translated in 30 languages (including English) The Fatigue Impact Scale can be used free of charge for clinical purposes[11]
Fatigue Impact Scale – Turkish	Intraclass correlation coefficient for FIS total =	Cronbach alpha for FIS total = 0.946	Not available	See Fatigue Impact Scale – Chinese version

Tool name	Test-retest reliability	Internal consistency	Inter-rater reliability	Other factors
version (FIS-T)[2]	0.830 (0.731-0.895) (p <0.001) Intraclass correlation coefficient for FIS cognitive = 0.787 (0.667-0.867) (p <0.001) Intraclass correlation coefficient for FIS physical = 0.734 (0.591-0.832) (p <0.001) Intraclass correlation coefficient for FIS psychosocial = 0.800 (0.686-0.879) (p <0.001)	Cronbach alpha for FIS cognitive = 0.917 Cronbach alpha for FIS physical = 0.803 Cronbach alpha for FIS psychosocial = 0.929		
Fatigue Impact Scale – Persian version (FIS-P)[25]	Intra-class correlation coefficient: • FIS-P Total = 0.991 (0.983-0.995), p=<0.001 • FIS-P Physical = 0.961 (0.979-0.926), p=<0.001 • FIS-P Cognitive = 0.987 (0.993-0.976), p=<0.001 • FIS-P Social = 0.987 (0.976-0.993), p=<0.001	Cronbach's alpha: • FIS-P Total = 0.895 • FIS-P Physical = 0.87 • FIS-P Cognitive = 0.90 • FIS-P Social = 0.95	Intra-class correlation coefficient: • FIS-P Total = 0.984 (0.848-0.848), p=0.001 (note: the confidence interval provided by the study is reported here, this is not a valid confidence interval) • FIS-P Physical = 0.911 (0.142-0.991), p=0.019 • FIS-P Cognitive = 0.987 (0.879-0.999), p=<0.001 • FIS-P Social = 0.987 (0.876-0.999), p=<0.001	See Fatigue Impact Scale – Chinese version

Tool name	Test-retest reliability	Internal consistency	Inter-rater reliability	Other factors
Fatigue Severity Scale – English version (FSS) [21]	Intraclass correlation coefficient: • Stroke survivors = 0.93 (0.88 to 0.96) • Healthy participants = 0.90 (0.84 to 0.94)	Cronbach's alpha: • Stroke survivors = 0.93 (0.89 to 0.96) • Healthy participants = 0.96 (0.94 to 0.98)	Not available	9 items, scored 1-7 (1 = strongly disagree, 7 = strongly agree) Time to administer: Less than 5 minutes[27] Language: English. Translations for German, Turkish and Norwegian. The Fatigue Severity Scale can be used free of charge for clinical purposes[16]
Fatigue Severity Scale – Arabic version [1]	Intra-class correlation coefficient (total score) (95% CI): 0.92 (0.85 to 0.96). p = 0.644.	Cronbach's alpha – total score = 0.934 Item-total correlations ranging from 0.64 to 0.82.	Not available	See Fatigue Severity Scale – English version
Fatigue Severity Scale – German version (FSS-G) [34]	Lin's concordance measure rho • Healthy subjects (N=104) = 0.88 (0.84 to 0.92)	Cronbach's alpha • Healthy subjects = 0.85 • Multiple sclerosis = 0.94 • Previous ischaemic stroke = 0.96 • Sleep-wake disorders = 0.94 • Total = 0.93	Not available	See Fatigue Severity Scale – English version
Fatigue Severity Scale – Norwegian version (FSS-N) [17]	Not available	Cronbach's alpha • Baseline = 0.87 • 6 months = 0.92 • 12 months = 0.93 • 18 months = 0.92	Not available	See Fatigue Severity Scale – English version
Fatigue Severity Scale – Turkish version (FSS-T) [24]	Intra-class correlation coefficient (all items): • Stroke survivors = 0.742 (0.512-0.863), p = <0.001	Cronbach's alpha: • Stroke survivors = 0.918 • Control = 0.874	Not available	See Fatigue Severity Scale – English version

Tool name	Test-retest reliability	Internal consistency	Inter-rater reliability	Other factors
Functional Assessment of Chronic Illness Therapy – Fatigue (FACIT-F)[6]	Not available	Stroke: Cronbach alpha (overall) = 0.91 (smaller than the value when compared to people with cancer [0.96] and HIV [0.97])	Not available	40 items, scored on a 5 point Likert-type scale Time to administer: 10-15 minutes Languages: Afrikaans, Albanian, Arabic, Bengali, Bosnian, Bulgarian, Catalan, Cebuano, Chinese – Simplified and Tradition, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Georgian, German, Greek, Gujarati, Hebrew, Hiligaynon, Hindi, Hungarian, Ilokano, Indonesian, Italian, Japanese, Kannada, Kazakh, Korean, Latvian, Lithuanian, Macedonian, Malay, Malayalam, Marathi, Montenegrin, Norwegian, Odia, Polish, Portuguese, Punjabi, Romanian, Serbian, Sesotho, Setswana, Slovak, Slovene, Spanish, Swedish, Tagalog, Tamil, Telugu, Thai, Turkish, Ukranian, Urdu, Vietnamese, Xhosa, Zulu Non-commercial use is assessed on a case-by-case basis. Licensing fees are typically not applied to investigator-initiated research, students or clinical use.[10]
Lee Fatigue Scale (LFS)[4]	Not available	Cronbach's alpha coefficient: LFS-5: 0.91	Not available	Is equivalent to the Visual Analogue Scale-Fatigue 3-5 items (in this version), rated on a numeric rating scale of 0-10 (higher scores indicate higher fatigue).
Modified Fatigue Impact Scale – Cantonese Version (MFIS-C)[23]	Intraclass correlation coefficient (95% confidence interval) overall MFIS-C = 0.84 (0.74-0.91) Intraclass correlation coefficient (95% confidence interval) cognitive subscale of MFIS-C = 0.83 (0.72-0.90) Intraclass correlation coefficient (95% confidence interval) physical/social	Cronbach's alpha for overall MFIS-C = 0.92 Cronbach's alpha for the cognitive subscale of MFIS-C = 0.85 Cronbach's alpha for the physical/social subscale of MFIS-C = 0.89	Not available	21 items, scored on a 5 point Likert-type scale. Time to administer: 2-10 minutes[28]

Tool name	Test-retest reliability	Internal consistency	Inter-rater reliability	Other factors
	subscale of MFIS-C = 0.81 (0.70-0.89)			
Multidimensional Fatigue Symptom Inventory-general subscale (MFSI-general)[19]	Between questions: • Percentage agreement (range): 39-60% • Kappa (95% CI) (range): 0.48 (0.27 to 0.69) – 0.69 (0.53 to 0.85)	Intraclass correlation coefficient: • Time 1 to time 2 = 0.76 (95% CI: 0.55 to 0.87) • Rater 1 to rater 2 = 0.88 (0.78 to 0.93)	Between questions: • Percentage agreement (range): 84-93% • Kappa (95% CI) (range): 0.81 (0.60 to 1.00) – 0.94 (0.87 to 1.00)	83 item scale (for total scale), 0-4 scale (0 = not at all, 4 = extremely)
Neurological Fatigue Index-Stroke English version (NFI-Stroke)[20]	• Physical subscale = 0.903 • Cognitive subscale = 0.786 • Summary = 0.896	Person-item separation index: • Physical subscale = 0.89 • Cognitive subscale = 0.78 • Summary = 0.89	Not available	22 items, 3 subscales (physical, cognitive and summary)[33]. Free for use in all public health and not-for-profit agencies[33].
Neurological Fatigue Index-Stroke Chinese version (C-NFI-Stroke)[13]	Intra-class correlation coefficient: • Summary scale = 0.93 • Physical subscale = 0.92 • Cognitive subscale = 0.88	Cronbach's alpha: • Summary scale = 0.88 • Physical subscale = 0.87 • Cognitive subscale = 0.69	Between questions: • Percentage agreement (range): 70.4-88.9% • Weighted kappa (95%) (range): 0.47 (0.17 to 0.77) – 0.79 (0.57 to 1.01)	See Neurological Fatigue Index-Stroke English version
Neurological Fatigue Index – Stroke Norwegian version (N-NFI-Stroke)[31]	Weighted Kappa: Between questions: Varied between 0.55 (0.40-0.71) to 0.78 (0.67-0.89). Subgroup analysis was completed with chronic stroke and subacute stroke populations that showed	Cronbach's alpha: • Total = 0.90 (corrected correlation lowest and highest values: 0.50-0.78) • Physical items = 0.89 (corrected correlation lowest and highest values: 0.55-0.79)	Not available	See Neurological Fatigue Index-Stroke English version

Tool name	Test-retest reliability	Internal consistency	Inter-rater reliability	Other factors
	higher values in the chronic stroke population than the subacute stroke population. Between questions (chronic stroke): Varied between 0.61 (0.40-0.83) to 0.91 (0.82-0.99). Between questions (subacute stroke): Varied between 0.12 (-0.21-0.48) to 0.54 (0.27-0.82).	• Cognitive items = 0.74 (corrected correlation lowest and highest values: 0.46-0.60)		
Numeric Rating Scale – Faces Rating Scale (NRS-FRS)[8]	Intraclass correlation coefficient: 0.95 (0.92-0.96)	Not available	Not available	1 item, 0-10 scale with 6 facial expressions of Wong-Baker FRS. The higher the score, the higher the fatigue.
Profile of Mood States – Fatigue subscale (POMS-fatigue)[7, 19]	Between questions: • Percentage agreement (range): 55-65% • Kappa (95% CI) (range): 0.45 (0.19 to 0.72) – 0.61 (0.42 to 0.80) ICC of POMS-F (95% CI) = 0.93 (0.88-0.95)	Intraclass correlation coefficient: • Time 1 to time 2 = 0.74 (95% CI: 0.56 to 0.85) • Rater 1 to rater 2 = 0.84 (0.72 to 0.91) Cronbach's alpha for POMS-F = 0.81	Between questions: • Percentage agreement (range): 85-92% • Kappa (95% CI) (range): 0.71 (0.45 to 0.97) – 0.89 (0.75 to 1.00)	7 items. 4 point Likert scale (0 = 'not at all', 3 = 'extremely')[33]
SF-36v2 vitality subscale (SF-36v2 vitality)[7, 19]	Between questions: Percentage agreement (range): 43-59% Kappa (95% CI) (range): 0.36 (0.07 to 0.63) – 0.47 (0.25 to 0.70)	Intraclass correlation coefficient: Time 1 to time 2 = 0.51 (95% CI: 0.27 to 0.69) Rater 1 to rater 2 = 0.92 (0.86 to 0.96)	Between questions: Percentage agreement (range): 86-90% Kappa (95% CI) (range): 0.72 (0.45 to 0.99) – 0.89 (0.75 to 1.00)	4 items. 5 point Likert (scoring 1-5).[9]

Tool name	Test-retest reliability	Internal consistency	Inter-rater reliability	Other factors
		Cronbach's alpha for SF-36-VT = 0.86		
Visual Analogue Fatigue Scale (VAFS)[32]	Intraclass correlation coefficient • VAFS at rest = 0.851 • VAFS post exercise = 0.846 • VAFS post recovery = 0.888 • Exertion fatigue (VAFS post exercise – VAFS at rest) = 0.829 • Recovery rate ([VAFS post exercise – VAFS post recovery]/[VAFS post exercise – VAFS at rest] x 100) = 0.893	Not available	VAFS at rest: • Visit 1 = 7.2 (4.3) • Visit 2 = 8.3 (4.5) VAFS post exercise: • Visit 1 = 69.4 (30.5) • Visit 2 = 65.8 (31.9) VAFS post recovery: • Visit 1 = 48.5 (25.4) • Visit 2 = 47.5 (26.9) Exertion fatigue: • Visit 1 = 62.4 (29.3) • Visit 2 = 57.5 (30.9) Recovery rate: • Visit 1 = 37.0 (17.3) • Visit 2 = 37.7 (15.9)	Is equivalent to the Lee Fatigue Scale 18 items, 0-100mm line with questions falling into two subscale: fatigue (items 1-5 and 11-18) and energy (items 6-10).[30]

Table 5: Summary of the quality assessment of studies reporting tools to assess fatigue for people after a stroke

Tool name	Population	Control population	Recruitment/selection bias	Sample size	Appropriateness of metrics	Limitations considered by the study authors
A case definition for fatigue Lynch 2007[18]	Some concerns (does not include information about all subgroups)	No concerns	No concerns	Some concerns (N=55)	No concerns	• A larger sample size would have provided a more precise estimate of reliability • Ethical approval stipulated that clinical staff had to make the initial approach to patients, so patients known to be tired may have been approached more often and therefore overrepresented, reducing

Tool name	Population	Control population	Recruitment/selection bias	Sample size	Appropriateness of metrics	Limitations considered by the study authors
						the range of fatigue severity in the sample • Relied on face validity and concurrent validity as there is currently no test for identifying fatigue after stroke, making criterion validity impossible • When assessing test-retest reliability, the interviewer may have remembered the outcome of the first interview when performing the second. • The assessment of inter-rater agreement would have been more rigorous if the second rater had repeated a face-to-face case definition interview and not merely listened to a recording • Limitations to the generalisability of the data. The sample was mainly inpatients, and the nature of fatigue may change after discharge from hospital as the patients' activities become more complex. They excluded people with dysphasia or confusion for whom visual analogue scales or pictorial representations of fatigue may be appropriate.
Detection List Fatigue (DLF) Kruithof 2016[15]	Some concerns (does not include information about all subgroups)	No concerns	No concerns	No concerns	No concerns	• HADS was not administered in each patient because not every patient received a standard neuropsychological examination • The HADS, CIS-f and FSS-7 scores were not assessed at both time points, but the FRS score was

Tool name	Population	Control population	Recruitment/selection bias	Sample size	Appropriateness of metrics	Limitations considered by the study authors
Dutch Multifactor Fatigue Scale (DMFS) Visser-Keizer 2015[35]	Some concerns (results merge people with traumatic brain injury/other acute brain injuries with people after stroke)	No concerns	No concerns	No concerns	No concerns	• Only people from outpatient neurorehabilitation. This may have selected those who suffer from fatigue or do not cope with fatigue adequately • Excluded people with complaints of chronic and ongoing fatigue in life before brain injury, but still had people who experienced an episode of burnout before their brain injury which could affect results • Research in more people would be useful to increase the certainty in the results
Fatigue Assessment Scale – English version (FAS) Mead 2007[19]	Some concerns (does not include information about all subgroups)	No concerns	No concerns	Some concerns (N=55)	No concerns	• Patients were not consecutive and may not be representative of stroke patients as a whole. • Smaller sample size. However, while a larger sample size would have given more precise estimates, a sample size of 50 is usually considered sufficient for studies of agreement. • The absence of any 'gold standard' for fatigue after stroke, means they could not assess criterion validity • When assessing test-retest reliability, the interviewer may have remembered the results of the first interview when performing the second, thereby artificially increasing apparent reliability • When assessing interrater reliability, audio recordings were used rather than repeat interviews, which may potentially increase apparent reliability.

Tool name	Population	Control population	Recruitment/selection bias	Sample size	Appropriateness of metrics	Limitations considered by the study authors
						• Not all interviews could be analysed for interrater reliability because of poor quality of some recordings, mainly due to background noise on hospital wards.
Fatigue Assessment Scale – English version (FAS) Smith 2008[29]	Some concerns (pools data from all populations together, does not include information about all subgroups)	No concerns	Some concerns (baseline characteristics not balanced between groups which could influence the results)	Some concerns (N stroke population = 80)	No concerns	• The relatively small sample size • The generalisability to the stroke population may be limited due to the exclusion of people suffering, among other things, from a reduced level of consciousness
Fatigue Assessment Scale – Chinese version (FAS-C) Ho 2020[14]	Some concerns (does not include information about all subgroups)	No concerns	Some concerns (only included a small subset of the population when testing test-retest reliability)	No concerns	No concerns	• Participants needed to come to the university for the assessment, so those with a very high level of fatigue or with poorer functional mobility might not have been willing to take part in the study • Cross-sectional study and so could not show changes in fatigue over time • The data collection period went through the national public holiday, which might have affected the results
Fatigue Assessment Scale – Language unclear (FAS-I) Cheraghifard 2022 [7]	Some concerns (does not include information about all subgroups)	No concerns	No concerns	No concerns	No concerns	• Absence of control group. • Small sample size and selection bias.

Tool name	Population	Control population	Recruitment/selection bias	Sample size	Appropriateness of metrics	Limitations considered by the study authors
Fatigue Assessment Scale – Swedish version (FAS-S) Brandal 2016[5]	Some concerns (does not include information about all subgroups)	No concerns	No concerns	Some concerns (N=90)	No concerns	• The time of the day for completion of the questionnaire was not specified. • The time span between test and retest was not fixed. • The exclusion of people with severe stroke. • Recall bias when completing questionnaires.
Fatigue Impact Scale – Chinese version (FIS-C) Wu 2008[36]	Some concerns (does not include information about all subgroups)	No concerns	No concerns	No concerns	No concerns	• None provided
Fatigue Impact Scale – Turkish version (FIS-T) Batur 2022[2]	Some concerns (does not include information about all subgroups)	Some concerns (control population may also experience fatigue)	Some concerns (population recruited from a physical therapy department so may miss people who may not attend the department who may have higher levels of fatigue)	Some concerns (a limited number of participants were included	No concerns	• Small sample size which may reduce the generalization of the results • 33% missing data. It might be related to many people becoming bored due to the long form of the FIS.
Fatigue Impact Scale – Persian version (FIS-P) Saneii 2020[25]	Some concerns (does not include information about all subgroups)	No concerns	Some concerns (unclear where control population was recruited from)	No concerns	No concerns	• The use of convenience sampling method • The lack of a previous Persian version of a similar instrument

Tool name	Population	Control population	Recruitment/selection bias	Sample size	Appropriateness of metrics	Limitations considered by the study authors
Fatigue Severity Scale – English version (FSS) Nadarajah 2017[21]	Some concerns (does not include information about all subgroups)	No concerns	No concerns	No concerns	No concerns	• Patient characteristics could have biased the results (the stroke group had a greater number of males than females than the control group) • Did not examine the responsiveness of FSS cover a longer period of time or after interventions • Found a weak correlation between FSS and SF36-v2 vitality, which could be resulted from the inconsistency in modes of administration
Fatigue Severity Scale – Arabic version Abdulla 2019[1]	Some concerns (limited information about all subgroups)	No concerns	No concerns	Some concerns (a limited number of participants were included in the retest portion only)	No concerns	• All questionnaires were self-administered. Therefore, misunderstanding of the questions could have occurred. • Almost all people had mild to moderate stroke, with mild to moderate depression. Therefore, the findings cannot be extended to severe stroke or severe depression. • People were not followed up over time to investigate changes in fatigue.
Fatigue Severity Scale – German version (FSS-G) Valko 2008[34]	Some concerns (only includes people with ischaemic stroke, limited information about all subgroups)	No concerns	Some concerns (only a subset of healthy participants were included in the evaluation of test-retest reliability, baseline values not comparable)	No concerns	Some concerns (convergent validity, absence of Pearson's correlations)	• The presence of depression was not assessed

Tool name	Population	Control population	Recruitment/selection bias	Sample size	Appropriateness of metrics	Limitations considered by the study authors
Fatigue Severity Scale – Norwegian version (FSS-N) Lerdal 2011[17]	Some concerns (only includes people with ischaemic stroke, limited information about all subgroups)	No concerns	No concerns	No concerns	No concerns	• The samples in the study consisted of a convenience sample of people after their first stroke • Generalisation of findings to people with multiple stroke events and severe disability should be done with careful consideration.
Fatigue Severity Scale – Turkish version (FSS-T) Ozyemisci-Taskiran 2019[24]	Some concerns (limited information about all subgroups)	Some concerns (control population includes populations that may experience fatigue)	Some concerns (people recruited may be more predisposed to experiencing fatigue than the general population)	Some concerns (N=98, N=46 stroke survivors)	No concerns	• Small sample size, which reduced the generalizability of the results to all Turkish stroke survivors • Control group was recruited among the individuals admitted to the outpatient clinic, which may explain the presence of fatigue being higher than the general population
Functional Assessment of Chronic Illness Therapy – Fatigue (FACIT-F) Butt 2013[6]	Some concerns (limited information about all subgroups)	Some concerns (control population includes populations that would also experience significant fatigue)	No concerns	Some concerns (N=51 stroke patients)	No concerns	• The study did not ensure that the items of FACIT-F capture every aspect of fatigue experienced by patients with HIV and stroke, but it did provide reassurance that the set of questions are perceived as relevant and responsive to fatigue caused by a variety of conditions. • While the cancer sample was relatively large, this study included a limited number of patients with stroke or HIV, and so may be less generalisable to that population. However, it is unlikely that a larger sample would result in significant changes.

Tool name	Population	Control population	Recruitment/selection bias	Sample size	Appropriateness of metrics	Limitations considered by the study authors
Lee Fatigue Scale (LFS) Bragstad 2020[4]	Some concerns (limited information about all subgroups)	Some concerns (control population includes people that may experience some fatigue)	Some concerns (selection of trial participants from previous studies)	No concerns	No concerns	• The study evaluated only five of the original LFS items, so it remains unclear whether the three items retained in the analysis represent the best three items for inclusion in a brief fatigue severity patient-reported outcome measure. • Evaluated a Norwegian version of the LFS and so may not be applicable to English language (some words and phrases do not translate directly). • The mode of data collection was not identical in the stroke and osteoarthritis samples. The stroke population was interviewed in person, while the osteoarthritis group completed the questionnaire. • Although no difference was found based on diagnostic groups, the different data collection mode for the two samples may have introduced bias in the interpretation of items.
Modified Fatigue Impact Scale – Cantonese Version (MFIS-C) Ng 2022[23]	Some concerns (does not include information about all subgroups)	No concerns	No concerns	Some concerns (limited sample size for test-retest reliability)	No concerns	• The translation of the questionnaire may be appropriate to people in the Hong Kong Environment but not to other Chinese populations such as those in Mainland China due to differences in culture. • The sample size calculation is based on reliability, and so may not be sufficient to detect the significant correlations between MFIS-C scores and other outcome measure scores

Tool name	Population	Control population	Recruitment/selection bias	Sample size	Appropriateness of metrics	Limitations considered by the study authors
						• The findings could only be generalised to those fulfilling the inclusion and exclusion criteria • The expert panel did not consist of a member with occupational training background. • The stroke participants were not checked for a history of chronic fatigue syndrome before entering the trial. • They did not examine the construct validity of the MFIS-C due to the insufficient number of subjects.
Multidimensional Fatigue Symptom Inventory-general subscale (MFSI-general) **Mead 2007[19]**	Some concerns (does not include information about all subgroups)	No concerns	No concerns	Some concerns (N=55)	No concerns	See Fatigue Assessment Scale (FAS)
Neurological Fatigue Index-Stroke English version (NFI-Stroke) **Mills 2012[20]**	Some concerns (does not include information about all subgroups)	No concerns	No concerns	No concerns	No concerns	• The non-response level of the study was high and older patients with multiple strokes were underrepresented. People with low levels of disability were well represented. However, some respondents had very high SIS scores suggesting that those with higher disability were not wholly excluded. • There was one item in the cognitive scale (coordination gets worse) with a slightly high negative fit residual which indicated a degree of redundancy and accounted for the inflated overall item residual standard deviation. They did not discard the scale because all other fit

Tool name	Population	Control population	Recruitment/selection bias	Sample size	Appropriateness of metrics	Limitations considered by the study authors
						statistics were acceptable and the retention of a comparative cognitive fatigue scale between stroke and MS was felt to be desirable.
Neurological Fatigue Index-Stroke Chinese version (C-NFI-Stroke) **Ho 2021**[13]	Some concerns (does not include information about all subgroups)	Some concerns (healthy participants are caregivers of stroke survivors, so may not represent the general population)	No concerns	No concerns	No concerns	• Participants came from only a few local self-help groups and a non-governmental organisation, limiting the generalisability of this study. • Some potential participants were unwilling to join because of the need to travel long distances to the assessment venue. • The sample size might not be adequate for comparisons to be made between stroke survivors and healthy older people with and without depressive symptoms • Since some healthy older people were caregivers of stroke survivors, they likely were not representative of the general healthy population.
Neurological Fatigue Index – Stroke Norwegian version (N-NFI-Stroke) **Taasen 2020**[31]	Some concerns (does not include information about all subgroups)	No concerns	No concerns	Some concerns (N=63)	No concerns	• The sample was of an adequate size according to the requirements for analysis of categorical data, but a larger sample would have been preferable to generalise results. • People with chronic stroke were predominant, with few people with subacute stroke. • The participants were slightly younger than the general stroke population and to a high degree independent with activities of daily living, which may limit generalisation of the results.

Tool name	Population	Control population	Recruitment/selection bias	Sample size	Appropriateness of metrics	Limitations considered by the study authors
						• The test procedures were performed under different circumstances, which could affect how answers were given between participants. • The study used classical test theory rather than modern psychometric techniques. • It may have added to the results if invariance in item difficulty between language versions was evaluated.
Numeric Rating Scale – Faces Rating Scale (NRS-FRS) Chuang 2015[8]	Some concerns (does not include information about all subgroups)	No concerns	No concerns	No concerns	No concerns	• All participants completed the NRS-FRS at two assessments, at the same time of day to minimize diurnal variation in fatigue. Fatigue was measured as a single time-point assessment; this is, current fatigue intensity, which might not reflect overall fatigue intensity, which might not reflect overall fatigue on the testing day. Testing at different times of the day may help to show whether there are changes in daily fluctuation that may be used to improve the psychometric properties of the test. • Future studies need to identify predictors of poststroke fatigue to address fatigue issues with an intervention in people with stroke. To explore the effectiveness, the ability to detect NRS-FRS to detect change over time requires further development.
Profile of Mood States – Fatigue subscale (POMS-	Some concerns (does not include	No concerns	No concerns	Some concerns (N=55)	No concerns	See Fatigue Assessment Scale (FAS)

Tool name	Population	Control population	Recruitment/selection bias	Sample size	Appropriateness of metrics	Limitations considered by the study authors
fatigue) – English version Mead 2007[19]	information about all subgroups)					
Profile of Mood States – Fatigue subscale – Language unclear (POMS-F) Cheraghifard 2022[7]	No concerns	No concerns	No concerns	No concerns	No concerns	• Absence of control group. • Small sample size and selection bias.
SF-36v2 vitality subscale (SF-36v2 vitality) – English version Mead 2007[19]	Some concerns (does not include information about all subgroups)	No concerns	No concerns	Some concerns (N=55)	No concerns	See Fatigue Assessment Scale (FAS)
SF-36 vitality subscale – Language unclear (SF-36-VT) Cheraghifard 2022[7]	No concerns	No concerns	No concerns	No concerns	No concerns	• Absence of control group. • Small sample size and selection bias.
Visual Analogue Fatigue Scale (VAFS) Tseng 2010[32]	Some concerns (does not include information about all subgroups)	No concerns	Some concerns (limited information about where participants came from, small number of people fulfilling subgroups)	Some concerns (N=21)	No concerns	• Subgroup analysis was not possible to see if the reliability and validity would be different between gender, different types of stroke and different times post stroke. • Although baseline fatigue was detected, the study was not designed to distinguish between other types of fatigue.

1.1.9 Economic evidence

1.1.9.1 Included studies

No health economic studies were included.

1.1.9.2 Excluded studies

No relevant health economic studies were excluded due to assessment of limited applicability or methodological limitations.

See also the health economic study selection flow chart in Appendix G.

1.1.10 Summary of included economic evidence

This area was not prioritised for new cost-effectiveness analysis.

1.1.11 Unit costs

Fatigue assessments require additional resource use compared to not providing such assessments related to staff time and questionnaires (printing and potential licensing costs). Studies included in the clinical review reported varied resource use due to either the staff type arranging the assessment, time taken to administer the test (for instruments where this information was available, ranging from 2-15 minutes) and whether communication with patients was conducted via face to face, phone or written correspondences. For instance, Taasen[31] reported that face-to-face appointments allowed for someone to read the documents aloud for patients when appropriate, while Brandal[5] used self-administered tests and asked patients to post the completed responses to lower costs associated with personal interviews. Generally, there are no charges for using the assessment questionnaires identified in the clinical review (although sometimes information regarding this was not identified) – see Table 4.

Relevant unit costs are provided below to aid consideration of cost effectiveness.

Table 4: Unit costs of health care professionals who may be involved in providing fatigue assessments

Resource	Cost per working hour (hospital/community)[(a)]	Cost for time taken to administer questionnaire (2 / 15 minutes)		Source
		hospital	community	
Band 6 PT/OT	£52/£50	£1.73/£13	£1.67/£12.50	PSSRU 2020[3]
Band 7 PT/OT	£62/£60	£2.07/£15.50	£2.00/£15	
Band 6 nurse	£53/£52	£1.77/£13.25	£1.73/£13	
Band 7 nurse	£62/£61	£2.07/ £15.50	£2.03/£15.25	

(a) Note: Costs per working hour include salary, salary oncosts, overheads (management and other non-care staff costs including administration and estates staff), capital overheads and qualification costs.

1.1.13 Evidence statements

Effectiveness/Qualitative

Economic

No relevant economic evaluations were identified.

1.1.14 The committee's discussion and interpretation of the evidence

1.1.14.1. The outcomes that matter most

This review included an assess-and-treat review and a validity and reliability review. No evidence was found for the assess-and-treat review. For the validity and reliability review, the committee took into account the validity (including face/content/construct validity, criterion/concurrent validity, discriminant/convergent validity), the reliability (including test-retest reliability, internal consistency and inter-rater reliability), responsiveness to change and the dimensions of fatigue considered. The committee gave equal weighting to all these outcomes.

The reporting of outcomes for different tools varied, with some reporting all aspects of validity and reliability, while others only reporting one factor for each or less. The reporting of responsiveness to change was also inconsistent, with it only being reported for a small number of tools.

1.1.14.2 The quality of the evidence

The review itself considers the quality of the tools and this aspect is discussed further in the section about advantages and disadvantages. However, further assessment of the risk of bias for studies to help assess the quality was conducted by analysing the following factors: population, control population, recruitment/selection bias, sample size and appropriateness of metrics. All studies had elements of population bias as they did not consider all of the population subgroups that the committee considered would be important. Otherwise, the risk of bias varied between studies. A common reason for concern was a small sample size (with less than 100 people in the study), recruitment of control populations that may experience fatigue and so may affect interpretation of the results (for example, people with osteoarthritis), and for unclear or potentially biased recruitment of participants.

1.1.14.3 *Key uncertainties*

No evidence was found for the assess-to-treat review. Due to this, the committee acknowledged that any recommendations they made based on the validity and reliability evidence would be missing evidence on whether the tool had an effect in clinical practice. However, an assess-to-treat study would be difficult to design for fatigue, as expert opinion is that there is no proven treatment for fatigue and so it would be difficult to know whether the use of an assessment tool led to any difference since any changes, or lack of them, might be a result of an inconsistent response to treatment.

While the committee acknowledged the limitations of basing their recommendations on validity and reliability studies only, they considered that any recommendation would help identify people with fatigue so that appropriate allowances could be made (for example: setting goals for rehabilitation while considering fatigue levels, providing coping strategies). However, given the lack of evidence of clinical effectiveness, the committee agreed a research recommendation to further investigate this.

The committee acknowledged that the experiences of fatigue are likely to overlap with other conditions, such as depression. The committee agreed that all causes of fatigue should be considered when assessing post-stroke fatigue, including any reversible causes such as anaemia. Anyone using these tools should be aware that fatigue can be caused by a wide range of conditions, and so should take a holistic view of the person after stroke and consider all possible causes.

The committee acknowledged that the tools mentioned in the recommendation are not adapted for people with communication or cognitive difficulties. Given this, they recommended further research into developing versions of these tools that are adapted to these populations where their validity and reliability could be assessed.

1.1.14.4 Benefits and harms

Twenty four reliability and validity studies considered 14 tools. These tools were:

- A case definition for fatigue
- Detection List Fatigue
- Dutch Multifactor Fatigue Scale (DMFS)
- Fatigue Assessment Scale (FAS) – including the English, Chinese and Swedish versions
- Fatigue Impact Scale (FIS) and the Modified Fatigue Impact Scale (MFIS) – including the Chinese, Persian and Turkish versions
- Fatigue Severity Scale (FSS) – including the English, Arabic, German, Norwegian and Turkish versions
- Functional Assessment of Chronic Illness Therapy – Fatigue (FACIT-F)
- Lee Fatigue Scale (LFS)
- Multidimensional Fatigue Symptom Inventory-general subscale (MFSI-general)
- Neurological Fatigue Index-Stroke (NFI-S) – including the English, Chinese and Norwegian version
- Numeric Rating Scale – Faces Rating Scale (NRS-FRS)
- Profile of Mood States – Fatigue subscale (POMS-fatigue)
- SF-36v2 vitality subscale (SF-36v2 vitality)
- Visual Analogue Fatigue Scale (VAFS)

The committee considered the appropriateness for each tool and came to the following conclusions:

- The following tools may have sufficient parameters for both validity and reliability (reporting at least one parameter for validity and for reliability that was of sufficient quality for use for people after stroke):
  - A case definition for fatigue
  - Fatigue Assessment Scale
  - Fatigue Impact Scale and Modified Fatigue Impact Scale
  - Fatigue Severity Scale
  - Lee Fatigue Scale
  - Multidimensional Fatigue Symptom Inventory-general subscale
  - Numeric Rating Scale – Faces Rating Scale
  - Profile of Mood States – Fatigue subscale
  - Visual Analogue Fatigue Scale

- The following tools either did not report 1 of either validity or reliability, or had at least 1 parameter for validity and reliability where the quality was insufficient for use for people after stroke:
  - Detection List Fatigue – no evidence for reliability
  - Dutch Multifactor Fatigue Scale – evidence of insufficient discriminant/convergent validity compared to people with traumatic brain injury and only indirect evidence for internal consistency that includes people with traumatic brain injury, which made it difficult for the committee to make a conclusion
  - Functional Assessment of Chronic Illness Therapy-Fatigue – no direct evidence for validity (while the study reported discriminant validity for different conditions, there were no formal comparisons and the evidence provided did not indicate a substantial difference between the populations)
  - SF-36v2 vitality subscale – evidence of insufficient test-retest reliability for people after a stroke.
- The following tools were only validated in non-English language and so the committee concluded that it would be difficult to assess their quality for a United Kingdom population based on this evidence:
  - Detection List Fatigue
  - Dutch Multifactor Fatigue Scale
  - Fatigue Impact Scale and Modified Fatigue Impact Scale – However, the committee agreed that this tool was widely used in English language for other conditions such as Multiple Sclerosis and was widely available. Therefore, they agreed to consider it based on the evidence of reliability and validity in other languages.
- The following were subscales of larger tools where either the scores may not have been designed to specifically investigate fatigue, or the tool as a whole may be appropriate but only a small section has been appraised. Based on this, the committee agreed that a tool designed to be used specifically to examine fatigue may be more appropriate:
  - Multidimensional Fatigue Symptom Inventory-general subscale
  - SF-36v2 vitality subscale
  - Profile of Mood States – Fatigue subscale

Based on this, the committee further examined the following tools for their applicability:

- A case definition for fatigue
- Fatigue Assessment Scale
- Fatigue Impact Scale and Modified Fatigue Impact Scale
- Fatigue Severity Scale
- Lee Fatigue Scale
- Neurological Fatigue Index – Stroke
- Numeric Rating Scale – Faces Rating Scale
- Visual Analogue Fatigue Scale

The committee weighed up the advantages and disadvantages of each tool. They agreed that for a tool to be useful in the NHS it should be a tool that: is easy for the stroke survivor and professional to use and interpret; is able to assess fatigue over multiple domains (for example: physical, cognitive, social); is able to provide information on the amount of fatigue being experienced so that the effect of any coping strategies or interventions can be monitored; can prompt further discussions around the causes of fatigue to identify any additional strategies; and is available for use without a significant cost.

With the evidence available, the committee were unable to discern the questions used in the Lee Fatigue Scale. Therefore, with this lack of information and limited evidence of validity and reliability when compared to other tools, the committee did not recommend the use of this tool. Similar findings were seen with the Neurological Fatigue Index – Stroke, where there was limited information about the usability of the scale and the values for reliability varied significantly between questions, with some falling below an acceptable range. While there was limited information about the usability, the committee noted that there was a significantly larger number of items for the scale (22 items) compared to others included for consideration, so the scale may take longer and be more difficult to complete. Therefore, the committee did not recommend the use of this tool. The committee agreed the evidence for the Visual Analogue Fatigue Scale was also limited. While they could assess the questions included in the tool, they agreed that the results were more general rather than specifically looking at domains of fatigue. The evidence for validity was limited to comparisons against tools that may not measure fatigue, and while the evidence of test-retest and inter-rater reliability was sufficient, there was no evidence for internal consistency. Given this, the committee did not recommend the use of this tool.

The committee discussed the Case definition for fatigue. While the interview for this was noted to be useful, providing a lot of insight into the nature of fatigue a person may be experiencing, the tool did not have complete information to confirm its validity and reliability. Given that this tool would likely take longer to complete, the committee agreed that this tool may not be appropriate to initially determine if someone had fatigue, and the nature of the answers may make it less objective when assessing the response to management strategies. Given this, the committee did not recommend the use of this tool for the purposes of this question.

The committee considered the Fatigue Assessment Scale, Fatigue Impact Scale/Modified Fatigue Impact Scale and the Fatigue Severity Scale. The committee noted that the Fatigue Impact Scale is a longer form version of the Modified Fatigue Impact Scale and agreed that in practice the Modified Fatigue Impact Scale would be used, but that the results could be applied between the two tools. They agreed that each tool appeared to fulfil the criteria for a good tool to assess fatigue. Each provided information on multiple domains of fatigue, with the Fatigue Assessment Scale and Fatigue Impact Scale/Modified Fatigue Impact Scale providing this information in subscales, which would support users to better interpret the results. While the studies investigating the Fatigue Severity Scale did not report the inter-rater reliability, it appeared to have superior test-retest reliability and internal consistency than the Fatigue Assessment Scale. All of the tools could be completed in a short amount of time. They are also available in a range of different languages and so could be used by the diverse population served by the NHS. Given these factors, the committee recommended the three tools, agreeing that they would be appropriate for assessing fatigue. However, given the lack of evidence for the clinical effectiveness of the tools, the committee agreed a research recommendation to investigate whether these tools would provide a useful change to current practice and would improve outcomes for stroke survivors.

The committee discussed the value of the Numeric Rating Scale – Faces Rating Scale. They acknowledged that this scale is the only tool that has been adapted to provide a non-verbal method of communicating fatigue. However, the tool assessed 1 question rather than multiple domains of fatigue. The committee thought that face symbols may not accurately match the experience of fatigue and so the validity of the tool was questioned. Given this, the committee did not recommend this tool. However, they agreed that tools specifically adapted and validated for people with communication and cognitive difficulties were required to ensure equity across stroke services. Given this, the committee agreed a research recommendation to investigate this

1.1.14.5 Cost effectiveness and resource use

No economic evidence was identified for this question. Unit costs of healthcare professionals who may be involved in providing post-stroke fatigue assessments were presented to the committee to inform consideration of cost-effectiveness.

Although there is limited information for the clinical benefit of fatigue assessments, the committee emphasised that identification and assessment of fatigue early in the rehabilitation process can significantly improve how a person responds to their treatment plan. This allows health care professionals to set more appropriate goals and prevent people from becoming discouraged if they do not meet targets that would have been otherwise set without the fatigue assessment. It was stressed that an individual's motivation to adhere to treatment and general morale are important factors that healthcare professionals need to consider when designing a recovery plan. Therefore, the committee agreed that any additional resource use that was required was justified by the benefit to patients.

Additional resource use associated with fatigue assessment will largely relate to staff time. Based on the clinical evidence summarised above, the most appropriate assessments of fatigue were the Fatigue Severity Scale, Fatigue Assessment Scale and the Fatigue Impact Scale/Modified Fatigue Impact Scale (Modified Fatigue Impact Scale is the longer form of the Fatigue impact Scale, producing results that can be applied between the two tools). The staff time and therefore, cost associated with administering these assessments is expected to be fairly low as they are quick to complete, approximately taking between 2-10 minutes to complete. Furthermore, they can be completed by the patient unsupervised if they are able to do so. Finally, there does not appear to be charges for using either questionnaire, however there will be some costs associated with printing the questionnaires.

The other tools identified in the clinical review were not considered for a clinical recommendation based on the following reasons: they did not report either validity or reliability or had at least 1 parameter for validity or reliability where the quality was insufficient for use for people after stroke; validation was done in non-English language which limited interpretation of their quality for a UK population; or the tools were subscales of larger tools where either the scores may not have been designed to specifically investigate fatigue, or the tool as a whole may be appropriate but only a small section had been appraised.

The committee agreed to make a 'consider' recommendation for including a standardised written assessment of fatigue as part of a 6-month stroke review. They specified that the Fatigue Assessment Scale, Fatigue Severity Scale and the Modified Fatigue Impact Scale could be considered as they were deemed to be the most appropriate for assessing fatigue. A stronger recommendation was not possible given the limited clinical evidence and lack of cost-effective evidence. The committee agreed that in addition to the consider recommendations, given the limited clinical evidence, a research recommendation to investigate whether these tools would provide a useful change to current practice and would improve outcomes for stroke survivors was warranted. The committee agreed that assessing and investigating post-stroke fatigue is part of current best practice but is not necessarily common practice in the NHS currently. However, given the low cost associated with providing these assessments it is not expected to lead to a significant resource impact to the NHS in England. An additional research recommendation was made to help identify the optimal tool for assessing fatigue in people after stroke with communication difficulties.

1.1.14.6 Other factors the committee took into account

The committee discussed the nature of fatigue. Fatigue is a common and disabling symptom for people after a stroke. In one study, 51% of people were found to have significant fatigue at 6 months[12]. Of those people, some of them experienced fatigue earlier after their stroke, while 38% reported new fatigue at 6 months. The committee agreed that, while there is not a

consistent treatment for post-stroke fatigue, there are strategies that can be implemented to help manage fatigue. If fatigue is considered earlier during rehabilitation, then the method of providing the rehabilitation can be adapted to the person's fatigue. Fatigue can be used to help guide goal setting and coping strategies can be taught. The committee recommended that fatigue should be considered throughout the person's recovery after stroke and should be assessed whenever it is relevant. However, in order to help with planning rehabilitation an assessment should be made at a reasonably early stage although the committee agreed that assessing fatigue during the early stages of hospital care (for example, in hyperacute stroke units) may not be appropriate as a lot of people may be experiencing fatigue that will change over the next days and weeks. They also agreed that an opportune moment to assess fatigue could be the 6-month assessment that should be provided for people after stroke (and is assessed for in the SSNAP audit) as this could help guide future management.

The accessibility of the tools was considered. For the most part, the tools consisted of verbal information. These tools could be completed by the person but some who find it difficult to understand and respond to the questions may require support. This could be provided by carers or healthcare professionals. Some tools may be difficult for people with aphasia to use. One tool used a pictorial representation for fatigue. The value of this tool for people where a verbal questionnaire may not be appropriate was considered by the committee when making recommendations. However, given the limited questions in the tool, the committee did not recommend its use as they did not want a worse quality of service to be provided for people with communication or cognitive difficulties They agreed that if recommended tools could be adapted for people in this population, they would provide a lot more information to help support them. Given this the committee made a research recommendation in Appendix K.

The committee recognised the value of completion of assessment tools for fatigue face-to-face with the person after stroke. This gives an opportunity for the person to discuss any concerns and anxiety about having fatigue, and for the healthcare professional to listen and give reassurance that fatigue is a typical part of stroke recovery.

1.1.15 Recommendations supported by this evidence review

This evidence review supports recommendations 1.7.1 and 1.7.2 and the research recommendations on tools for assessing fatigue in people with communication difficulties and tools for fatigue in Appendix K.

1.1.16 References

1. Abdulla FA, Al-Khamis FA, Alsulaiman AA, Alshami AM. Psychometric properties of an Arabic version of the fatigue severity scale in patients with stroke. Topics in Stroke Rehabilitation. 2019; 26(6):448-455

2. Batur EB, Ozyemisci-Taskiran O, Yuksel S, Cengiz M, Karatas GK. Validity and reliability of the fatigue impact scale in stroke. Topics in Stroke Rehabilitation. 2022; 29(7):526-537

3. Beecham J, Curtis L. Unit costs of health and social care 2020. Canterbury. Personal Social Services Research Unit University of Kent, 2020. Available from: https://www.pssru.ac.uk/project-pages/unit-costs/

4. Bragstad LK, Lerdal A, Gay CL, Kirkevold M, Lee KA, Lindberg MF et al. Psychometric properties of a short version of Lee Fatigue Scale used as a generic PROM in persons with stroke or osteoarthritis: assessment using a Rasch analysis approach. Health & Quality of Life Outcomes. 2020; 18(1):168

5. Brandal A, Eriksson M, Wester P, Lundin-Olsson L. Reliability and validity of the Swedish Fatigue Assessment Scale when self-administrered by persons with mild to moderate stroke. Topics in Stroke Rehabilitation. 2016; 23(2):90-97

6. Butt Z, Lai JS, Rao D, Heinemann AW, Bill A, Cella D. Measurement of fatigue in cancer, stroke, and HIV using the Functional Assessment of Chronic Illness Therapy - Fatigue (FACIT-F) scale. Journal of Psychosomatic Research. 2013; 74(1):64-68

7. Cheraghifard M, Sarlak N, Taghizadeh G, Azad A, Fallah S, Akbarfahimi M. Minimal and robust clinically important difference of three fatigue measures in chronic stroke survivors. Topics in Stroke Rehabilitation. 2022:1-10

8. Chuang LL, Lin KC, Hsu AL, Wu CY, Chang KC, Li YC et al. Reliability and validity of a vertical numerical rating scale supplemented with a faces rating scale in measuring fatigue after stroke. Health & Quality of Life Outcomes. 2015; 13:91

9. Elbers RG, Rietberg MB, van Wegen EE, Verhoef J, Kramer SF, Terwee CB et al. Self-report fatigue questionnaires in multiple sclerosis, Parkinson's disease and stroke: a systematic review of measurement properties. Quality of Life Research. 2012; 21(6):925-944

10. FACIT Group. Functional Assessment of Chronic Illness Therapy – Fatigue: A 13-item FACIT Fatigue Scale. Available from: https://www.facit.org/measures/FACIT-F Last accessed: 13/02/2023.

11. Frith J, Newton J. Fatigue Impact Scale. Occupational Medicine (Oxford, England). 2010; 60(2):159

12. Hawkins L, Lincoln NB, Sprigg N, Ward NS, Mistri A, Tyrrell P et al. The Nottingham Fatigue After Stroke (NotFAST) study: results from follow-up six months after stroke. Topics in Stroke Rehabilitation. 2017; 24(8):592-596

13. Ho LY, Lai CK, Ng SS. Testing the psychometric properties of the Chinese version of the Neurological Fatigue Index-Stroke. Clinical Rehabilitation. 2021; 35(9):1329-1340

14. Ho LYW, Lai CKY, Ng SSM. Measuring fatigue following stroke: the Chinese version of the Fatigue Assessment Scale. Disability and Rehabilitation. 2020:1-8

15. Kruithof N, Van Cleef MH, Rasquin SM, Bovend'Eerdt TJ. Screening Poststroke Fatigue; Feasibility and Validation of an Instrument for the Screening of Poststroke

Fatigue throughout the Rehabilitation Process. Journal of Stroke and Cerebrovascular Diseases. 2016; 25(1):188-196

16. Krupp LB. Fatigue Severity Scale (FSS). 1989. Available from: https://www.thoracic.org/members/assemblies/assemblies/srn/questionaires/fss.php Last accessed: 13/02/2023.

17. Lerdal A, Kottorp A. Psychometric properties of the Fatigue Severity Scale-Rasch analyses of individual responses in a Norwegian stroke cohort. International Journal of Nursing Studies. 2011; 48(10):1258-1265

18. Lynch J, Mead G, Greig C, Young A, Lewis S, Sharpe M. Fatigue after stroke: the development and evaluation of a case definition. Journal of Psychosomatic Research. 2007; 63(5):539-544

19. Mead G, Lynch J, Greig C, Young A, Lewis S, Sharpe M. Evaluation of fatigue scales in stroke patients. Stroke. 2007; 38(7):2090-2095

20. Mills RJ, Pallant JF, Koufali M, Sharma A, Day S, Tennant A et al. Validation of the Neurological Fatigue Index for stroke (NFI-Stroke). Health & Quality of Life Outcomes. 2012; 10:51

21. Nadarajah M, Mazlan M, Abdul-Latif L, Goh HT. Test-retest reliability, internal consistency and concurrent validity of Fatigue Severity Scale in measuring post-stroke fatigue. European journal of physical & rehabilitation medicine. 2017; 53(5):703-709

22. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual [updated January 2022]. London. National Institute for Health and Care Excellence, 2014. Available from: https://www.nice.org.uk/process/pmg20

23. Ng SSM, Liu TW, Tsoh J. Translation and initial validation of Chinese (Cantonese) version of Modified Fatigue Impact Scale (MFIS-C) in people with stroke. BMC Neurology. 2022; 22(1):300

24. Ozyemisci-Taskiran O, Batur EB, Yuksel S, Cengiz M, Karatas GK. Validity and reliability of fatigue severity scale in stroke. Topics in Stroke Rehabilitation. 2019; 26(2):122-127

25. Saneii SH, Heidari M, Zaree M, Akbarfahimi M. Psychometric features of the persian version of the fatigue impact scale in iranian stroke patients. Journal of Advances in Medical and Biomedical Research. 2020; 28(127):111-110

26. Shahid A, Wilkinson K, Marcu S, Shapiro CM. Fatigue Assessment Scale (FAS). 'In:' Shahid A, Wilkinson K, Marcu S, Shapiro CM, editors. STOP, THAT and One Hundred Other Sleep Scales. New York, NY: Springer New York. 2012. p. 161-162.

27. Shirley Ryan Ability Lab. Fatigue Severity Scale. 2016. Available from: https://www.sralab.org/rehabilitation-measures/fatigue-severity-scale Last accessed: 13/02/2023.

28. Shirley Ryan Ability Lab. Modified fatigue impact scale. 2012. Available from: https://www.sralab.org/rehabilitation-measures/modified-fatigue-impact-scale Last accessed: 13/02/2023.

29. Smith OR, van den Broek KC, Renkens M, Denollet J. Comparison of fatigue levels in patients with stroke and patients with end-stage heart failure: application of the fatigue assessment scale. Journal of the American Geriatrics Society. 2008; 56(10):1915-1919

30. Stein KD, Jacobsen PB. Information about the fatigue symptom inventory (FSI) and the multidimensional fatigue symptom inventory (MFSI). Available from: http://www.cas.usf.edu/~jacobsen/HANDOUT.FSI&MFSI.pdf

31. Taasen I, Loureiro AP, Langhammer B. The neurological fatigue index for stroke. Reliability of a Norwegian version. Physiotherapy Theory & Practice. 2020:1-8

32. Tseng BY, Gajewski BJ, Kluding PM. Reliability, responsiveness, and validity of the visual analog fatigue scale to measure exertion fatigue in people with chronic stroke: a preliminary study. Stroke research and treatment. 2010;

33. Tyson SF, Brown P. How to measure fatigue in neurological conditions? A systematic review of psychometric properties and clinical utility of measures used so far. Clinical Rehabilitation. 2014; 28(8):804-816

34. Valko PO, Bassetti CL, Bloch KE, Held U, Baumann CR. Validation of the fatigue severity scale in a Swiss cohort. Sleep. 2008; 31(11):1601-1607

35. Visser-Keizer AC, Hogenkamp A, Westerhof-Evers HJ, Egberink IJ, Spikman JM. Dutch multifactor fatigue scale: a new scale to measure the different aspects of fatigue after acquired brain injury. Archives of Physical Medicine and Rehabilitation. 2015; 96(6):1056-1063

36. Wu CW, Liu ZD, Zhang YB, Li JM, Wang DX. Validity and reliability of Chinese version of Fatigue Impact Scale in cerebral infarction patients. Neural Regeneration Research. 2008; 3(2):177-181

Appendices

Appendix A – Review protocols

Review protocol for the optimal tool for assessment of fatigue in people after a stroke

ID	Field	Content
0.	PROSPERO registration number	CRD42021272568
1.	Review title	In people after stroke, what is the optimal tool for assessment of fatigue?
2.	Review question	2.1 In people after stroke, what is the optimal tool for assessment of fatigue?
3.	Objective	To determine the optimal tool for assessment of fatigue in people after a stroke.
4.	Searches	Key papers: Ozyemisci-Taskiran O, Batur EB, Yuksel S, Cengiz M, Karatas GK. Validity and reliability of fatigue severity scale in stroke. Top Stroke Rehabil. 2019 Mar;26(2):122-127 Mills, R.J., Pallant, J.F., Koufali, M. et al. Validation of the Neurological Fatigue Index for stroke (NFI-Stroke). Health Qual Life Outcomes 10, 51 (2012). Mead G, Lynch J, Greig C, Young A, Lewis S, Sharpe M. Evaluation of fatigue scales in stroke patients. Stroke. 2007 Jul;38(7):2090-5. Bråndal A, Eriksson M, Wester P, Lundin-Olsson L. Reliability and validity of the Swedish Fatigue Assessment Scale when self-administrered by persons with mild to moderate stroke. Top Stroke Rehabil. 2016 Apr;23(2):90-7 Smith, O.R.F., Van Den Broek, K.C., Renkens, M. and Denollet, J. (2008), Comparison of Fatigue Levels in Patients with Stroke and Patients with End-Stage Heart Failure: Application of the Fatigue Assessment Scale. Journal of the American Geriatrics Society, 56: 1915-1919 Ho LYW, Lai CKY, Ng SSM. Measuring fatigue following stroke: the Chinese version of the Fatigue Assessment Scale. Disabil Rehabil. 2020 Mar 6:1-8. Ho LY, Lai CK, Ng SS. Testing the psychometric properties of the Chinese version of the Neurological Fatigue Index-Stroke. Clin Rehabil. 2021 Mar 16:2692155211001684. Taasen I, Loureiro AP, Langhammer B. The neurological fatigue index for stroke. Reliability of a Norwegian version. Physiother Theory Pract. 2020 Sep 24:1-8.

		Abdulla FA, Al-Khamis FA, Alsulaiman AA, Alshami AM. Psychometric properties of an Arabic version of the fatigue severity scale in patients with stroke. Top Stroke Rehabil. 2019 Sep;26(6):448-455. The following databases (from inception) will be searched: • Cochrane Central Register of Controlled Trials (CENTRAL) • Cochrane Database of Systematic Reviews (CDSR) • Embase • MEDLINE • Epistemonikas • PsycINFO Searches will be restricted by: • English language studies • Human studies Other searches: • Inclusion lists of systematic reviews The searches may be re-run 6 weeks before the final committee meeting and further studies retrieved for inclusion if relevant. The full search strategies will be published in the final review. Medline search strategy to be quality assured using the PRESS evidence-based checklist (see methods chapter for full details).
5.	Condition or domain being studied	Adults and young people (16 or older) after a stroke
6.	Population	Inclusion: • Adults (age ≥16 years) who have had a first or recurrent stroke (including people after subarachnoid haemorrhage) Exclusion: • Children (age <16 years)

		• People who have had a transient ischaemic attack

7.	Intervention/Exposure/Test	Assess-and-treat review
		• Tools for assessment of fatigue after a stroke:
		  ○ Fatigue Assessment Scale
		  – Cut off: 23
		  ○ Fatigue Severity Scale
		  – Cut off: 36
		  ○ Brief Fatigue Inventory
		  – Cut offs:
		  • 1-3 (mild)
		  • 4-7 (moderate)
		  • 8-10 (severe)
		  ○ Combinations of the above
		Where studies include a mixture of the above categories studies will be included if at least 80% satisfy the criteria for one category. If <10% of participants are in a different category (for example: 9% have a Fatigue Severity Scale assessment, 91% have a Modified Fatigue Impact Scale assessment this study will be included in the majority category without downgrading for indirectness. If 10-20% are in a different category, this study will be included in the majority category and downgraded for intervention indirectness.
		
		Validity and Reliability review
		• Any tools for assessment of fatigue after a stroke (either designed for a stroke survivor population, or later validitated for stroke survivors)

8.	Comparator/Confounding factors	Assess-and-treat review • Each other Confounding factors (for non-randomised studies only): • Presence of comorbidities • Stroke severity • Time period since stroke • Medication usage • Age • Presence of communication difficulties • Baseline psychological distress scores Confounding factors to be considered in the inclusion criteria (studies will not be excluded if they do not adjust for this in a multivariate/univariate analysis or with matched groups): • Time of day of test administration (to consider in the inclusion criteria)

		Validity and reliability review Compared to the validity and reliability in a healthy population (or other non-stroke survivor population) or to itself (for reliability)

9.	Types of study to be included	Assess-and-treat review
		• Systematic reviews of RCTs
		• Parallel RCTs (test and treat)
		• Non-randomised studies (if insufficient evidence from parallel RCTs)
		  ○ Prospective cohort study
		  ○ Retrospective cohort study
		
		Published NMAs and IPDs will be considered for inclusion.
		
		Non-randomised studies will only be included if all of the key confounders have been accounted for in a multivariate analysis. In the absence of multivariate analysis, studies that account for key confounders with univariate analysis or matched groups will be considered.
		
		Validity and reliability review
		
		Cohort studies and cross sectional studies (investigating tool validation)

10.	Other exclusion criteria	• Non-English language studies • Conference abstracts will be excluded as it is expected there will be sufficient full text published studies available.
11.	Context	People with fatigue after a stroke. This may include people in an acute (<7 days), subacute (7 days – 6 months) or chronic (>6 months) time horizon.

12.	Primary outcomes (critical outcomes)	All outcomes are considered equally important for decision making and therefore have all been rated as critical:
		
		Clinical effectiveness (assess-and-treat) outcomes:
		At time period
		• <1 year
		• ≥1 year
		
		• Person/participant generic health-related quality of life (continuous outcomes will be prioritised)
		  ○ EQ-5D
		  ○ SF-6D
		  ○ SF-36
		  ○ SF-12
		  ○ Other utility measures (AQOL, HUI, 15D, QWB)

		• Carer generic health-related quality of life (continuous outcomes will be prioritised)
		  ○ EQ-5D
		  ○ SF-6D
		  ○ SF-36
		  ○ SF-12
		  ○ Other utility measures (AQOL, HUI, 15D, QWB)
		• Activities of daily living (continuous outcomes will be prioritised)
		  ○ Barthel Index
		  ○ National Institutes of Health Stroke Scale
		  ○ Orpington Prognostic Scale
		  ○ Canadian Occupational Performance Measure
		  ○ Extended activities of daily living
		• Psychological distress (continuous outcomes will be prioritised)
		  ○ Depression
		  – PHQ-9
		  – Hospital Anxiety and Depression scale - depression subscale
		  – Beck Depression Inventory
		  – Hamilton Depression Scale
		  – Centre of Epidemiologic Studies Depression
		  – GHQ-28
		  – Geriatric Depression Scale
		  ○ Anxiety
		  – GAD-7
		  – Hospital Anxiety and Depression scale - anxiety subscale
		  – The Geriatric Anxiety Inventory
		  – GHQ-28
		  – Beck Anxiety Inventory
		  ○ Distress
		  – The Distress Management System for Stroke (DMSS)
		• Stroke-specific Patient-Reported Outcome Measures (continuous outcomes will be prioritised)
		  ○ Stroke-Specific Quality of Life (SS-QOL)
		  ○ Stroke Impact Scale (SIS)
		  ○ Stroke-specific Sickness Impact Profile (SA-SIP30)
		  ○ Neuro-QOL
		  ○ PROMIS-10
		  ○ Satisfaction with International Classification of Functioning, Disability and Health – Stroke (SATIS-Stroke)

		• Participation in leisure activities/social groups scores (continuous outcomes will be prioritised)
		  ○ Mayo-Portland Adaptability Inventory 4 (MPAI-4) part C (participation)
		  ○ Frenchay Activities Index
		• Withdrawal due to adverse events (dichotomous outcome)
		
		If not mentioned above, other validated scores will be considered and discussed with the committee to deliberate on their inclusion.
		
		Validity and reliability outcomes:
		Validity:
		• Face/content/construct validity
		• Criterion/Concurrent validity
		• Discriminant/convergent validity
		Reliability:
		• Test-retest reliability
		• Internal consistency (including Cronbach's alpha and intraclass correlation coefficient):
		  ○ Intertest reliability
		  ○ Intratest reliability
		• Inter-rater reliability
		Responsiveness to change
		Dimensions of fatigue considered (for example: physical, emotional, cognitive)

14.	Data extraction (selection and coding)	All references identified by the searches and from other sources will be uploaded into EPPI reviewer and de-duplicated. 10% of the abstracts will be reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third independent reviewer. The full text of potentially eligible studies will be retrieved and will be assessed in line with the criteria outlined above. A standardised form will be used to extract data from studies (see Developing NICE guidelines: the manual section 6.4). 10% of all evidence reviews are quality assured by a senior research fellow. This includes checking: • papers were included /excluded appropriately • a sample of the data extractions • correct methods are used to synthesise data

		• a sample of the risk of bias assessments Disagreements between the review authors over the risk of bias in particular studies will be resolved by discussion, with involvement of a third review author where necessary. Study investigators may be contacted for missing data where time and resources allow.
15.	Risk of bias (quality) assessment	Risk of bias will be assessed using the appropriate checklist as described in Developing NICE guidelines: the manual. • Systematic reviews: Risk of Bias in Systematic Reviews (ROBIS) • Randomised Controlled Trial: Cochrane RoB (2.0) • Non randomised study, including cohort studies: Cochrane ROBINS-I • Case control study: CASP case control checklist Validity and reliability studies risk of bias will be assessed considering: • Population • Control population • Recruitment/selection bias • Sample size • Whether the metrics used are appropriate for the objective of the study
16.	Strategy for data synthesis	For assess-and-treat evidence: • Pairwise meta-analyses will be performed using Cochrane Review Manager (RevMan5). Fixed-effects (Mantel-Haenszel) techniques will be used to calculate risk ratios for the binary outcomes where possible. Continuous outcomes will be analysed using an inverse variance method for pooling weighted mean differences. Heterogeneity between the studies in effect measures will be assessed using the I^2 statistic and visually inspected. An I^2 value greater than 50% will be considered indicative of substantial heterogeneity. Sensitivity analyses will be conducted based on pre-specified subgroups using stratified meta-analysis to explore the heterogeneity in effect estimates. If this does not explain the heterogeneity, the results will be presented pooled using random-effects. • GRADEpro will be used to assess the quality of evidence for each outcome, taking into account individual study quality and the meta-analysis results. The 4 main quality elements (risk of bias, indirectness, inconsistency and imprecision) will be appraised for each outcome. Publication bias is

		tested for when there are more than 5 studies for an outcome. The risk of bias across all available evidence was evaluated for each outcome using an adaptation of the 'Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox' developed by the international GRADE working group http://www.gradeworkinggroup.org/ • Where meta-analysis is not possible, data will be presented and quality assessed individually per outcome. • WinBUGS will be used for network meta-analysis, if possible given the data identified. For validity and reliability evidence: Findings will be presented in a table to summarise the key findings and limitations.
17.	Analysis of sub-groups	Subgroups that will be investigated if heterogeneity is present: Type of stroke (haemorrhagic compared to ischaemic) • Haemorrhagic (subarachnoid haemorrhage) • Haemorrhagic (non-subarachnoid haemorrhage) • Ischaemic Type of stroke (location, using the Bamford scale): • Total anterior circulation stroke (TACS) • Partial anterior circulation stroke (PACS) • Lacunar stroke (LACS) • Posterior circulation stroke (POCS) Initial stroke treatment • Thrombolysis • Thrombolectomy Physical activity prior to stroke • Low • Moderate • High Gender • Male • Female • Non-binary Severity (as stated by category or as measured by NIHSS scale): • Mild (or NIHSS 1-5)

<table>
<tr><td></td><td></td><td colspan="3">• Moderate (or NIHSS 5-14)
• Severe (or NIHSS 15-24)
• Very severe (or NIHSS >25)</td></tr>
<tr><td rowspan="8">18.</td><td rowspan="8">Type and method of review</td><td>☐</td><td colspan="2">Intervention</td></tr>
<tr><td>☐</td><td colspan="2">Diagnostic</td></tr>
<tr><td>☐</td><td colspan="2">Prognostic</td></tr>
<tr><td>☐</td><td colspan="2">Qualitative</td></tr>
<tr><td>☐</td><td colspan="2">Epidemiologic</td></tr>
<tr><td>☐</td><td colspan="2">Service Delivery</td></tr>
<tr><td>☒</td><td colspan="2">Other (please specify)
Assess-and-treat and Validity and Reliability review</td></tr>
<tr><td></td><td colspan="3"></td></tr>
<tr><td>19.</td><td>Language</td><td colspan="3">English</td></tr>
<tr><td>20.</td><td>Country</td><td colspan="3">England</td></tr>
<tr><td>21.</td><td>Anticipated or actual start date</td><td colspan="3">24/02/2021</td></tr>
<tr><td>22.</td><td>Anticipated completion date</td><td colspan="3">14/12/2022</td></tr>
<tr><td rowspan="8">23.</td><td rowspan="8">Stage of review at time of this submission</td><td>Review stage</td><td>Started</td><td>Completed</td></tr>
<tr><td>Preliminary searches</td><td>☐</td><td>☐</td></tr>
<tr><td>Piloting of the study selection process</td><td>☐</td><td>☐</td></tr>
<tr><td>Formal screening of search results against eligibility criteria</td><td>☐</td><td>☐</td></tr>
<tr><td>Data extraction</td><td>☐</td><td>☐</td></tr>
<tr><td>Risk of bias (quality) assessment</td><td>☐</td><td>☐</td></tr>
<tr><td>Data analysis</td><td>☐</td><td>☐</td></tr>
<tr><td></td><td colspan="3"></td></tr>
<tr><td>24.</td><td>Named contact</td><td colspan="3">5a. Named contact
National Guideline Centre

5b Named contact e-mail
StrokeRehabUpdate@nice.nhs.uk

5e Organisational affiliation of the review
National Institute for Health and Care Excellence (NICE) and National Guideline Centre</td></tr>
<tr><td>25.</td><td>Review team members</td><td colspan="3">From the National Guideline Centre:</td></tr>
</table>

		Bernard Higgins (Guideline lead) George Wood (Senior systematic reviewer) Madelaine Zucker (Systematic reviewer) Kate Lovibond (Health economics lead) Claire Sloan (Health economist) Joseph Runicles (Information specialist) Nancy Pursey (Senior project manager)
26.	Funding sources/sponsor	This systematic review is being completed by the National Guideline Centre which receives funding from NICE.
27.	Conflicts of interest	All guideline committee members and anyone who has direct input into NICE guidelines (including the evidence review team and expert witnesses) must declare any potential conflicts of interest in line with NICE's code of practice for declaring and dealing with conflicts of interest. Any relevant interests, or changes to interests, will also be declared publicly at the start of each guideline committee meeting. Before each meeting, any potential conflicts of interest will be considered by the guideline committee Chair and a senior member of the development team. Any decisions to exclude a person from all or part of a meeting will be documented. Any changes to a member's declaration of interests will be recorded in the minutes of the meeting. Declarations of interests will be published with the final guideline.
28.	Collaborators	Development of this systematic review will be overseen by an advisory committee who will use the review to inform the development of evidence-based recommendations in line with section 3 of Developing NICE guidelines: the manual. Members of the guideline committee are available on the NICE website: https://www.nice.org.uk/guidance/indevelopment/gid-ng10175
29.	Other registration details	N/A
30.	Reference/URL for published protocol	N/A
31.	Dissemination plans	NICE may use a range of different methods to raise awareness of the guideline. These include standard approaches such as: • notifying registered stakeholders of publication • publicising the guideline through NICE's newsletter and alerts • issuing a press release or briefing as appropriate, posting news articles on the NICE website, using social media channels, and publicising the guideline within NICE.
32.	Keywords	Adults; Assessment tools; Diagnostic; Fatigue; Intervention; Rehabilitation; Stroke

33.	Details of existing review of same topic by same authors	N/A	
34.	Current review status	☐	Ongoing
		☐	Completed but not published
		☒	Completed and published
		☐	Completed, published and being updated
		☐	Discontinued
35..	Additional information	N/A	
36.	Details of final publication	www.nice.org.uk	

Review protocol for health economic literature review

Review question	All questions – health economic evidence
Objectives	To identify health economic studies relevant to any of the review questions.
Search criteria	• Populations, interventions and comparators must be as specified in the clinical review protocol above. • Studies must be of a relevant health economic study design (cost–utility analysis, cost-effectiveness analysis, cost–benefit analysis, cost–consequences analysis, comparative cost analysis). • Studies must not be a letter, editorial or commentary, or a review of health economic evaluations. (Recent reviews will be ordered although not reviewed. The bibliographies will be checked for relevant studies, which will then be ordered.) • Unpublished reports will not be considered unless submitted as part of a call for evidence. • Studies must be in English.
Search strategy	A health economic study search will be undertaken using population-specific terms and a health economic study filter – see appendix B below. Databases searched: • Centre for Reviews and Dissemination NHS Economic Evaluations Database (NHS EED) – all years (closed to new records April 2015) • Centre for Reviews and Dissemination Health Technology Assessment database – all years (closed to new records March 2018) • International HTA database (INAHTA) – all years • Medline and Embase – from 2014 (due to NHS EED closure)
Review strategy	Studies not meeting any of the search criteria above will be excluded. Studies published before 2006 (including those included in the previous guideline), abstract-only studies and studies from non-OECD countries or the USA will also be excluded. Each remaining study will be assessed for applicability and methodological limitations using the NICE economic evaluation checklist which can be found in appendix H of Developing NICE guidelines: the manual (2014).[22] Studies published in 2006 or later that were included in the previous guideline will be reassessed for inclusion and may be included or selectively excluded based on their relevance to the questions covered in this update and whether more applicable evidence is also identified. **Inclusion and exclusion criteria** • If a study is rated as both 'Directly applicable' and with 'Minor limitations' then it will be included in the guideline. A health economic evidence table will be completed and it will be included in the health economic evidence profile. • If a study is rated as either 'Not applicable' or with 'Very serious limitations' then it will usually be excluded from the guideline. If it is excluded then a health economic evidence table will not be completed and it will not be included in the health economic evidence profile. • If a study is rated as 'Partially applicable', with 'Potentially serious limitations' or both then there is discretion over whether it should be included. **Where there is discretion** The health economist will make a decision based on the relative applicability and quality of the available evidence for that question, in discussion with the guideline committee if required. The ultimate aim is to include health economic studies that are helpful for decision-making in the context of the guideline and the current NHS setting. If several studies are considered of sufficiently high applicability and methodological quality that they could all be included, then the health economist, in

	discussion with the committee if required, may decide to include only the most applicable studies and to selectively exclude the remaining studies. All studies excluded on the basis of applicability or methodological limitations will be listed with explanation in the excluded health economic studies appendix below. The health economist will be guided by the following hierarchies. *Setting:* • UK NHS (most applicable). • OECD countries with predominantly public health insurance systems (for example, France, Germany, Sweden). • OECD countries with predominantly private health insurance systems (for example, Switzerland). • Studies set in non-OECD countries or in the USA will be excluded before being assessed for applicability and methodological limitations. *Health economic study type:* • Cost–utility analysis (most applicable). • Other type of full economic evaluation (cost–benefit analysis, cost-effectiveness analysis, cost–consequences analysis). • Comparative cost analysis. • Non-comparative cost analyses including cost-of-illness studies will be excluded before being assessed for applicability and methodological limitations. *Year of analysis:* • The more recent the study, the more applicable it will be. • Studies published in 2006 or later (including any such studies included in the previous guideline) but that depend on unit costs and resource data entirely or predominantly from before 2006 will be rated as 'Not applicable'. • Studies published before 2006 (including any such studies included in the previous guideline) will be excluded before being assessed for applicability and methodological limitations. *Quality and relevance of effectiveness data used in the health economic analysis:* • The more closely the clinical effectiveness data used in the health economic analysis match with the outcomes of the studies included in the clinical review the more useful the analysis will be for decision-making in the guideline.

Appendix B – Literature search strategies

B.1 Clinical search literature search strategy

Searches were constructed using a PICO framework where population (P) terms were combined with Intervention (I) and in some cases Comparison (C) terms. Outcomes (O) are rarely used in search strategies as these concepts may not be indexed or described in the title or abstract and are therefore difficult to retrieve. Search filters were applied to the search where appropriate.

Table 6: Database parameters, filters and limits applied

Database	Dates searched	Search filter used
Medline (OVID)	1946 – 08 January 2023	Exclusions (animal studies, letters, comments, editorials, case studies/reports) English language
Embase (OVID)	1974 – 08 January 2023	Exclusions (animal studies, letters, comments, editorials, case studies/reports, conference abstracts) English language
The Cochrane Library (Wiley)	Cochrane Reviews to 2023 Issue 1 of 12 CENTRAL to 2023 Issue 1 of 12	Exclusions (clinical trials, conference abstracts)
PsycINFO (OVID)	Inception – 08 January 2023	Exclusions (animal studies, letters, case reports) Human English language
Epistemonikos (The Epistemonikos Foundation)	Inception – 08 January 2023	Exclusions (Cochrane reviews) English language

Medline (Ovid) search terms

1.	exp Stroke/
2.	Stroke Rehabilitation/
3.	exp Cerebral Hemorrhage/
4.	(stroke or strokes or cva or poststroke* or apoplexy or "cerebrovascular accident").ti,ab.
5.	((cerebro* or brain or brainstem or cerebral*) adj3 (infarct* or accident*)).ti,ab.
6.	"brain attack*".ti,ab.
7.	or/1-6
8.	letter/
9.	editorial/

10.	news/
11.	exp historical article/
12.	Anecdotes as Topic/
13.	comment/
14.	case report/
15.	(letter or comment*).ti.
16.	or/8-15
17.	randomized controlled trial/ or random*.ti,ab.
18.	16 not 17
19.	animals/ not humans/
20.	exp Animals, Laboratory/
21.	exp Animal Experimentation/
22.	exp Models, Animal/
23.	exp Rodentia/
24.	(rat or rats or mouse or mice or rodent*).ti.
25.	or/18-24
26.	7 not 25
27.	limit 26 to English language
28.	(fatigue* adj4 (assess* or tool* or index or indices or scale* or test* or retest* or psychometr* or inventor*)).ti,ab.
29.	"profile of mood states".ti,ab.
30.	(VAS-F or NFI-Stroke or C-NFI-Stroke or FSS or FSS-A or SSQOL-A or SSQOL-A-E or BDI-II).ti,ab.
31.	(fatigue* adj3 (mental* or physical*)).ti,ab.
32.	(SF-36v or SF-36v2).ti,ab.
33.	(fatigue* and (Beck depression inventory or stroke specific quality of life)).ti,ab.
34.	or/28-33
35.	27 and 34

Embase (Ovid) search terms

1.	exp Cerebrovascular accident/
2.	exp Brain infarction/
3.	Stroke Rehabilitation/
4.	(stroke or strokes or cva or poststroke* or apoplexy or "cerebrovascular accident").ti,ab.
5.	((cerebro* or brain or brainstem or cerebral*) adj3 (infarct* or accident*)).ti,ab.
6.	"brain attack*".ti,ab.
7.	Intracerebral hemorrhage/
8.	or/1-7
9.	letter.pt. or letter/
10.	note.pt.
11.	editorial.pt.
12.	case report/ or case study/
13.	(letter or comment*).ti.
14.	(conference abstract or conference paper).pt.

15.	or/9-14
16.	randomized controlled trial/ or random*.ti,ab.
17.	15 not 16
18.	animal/ not human/
19.	nonhuman/
20.	exp Animal Experiment/
21.	exp Experimental Animal/
22.	animal model/
23.	exp Rodent/
24.	(rat or rats or mouse or mice or rodent*).ti.
25.	or/17-24
26.	8 not 25
27.	limit 26 to English language
28.	(fatigue* adj4 (assess* or tool* or index or indices or scale* or test* or retest* or psychometr* or inventor*)).ti,ab.
29.	"profile of mood states".ti,ab.
30.	(VAS-F or NFI-Stroke or C-NFI-Stroke or FSS or FSS-A or SSQOL-A or SSQOL-A-E or BDI-II).ti,ab.
31.	(fatigue* adj3 (mental* or physical*)).ti,ab.
32.	(SF-36v or SF-36v2).ti,ab.
33.	(fatigue* and (Beck depression inventory or stroke specific quality of life)).ti,ab.
34.	or/28-33
35.	27 and 34

Cochrane Library (Wiley) search terms

#1.	MeSH descriptor: [Stroke] explode all trees
#2.	MeSH descriptor: [Stroke Rehabilitation] explode all trees
#3.	MeSH descriptor: [Cerebral Hemorrhage] explode all trees
#4.	(stroke or strokes or cva or poststroke* or apoplexy or "cerebrovascular accident"):ti,ab
#5.	((cerebro* or brain or brainstem or cerebral*) near/3 (infarct* or accident*)):ti,ab
#6.	brain attack*:ti,ab
#7.	(or #1-#6)
#8.	conference:pt or (clinicaltrials or trialsearch):so
#9.	#7 not #8
#10.	(fatigue* near/4 (assess* or tool* or index or indices or scale* or test* or retest* or psychometr* or inventor*)):ti,ab
#11.	profile of mood states:ti,ab
#12.	(VAS-F or NFI-Stroke or C-NFI-Stroke or FSS or FSS-A or SSQOL-A or SSQOL-A-E or BDI-II):ti,ab
#13.	(fatigue* near/3 (mental* or physical*)):ti,ab
#14.	(SF-36v or SF-36v2):ti,ab
#15.	(fatigue* and (Beck depression inventory or stroke specific quality of life)):ti,ab
#16.	(or #10-#15)
#17.	#9 and #16

PsycINFO (OVID) search terms

1.	exp Stroke/
2.	exp Cerebral hemorrhage/
3.	(stroke or strokes or cva or poststroke* or apoplexy or "cerebrovascular accident").ti,ab.
4.	((cerebro* or brain or brainstem or cerebral*) adj3 (infarct* or accident*)).ti,ab.
5.	"brain attack*".ti,ab.
6.	Cerebrovascular accidents/
7.	exp Brain damage/
8.	(brain adj2 injur*).ti.
9.	or/1-8
10.	Letter/
11.	Case report/
12.	exp rodents/
13.	or/10-12
14.	9 not 13
15.	limit 14 to (human and English language)
16.	(fatigue* adj4 (assess* or tool* or index or indices or scale* or test* or retest* or psychometr* or inventor*)).ti,ab.
17.	"profile of mood states".ti,ab.
18.	(VAS-F or NFI-Stroke or C-NFI-Stroke or FSS or FSS-A or SSQOL-A or SSQOL-A-E or BDI-II).ti,ab.
19.	(fatigue* adj3 (mental* or physical*)).ti,ab.
20.	(SF-36v or SF-36v2).ti,ab.
21.	(fatigue* and (Beck depression inventory or stroke specific quality of life)).ti,ab.
22.	or/16-21
23.	15 and 22

Epistemonikos search terms

1.	(title:((title:((stroke OR strokes OR cva OR poststroke* OR apoplexy OR "cerebrovascular accident")) OR abstract:((stroke OR strokes OR cva OR poststroke* OR apoplexy OR "cerebrovascular accident")))) OR abstract:((title:((stroke OR strokes OR cva OR poststroke* OR apoplexy OR "cerebrovascular accident")) OR abstract:((stroke OR strokes OR cva OR poststroke* OR apoplexy OR "cerebrovascular accident"))))) AND (title:((title:(Fatigue) OR abstract:(Fatigue)) AND (title:(assess* OR tool* OR index OR indices OR scale* OR test* OR retest* OR psychometr* OR inventor* OR mental OR physical) OR abstract:(assess* OR tool* OR index OR indices OR scale* OR test* OR retest* OR psychometr* OR inventor* OR mental OR physical))) OR abstract:((title:(Fatigue) OR abstract:(Fatigue)) AND (title:(assess* OR tool* OR index OR indices OR scale* OR test* OR retest* OR psychometr* OR inventor* OR mental OR physical) OR abstract:(assess* OR tool* OR index OR indices OR scale* OR test* OR retest* OR psychometr* OR inventor* OR mental OR physical))))

B.2 Health Economics literature search strategy

Health economic evidence was identified by conducting searches using terms for a broad Stroke Rehabilitation population. The following databases were searched: NHS Economic Evaluation Database (NHS EED - this ceased to be updated after 31st March 2015), Health Technology Assessment database (HTA - this ceased to be updated from 31st March 2018) and The International Network of Agencies for Health Technology Assessment (INAHTA). Searches for recent evidence were run on Medline and Embase from 2014 onwards for

health economics, and all years for quality-of-life studies. Additional searches were run in CINAHL and PsycInfo looking for health economic evidence.

Table 2: Database parameters, filters and limits applied

Database	Dates searched	Search filters and limits applied
Medline (OVID)	Health Economics 1 January 2014 – 08 January 2023	Health economics studies Quality of life studies Exclusions (animal studies, letters, comments, editorials, case studies/reports,) English language
	Quality of Life 1946 – 08 January 2023	
Embase (OVID)	Health Economics 1 January 2014 – 08 January 2023	Health economics studies Quality of life studies Exclusions (animal studies, letters, comments, editorials, case studies/reports, conference abstracts) English language
	Quality of Life 1974 – 08 January 2023	
NHS Economic Evaluation Database (NHS EED) (Centre for Research and Dissemination - CRD)	Inception –31st March 2015	
Health Technology Assessment Database (HTA) (Centre for Research and Dissemination – CRD)	Inception – 31st March 2018	
The International Network of Agencies for Health Technology Assessment (INAHTA)	Inception - 08 January 2023	English language
PsycINFO (OVID)	1 January 2014 – 08 January 2023	Health economics studies Exclusions (animal studies, letters, case reports) Human English language
Current Nursing and Allied Health Literature - CINAHL (EBSCO)	1 January 2014 – 08 January 2023	Health economics studies Exclusions (Medline records, animal studies, letters, editorials, comments, theses) Human

Database	Dates searched	Search filters and limits applied
		English language

Medline (Ovid) search terms

1.	exp Stroke/
2.	exp Cerebral Hemorrhage/
3.	(stroke or strokes or cva or poststroke* or apoplexy or "cerebrovascular accident").ti,ab.
4.	((cerebro* or brain or brainstem or cerebral*) adj3 (infarct* or accident*)).ti,ab.
5.	"brain attack*".ti,ab.
6.	or/1-5
7.	letter/
8.	editorial/
9.	news/
10.	exp historical article/
11.	Anecdotes as Topic/
12.	comment/
13.	case report/
14.	(letter or comment*).ti.
15.	or/7-14
16.	randomized controlled trial/ or random*.ti,ab.
17.	15 not 16
18.	animals/ not humans/
19.	exp Animals, Laboratory/
20.	exp Animal Experimentation/
21.	exp Models, Animal/
22.	exp Rodentia/
23.	(rat or rats or mouse or mice or rodent*).ti.
24.	or/17-23
25.	6 not 24
26.	Economics/
27.	Value of life/
28.	exp "Costs and Cost Analysis"/
29.	exp Economics, Hospital/
30.	exp Economics, Medical/
31.	Economics, Nursing/
32.	Economics, Pharmaceutical/
33.	exp "Fees and Charges"/
34.	exp Budgets/

35.	budget*.ti,ab.
36.	cost*.ti.
37.	(economic* or pharmaco?economic*).ti.
38.	(price* or pricing*).ti,ab.
39.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
40.	(financ* or fee or fees).ti,ab.
41.	(value adj2 (money or monetary)).ti,ab.
42.	or/26-41
43.	quality-adjusted life years/
44.	sickness impact profile/
45.	(quality adj2 (wellbeing or well being)).ti,ab.
46.	sickness impact profile.ti,ab.
47.	disability adjusted life.ti,ab.
48.	(qal* or qtime* or qwb* or daly*).ti,ab.
49.	(euroqol* or eq5d* or eq 5*).ti,ab.
50.	(qol* or hql* or hqol* or h qol* or hrqol* or hr qol*).ti,ab.
51.	(health utility* or utility score* or disutilit* or utility value*).ti,ab.
52.	(hui or hui1 or hui2 or hui3).ti,ab.
53.	(health* year* equivalent* or hye or hyes).ti,ab.
54.	discrete choice*.ti,ab.
55.	rosser.ti,ab.
56.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.
57.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
58.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
59.	(sf12* or sf 12* or short form 12* or shortform 12* or shortform12*).ti,ab.
60.	(sf8* or sf 8* or short form 8* or shortform 8* or shortform8*).ti,ab.
61.	(sf6* or sf 6* or short form 6* or shortform 6* or shortform6*).ti,ab.
62.	or/43-61
63.	25 and 42
64.	25 and 62
65.	limit 63 to English language
66.	limit 64 to English language

Embase (Ovid) search terms

1.	exp Cerebrovascular accident/
2.	exp Brain infarction/
3.	(stroke or strokes or cva or poststroke* or apoplexy or "cerebrovascular accident").ti,ab.
4.	((cerebro* or brain or brainstem or cerebral*) adj3 (infarct* or accident*)).ti,ab.
5.	"brain attack*".ti,ab.

6.	Intracerebral hemorrhage/
7.	or/1-6
8.	letter.pt. or letter/
9.	note.pt.
10.	editorial.pt.
11.	case report/ or case study/
12.	(letter or comment*).ti.
13.	or/8-12
14.	randomized controlled trial/ or random*.ti,ab.
15.	13 not 14
16.	animal/ not human/
17.	nonhuman/
18.	exp Animal Experiment/
19.	exp Experimental Animal/
20.	animal model/
21.	exp Rodent/
22.	(rat or rats or mouse or mice).ti.
23.	or/15-22
24.	7 not 23
25.	health economics/
26.	exp economic evaluation/
27.	exp health care cost/
28.	exp fee/
29.	budget/
30.	funding/
31.	budget*.ti,ab.
32.	cost*.ti.
33.	(economic* or pharmaco?economic*).ti.
34.	(price* or pricing*).ti,ab.
35.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
36.	(financ* or fee or fees).ti,ab.
37.	(value adj2 (money or monetary)).ti,ab.
38.	or/25-37
39.	quality adjusted life year/
40.	"quality of life index"/
41.	short form 12/ or short form 20/ or short form 36/ or short form 8/
42.	sickness impact profile/
43.	(quality adj2 (wellbeing or well being)).ti,ab.
44.	sickness impact profile.ti,ab.
45.	disability adjusted life.ti,ab.
46.	(qal* or qtime* or qwb* or daly*).ti,ab.

47.	(euroqol* or eq5d* or eq 5*).ti,ab.
48.	(qol* or hql* or hqol* or h qol* or hrqol* or hr qol*).ti,ab.
49.	(health utility* or utility score* or disutilit* or utility value*).ti,ab.
50.	(hui or hui1 or hui2 or hui3).ti,ab.
51.	(health* year* equivalent* or hye or hyes).ti,ab.
52.	discrete choice*.ti,ab.
53.	rosser.ti,ab.
54.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.
55.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
56.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
57.	(sf12* or sf 12* or short form 12* or shortform 12* or shortform12*).ti,ab.
58.	(sf8* or sf 8* or short form 8* or shortform 8* or shortform8*).ti,ab.
59.	(sf6* or sf 6* or short form 6* or shortform 6* or shortform6*).ti,ab.
60.	or/39-59
61.	limit 24 to English language
62.	38 and 61
63.	60 and 61

NHS EED and HTA (CRD) search terms

#1.	MeSH DESCRIPTOR Stroke EXPLODE ALL TREES
#2.	MeSH DESCRIPTOR Cerebral Hemorrhage EXPLODE ALL TREES
#3.	(stroke* or cva or poststroke* or apoplexy or "cerebrovascular accident")
#4.	(((cerebro* or brain or brainstem or cerebral*) adj3 (infarct* or accident*)))
#5.	("brain attack*")
#6.	#1 OR #2 OR #3 OR #4 OR #5

INAHTA search terms

1.	(brain attack*) OR (((cerebro* or brain or brainstem or cerebral*) and (infarct* or accident*))) OR ((stroke or strokes or cva or poststroke* or apoplexy or "cerebrovascular accident")) OR ("Cerebral Hemorrhage"[mhe]) OR ("Stroke"[mhe])

CINAHL search terms

1.	MH "Economics+"
2.	MH "Financial Management+"
3.	MH "Financial Support+"
4.	MH "Financing, Organized+"
5.	MH "Business+"
6.	S2 OR S3 or S4 OR S5
7.	S1 not S6
8.	MH "Health Resource Allocation"
9.	MH "Health Resource Utilization"
10.	S8 OR S9
11.	S7 OR S10
12.	(cost or costs or economic* or pharmacoeconomic* or price* or pricing*) OR AB (cost or costs or economic* or pharmacoeconomic* or price* or pricing*)

13.	S11 OR S12
14.	PT editorial
15.	PT letter
16.	PT commentary
17.	S14 or S15 or S16
18.	S13 NOT S17
19.	MH "Animal Studies"
20.	(ZT "doctoral dissertation") or (ZT "masters thesis")
21.	S18 NOT (S19 OR S20)
22.	PY 2014-
23.	S21 AND S22
24.	MW Stroke or MH Cerebral Hemorrhage
25.	stroke* or cva or poststroke* or apoplexy or "cerebrovascular accident"
26.	(cerebro* OR brain OR brainstem OR cerebral*) AND (infarct* OR accident*)
27.	"brain attack*"
28.	S24 OR S25 OR S26 OR S27
29.	S23 AND S28

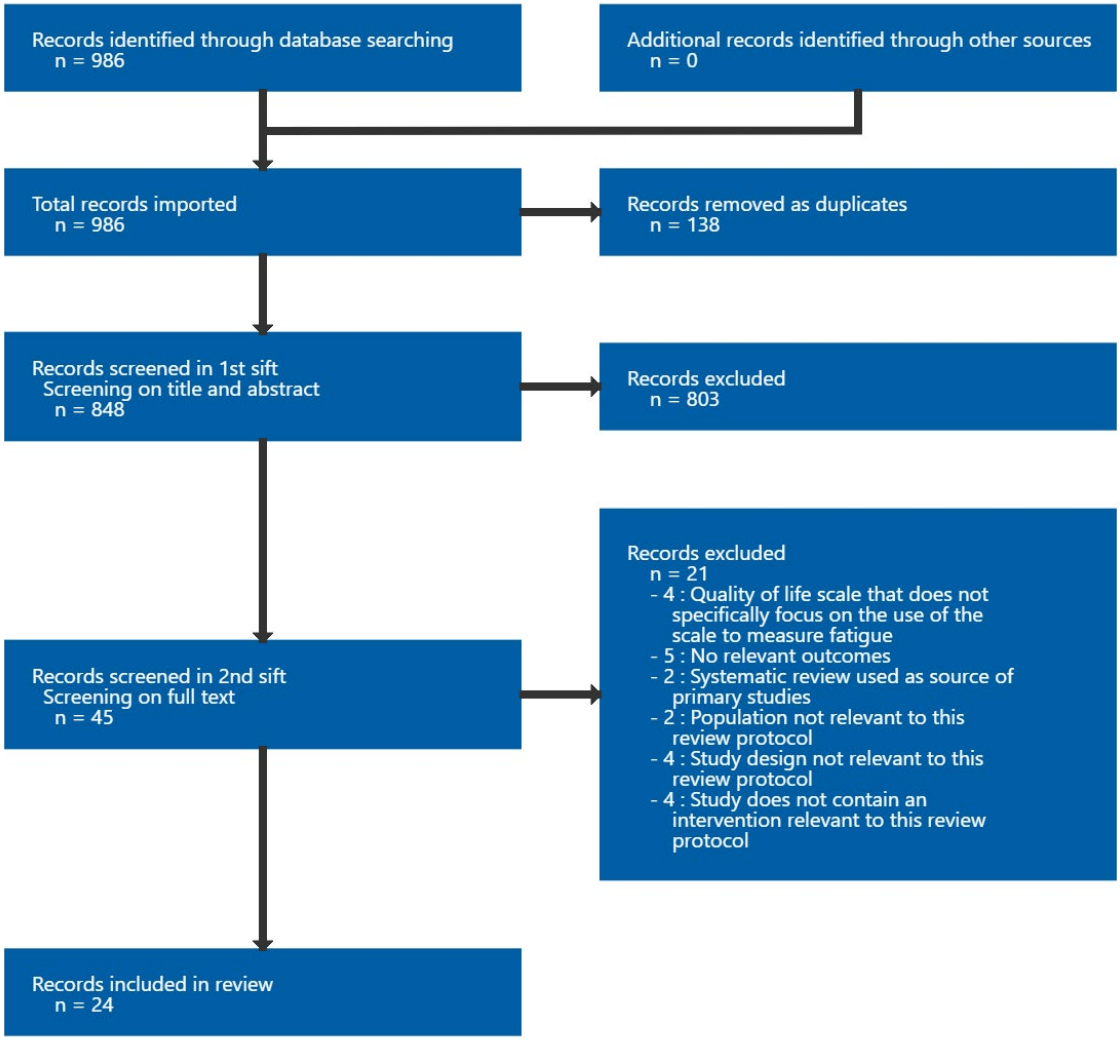
PsycINFO search terms

1.	exp Stroke/
2.	exp Cerebral hemorrhage/
3.	(stroke or strokes or cva or poststroke* or apoplexy or "cerebrovascular accident").ti,ab.
4.	((cerebro* or brain or brainstem or cerebral*) adj3 (infarct* or accident*)).ti,ab.
5.	"brain attack*".ti,ab.
6.	Cerebrovascular accidents/
7.	exp Brain damage/
8.	(brain adj2 injur*).ti.
9.	or/1-8
10.	Letter/
11.	Case report/
12.	exp Rodents/
13.	or/10-12
14.	9 not 13
15.	limit 14 to (human and english language)
16.	First posting.ps.
17.	15 and 16
18.	15 or 17
19	"costs and cost analysis"/
20.	"Cost Containment"/
21.	(economic adj2 evaluation$).ti,ab.
22.	(economic adj2 analy$).ti,ab.

23.	(economic adj2 (study or studies)).ti,ab.
24.	(cost adj2 evaluation$).ti,ab.
25.	(cost adj2 analy$).ti,ab.
26.	(cost adj2 (study or studies)).ti,ab.
27.	(cost adj2 effective$).ti,ab.
28.	(cost adj2 benefit$).ti,ab.
29.	(cost adj2 utili$).ti,ab.
30.	(cost adj2 minimi$).ti,ab.
31.	(cost adj2 consequence$).ti,ab.
32.	(cost adj2 comparison$).ti,ab.
33.	(cost adj2 identificat$).ti,ab.
34.	(pharmacoeconomic$ or pharmaco-economic$).ti,ab.
35.	or/19-34
36.	(0003-4819 or 0003-9926 or 0959-8146 or 0098-7484 or 0140-6736 or 0028-4793 or 1469-493X).is.
37.	35 not 36
38.	18 and 37

Appendix C – Effectiveness evidence study selection

Figure 1: Flow chart of clinical study selection for the review of the optimal tool to assess fatigue in people after stroke



Appendix D – Effectiveness (assess-to-treat) and validity and reliability evidence

D.1 Effectiveness (assess-to-treat) evidence

No assess-to-treat evidence was identified.

D.2 Validity and reliability evidence

Reference	Abdulla 2019[1]
Study type Setting/Location	Cross-sectional study Saudi Arabia, three general hospitals in the area
Number of participants and characteristics	N=217 (186 people with stroke and 90 healthy participants assessed for eligibility. Of the stroke survivors, 8 not first stroke, 10 refused to participate, 1 cannot read Arabic, 2 have other neurological problems, 5 have cognitive dysfunction/dementia. Of the healthy participants, 8 had chronic diseases, 7 refused to participate. This left 160 stroke survivors and 75 healthy participants.12 did not return the questionnaires, 6 less than 50% complete. This left 147 stroke survivors and 70 healthy participants). **Inclusion criteria:** First-ever stroke diagnosis by a neurologist; at sub-acute stage (three to six months) or chronic stage (>six months); presence of no other comorbid neurological problems which may affect fatigue. Also healthy participants were recruited – these were people with no chronic illness or regular medications for the last three months, who were randomly recruited from visitors to the three hospitals. **Exclusion criteria:** Stroke: Having a stroke beyond first stroke, other neurological problems. Healthy participants: Having chronic illness or regular medications for the last three months All: Cannot read Arabic. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise**

Reference	Abdulla 2019[1]

	Stroke survivors
	• Mean age (SD): 59.63 (10.97) years
	• Males:Females = 69:78 (46.9%:53.1%)
	• Haemorrhagic:Ischaemic = 23:122 (15.7%:83.0%)
	• Left/Right sided stroke: 82:65 (55.8%:44.2%)
	• Living status (family:alone) = 120:27 (81.6%:18.5%)
	• Mean time since stroke (SD) = 12.0 (15.6) months
	  ○ Sub-acute = 68 (46.3%)
	  ○ Chronic = 79 (53.7%)
	• NIHSS groups
	  ○ 0 = 3 (2.0%)
	  ○ 1-4 = 79 (53.7%)
	  ○ 5-15 = 63 (42.9%)
	  ○ 16-20 = 2 (1.4%)
	  ○ 21-24 = 0 (0%)
	
	Healthy participants
	• Mean age (SD): 57.830 (11.52) years
	• Males:Females = 28:42 (40.0%:60.0%)

Intervention	**Fatigue severity scale – Arabic (FSS-A)** Fatigue Visual Analogue Scale (VAS-F) Stroke Specific quality of Life (SSQOL-A) including an energy domain (SSQOL-A-E) SF-36, including a vitality domain (SF-36v) Beck Depression Inventory-II (BDI-II) National Institutes of Health Stroke Scale (NIHSS)
Comparison	Itself, other scales, different time points
Length of follow-up	7 days
Outcome measures	Face/content/construct validity

<table>
<tr><th>Reference</th><th>Abdulla 2019[1]</th></tr>
<tr><td></td><td>Test-retest reliability
Internal consistency (Cronbach’s alpha)
Responsiveness to change</td></tr>
<tr><td>Source of funding</td><td>No additional information</td></tr>
<tr><td>Outcomes</td><td>Face/content/construct validity
91.6% of the patients and 93.3% of the healthy participants returned acceptable questionnaires. The Shapiro-Wilk test and visual inspection of frequency distribution histograms were relatively symmetrical with a skewness of -0.59. The scale item skewness fell between -0.19 to -0.76. 7.5% or fewer of participants selected the minimum score while 10.9% or fewer respondents selected the maximum score.

Construct validity
<table>
<tr><td>Scale/item</td><td>R_s (correlation coefficient) (95% CI)</td><td>P-value</td></tr>
<tr><td>VAS-F</td><td>0.693 (0.598 to 0.768)</td><td><0.001</td></tr>
<tr><td>SSQOL-A</td><td>-0.587 (-0.683 to -0.429)</td><td><0.001</td></tr>
<tr><td>SSQOL-A-E</td><td>-0.632 (-0.720 to -0.523)</td><td><0.001</td></tr>
<tr><td>SF-36</td><td>-0.531 (-0.638 to -0.404)</td><td><0.001</td></tr>
<tr><td>SF-36v</td><td>-0.558 (-0.660 to -0.435)</td><td><0.001</td></tr>
<tr><td>BDI-II</td><td>0.475 (0.339 to 0.591)</td><td><0.001</td></tr>
</table>
Exploratory factor analysis showed that the FSS-A has one factor with an eigenvalue >1. This explained 65.91% of the variance with factor loading between 0.713 (Item 1 and 0.870 (Item 8), indicating that the scale measures one construct (fatigue). It showed good fit indexes with KMO = 0.925, and Bartlett’s test of sphericity = 920.90, df = 36, p <0.001. The total score showed a strong positive correlation with VAS-F, strong negative correlations with SSQOL-A-E, moderate negative correlations with SSQOL-A, SF-36 and SF-36v and a moderate positive correlation to BDI-II.

Test-retest reliability
<table>
<tr><td>FSS-A Item</td><td>ICC (95% CI)</td><td>P-value</td></tr>
<tr><td>1</td><td>0.79 (0.61 to 0.89)</td><td>0.612</td></tr>
<tr><td>2</td><td>0.77 (0.57 to 0.88)</td><td>0.432</td></tr>
<tr><td>3</td><td>0.88 (0.78 to 0.94)</td><td>0.477</td></tr>
<tr><td>4</td><td>0.92 (0.84 to 0.96)</td><td>0.243</td></tr>
<tr><td>5</td><td>0.92 (0.85 to 0.96)</td><td>0.465</td></tr>
</table></td></tr>
</table>

Reference	Abdulla 2019[1]

6	0.88 (0.78 to 0.94)	0.230
7	0.81 (0.64 to 0.90)	0.952
8	0.87 (0.76 to 0.93)	0.240
9	0.89 (0.79 to 0.94)	0.149
Total score	0.92 (0.85 to 0.96)	0.644

Internal consistency (Cronbach's alpha)

FSS-A Item	Cronbach's alpha if item is deleted
1	0.93
2	0.93
3	0.93
4	0.93
5	0.93
6	0.92
7	0.93
8	0.92
9	0.92

Total score = 0.934

Item-total correlations ranging from 0.64 to 0.82.

Responsiveness to change

FSS-A Item	Minimal detectable change at 95% confidence interval
1	1.77
2	1.96
3	1.32
4	1.11
5	1.25
6	1.63
7	1.91

Reference	Abdulla 2019[1]
	8 | 1.49 9 | 1.58 Total score | 1.02
Risk of bias assessment	**Population** Stroke survivors. Has a mix of different genders, time after stroke, type of stroke. Does not provide information on the location of stroke, initial stroke treatment, physical activity prior to stroke and severity. **Control population** Healthy participants. No indirectness. **Recruitment/selection bias** People recruited from three general hospitals. No obvious problems with recruitment. **Sample size** Larger sample size = 217 (147 stroke survivors). However, only 42 people were randomly selected to complete the FSS-A a second time after one week, which reduces the sample size for the repeat testing. **Appropriateness of metrics** Metrics used were appropriate.
Limitations	Author provided limitations: • All questionnaires were self-administered. Therefore, misunderstanding of the questions could have occurred. • Almost all people had mild to moderate stroke, with mild to moderate depression. Therefore, the findings cannot be extended to severe stroke or severe depression. • People were not followed up over time to investigate changes in fatigue.
Summary/author's conclusion	**Author's conclusion** The FSS-A is a valid and reliable scale that can differentiate between different levels of fatigue. It is a useful tool to measure fatigue in patients with stroke in Arabic-speaking populations.

Reference	Batur 2022[2]
Study type Setting/Location	Cross-sectional study Turkey, people with stroke who had attended the Physical Medicine and Rehabilitation Department.
Number of participants and characteristics	N=108 (initial sample, of these 67 were excluded from the study due to not fulfilling the inclusion criteria: 34 had aphasia or cognitive dysfunction, 1 unable to speak Turkish, 7 recent medication change, 5 serious medical conditions such as cardiac events, hip fracture, 4 dementia, 3 schizophrenia, 3 premature discharge, 10 not consented to participate. 41 met the inclusion criteria. Of these 34 returned for the retest at 7 days). 41 age and gender matched people who were admitted to the Physical Medicine and Rehabilitation Department for musculoskeletal reasons were used as the control subjects. None were excluded, 25 returned for the retest. **Inclusion criteria:** People after stroke: Age above 18 years; medical stability for at least 4 weeks; ability to communicate. Control group: Healthy participants with no history of stroke or neurological disorders, heart failure, uncontrolled hypertension or malignities. **Exclusion criteria:** Disorders other than stroke such as tumour, thyroid, rheumatological or neurological disorders; if they had a medication change in the preceding two weeks that could cause fatigue; aphasia; severe cognitive deficits, which could prevent them from completing the questionnaires. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** **Stroke survivors** • Mean age (SD): 56.4 (12.9) years • Male/female: 19:22 • Education >8 years/<8 years: 14:27 • Employed/unemployed: 13:28 • Married/single: 34:7 • Coronary artery disease present/absent: 13:28 • Hypertension present/absent: 30:11 • Diabetes mellitus present/absent: 12:29 • Chronic obstructive pulmonary disease present/absent: 7:34 • Charlson comorbidity index (median [range]): 2 (2-6)

Reference	Batur 2022[2]
	• HADS depression (median [range]): 7 (0-15) • HADS anxiety (median [range]): 7 (0-14) • FIM Motor (median [range]): 62 (26-90) • FIM Cognitive (median [range]): 35 (20-35) • FIM total (median [range]): 97 (55-125) **Control group** • Mean age (SD): 55.2 (14.6) years • Male/female: 19:22 • Education >8 years/<8 years: 19:22 • Employed/unemployed: 19:22 • Married/single: 34:5 • Coronary artery disease present/absent: 1:40 • Hypertension present/absent: 13:28 • Diabetes mellitus present/absent: 5:36 • Chronic obstructive pulmonary disease present/absent: 3:38 • Charlson comorbidity index (median [range]): 0 (0-2) • HADS depression (median [range]): 5.5 (0-17) • HADS anxiety (median [range]): 5.5 (0-16) • FIM Motor (median [range]): 91 (77-91) • FIM Cognitive (median [range]): 35 (35-35) • FIM total (median [range]): 126 (112-126)
Intervention	**Fatigue impact scale (FIS) – Turkish version**
Comparison	Between people with stroke and control subjects. Between different scales, including: SF-36 vitality (SF-36 v) and the fatigue severity scale (FSS) for convergent validity. Against the Hospital Anxiety and Depression Scale (HADS) for divergent validity.
Length of follow-up	Re-test at 7 days.
Outcome measures	Face/content/construct validity Criterion/concurrent validity

Reference	Batur 2022[2]
	Discriminant/convergent validity Test-retest reliability Internal consistency Dimensions of fatigue considered
Source of funding	No funds were received in support of this work.
Outcomes	**Face/content/construct validity** Content validity: 82% of people with stroke found the questions understandable, and 65% reported the items represented the impact of their fatigue on their health. Internal construct validity: The Kaiser-Meyer Olkin Measures of sampling adequacy were 0.82, 0.73 and 0.63, which confirm the appropriateness of factor analysis for the cognitive, physical and psychosocial dimensions respectively. Each component explained the 58%, 45% and 38% variances of the cognitive, psychosocial and physical dimensions respectively. The factor analysis supported all items except items 18 and 34. **Criterion/concurrent validity** In the ROC analysis, the area under the curve was 0.730 (standard error = 0.055) (p <0.001) for the total DIS score. The sensitivity and specificity values were 76% and 68% respectively for an optimal cutoff value of 43.5. **Discriminant/convergent validity** A negative significant correlation was detected between the FIS scores and the SF-36 mental and vitality survey results. There was a positive but weak correlation between the total FIS score, the FIS cognitive and psychosocial scores and the FSS score. HAD depression was not correlated with the FIS scores, except for a moderate correlation with the FIS cognitive scores. The FIS psychosocial and FIS total scores were weakly correlated with the HAD anxiety score. **Test-retest reliability** Intraclass correlation coefficient for FIS total = 0.830 (0.731-0.895) (p <0.001) Intraclass correlation coefficient for FIS cognitive = 0.787 (0.667-0.867) (p <0.001) Intraclass correlation coefficient for FIS physical = 0.734 (0.591-0.832) (p <0.001) Intraclass correlation coefficient for FIS psychosocial = 0.800 (0.686-0.879) (p <0.001) **Internal consistency**

Reference	Batur 2022[2]
	Cronbach alpha for FIS total = 0.946 Cronbach alpha for FIS cognitive = 0.917 Cronbach alpha for FIS physical = 0.803 Cronbach alpha for FIS psychosocial = 0.929 **Dimensions of fatigue considered** Cognitive, physical and psychosocial.
Risk of bias assessment	**Population** Appropriately selected population. Excludes people with communication and cognitive difficulties which limits the scope of the results and applicability to the stroke population. However, in general balanced between groups. Mixture of people with different characteristics. Limited information about some subgroups (no information about physical activity prior to stroke, about initial stroke treatment etc.). **Control population** Appropriate population used (people with musculoskeletal conditions seems like an appropriate matched control). Though they may also experience fatigue so may not necessarily be a good match. **Recruitment/selection bias** Selection of people attending the department may introduce some bias of selecting only people with a certain level of fatigue (may not include people with severe fatigue). But in general seems appropriate. **Sample size** Small sample size (n=34) people after stroke who attended the retest). Identified as a weakness by the study authors. **Appropriateness of metrics** Appropriate metrics used. Good thorough analysis.
Limitations	Author provided limitations: • Small sample size which may reduce the generalization of the results • 33% missing data. It might be related to many people becoming bored due to the long form of the FIS.
Summary/author's conclusion	**Author's conclusion** The study provided evidence of the validity and reliability of the FIS as a multidimensional measurement scale of the influence of fatigue on the cognitive, physical and psychosocial functions of post-stroke patients.

Reference	Bragstad 2020[4]
Study type Setting/Location	Cross-sectional study, including initial pre-intervention data from two longitudinal studies, a multicentre randomised controlled trial evaluating the effect of an intervention promoting psychosocial wellbeing following a stroke and a longitudinal study investigating pain and other symptoms in patients with osteoarthritis undergoing total knee arthroplasty. Norway, acute stroke or rehabilitation units in university hospitals and other local hospitals providing acute care in Norway (for people after stroke). Inpatient or surgical clinic attendees (for people with osteoarthritis).

Number of participants and characteristics	N=525 (322 adult stroke survivors, 203 people with osteoarthritis).
	
	Inclusion criteria:
	Stroke survivors: Adults ≥18 years of age; acute stroke within 4 weeks prior to inclusion; medically stable; sufficient cognitive functioning to participate (assessed by their physician/stroke team); able to understand and speak Norwegian; able to give informed consent.
	People with osteoarthritis: Adults ≥18 years of age; ability to read, write and understand Norwegian; scheduled for unilateral primary total knee arthroplasty.
	
	Exclusion criteria:
	Stroke survivors: Moderate to severe dementia; serious somatic or psychiatric disease.
	People with osteoarthritis: People undergoing unicompartmental or revision surgery.
	
	Values listed below are presented as mean (SD) or number (%) unless stated otherwise
	
	Stroke survivors
	• Mean age (SD): 66.4 (12.8) years
	• Male/female: 190:132 (59.0%:41.0%)
	• Level of education
	  ○ Low (7-13 years): 217 (67.8%)
	  ○ High (14+ years): 103 (32.2%)
	• Marital status (married or partner/not married): 170:152 (52.8%:47.2%)
	• Lives alone/lives with someone: 104:218 (32.3%:67.7%)
	• Paid employment: 12 (3.7%)

Reference	Bragstad 2020[4]

	People with osteoarthritis
	• Mean age (SD): 68.2 (9.2) years
	• Male/female: 64:139 (31.5%:68.5%)
	• Level of education
	  ○ Low (7-13 years): 96 (48.5%)
	  ○ High (14+ years): 102 (51.5%)
	• Marital status (married or partner/not married): 115:84 (57.8%:42.2%)
	• Lives alone/lives with someone: 79:122 (39.3%:60.7%)
	• Paid employment: 70 (35.0%)

Intervention	**Lee Fatigue Scale (LFS) – compares the initial five item score and the final three item score.**
Comparison	Between scale types (initial five item scale and final three item scale).
Length of follow-up	No follow up
Outcome measures	Face/content/construct validity Discriminant/convergent validity Internal consistency
Source of funding	The stroke study was supported by a grant from the South-Eastern Norway Regional Health Authority (grant #2013086), a grant from the Extra Foundation (grant # 2015/FO13753), and funding from the European Union Seventh Framework Programme (FP7-PEOPLE-2013-COFUND) under grant agreement #609020 - Scientia Fellows. The University of Oslo, Oslo University Hospital, and the Inland Norway University of Applied Sciences have provided research time, administrative and organizational support and additional funding for the stroke study. The osteoarthritis study was funded by Lovisenberg Diaconal Hospital, the US-Norway Fulbright Foundation and the Norwegian Nurses Organization. Anners Lerdal received funding from the Norwegian Research Council of Norway (grant #287816). The postdoctoral fellowship for Maren. F. Lindberg was funded by the South-Eastern Norway Regional Health Authority (grant #2018060). The Norwegian non-profit National Association for Public Health's doctoral scholarship funds Ingrid Johansen Skogestad. The funding source had no involvement in conducting and reporting of this study.
Outcomes	**Face/content/construct validity** Meets criteria for rating scale functioning. Final LFS short form scale has adequate item goodness-of-fit statistics (between 0.7 and 1.3). Variance could be explained by the first latent variable 81.6% of the time. Person goodness-of-fit statistics were above the criterion of <5% (5.6%). Ceiling effects were less than 10% (0.6%) while floor effects were more than 10% (16.9%). Person separation

<table>
<tr><th>Reference</th><th>Bragstad 2020[4]</th></tr>
<tr><td></td><td>reliability was above 2.0 (2.49) and internal consistency was above 0.80 for all variable (Rasch 0.89, correlation between sum scores and Rasch measures 0.98).

Discriminant/convergent validity
<table>
<tr><th>Fatigue scores</th><th>Stroke sample (N=322)</th><th>Osteoarthritis sample (N=203)</th><th>P-value</th></tr>
<tr><td>LFS item 1</td><td>4.58 (2.33)</td><td>3.85 (2.45)</td><td>0.001</td></tr>
<tr><td>LFS item 4</td><td>3.54 (2.30)</td><td>2.90 (2.17)</td><td>0.002</td></tr>
<tr><td>LFS item 5</td><td>3.54 (2.32)</td><td>2.92 (2.14)</td><td>0.002</td></tr>
<tr><td>LFS item 16</td><td>3.06 (2.38)</td><td>2.26 (1.95)</td><td><0.001</td></tr>
<tr><td>LFS item 17</td><td>2.91 (2.56)</td><td>3.16 (2.62)</td><td>0.286</td></tr>
<tr><td>LFS-5</td><td>3.52 (1.89)</td><td>3.01 (1.92)</td><td>0.003</td></tr>
<tr><td>LFS-3</td><td>3.88 (2.10)</td><td>3.23 (2.13)</td><td>0.001</td></tr>
</table>
Conclusion: Similar values for LFS item 17, but otherwise the values appear to be distinctly different between the populations.

Internal consistency
• Cronbach’s alpha coefficient:
  ○ 0.87 (Initial LFS short form scale)
  ○ 0.91 (Final LFS short form scale)</td></tr>
<tr><td>Risk of bias assessment</td><td>Population
Stroke population appropriate. People at the subacute stage after stroke. Mixture of genders. Does not discuss type of stroke, initial stroke treatment, physical activity prior to stroke and severity. Baseline values provided.

Control population
People with osteoarthritis. May be a source of indirectness, as people with osteoarthritis may experience fatigue (albeit in a different way to stroke survivors).

Recruitment/selection bias
Selecting people included in two different trials (with different populations in each trial).

Sample size</td></tr>
</table>

Reference	Bragstad 2020[4]
	Larger sample size (N=525). **Appropriateness of metrics** Metrics appropriate.
Limitations	Author provided limitations: • The study evaluated only five of the original LFS items, so it remains unclear whether the three items retained in the analysis represent the best three items for inclusion in a brief fatigue severity patient-reported outcome measure. • Evaluated a Norwegian version of the LFS and so may not be applicable to English language (some words and phrases do not translate directly). • The mode of data collection was not identical in the stroke and osteoarthritis samples. The stroke population was interviewed in person, while the osteoarthritis group completed the questionnaire. • Although no difference was found based on diagnostic groups, the different data collection mode for the two samples may have introduced bias in the interpretation of items.
Summary/author's conclusion	**Author's conclusion** The results indicate that a 3-item version of the Lee Fatigue Scale has acceptable psychometric properties and is sufficiently generic for use as a patient-reported outcome measure for fatigue severity with patients post-stroke and patients living with osteoarthritis.

Reference	Brandal 2016[5]
Study type Setting/Location	Cross-sectional study Sweden, stroke unit
Number of participants and characteristics	N=90. Recruited between 1 April 2012 to 13 December 2012. **Inclusion criteria:** Diagnosis of mild to moderate ischaemic or haemorrhagic stroke **Exclusion criteria:** Living in residential care facilities due to severe cognitive dysfunction or if they had suffered from severe stroke (i.e., with modified Rankin Scale >3). **Values listed below are presented as mean (SD) or number (%) unless stated otherwise**

Reference	Brandal 2016[5]

	• Mean age (range): 68 (39 to 87) years
	• Male:Female = 49:22 (69%:31%)
	• Diagnosis of stroke:
	  ○ Ischaemic stroke (anterior circulation): 60 (83%)
	  ○ Ischaemic stroke (posterior circulation): 6 (8%)
	  ○ Intracerebral haemorrhage (anterior circulation): 5 (7%)
	  ○ Intracerebral haemorrhage (posterior circulation): 1 (1%)
	• Mean NIH Stroke Scale (SD) (range) = 2 (2.8) (0-13)
	• Functional status in terms of mobility, RMI, n (%) – walks independently = 67 (93%)
	• Median time between onset of stroke and first assessment with FAS-S (IQR): 132 (115-148) days

Intervention	**Functional Assessment Scale – Swedish version (FAS-S)** SF-36 vitality subscale (SF-36v) Geriatric Depression Scale (GDS-15) Rivermead Mobility Index (RMI)
Comparison	Itself, over time, different scales
Length of follow-up	Mean (SD): 9.6 (3.3) days
Outcome measures	Face/criterion/construct validity Test-retest reliability Internal consistency (Cronbach's alpha)
Source of funding	This study was supported by the Swedish Heart and Lung foundation, the Swedish Stroke foundation, the Northern Swedish Stroke Fund, the County of Västerbotten, and the medical faculty of Umeå University.
Outcomes	**Face/criterion/construct validity** Floor/ceiling effects: None of the participants received the highest or lowest scores, which implied that there were no floor or ceiling effects. Construct validity: The FAS correlated with the SF-36 subscale for vitality (r_s = -0.73) and with the GDS-15 (r_s = 0.62), indicating that the Swedish FAS had convergent construct validity, but not divergent construct validity.

Reference	Brandal 2016[5]
	Test-retest reliability Interclass correlation coefficient = 0.73 (see table below) **Internal consistency (Cronbach's alpha)** Cronbach's alpha = 0.82
Risk of bias assessment	**Population** People after stroke. Considers the time after stroke, diagnosis of stroke, severity and gender. Does not consider initial stroke treatment and physical activity prior to stroke. **Control population** Not applicable. **Recruitment/selection bias** People consecutively admitted to the stroke unit at the University Hospital of Umea, Sweden. Retrospectively identified. No additional information.

Item	Exact agreement n(%)	Weighted Kappa value (95% CI)
1. I am troubled by fatigue	43 (68%)	0.74 (0.61-0.87)
2. I get tired very quickly	38 (60%)	0.74 (0.61-0.86)
3. I only do a few things during the day	33 (52%)	0.57 (0.39-0.75)
4. I have enough energy to manage my everyday life	38 (60%)	0.49 (0.24-0.73)
5. I feel physically exhausted	35 (55%)	0.61 (0.45-0.78)
6. I find it difficult to make a start on things	42 (63%)	0.40 (0.22-0.58)
7. I experience problems with thinking things through	42 (67%)	0.64 (0.50-0.77)
8. I feel no desire to do anything	43 (70%)	0.49 (0.25-0.73)
9. I feel mentally exhausted	44 (70%)	0.64 (0.45-0.84)
10. I can concentrate quite well when I do something	31 (49%)	0.22 (-0.02-0.45)

Reference	Brandal 2016[5]
	Sample size Sample size (n=90). **Appropriateness of metrics** Metrics used were appropriate.
Limitations	Author provided limitations: • The time of the day for completion of the questionnaire was not specified. • The time span between test and retest was not fixed. • The exclusion of people with severe stroke. • Recall bias when completing questionnaires.
Summary/author's conclusion	**Author's conclusion** The Swedish FAS used at home as a self-administered questionnaire is a reliable and valid questionnaire for measuring fatigue in people with mild to moderate stroke. The FAS showed no floor or ceiling effects and, when used by an individual patient, a total score difference between two assessments of at least nine points indicates a real change of fatigue level.

Reference	Butt 2013[6]
Study type Setting/Location	Cross sectional study. United States of America, outpatients (recruited from rehabilitation institutes or other outpatient departments).
Number of participants and characteristics	N=399 (cancer = 297, stroke = 51, HIV/AIDS = 51). **Inclusion criteria:** Aged 18 or older, with a diagnosis of cancer, stroke or HIV/AIDS; understand and speak English; could interact with a touch screen computer with minimal assistance; stroke patients required to have a score of 24 or higher on the Folstein Mini-Mental Status Exam. **Exclusion criteria:** No additional information **Values listed below are presented as mean (SD) or number (%) unless stated otherwise**

Reference	Butt 2013[6]

	Cancer
	• Female: 64.3%
	• Mean age (SD): 58.1 (13.5) years
	• Race
	  ○ Caucasian = 82.5%
	  ○ Hispanic = 3.7%
	  ○ African-American = 9.4%
	  ○ Asian = 4.0%
	  ○ Pacific Island = 0.7%
	  ○ Other = 1.0%
	• Education
	  ○ Less than or equal to high school diploma = 18.2%
	  ○ Some college = 28.6%
	  ○ College degree = 31.0%
	  ○ Advanced degree = 22.2%
	
	Stroke
	• Female: 51.0%
	• Mean age (SD): 62.6 (13.9) years
	• Race
	  ○ Caucasian = 62.7%
	  ○ Hispanic = 2.0%
	  ○ African-American = 31.4%
	  ○ Asian = 3.9%
	  ○ Pacific Island = 0%
	  ○ Other = 4.0%
	• Education
	  ○ Less than or equal to high school diploma = 27.5%
	  ○ Some college = 29.4%
	  ○ College degree = 29.4%

Reference	Butt 2013[6]
	○ Advanced degree = 13.7% • Type of stroke – infarct: 70.0% • Subtype of stroke ○ Intracerebral haemorrhage = 36.4% ○ Subarachnoid haemorrhage = 12.1% ○ Thrombotic = 36.4% ○ Embolic = 15.2% • Location ○ Superficial/cortical = 27.3% ○ Subcortical = 56.8% ○ Combination or Other = 15.9% ○ % with previous stroke = 27.5% • Current stroke treatment ○ Physical therapy = 51.0% ○ Speech therapy = 17.6% ○ Vocational therapy = 3.9% ○ Psychological intervention = 3.9% ○ Occupational Therapy = 35.3% **HIV/AIDS** • Female: 11.8% • Mean age (SD): 40.2 (6.9) years • Race ○ Caucasian = 50.0% ○ Hispanic = 17.6% ○ African-American = 43.1% ○ Asian = 0% ○ Pacific Island = 0% ○ Other = 2.0% • Education ○ Less than or equal to high school diploma = 25.5%

<table>
<tr><td>Reference</td><td>Butt 2013[6]</td></tr>
<tr><td></td><td>o Some college = 29.4%
o College degree = 35.3%
o Advanced degree = 9.8%</td></tr>
<tr><td>Intervention</td><td>Functional Assessment of Chronic Illness Therapy – Fatigue (FACIT-F)

ECOG performance status rating</td></tr>
<tr><td>Comparison</td><td>Different populations, other scales</td></tr>
<tr><td>Length of follow-up</td><td>No follow up</td></tr>
<tr><td>Outcome measures</td><td>Discriminant/convergent validity

Internal consistency (Cronbach’s alpha)</td></tr>
<tr><td>Source of funding</td><td>NIH grant #K23 MH 084551.</td></tr>
<tr><td>Outcomes</td><td>Discriminant/convergent validity
<table>
<tr><th>Item</th><th>Stroke (N=51) (mean [SD])</th><th>Cancer (N=297) (mean [SD])</th><th>HIV (N=51) (mean [SD])</th></tr>
<tr><td>HI7: I feel fatigued</td><td>2.78 (1.08)</td><td>2.46 (1.18)</td><td>2.14 (1.10)</td></tr>
<tr><td>HI12: I feel weak all over</td><td>3.28 (1.01)</td><td>2.98 (1.20)</td><td>2.80 (1.15)</td></tr>
<tr><td>An1: I feel listless (“washed out”)</td><td>3.06 (1.18)</td><td>2.80 (1.21)</td><td>2.80 (1.11)</td></tr>
<tr><td>An2: I feel tired</td><td>2.76 (1.08)</td><td>2.63 (1.16)</td><td>2.29 (1.12)</td></tr>
<tr><td>An3: I have trouble starting things because I am tired</td><td>3.02 (1.12)</td><td>2.82 (1.18)</td><td>2.55 (1.15)</td></tr>
<tr><td>An4: I have trouble finishing things because I am tired</td><td>2.96 (1.12)</td><td>2.76 (1.18)</td><td>2.53 (1.12)</td></tr>
<tr><td>An5: I have energy</td><td>2.10 (1.03)</td><td>2.18 (1.12)</td><td>2.18 (1.01)</td></tr>
<tr><td>An7: I am able to do my usual activities</td><td>2.06 (1.27)</td><td>2.38 (1.21)</td><td>2.55 (1.10)</td></tr>
<tr><td>An8: I need to sleep during the day</td><td>2.96 (1.23)</td><td>2.89 (1.06)</td><td>2.63 (1.25)</td></tr>
<tr><td>An12: I am too tired to eat</td><td>3.72 (0.67)</td><td>3.58 (0.76)</td><td>3.33 (0.91)</td></tr>
<tr><td>An14: I need help doing my usual activities</td><td>3.04 (1.03)</td><td>3.20 (1.07)</td><td>3.12 (1.03)</td></tr>
<tr><td>An15: I am frustrated by being too tired to do the things I want to do</td><td>3.26 (0.94)</td><td>2.73 (1.27)</td><td>2.57 (1.38)</td></tr>
<tr><td>An16: I have to limit my social activity because I am tired</td><td>3.12 (1.06)</td><td>2.72 (1.28)</td><td>2.47 (1.21)</td></tr>
</table></td></tr>
</table>

Reference	Butt 2013[6]

Total	38.1 (9.6) Cronbach alpha: 0.91	36.0 (12.1) Cronbach alpha: 0.96	34.0 (12.6) Cronbach alpha: 0.97

Conclusion: Values are reasonably similar for some of the questions between groups and may be measuring a similar phenomenon experienced between the conditions.

Internal consistency

Cronbach alpha (overall) = 0.91 (smaller than the value when compared to people with Cancer [0.96] and HIV [0.97]).

Item	Item-total correlation	Cronbach alpha if item deleted
HI7: I feel fatigued	0.64	0.90
HI12: I feel weak all over	0.78	0.90
An1: I feel listless (“washed out”)	0.74	0.90
An2: I feel tired	0.84	0.89
An3: I have trouble starting things because I am tired	0.85	0.89
An4: I have trouble finishing things because I am tired	0.85	0.89
An5: I have energy	0.60	0.90
An7: I am able to do my usual activities	0.49	0.91
An8: I need to sleep during the day	0.15	0.92
An12: I am too tired to eat	0.46	0.91
An14: I need help doing my usual activities	0.38	0.91
An15: I am frustrated by being too tired to do the things I want to do	0.75	0.90
An16: I have to limit my social activity because I am tired	0.77	0.90

The questions examined here are only the ‘additional concerns’ section of the FACIT-F scale (while the full scale includes physical wellbeing, social/family wellbeing, emotional wellbeing and functional wellbeing). Questions are scored 0-4.

Risk of bias assessment	**Population** Population is appropriate. Includes people with stroke in a range of locations, different types of stroke and gender. No information about the time since stroke, initial stroke treatment, physical activity prior to stroke and severity.

Reference	Butt 2013[6]
	Control population People with cancer and HIV/AIDS. Both of these population can experience fatigue. This would be a source of indirectness. **Recruitment/selection bias** People from a recruitment site. Baseline values were comparable between groups. **Sample size** Small sample of stroke patients (N=51) **Appropriateness of metrics** The metrics used appear appropriate.
Limitations	Author provided limitations: • The study did not ensure that the items of FACIT-F capture every aspect of fatigue experienced by patients with HIV and stroke, but it did provide reassurance that the set of questions are perceived as relevant and responsive to fatigue caused by a variety of conditions. • While the cancer sample was relatively large, this study included a limited number of patients with stroke or HIV, and so may be less generalisable to that population. However, it is unlikely that a larger sample would result in significant changes.
Summary/author's conclusion	**Author's conclusion** The tool is a brief, easy to administer, patient-reported instrument to assess fatigue. While there are other tools that can be used to compare people after stroke, it shows that a tool that is usable in other populations can be used in this population, instead of changing tools each time by examining the common elements of fatigue rather than the unique elements of each condition.

Reference	Cheraghifard 2022[7]
Study type Setting/Location	Cross-sectional study Iran, community dwelling people who attended four rehabilitation centers.
Number of participants and characteristics	N=124. **Inclusion criteria:** Stroke at least 6 months before the study, which was diagnosed by a neurologist according to the International Classification of Diseases-Tenth Revision; moderate or severe fatigue (i.e. score greater than 2 on a single-item self-report questionnaire of "over the

Reference	Cheraghifard 2022[7]
	past week, have you usually felt fatigued? This feeling is not part of being sleepy or sad?"); no cognitive impiarments (MMSE at least 24); verbal and communicative ability to understand and follow instructions; no neurologic, orthopedic, rheumatologic, or psychological comorbidity; no cardiovascular instability; no unilateral visuospatial neglect. **Exclusion criteria:** Recurrent stroke during the study. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** **Stroke survivors** • Mean age (SD): 59.48 (11.78) years • Male/female: 73:51 (58.9%:41.1%) • Months after onset: 25.51 (21.24) months • Married/single/windowed or divorced: 99:9:16 (79.8%:7.3%:12.9%) • Jobless/retired/employment: 55:48:21 (44.4%:38.7%:16.9%) • Sub-diploma/diploma/academic: 43:37:44 (34.7%:29.8%:35.5%) • Dominant hand – right/left: 111:13 (89.5%:10.5%) • Affected side – right/left: 58:66 (46.8%:53.2%) • Ischaemic/haemorrhagic: 111:12 (89.5%:9.7%) • Pain present/absent: 69:55 (55.6%:44.4%) • Routine sport activity: 39 (31.5%) • Assistive device for gait: 59 (47.6%) • Falls in the past 6 months – 0 = 65 (52.4%), 1-3 = 44 (35.5%), >3 = 15 (12.1%). • Level of disability: no symptoms = 9 (7.3%), no significant disability = 32 (25.8%), slight disability = 29 (23.4%), moderate disability = 30 (24.2%), moderate severe disability = 24 (19.3%). • Height = 168.22 (8.61) cm • Weight = 72.48 (12.52) kg • Mini-Mental State Examination = 26.83 (2.43)
Intervention	**Fatigue Assessment Scale (FAS)** **Profile of Mood States – Fatigue subscale (POMS-F)** **SF-36 Vitality subscale (SF-36 VT)**

Reference	Cheraghifard 2022[7]
Comparison	Between scales
Length of follow-up	No follow up
Outcome measures	Test-retest reliability Internal consistency Responsiveness to change
Source of funding	This work was supported by the Iran University of Medical Sciences.
Outcomes	**Test-retest reliability** ICC of FAS (95% CI) = 0.87 (0.78-0.92) ICC of POMS-F (95% CI) = 0.93 (0.88-0.95) **Internal consistency** Cronbach's alpha for FAS = 0.86 Cronbach's alpha for POMS-F = 0.81 Cronbach's alpha for SF-36-VT = 0.86 **Responsiveness to change** MCID of FAS = 3.16 to 8.76 (6.3-17.5% of the total score) MCID of POMS-F = 1.49 to 5.63 (5.3-20.1% of the total score) MCID of SF-36-VT = -5.58 to -15.43
Risk of bias assessment	**Population** Varied population from a number of different groups. Limited information about some subgroups (no information about physical activity prior to stroke, about initial stroke treatment etc.). **Control population** Not applicable. **Recruitment/selection bias** Recruitment from four centers. Excluded people with cognitive impairment. Did not specify whether people with communication difficulties were included.

Reference	Cheraghifard 2022[7]
	Sample size Good sample size (n=124). Reported to be small by the study authors. **Appropriateness of metrics** The metrics used seem appropriate.
Limitations	Author provided limitations: Absence of control group. Small sample size and selection bias.
Summary/author's conclusion	**Author's conclusion** Fatigue is the most common symptom in stroke survivors that should be considered. This study provided the MCID of the FAS, POMS-F and SF-36-VT in chronic stroke survivors. Using these, the results of interventions can be interpreted more appropriately.

Reference	Chuang 2015[8]
Study type Setting/Location	Cross sectional study. Taiwan, outpatient rehabilitation program
Number of participants and characteristics	N=106 **Inclusion criteria:** Stroke patients diagnosed between December 2013 and January 2015 recruited at three medical centres. First-ever stroke onset of at least 3 months before recruitment; enrolment in an outpatient rehabilitation program; ability to follow study instructions and complete the scale (Mini-Mental State Examination score of at least 22); no participation in experimental rehabilitation or drug studies during the study period. **Exclusion criteria:** No additional information. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** • Male:Female = 77:29 (72.6%:27.4%) • Mean age (SD): 53.63 (11.25) years

Reference	Chuang 2015[8]

	• Localisation, right:left: 48:58 (45.3%:54.7%)
	• Interval after stroke onset (SD): 24.40 (24.11) months
	• Brunnstrom stage of upper limb
	  ○ Proximal = 4 (1-6)
	  ○ Distal = 3 (1-6)
	• Brunnstrom stage of lower limb: 4 (3-5)
	• Fugl-Meyer assessment of upper limb (SD): 33.74 (17.62)
	• Fugl-Meyer assessment of lower limb (SD): 21.18 (6.94)
	• Mini Mental Stage Exam scores: 27.56 (2.43)

Intervention	**Numeric Rating Scale – Faces Rating Scale (NRS-FRS)** Numeric Rating Scale
Comparison	Each other and different time point
Length of follow-up	1 week
Outcome measures	Criterion/concurrent validity (Spearman Rank Correlation Coefficient) Test-retest reliability Responsiveness to change
Source of funding	This work was supported by the Ministry of Science and Technology (MOST-102-2314-B-182-003, 102-2314-B-002-154-MY2, 102-2628-B-182 -005 -MY3, and 103-2314-B-182-004-MY3), the National Health Research Institutes (NHRI-EX104-10403PI), Healthy Ageing Research Center at Chang Gung University (EMRPD1E1711), and Chang Gung Memorial Hospital (CMRPD3E0331, CMRPD1B0332, CMRPD1C0403) in Taiwan
Outcomes	**Criterion/concurrent validity** Test NRS-FRS = 0.85 (0.75-0.91) Retest NRS-FRS = 0.84 (0.73-0.91) Conclusion: There is a similar amount of agreement between the two measures that was consistent between the two time periods. **Test-retest reliability** Intraclass correlation coefficient = 0.95 (0.92-0.96)

Reference	Chuang 2015[8]
	Responsiveness to change Minimal detectable change at the 95% confidence interval level: 1.39
Risk of bias assessment	**Population** Population included stroke patients in the chronic phase. Mixture of localisation of stroke and gender. However, no information on type of stroke, initial stroke treatment, physical activity prior to stroke and severity. **Control population** No additional population. No indirectness. **Recruitment/selection bias** Recruited at three medical centers. Baseline values appear comparable for the cohort and include a range of people. However, while the criteria state that people have to have had a stroke of at least 3 months before recruitment, the number of months on average is significantly more than this (may indicate that less people in the subacute phase were included). **Sample size** Sample size = 106. **Appropriateness of metrics** Metrics used appear appropriate.
Limitations	Author provided limitations: • All participants completed the NRS-FRS at two assessments, at the same time of day to minimize diurnal variation in fatigue. Fatigue was measured as a single time-point assessment; this is, current fatigue intensity, which might not reflect overall fatigue intensity, which might not reflect overall fatigue on the testing day. Testing at different times of the day may help to show whether there are changes in daily fluctuation that may be used to improve the psychometric properties of the test. • Future studies need to identify predictors of poststroke fatigue to address fatigue issues with an intervention in people with stroke. To explore the effectiveness, the ability to detect NRS-FRS to detect change over time requires further development.
Summary/author's conclusion	**Author's conclusion** The vertical NRS-FRS has good test-retest reliability and validity in measuring physical fatigue after stroke, with good agreement, low measurement error and high sensitivity and specificity.

Reference	Ho 2020[14]
Study type Setting/Location	Cross-sectional study Hong Kong, local self-help groups
Number of participants and characteristics	N=112 (recruited from January to April 2019). Only the first 27 participants had a reassessment. **Inclusion criteria:** Age 55 or above; able to understand Cantonese; diagnosed with stroke for at least 1 year; able to walk 10m independently with or without walking aids; living in the community. **Exclusion criteria:** Any other confirmed neurological diseases; unstable medical conditions leading to fatigue; suffered a transient ischaemic attack. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** • Mean age (SD): 64.15 (5.79) years • Female:Male = 38:74 (33.9%:66.1%) • Number of strokes (SD): 1.13 (0.41) • Time since stroke (SD): 6.13 (4.79) years • Geriatric depression scale (SD): 5.39 (3.74) • Marital status (single:married:widow/widower:divorced/separated): 7:89:8:8 • Employment status (full-time:part-time:unemployed:retired): 4:2:65:41
Intervention	**Fatigue Assessment Scale – Chinese version (FAS-C)** Fatigue Severity Scale – Chinese version (FSS-C) Mental Fatigue Scale – Chinese version (MFS-C) Epworth Sleepiness Scale – Chinese version (ESS-C) Five times Sit-To-Stand Test (FTSTS) Fugl-Meyer Assessment (FMA) 15-Item Geriatric Depression Scale – Chinese Version (GDS-15-C)
Comparison	Other scales, different times, different duration of times since stroke

<table>
<tr><th>Reference</th><th>Ho 2020[14]</th></tr>
<tr><td>Length of follow-up</td><td>1 week</td></tr>
<tr><td>Outcome measures</td><td>Face/content/construct validity

Test-retest reliability
Internal consistency (Cronbach’s alpha)
Responsiveness to change

Dimensions of fatigue considered</td></tr>
<tr><td>Source of funding</td><td>This work was supported by the Hong Kong Polytechnic University, and the Departmental Research Grant (ref: 90013897) from the Department of Rehabilitation Sciences, The Hong Kong Polytechnic University.</td></tr>
<tr><td>Outcomes</td><td>Face/content/construct validity

Content validity index:
The item level-content validity index (assessed by five expert panel members) ranged from 0.80 to 1.00. The values for items 1, 2 and 9 were 0.80 due to the use of Chinese wording. A consensus on the use of word was reached in an expert panel meeting. Minor amendments were made to the sentence structure of items 4 and 10. The scale-content validity index was 0.94.

<table>
<tr><th>Item</th><th>Content validity index</th></tr>
<tr><td>1</td><td>0.80</td></tr>
<tr><td>2</td><td>0.80</td></tr>
<tr><td>3</td><td>1.00</td></tr>
<tr><td>4</td><td>1.00</td></tr>
<tr><td>5</td><td>1.00</td></tr>
<tr><td>6</td><td>1.00</td></tr>
<tr><td>7</td><td>1.00</td></tr>
<tr><td>8</td><td>1.00</td></tr>
<tr><td>9</td><td>0.80</td></tr>
<tr><td>10</td><td>1.00</td></tr>
<tr><td>Overall</td><td>0.94</td></tr>
</table>
</td></tr>
</table>

Reference	Ho 2020[14]

Ceiling and floor effects:
None of the participants received the higher and lowest sum score. One person received the highest physical score and 3 received the lowest physical score. None got the highest mental health score and 2 got the lowest.

Construct validity:
Kaiser-Meyer-Olkin measure = 0.84, indicating sufficient items for each factor. The Bartlett's test of sphericity was significant. Two factors, physical fatigue and mental fatigue, were identified from the scree plot. The loadings of all 10 items ranged from 0.32 to 0.78, with 53.72% of the total variance explained.

The FAS-C was correlated with:

- MFS-C (r_s = 0.68, p <0.001)
- FSS-C (r_s = 0.57, p <0.001)
- ESS-C (r_s = 0.36, p <0.001)
- FMA upper extremities (r_s = 0.24, p 0.011)
- FMA lower extremities (r_s = 0.24, p 0.012)

It did not correlate with FTSTS time (r_s = 0.13, p 0.170)

Level of physical and mental fatigue:
The median physical fatigue score was significantly higher than the mental fatigue score (12.00 vs. 10.00, p<0.001). There was a statistically insignificant difference between the groups across 3 stroke durations (1-5 years, >5-10 years, >10 years).

Test-retest reliability
Intra-class correlation coefficient:

- Summary score = 0.92 (p <0.001)
- Physical score = 0.95 (p <0.001)
- Mental score = 0.77 (p <0.001)

Items	Exact agreement (%)	Weighted kappa value (95% CI)
1	88.9	0.83 (0.64-1.01)
2	70.4	0.54 (0.27-0.80)

Reference	Ho 2020[14]

3	66.7	0.54 (0.30-0.78)
4	55.6	0.44 (0.21-0.68)
5	74.1	0.53 (0.24-0.82)
6	63.0	0.38 (0.06-0.69)
7	74.1	0.56 (0.27-0.85)
8	74.1	0.58 (0.30-0.86)
9	63.0	0.45 (0.20-0.70)
10	63.0	0.51 (0.27-0.75)

Internal consistency (Cronbach's alpha)
Cronbach's alpha:

- Summary score = 0.82
- Physical score = 0.78
- Mental score = 0.71

Responsiveness to change
Minimal detectable change:

- Summary score = 4.69
- Physical score = 2.44
- Mental score = 4.10

Dimensions of fatigue considered
Physical and mental

Risk of bias assessment

Population
People after stroke. Mixture of men and women. Most people would have a chronic time period after stroke. Does not provide information on type of stroke, location of stroke, initial stroke treatment, physical activity prior to stroke, gender and severity.

Control population
Not applicable.

Reference	Ho 2020[14]
	Recruitment/selection bias Convenience samplings of participants. Taken from a community setting. Only includes a small number of participants when testing test-retest reliability. **Sample size** Sample size (N=112) **Appropriateness of metrics** Metrics used were appropriate
Limitations	Author provided limitations: • Participants needed to come to the university for the assessment, so those with a very high level of fatigue or with poorer functional mobility might not have been willing to take part in the study • Cross-sectional study and so could not show changes in fatigue over time • The data collection period went through the national public holiday, which might have affected the results
Summary/author's conclusion	**Author's conclusion** The FAS-C captured the characteristics of fatigue following stroke and is a reliable and valid tool. The minimal detectable change was good. The scale had two factors, physical and mental fatigue, without ceiling and floor effects. The FAS-C significantly correlated with the MFS-C, FSS-C, ESS-C and FMA of upper and lower extremities. The level of physical fatigue was higher than that of mental fatigue in community-dwelling stroke survivors, and those with depressive symptoms had higher fatigue scores than those without.

Reference	Ho 2021[13]
Study type Setting/Location	Validation study. Hong Kong, self-help groups and a community center.
Number of participants and characteristics	N=177 (112 stroke survivors, 65 healthy older people). Stroke survivors and healthy older people recruited by convenience sampling from self-help groups for people with neurological diseases and from a non-governmental organisation in the Hong Kong SAR. **Inclusion criteria:**

Reference	Ho 2021[13]
	Community-dwelling residents; who are ethnic Chinese; able to speech Chinese; aged at least 55 years; who had a confirmed diagnosis of stroke for at least 1 year (stroke group only). **Exclusion criteria:** People with any other neurological diseases; people who had experienced a transient ischaemic attack; people with unstable medical conditions. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** • Mean age (SD): 64.15 (5.79) years • Male: 66.1% • Mean time since diagnosis (SD): 73.60 (57.43) months
Intervention	**Chinese version of the Neurological Fatigue Index-Stroke (C-NFI-Stroke)** Fatigue Severity Scale (FSS) Mental Fatigue Scale (MFS) General Self-Efficacy Scale (GSES) Geriatric Depression Scale (GDS)
Comparison	Different populations
Length of follow-up	1 week
Outcome measures	Face/content/construct validity Criterion/Concurrent validity (Spearman's correlation coefficients) Discriminant/convergent validity Test-retest reliability Internal consistency (Cronbach's alpha, intra-class correlation coefficient) Inter-rater reliability Responsiveness to change Dimensions of fatigue considered

<table>
<tr><th>Reference</th><th>Ho 2021[13]</th></tr>
<tr><td>Source of funding</td><td>The work was financially supported by 1) The Hong Kong Polytechnic University; 2) The Departmental Research Grant from the Department of Rehabilitation Sciences, The Hong Kong Polytechnic University (Grant number: 90013897).</td></tr>
<tr><td>Outcomes</td><td>Face/content/construct validity
Item-level content validity index scores ranged from 0.6-1, with the scale-level content validity index score of 0.95 (a score of 0.78 was considered good).
The Kaiser-Meyer-Olkin measure was 0.91, implying sufficient items for a factor analysis (a score of 0.90 was considered good).

Criterion/Concurrent validity
<table>
<tr><th></th><th>Fatigue Severity Scale</th><th>Mental Fatigue Scale</th><th>General Self-Efficacy Scale</th><th>Geriatric Depression Scale</th></tr>
<tr><td>C-NFI-Stroke summary scale</td><td>0.62</td><td>0.63</td><td>-0.35</td><td>0.60</td></tr>
<tr><td>C-NFI-Stroke physical subscale</td><td>0.60</td><td>0.61</td><td>-0.31</td><td>0.58</td></tr>
<tr><td>C-NFI-Stroke cognitive subscale</td><td>0.55</td><td>0.61</td><td>-0.37</td><td>0.54</td></tr>
</table>
Conclusion: The C-NFI-Stroke score correlates with the FSS, MFS and GDS. It correlates least with the GSES.

Discriminant/convergent validity
<table>
<tr><th></th><th>Median (interquartile range)</th><th>Mann-Whitney U</th><th>Z value</th><th>P value</th></tr>
<tr><td>Summary score</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Stroke survivors</td><td>16.00 (6.50)</td><td>1133.50</td><td>-4.67</td><td><0.001</td></tr>
<tr><td>Healthy older people</td><td>12.00 (10.50)</td><td></td><td></td><td></td></tr>
<tr><td>Physical score</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Stroke survivors</td><td>13.00 (5.00)</td><td>1128.50</td><td>-4.70</td><td><0.001</td></tr>
<tr><td>Healthy older people</td><td>9.00 (8.00)</td><td></td><td></td><td></td></tr>
<tr><td>Cognitive score</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Stroke survivors</td><td>6.00 (2.25)</td><td>987.00</td><td>-5.39</td><td><0.001</td></tr>
<tr><td>Healthy older people</td><td>4.00 (4.50)</td><td></td><td></td><td></td></tr>
</table></td></tr>
</table>

Reference	Ho 2021[13]

Conclusion: Statistically significantly different values were seen between the two groups.

Test-retest reliability

Intra-class correlation coefficient

- Summary scale = 0.93
- Physical subscale = 0.92
- Cognitive subscale = 0.88

Internal consistency

Cronbach's alpha:

- Summary scale = 0.88
- Physical subscale = 0.87
- Cognitive subscale = 0.69

Inter-rater reliability (Intratest reliability)

Item	Exact agreement	Weighted Kappa value (95% confidence interval)
1	88.9%	0.79 (0.57-1.01)
2	70.4%	0.47 (0.17-0.77)
3	70.4%	0.48 (0.20-0.77)
4	81.5%	0.58 (0.27-0.89)
5	77.8%	0.58 (0.28-0.87)
6	77.8%	0.60 (0.33-0.87)
7	85.2%	0.69 (0.44-0.95)
8	77.8%	0.47 (0.14-0.80)
9	77.8%	0.53 (0.21-0.86)
10	85.2%	0.73 (0.49-0.97)
11	81.5%	0.65 (0.39-0.91)
12	74.1%	0.52 (0.22-0.81)

Responsiveness to change

Reference	Ho 2021[13]
	Minimal detectable change: • Summary scale = 2.92 • Physical subscale = 2.68 • Cognitive subscale = 1.57 **Dimensions of fatigue considered** Includes physical and cognitive subscales.
Risk of bias assessment	**Population** People after stroke. Mixture of genders. Chronic time horizon. No information on type of stroke, initial stroke treatment, physical activity prior to stroke and severity. **Control population** Healthy participants. Appropriate comparison. However, the limitations note that this population included caregivers of stroke survivors, and so may not be representative of the general population. **Recruitment/selection bias** Convenience sample. Baseline values appear comparable. **Sample size** Sample size (N=177) with a small number in each arm (N=112 and N=65 respectively). **Appropriateness of metrics** Metrics are appropriate.
Limitations	Author provided limitations: • Participants came from only a few local self-help groups and a non-governmental organisation, limiting the generalisability of this study. • Some potential participants were unwilling to join because of the need to travel long distances to the assessment venue. • The sample size might not be adequate for comparisons to be made between stroke survivors and healthy older people with and without depressive symptoms • Since some healthy older people were caregivers of stroke survivors, they likely were not representative of the general healthy population.

Reference	Ho 2021[13]
Summary/author's conclusion	**Author's conclusion** The C-NFI-Stroke is reliable and valid for measuring fatigue in both clinical practice and research. Fatigue was correlated with self-efficacy and depressive symptoms. The C-NFI-Stroke may help clinicians and researchers to evaluate the effectiveness of interventions designed to alleviate fatigue in stroke survivors.

Reference	Kruithof 2016[15]
Study type Setting/Location	Prospective cohort study The Netherlands, the stroke department of Adelante Rehabilitation Centre (people following an inpatient or outpatient rehabilitation program)

Number of participants and characteristics	N=107 (121 approached, 13 excluded as they did not fulfil the inclusion criteria, 1 refused to participate for unknown reasons).
	
	Inclusion criteria:
	People (18 years or older) clinically diagnosed with stroke and following an inpatient or outpatient rehabilitation program between May 2013 and October 2014. People in the postacute phase of stroke (5-26 weeks) who were receiving physical therapy or occupational therapy at the time of inclusion.
	
	Exclusion criteria:
	Psychiatric disorder was present for which medication was prescribed at the time of inclusion; poor understanding of the Dutch language; people who were not able to give adequate verbal informed consent as a result of severe receptive aphasia or severe cognitive impairments.
	
	Values listed below are presented as mean (SD) or number (%) unless stated otherwise
	
	• Mean age (SD): 60 (10) years
	• Male = 72 (67.3%)
	• Type of lesion:
	  ○ Infarct = 74 (69.2%)
	  ○ Haemorrhage = 24 (22.4%)
	  ○ Subarachnoid haemorrhage = 6 (5.6%)
	  ○ Unknown = 3 (2.8%)

Reference	Kruithof 2016[15]

	• Lesion location:
	  ○ Left = 41 (38.3%)
	  ○ Right = 57 (53.3%)
	  ○ Diffuse = 9 (8.4%)
	• Single stroke = 97 (90.7%)
	• 2 strokes = 9 (8.4%)
	• More than 2 strokes = 1 (0.9%)

Intervention	**Detection List Fatigue (tested in Dutch) (DLF)** Hospital Anxiety and Depression Scale – Anxiety and Depression subscales (HADS-A/HADS-D) – T1 only Fatigue rating scale (FRS) – T1 and T2 Checklist Individual Strength subscale fatigue (CIS-f) – T2 only Fatigue Severity Scale – 7 item version (FSS) – T2 only
Comparison	Different scales, itself
Length of follow-up	6 weeks
Outcome measures	Face/content/construct validity Discriminant/convergent validity Dimensions of fatigue considered
Source of funding	No additional information
Outcomes	**Face/content/construct validity** According to the assessors, the DLF was generally easy to understand for patients and quick to administer. At T1, 1 person was not able to answer item 3 after it was repeated to them 3 times. Four people had no idea to what degree item 3 was true or not, and 1 person had no idea to what degree items 6 and 9 were true or not. At T2, 1 person had no idea to what degree item 5 was true ot not, and 2 people had no idea to what degree item 9 was true or not. In a small number of cases, it was difficult to know whether a patient comprehended an item properly because cognitive deficits, aphasia or a combination of both resulted in communication difficulties. People with light to moderate aphasia or cognitive deficits could relatively easily complete the DLF by pointing out the answer on a separate answer sheet. The mean time to administer the test including the instruction time was 6 minutes (+/- 2, range 3-14). **Discriminant/convergent validity** Spearman's rank correlation coefficients:

<table>
<tr><td>Reference</td><td>Kruithof 2016[15]</td></tr>
<tr><td></td><td>
<table>
<tr><td>Measures</td><td>N</td><td>Validity – r_s</td></tr>
<tr><td>T1: DLF-HADSA</td><td>66</td><td>0.45, p<0.001</td></tr>
<tr><td>T1: DLF-HADSD</td><td>66</td><td>0.31, p<0.010</td></tr>
<tr><td>T1: DLF-FRS</td><td>107</td><td>0.58, p<0.001</td></tr>
<tr><td>T2: DLF-FRS</td><td>72</td><td>0.63, p<0.001</td></tr>
<tr><td>T2: DLF-CIS-f</td><td>72</td><td>0.85, p<0.001</td></tr>
<tr><td>T2: DLF-FSS-7</td><td>72</td><td>0.79, p<0.001</td></tr>
</table>
Dimensions of fatigue considered
Physical and mental</td></tr>
<tr><td>Risk of bias assessment</td><td>Population
People after stroke. Mixture of gender, type and location of stroke. Does not provide information on initial stroke treatment, prior physical activity and severity.

Control population
Not applicable.

Recruitment/selection bias
Inpatients and outpatients. People attending for rehabilitation at one center. No obvious problems.

Sample size
Sample size (n=107). Less people at T2 (n=72), but no significant problems.

Appropriateness of metrics
Metrics appear appropriate.</td></tr>
<tr><td>Limitations</td><td>Author provided limitations:
HADS was not administered in each patient because not every patient received a standard neuropsychological examination
The HADS, CIS-f and FSS-7 scores were not assessed at both time points, but the FRS score was</td></tr>
<tr><td>Summary/author’s conclusion</td><td>Author’s conclusion</td></tr>
</table>

Reference	Kruithof 2016[15]
	It can be concluded that the DLF is suitable for clinical practice as it is valid, short and simple to administer for a wide variety of stroke patients.

Reference	Lerdal 2011[17]
Study type Setting/Location	Cross-sectional study Norway, participants in the post-stroke fatigue study (PS-Study) recruited between March 2007 and September 2008 at one hospital in the south-eastern region, and between September 2007 and June 2008 at the university hospital in Oslo.
Number of participants and characteristics	N=119 **Inclusion criteria:** First-ever clinical presentation of stroke according to the ICD-10; were 18 years or older; had satisfactory cognitive functions to participate. **Exclusion criteria:** No additional information **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** • Mean age (SD): 68.3 (13.1) years • Women:Men = 47:72 (39.5%:60.5%) • Level of education (13 years and above:11-12 years:7-10 years) = 38:47:34 (31.9%:39.5%:28.6%)
Intervention	**Fatigue Severity Scale – Norwegian version (FSS-N)** SF-36 vitality subscale (SF-36v) Energy-VAS (E-VAS)
Comparison	Itself, different time periods
Length of follow-up	Filled in at four time points: baseline (n=119), 6 months (n=106), 12 months (n=104), 18 months (n=99).

Reference	Lerdal 2011[17]
Outcome measures	Face/content/construct validity Criterion/concurrent validity Internal consistency (Cronbach's alpha)
Source of funding	This project is funded by the Research Council of Norway and Buskerud University College for 2006 to 1010 (Grant: 176503).
Outcomes	**Face/content/construct validity** Psychometric properties (assessed in 428 measurements): The 9-item version has 1 item not meeting the three criteria for rating scale and 1 item misfit. The 8-item version had 2 items that had misfit. The 7-item version (with items 1 and 2 removed) had no misfit and all items met the criteria for rating scale. First latent variable, % = 83.1% 2nd dimension, % = 3.6% Eigenvalue = 1.5 Person misfit, n (%) = 30 (7.0%) (maximum score achieved in 7, minimum score achieved in 23) Person-separation index (without extremes) = 2.40 Person-separation reliability = 0.85 **Criterion/concurrent validity** SF-36v: • Baseline = -0.56 • 6 months = -0.62 • 12 months = -0.72 • 18 months = -0.67 E-VAS: • Baseline = -0.40 • 6 months = -0.56 • 12 months = -0.61 • 18 months = -0.53 **Internal consistency** Cronbach's alpha:

Reference	Lerdal 2011[17]
	• Baseline = 0.87 • 6 months = 0.92 • 12 months = 0.93 • 18 months = 0.92
Risk of bias assessment	**Population** People after stroke. Mixture of genders. No information about type of stroke, location of stroke, initial stroke treatment, physical activity prior to stroke and severity. **Control population** Not applicable. **Recruitment/selection bias** People recruited from an existing study upon hospital admission. Convenience sample. **Sample size** Sample size (n=119 with repeated measures over time to make a larger number of tests). **Appropriateness of metrics** No problems noted.
Limitations	Author provided limitations: • The samples in the study consisted of a convenience sample of people after their first stroke • Generalisation of findings to people with multiple stroke events and severe disability should be done with careful consideration.
Summary/author's conclusion	**Author's conclusion** Misfitting items that were only detected using modern test theory and not with a traditional approach when analysing psychometric properties of the FSS were found. The FSS-7 demonstrates better validity and reliability and is probably more sensitive for measuring change in fatigue.

Reference	Lynch 2007[18]
Study type	Cross-sectional study
Setting/Location	

Reference	Lynch 2007[18]
	United Kingdom, two stroke units (one acute and one rehabilitation) and people in the community

Number of participants and characteristics	N=55
	
	Inclusion criteria:
	Inpatient/outpatient care: People admitted with a new stroke.
	Community: Community stroke nurses visiting people after stroke.
	
	Exclusion criteria:
	Medically unstable because of another condition; those with dysphasia or confusion severe enough to prevent them from understanding the requirements of participation.
	
	Values listed below are presented as mean (SD) or number (%) unless stated otherwise
	
	• Median age (IQR): 74 (66-81) years
	• Male = 31 (56%)
	• Right hemisphere stroke = 27 (49%)
	• Haemorrhagic stroke = 3 (6%)
	• Location of stroke:
	  ○ TACS = 11 (20%)
	  ○ PACS = 22 (40%)
	  ○ LACS = 16 (29%)
	  ○ POCS = 6 (11%)
	• Relevant lesion on computer tomography = 43 (81%)
	• Median HADS anxiety score (IQR) = 7 (4-11)
	• Median HADS depression score (IQR) = 7.5 (3-10)
	• Fatigued according to case definition = 20 (36%)

Intervention	**Case definition for fatigue** The final version of the case definition (made during the study) was: Community patients: Over the past month, there had been at least a 2 week period when patient has experienced fatigue, a lack of energy or an increased need to rest every day or nearly every day. This fatigue has led to difficulty taking part in everyday activities.

Reference	Lynch 2007[18]
	Hospital patients: Since their stroke, the patient has experienced fatigue, a lack of energy, or an increased need to rest every day or nearly every day. This fatigue has led to difficulty taking part in everyday activities (for inpatients this may include therapy and may include the need to terminate an activity early because of fatigue). Structured interview schedule for community patients (can be applied to hospital cases based on later writing in the study, the questions just need a change in wording from 'over the past month' to 'since your stroke'): 1a) Over the past month, have you experienced fatigue, a lack of energy or an increased need to rest? 1bi) Can you describe what fatigue feels like, in your own words? 1bii) Is it a sleepy feeling, or is it more a lack of energy? 1ci) Over the past month, how much of the time do you feel fatigued? 1cii) How much of the day do you feel fatigued? 2a) Do you feel that fatigue is a problem for you? 2b) Is there anything else about your experience of fatigue that you feel is important? Scoring: 1a) Patient must answer yes to fulfil the case definition. If they answer no, go straight to question 2a. 1bi) Patient should describe feelings which are consistent with fatigue or lack of energy or need to rest rather than motivation or boredom. 1bii) Patients should describe feelings of fatigue (or lack of energy or increased need to rest) rather than sleepiness 1ci) Fatigue should have been present everyday or nearly everyday for at least 2 weeks in the past month 1cii) Fatigue must be present for >50% of waking hours 2a) Fatigue must be perceived as a problem and affect everyday activities (e.g. activities of daily living, recreational activities such as reading or watching the television) and may or may not affect participation in therapy. If a person has put adaptations in place for this that they do not perceive as problematic, then it should be rated as no. If the fatigue affects their mood, this in turn affects their participation in activities and so should be scored yes. 2b) Note the patient's response. SF-36 (raw total) Fatigue Assessment Scale (FAS) Profile of Mood States (POMS) Modified Fatigue Severity Index (MFSI)

Reference	Lynch 2007[18]
Comparison	Other scales, itself, over time, between raters
Length of follow-up	4 days
Outcome measures	Face/content/construct validity Criterion/concurrent validity Test-retest reliability Inter-rater reliability Dimensions of fatigue considered
Source of funding	No additional information
Outcomes	**Face/content/construct validity** Feasibility: At both interviews, all participating patients provided satisfactory answers to all case definition probe questions. Clinical characteristics of cases: 20 (36%) patients fulfilled the case definition at the first interview. This was associated with female gender and greater emotional distress. **Criterion/concurrent validity** Assessed in graphs. The paper stated that patients fulfilling the case definition generally had substantially higher fatigue scores on all four fatigue scales, suggesting good concurrent validity ($p<0.001$, Mann-Whitney U test). **Test-retest reliability** Kappa (95% confidence interval) = 0.78 (0.60-0.96), n=51 **Inter-rater reliability** Kappa (95% confidence interval) = 0.82 (0.64-0.99), n=43 (participants excluded due to technical difficulties with the tape recordings [3], the tape being unclear [3] and the second observer deciding there was not enough information to make a decision about case definition fulfilment [2]). **Dimensions of fatigue considered**

Reference	Lynch 2007[18]
	Hollistic (by the nature of the interview).
Risk of bias assessment	**Population** People after stroke. Information on gender, location of stroke and type of stroke. No information about initial stroke treatment, physical activity prior to stroke and severity. **Control population** Not applicable. **Recruitment/selection bias** People from a range of settings. Convenience sample. **Sample size** Small sample size (n=55). **Appropriateness of metrics** Metrics used were appropriate.
Limitations	Author provided limitations: • A larger sample size would have provided a more precise estimate of reliability • Ethical approval stipulated that clinical staff had to make the initial approach to patients, so patients known to be tired may have been approached more often and therefore overrepresented, reducing the range of fatigue severity in the sample • Relied on face validity and concurrent validity as there is currently no test for identifying fatigue after stroke, making criterion validity impossible • When assessing test-retest reliability, the interviewer may have remembered the outcome of the first interview when performing the second. • The assessment of inter-rater agreement would have been more rigorous if the second rater had repeated a face-to-face case definition interview and not merely listened to a recording • Limitations to the generalisability of the data. The sample was mainly inpatients, and the nature of fatigue may change after discharge from hospital as the patients' activities become more complex. They excluded people with dysphasia or confusion for whom visual analogue scales or pictorial representations of fatigue may be appropriate.
Summary/author's conclusion	**Author's conclusion** A new case definition and interview for clinically significant post-stroke fatigue after stroke is proposed, which has face validity and concurrent validity, is feasible to administer and is reliable in practice at least in stroke inpatients without communication difficulties.

Reference	Mead 2007[19]
Study type Setting/Location	Cross sectional study. Fatigue scales were identified by a systematic search, and the 5 with the best face validity was identified by expert consensus. Face validity was assessed on the criteria that the scale 1) captured the phenomenon of poststroke fatigue, 2) was free from items indistinguishable from the effects of the stroke (e.g. “my limbs feel weak”. These scales were then pilot tested in 13 stroke inpatients. Patients were then interviewed, and each questionnaire was evaluated. United Kingdom. Inpatient and community settings.

Number of participants and characteristics	N=64 invited to participate, 55 consented
	
	Inclusion criteria:
	People were recruited from hospital stroke wards (at least 1 week after stroke onset) and from the community via stroke clinics or community nurses.
	
	Exclusion criteria:
	Patients with dysphasia or confusion severe enough to prevent them from understanding the rationale for the study or giving informed consent; medically unstable because of another condition.
	
	Values listed below are presented as mean (SD) or number (%) unless stated otherwise
	
	• Median age (IQR): 73 (66 to 81) years
	• Male: 31 (56%)
	• Inpatients/Community: 40/15
	• Type of stroke
	  ○ Right hemisphere stroke: 27 (49%)
	  ○ Haemorrhagic stroke: 3 (6%)
	  ○ Total anterior circulation syndromes: 11 (20%)
	  ○ Partial anterior circulation syndromes: 22 (40%)
	  ○ Lacunar syndromes: 16 (29%)
	  ○ Posterior circulation syndromes: 6 (11%)
	• Relevant stroke lesion on brain computed tomography scan: 43 (78%)
	• Median time (IQR) between stroke and first assessment:
	  ○ Inpatient: 23 (10 to 53) days

<table>
<tr><th>Reference</th><th>Mead 2007[19]</th></tr>
<tr><td></td><td>o Community patients: 137 (93 to 217) days</td></tr>
<tr><td>Intervention</td><td>SF-36v2 vitality subscale (SF-36v2 vitality)
Fatigue subscale of the Profile of Mood States (POMS-fatigue)
Fatigue Assessment Scale (FAS)
General subscale of Multidimensional Fatigue Symptom Inventory (MFSI-general)
Brief Fatigue Inventory (BFI)</td></tr>
<tr><td>Comparison</td><td>Test-retest
Interrater
Each other</td></tr>
<tr><td>Length of follow-up</td><td>For test-retest reliability – 3 days later</td></tr>
<tr><td>Outcome measures</td><td>Face/content/construct validity
Discriminant/convergent validity

Test-retest reliability
Internal consistency (intratest reliability)
Inter-rater reliability</td></tr>
<tr><td>Source of funding</td><td>The study received funding from the Chief Scientist Office of the Scottish Executive Health Department (reference CZG/2/161).</td></tr>
<tr><td>Outcomes</td><td>Face/content/construct validity
Five tools were included as they were deemed to have adequate face validity according to the previously explained criteria.
After pilot testing, one question on the MFSI-general (“I feel pooped”) was poorly understood and was changed to “I feel exhausted”.

Discriminant/convergent validity
The construct validity for FAS and MFSI-general was higher than for SF36v2 vitality and MFSI-general. POMS-fatigue and FAS correlate better with MFSI-general, but correlate similarly to SF36v1 vitality when compared against each other.
<table>
<tr><th>Scale</th><th>POMS-fatigue</th><th>FAS</th><th>MFSI-general</th></tr>
<tr><td>SF-36v2 vitality</td><td>-0.58 (<0.001)
N=55</td><td>-0.41 (0.03)
N=52</td><td>-0.47 (<0.001)
N=55</td></tr>
<tr><td>POMS-fatigue</td><td></td><td>0.59 (<0.001)
N=52</td><td>0.75 (<0.001)
N=55</td></tr>
</table>
</td></tr>
</table>

Reference	Mead 2007[19]

FAS			0.71 (<0.001) N=52

Test-retest reliability

Scale	Item	Percentage agreement	Kappa (95% CI)
SF36v2 vitality	In the past 4 weeks, did you feel full of life?	43%	0.36 (0.07-0.63) N=51
	In the past 4 weeks, did you have a lot of energy?	59%	0.46 (0.14-0.77) N=51
	In the past 4 weeks, did you feel worn out?	43%	0.47 (0.25-0.70) N=51
	In the past 4 weeks, did you feel tired?	43%	0.43 (0.22-0.63) N=51

Scale	Item	Percentage agreement	Kappa (95% CI)
POMS fatigue	In the past week, do you feel tired?	65%	0.51 (0.29-0.73) N=51
	In the past week, do you feel fatigued?	57%	0.52 (0.32-0.72) N=51
	In the past week, do you feel worn out?	55%	0.59 (0.41-0.76) N=51
	In the past week, do you feel sluggish?	56%	0.61 (0.42-0.80) N=50
	In the past week, do you feel weary?	57%	0.52 (0.30-0.73) N=51
	In the past week, do you feel sleepy?	55%	0.45 (0.19-0.72) N=51

Scale	Item	Percentage agreement	Kappa (95% CI)
FAS	I am bothered by fatigue	37%	0.37 (0.17-0.56) N=51
	I get tired very quickly	47%	0.49 (0.27-0.71) N=51
	I don’t do much during the day	44%	0.50 (0.30-0.71) N=50
	I have enough energy for everyday life	46%	0.44 (0.21-0.66) N=50
	Physically I feel exhausted	51%	0.64 (0.48-0.81) N=51
	I have problems starting things	52%	0.53 (0.31-0.75) N=50
	I have problems thinking clearly	71%	0.77 (0.61-0.92) N=51
	I feel no desire to do anything	49%	0.33 (0.05-0.61) N=51
	Mentally I feel exhausted	51%	0.65 (0.46-0.84) N=51

Reference	Mead 2007[19]

	When I am doing something, I can concentrate quite well	47%	0.48 (0.25-0.17) N=51

Scale	Item	Percentage agreement	Kappa (95% CI)
MFSI general	In the past week, I feel exhausted	41%	0.58 (0.41-0.74) N=51
	In the past week, I feel run down	49%	0.64 (0.47-0.81) N=51
	In the past week, I feel worn out	45%	0.48 (0.27-0.69) N=51
	In the past week, I feel fatigued	39%	0.56 (0.37-0.74) N=51
	In the past week, I feel sluggish	60%	0.66 (0.47-0.84) N=50
	In the past week, I feel tired	59%	0.69 (0.53-0.85) N=51

Conclusion: POMS fatigue and MFSI general have higher consistent result in kappa across all question. However, FAS has the highest value in a select question ("I have problems thinking clearly"). Overall, the scales appear broadly comparable.

Internal consistency (intratest reliability)

Intraclass correlation coefficient (time 1 to time 2)

- SF-36v2 vitality = 0.51 (0.27 to 0.69)
- POMS fatigue = 0.74 (0.56 to 0.85)
- FAS = 0.77 (0.62 to 0.86)
- MFSI general = 0.76 (0.55 to 0.87)

Intraclass correlation coefficient (rater 1 to rater 2)

- SF-36v2 vitality = 0.92 (0.86 to 0.96)
- POMS fatigue = 0.84 (0.72 to 0.91)
- FAS = 0.88 (0.77 to 0.94)
- MFSI general = 0.88 (0.78 to 0.93)

Conclusion: SF-36v2 vitality has a lower intraclass correlation coefficient when comparing time periods than the other tests. However, it has a higher intraclass correlation coefficient when comparing between raters than the other tests.

Reference	Mead 2007[19]

Interrater reliability

Scale	Item	Percentage agreement	Kappa (95% CI)
SF36v2 vitality	In the past 4 weeks, did you feel full of life?	90%	0.89 (0.77-1.00) N=49
	In the past 4 weeks, did you have a lot of energy?	89%	0.87 (0.74-0.99) N=47
	In the past 4 weeks, did you feel worn out?	86%	0.72 (0.45-0.99) N=47
	In the past 4 weeks, did you feel tired?	87%	0.88 (0.75-1.00) N=48

Scale	Item	Percentage agreement	Kappa (95% CI)
POMS fatigue	In the past week, do you feel tired?	92%	0.89 (0.75-1.00) N=48
	In the past week, do you feel fatigued?	85%	0.82 (0.66-0.98) N=46
	In the past week, do you feel worn out?	87%	0.71 (0.45-0.97) N=45
	In the past week, do you feel sluggish?	90%	0.87 (0.74-1.00) N=48
	In the past week, do you feel weary?	91%	0.85 (0.68-1.00) N=46
	In the past week, do you feel sleepy?	88%	0.74 (0.46-1.00) N=48

Scale	Item	Percentage agreement	Kappa (95% CI)
FAS	I am bothered by fatigue	91%	0.78 (0.51-1.00) N=46
	I get tired very quickly	84%	0.67 (0.37-0.97) N=48
	I don’t do much during the day	89%	0.87 (0.75-1.00) N=43
	I have enough energy for everyday life	91%	0.82 (0.58-1.00) N=44
	Physically I feel exhausted	94%	0.98 (0.95-1.00) N=48
	I have problems starting things	87%	0.84 (0.68-1.00) N=47
	I have problems thinking clearly	90%	0.80 (0.61-0.99) N=48
	I feel no desire to do anything	89%	0.90 (0.75-1.00) N=44
	Mentally I feel exhausted	91%	0.91 (0.81-1.00) N=44
	When I am doing something, I can concentrate quite well	91%	0.95 (0.90-1.00) N=44

<table>
<tr><td>Reference</td><td>Mead 2007[19]</td></tr>
<tr><td></td><td>
<table>
<tr><th>Scale</th><th>Item</th><th>Percentage agreement</th><th>Kappa (95% CI)</th></tr>
<tr><td>MFSI general</td><td>In the past week, I feel exhausted</td><td>93%</td><td>0.94 (0.87-1.00) N=45</td></tr>
<tr><td></td><td>In the past week, I feel run down</td><td>87%</td><td>0.81 (0.60-1.00) N=46</td></tr>
<tr><td></td><td>In the past week, I feel worn out</td><td>89%</td><td>0.90 (0.81-1.00) N=44</td></tr>
<tr><td></td><td>In the past week, I feel fatigued</td><td>84%</td><td>0.87 (0.75-0.97) N=44</td></tr>
<tr><td></td><td>In the past week, I feel sluggish</td><td>89%</td><td>0.82 (0.63-1.00) N=45</td></tr>
<tr><td></td><td>In the past week, I feel tired</td><td>89%</td><td>0.92 (0.83-1.00) N=45</td></tr>
</table>
Conclusion: All scales had a high degree of interrater agreement, with the highest values being seen with MFSI general.
</td></tr>
<tr><td>Risk of bias assessment</td><td>Population
People after stroke. Mixture of people after acute or chronic stroke. Mixture of people with different locations of stroke and different genders. No information about initial stroke treatment, physical activity prior to stroke and severity.

Control population
Controlled against other people after stroke. No indirectness.

Recruitment/selection bias
Determining the scales to be included was decided by the investigators. However, independent researchers had to agree by a consensus discussion reducing the chance of bias. People were recruited from hospital stroke wards and the community via stroke clinics or community nurses. Limited information about differences in baseline values between the groups.

Sample size
Smaller sample size (55 participants).

Appropriateness of metrics
Metrics used were appropriate for the objective of the study.</td></tr>
<tr><td>Limitations</td><td>Author provided limitations:
Patients were not consecutive and may not be representative of stroke patients as a whole.
Smaller sample size. However, while a larger sample size would have given more precise estimates, a sample size of 50 is usually considered sufficient for studies of agreement.
The absence of any ‘gold standard’ for fatigue after stroke, means they could not assess criterion validity</td></tr>
</table>

Reference	Mead 2007[19]
	• When assessing test-retest reliability, the interviewer may have remembered the results of the first interview when performing the second, thereby artificially increasing apparent reliability • When assessing interrater reliability, audio recordings were used rather than repeat interviews, which may potentially increase apparent reliability. • Not all interviews could be analysed for interrater reliability because of poor quality of some recordings, mainly due to background noise on hospital wards.
Summary/author's conclusion	**Author's conclusion** The four scales are all usable. Their recommendation depends on the intended used: the fatigue assessment scale has higher face validity, was feasible in most patients and had the best test-retest reliability and high construct validity. However, the other three scales had higher internal consistency.

Reference	Mills 2012[20]
Study type Setting/Location	Cross sectional study (with initial qualitative component). United Kingdom, outpatient departments (the Department of Medicine for the Elderly at University Hospitals Aintree, Liverpool and the Neurology Rehabilitation Unit, Walton Centre for Neurology and Neurosurgery, Liverpool, UK).
Number of participants and characteristics	N=282 responders (717 non-responders). **Inclusion criteria:** People with radiologically confirmed stroke who attended an outpatient clinic. **Exclusion criteria:** Marked impairment of communication; had another neurological condition. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** • Mean age (SD): 67.3 (13.4) years • Male = 61.3% • Mean months post stroke: 17.2 (11.4) months • Previous stroke = 9.6% • Previous TIA = 11.6%

<table>
<tr><th>Reference</th><th>Mills 2012[20]</th></tr>
<tr><td></td><td>

Ischaemic stroke = 78.7%
Working = 16.5%
Median Stroke Impact Scale (range): 17 (0-64)
Very difficult or unable to:

Climb one flight of stairs = 25.9%
To dress top half of body = 10.1%
Control bladder = 8.3%
Transfer from bed to chair = 5.2%

</td></tr>
<tr><td>Intervention</td><td>Neurological Fatigue Index-Stroke (NFI-Stroke)

Fatigue severity scale (FSS)
Visual analogue scale (VAS)
Stroke Impact Scale (SIS)</td></tr>
<tr><td>Comparison</td><td>To each other, to itself at different times</td></tr>
<tr><td>Length of follow-up</td><td>2-4 weeks (only on the first 80 respondants to the main mailout)</td></tr>
<tr><td>Outcome measures</td><td>Criterion/Concurrent validity

Test-retest reliability
Internal consistency (person-item separation index)

Dimensions of fatigue considered</td></tr>
<tr><td>Source of funding</td><td>No additional information.</td></tr>
<tr><td>Outcomes</td><td>Criterion/Concurrent validity
Spearman correlation coefficients:
<table>
<tr><th></th><th>FSS</th><th>VAS</th><th>SIS</th></tr>
<tr><td>NFI-Stroke Physical</td><td>0.604</td><td>0.556</td><td>0.615</td></tr>
<tr><td>NFI-Stroke Cognitive</td><td>0.509</td><td>0.385</td><td>0.532</td></tr>
<tr><td>NFI-Stroke Summary</td><td>0.622</td><td>0.534</td><td>0.628</td></tr>
</table>
</td></tr>
</table>

Reference	Mills 2012[20]
	Test-retest reliability • Physical = 0.903 • Cognitive = 0.786 • Summary = 0.896 **Internal consistency (person-item separation index)** • NFI-Stroke Physical = 0.89 • NFI-Stroke Cognitive = 0.78 • NFI-Stroke Summary = 0.89 **Dimensions of fatigue considered** Physical and cognitive subscales
Risk of bias assessment	**Population** Radiologically confirmed stroke in the previous 50 minutes. The type of stroke was known from the register, but the Oxfordshire Community Stroke Project subtype was not available. Range of gender and type of stroke. No information on location of stroke, initial stroke treatment, physical activity prior to stroke and severity. **Control population** Compared to themself. No indirectness. **Recruitment/selection bias** Random cross-section of stroke patients identified from the Aintree Stroke Register held at the University Hospital Aintree, Liverpool, UK. **Sample size** Good sample size (N = 282). A reasonably large number of non-responders. Additionally, the number of people where test-retest reliability was tested was lower than the sample size and is an area of concern. **Appropriateness of metrics** Metrics appear appropriate for the measure.
Limitations	Author provided limitations:

Reference	Mills 2012[20]
	• The non-response level of the study was high and older patients with multiple strokes were underrepresented. People with low levels of disability were well represented. However, some respondents had very high SIS scores suggesting that those with higher disability were not wholly excluded. • There was one item in the cognitive scale (coordination gets worse) with a slightly high negative fit residual which indicated a degree of redundancy and accounted for the inflated overall item residual standard deviation. They did not discard the scale because all other fit statistics were acceptable and the retention of a comparative cognitive fatigue scale between stroke and MS was felt to be desirable.
Summary/author's conclusion	**Author's conclusion** The NFI-Stroke provides a brief (12 item) and easy-to-use tool for measurement of a clearly defined concept of fatigue. The scales satisfies strict Rasch model measurement requirements and, as a result, interval level scaling is available for when change scores need to be calculated. The scales have specific validation for stroke and can be used on patients of, amongst other factors, any age or sex. The scale may be useful in both a clinical setting and as an outcome measure in clinical trials. The NFI-Stroke is free for use by all state-funded health-care organisations and not -for-profit agencies.

Reference	Nadarajah 2017[21]
Study type Setting/Location	Cross-sectional study United States of America, teaching hospital outpatient setting
Number of participants and characteristics	N=100 (50 individuals with stroke, 50 non-stroke) **Inclusion criteria:** Stroke survivors: People with stroke who were older than 21 years; over three months post-stroke; able to read and communicate in English Healthy participants: Older than 21 years; no history of stroke, Parkinson's disease, heart failure, uncontrolled hypertension or cancer; able to comprehend in English. **Exclusion criteria:** Stroke survivors: People who were medically unstable; depressed (scores >10 in Patient Health Questionnaire 9); had difficulty in language comprehension or expression; scored <26 in Mini Mental State Exam. Healthy participants: Medically unstable; depressed; had difficulty in language comprehension or expression; scored <26 in MMSE. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise**

Reference	Nadarajah 2017[21]
	Stroke survivors • Mean age (SD): 63.6 (10.3) years • Female:Male = 16:34 • Duration post-stroke: 35.1 (50.0) months • Fugl-Meyer Motor Score = 68.8 (31.9) • MMSE = 29.0 (1.6) • PHQ-9 = 1.54 (2.3) • Barthel Index = 15.6 (5.7) • FSS = 29.2 (11.3) • VAS-F = 4.3 (1.7) • SF-36v2 vitality = 55.1 (14.4) **Healthy participants** • Mean age (SD): 61.1 (7.4) years • Female:Male = 28:22 • MMSE = 30.0 (1.3) • PHQ-9 = 0.8 (1.2) • Barthel Index = 20.0 (0) • FSS = 16.9 (9.5) • VAS-F = 2.6 (1.7) • SF-36v2 vitality = 89.5 (14.1)
Intervention	**Fatigue Severity Scale** Visual analogue scale – Fatigue (VAS-F) SF-36v2 Vitality
Comparison	Itself, other scales, different populations
Length of follow-up	Unclear (test-retest reliability was tested, but unclear how long after the test was completed)

<table>
<tr><th>Reference</th><th>Nadarajah 2017[21]</th></tr>
<tr><td>Outcome measures</td><td>Criterion/Concurrent validity (Spearman correlation coefficient)
Discriminant/convergent validity

Test-retest reliability (intraclass correlation coefficient)
Internal consistency (Cronbach’s alpha)</td></tr>
<tr><td>Source of funding</td><td>This study was supported by a research fund from the University of Malaya (PG107-2014A).</td></tr>
<tr><td>Outcomes</td><td>Criterion/Concurrent validity
<table>
<tr><th>Group</th><th>VAS-F</th><th>P value</th><th>SF36-v2 Vitality</th><th>P value</th></tr>
<tr><td>Stroke (N=50)</td><td>0.68 (0.50-0.81)</td><td><0.01</td><td>-0.32 (-0.55 to -0.05)</td><td>0.02</td></tr>
<tr><td>Healthy (N=50)</td><td>0.66 (0.47-0.79)</td><td><0.01</td><td>-0.01 (-0.29 to 0.27)</td><td>0.97</td></tr>
</table>
Discriminant/convergent validity
<table>
<tr><th>Item</th><th>Stroke (N=50) (mean [SD])</th><th>Healthy (N=50) (mean [SD])</th></tr>
<tr><td>1</td><td>3.62 (1.54)</td><td>2.42 (1.51)</td></tr>
<tr><td>2</td><td>3.70 (1.53)</td><td>1.98 (1.12)</td></tr>
<tr><td>3</td><td>3.60 (1.58)</td><td>1.78 (0.99)</td></tr>
<tr><td>4</td><td>3.52 (1.53)</td><td>2.04 (1.24)</td></tr>
<tr><td>5</td><td>3.28 (1.47)</td><td>1.74 (1.10)</td></tr>
<tr><td>6</td><td>3.04 (1.51)</td><td>1.74 (1.21)</td></tr>
<tr><td>7</td><td>2.76 (1.62)</td><td>1.80 (1.32)</td></tr>
<tr><td>8</td><td>2.70 (1.53)</td><td>1.76 (1.27)</td></tr>
<tr><td>9</td><td>2.72 (1.53)</td><td>1.68 (1.19)</td></tr>
<tr><td>Total</td><td>29.2 (11.3)</td><td>16.9 (9.5) (p-value comparing between groups = <0.01)</td></tr>
</table>
Conclusion: Answers appear to be different between the groups indicating that there may be sufficient discriminant validity.

Test-retest reliability
Intraclass correlation coefficient
• Stroke survivors = 0.93 (0.88 to 0.96)</td></tr>
</table>

Reference	Nadarajah 2017[21]
	• Healthy participants = 0.90 (0.84 to 0.94) **Internal consistency** Cronbach's alpha • Stroke survivors = 0.93 (0.89-0.96) • Healthy participants = 0.96 (0.94-0.98)
Risk of bias assessment	**Population** Stroke survivors. People at least 3 months after stroke, but the majority were significantly far away from 3 months after stroke. Mixture of people of different gender. No information provided on: type of stroke, initial stroke treatment, physical activity prior to stroke and severity. **Control population** Healthy participants. Appropriate comparison. No indirectness noted. **Recruitment/selection bias** Recruitment of people from a teaching hospital where they received outpatient stroke rehabilitation. Unclear what the location of this hospital is. **Sample size** Sample size (N=100). 50 participants in each study arm. **Appropriateness of metrics** Metrics used are appropriate.
Limitations	Author provided limitations: • Patient characteristics could have biased the results (the stroke group had a greater number of males than females than the control group) • Did not examine the responsiveness of FSS cover a longer period of time or after interventions • Found a weak correlation between FSS and SF36-v2 vitality, which could be resulted from the inconsistency in modes of administration
Summary/author's conclusion	**Author's conclusion**

Reference	Nadarajah 2017[21]
	The Fatigue Severity Scale is a reliable and valid tool to measure post-stroke fatigue. It is also sensitive to distinguish fatigue in individuals with or without stroke. Findings suggest that Fatigue Severity Scale is readily to be used to measure post-stroke fatigue for both clinical and research purposes.

Reference	Ng 2022[23]
Study type Setting/Location	Cross-sectional study China, a convenience sample of people from a university-affiliated neurorehabilitation laboratory gathered through a poster advertisement.
Number of participants and characteristics	N=101 people with stroke and 50 healthy older people. **Inclusion criteria:** Chronic stroke with over 12 months post stroke with age 50 or older; cognitively intact with score in the Abbreviated Mental Test at least 7; capable of walking independently or with assistance or walking aids for 10m; understand Cantonese. **Exclusion criteria:** Unstable medical conditions such as coronary disease, comorbid psychiatric diseases, neurological problems (e.g. multiple sclerosis and Parkinson's disease); musculoskeletal problems (e.g. painful knee osteoarthritis) that might impede the assessment procedures. The same inclusion and exclusion criteria were used for the recruitment of healthy older people except for having a history of stroke. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** **Stroke survivors** • Mean age (SD): 63.82 (6.40) years • Male/female: 58:43 (57.4%:42.6%) • BMI: 24.13 (3.02) kg/m^2 • Education level – primary or below/secondary/college or above: 23:64:14 (22.8%:63.4%:13.8%) • Marital status – single/married/divorced or separated/widowed or widowered: 9:76:9:7 (8.9%:75.2%:8.9%:6.9%) • Living arrangement – alone/with others: 10:91 (9.9%:90.1%) • Ischaemic/haemorrhagic: 69:32 (68.3%:31.7%)

Reference	Ng 2022[23]
	• Years since stroke: 6.74 (4.42) years • Side of stroke – left/right: 46:55 (45.5%:54.5%) • Fugl Meyer Assessment Lower extremity: 26.12 (4.46) • Timed Up and Go: 17.64 (14.23) • Community Integration Measure: 40.38 (7.05) • Physical Component Summary: 39.25 (9.09) • Mental Component Summary: 48.62 (10.64) **Healthy participants** • Mean age (SD): 61.78 (7.41) years • Male/female: 15:35 (30%:70%) • BMI: 22.43 (3.19) kg/m^2 • Education level – primary or below/secondary/college or above: 9:33:8 (18%:66%:16%) • Marital status – single/married/divorced or separated/widowed or widowered: 4:45:1:0 (8%:90%:2%) • Living arrangement – alone/with others: 4:46 (8%:92%)
Intervention	**Modified Fatigue Impact Scale (MFIS) – Cantonese version**
Comparison	Between different scales (compared with the Fugl Meyer Assessment – Lower Extremity, Timed Up and Go completion time, Community Integration Measure, and 12-item Short Form Health Survey version 2 Physical and Mental Component scores)
Length of follow-up	1 week (52 of the 101 stroke participants were randomly selected for retesting)
Outcome measures	Face/content/construct validity Discriminant/convergent validity Test-retest reliability Internal consistency Responsiveness to change Dimensions of fatigue considered
Source of funding	Supported by Departmental Research Grant (ref P0013897) from Department of Rehabilitation Sciences, the Hong Kong Polytechnic University.

Reference	Ng 2022[23]
Outcomes	**Face/content/construct validity** Content validity: All the I-CVI, S-CVI/Ave and S-CVI/UA values are 1, suggesting that the validity of each individual item and the overall MFIS-C are satisfactory. **Discriminant/convergent validity** All MFIS-C subscale scores have no significant correlations with the FMA-LE score and TUG completion time, but significant weak to moderate negative correlations with the CIM-C, and SF-12 PCS and MCS scores. **Test-retest reliability** Intraclass correlation coefficient (95% confidence interval) overall MFIS-C = 0.84 (0.74-0.91) Intraclass correlation coefficient (95% confidence interval) cognitive subscale of MFIS-C = 0.83 (0.72-0.90) Intraclass correlation coefficient (95% confidence interval) physical/social subscale of MFIS-C = 0.81 (0.70-0.89) **Internal consistency** Cronbach's alpha for overall MFIS-C = 0.92 Cronbach's alpha for the cognitive subscale of MFIS-C = 0.85 Cronbach's alpha for the physical/social subscale of MFIS-C = 0.89 **Responsiveness to change** Minimally detectable change$_{95}$ (standard error) and minimally detectable change$_{95\%}$ overall score = 14.86 (5.38) and 38.3% Minimally detectable change$_{95}$ (standard error) and minimally detectable change$_{95\%}$ cognitive subscale = 7.49 (2.71) and 44.8% Minimally detectable change$_{95}$ (standard error) and minimally detectable change$_{95\%}$ physical and psychosocial subscale = 9.70 (3.51) and 44.0% **Dimensions of fatigue considered** Cognitive, physical and psychosocial.
Risk of bias assessment	**Population** Population well reported. Did not report much information about subgroups of the population. But overall appears to have a mixture of different people. **Control population** Healthy participants. Appears to be an appropriate control population to use.

Reference	Ng 2022[23]
	Recruitment/selection bias Recruitment from only one center. Convenience sampling. However, uses random sampling to choose the people for test-retest reliability. **Sample size** Good sample size in general. Only 52 of the 101 stroke participants were randomly selected for re-assessment. **Appropriateness of metrics** Metrics used appear appropriate.
Limitations	Author provided limitations: • The translation of the questionnaire may be appropriate to people in the Hong Kong Environment but not to other Chinese populations such as those in Mainland China due to differences in culture. • The sample size calculation is based on reliability, and so may not be sufficient to detect the significant correlations between MFIS-C scores and other outcome measure scores • The findings could only be generalised to those fulfilling the inclusion and exclusion criteria • The expert panel did not consist of a member with occupational training background. • The stroke participants were not checked for a history of chronic fatigue syndrome before entering the trial. • They did not examine the construct validity of the MFIS-C due to the insufficient number of subjects.
Summary/author's conclusion	**Author's conclusion** The study provides evidence that MFIS is a valid and reliable measure to assess and monitor fatigue in both clinical and research settings.

Reference	Ozyemisci-Taskiran 2019[24]
Study type Setting/Location	Cross-sectional study Turkey, inpatients admitted for rehabilitation to Physical Medicine and Rehabilitation Department in Gazi University School of Medicine.
Number of participants and characteristics	N=46 stroke survivors (108 assessed for eligibility, 34 excluded for aphasia and other cognitive deficits that interfered with answering questions, 7 excluded for recent medication change that may affect fatigue, 5 excluded for coexisting serious medical conditions, 4 for dementia, 3 for psychiatric disorders, 3 for premature discharge, 1 for inability to understand and speak Turkish, 5 excluded due to refusal to participant). 52 control subjects were recruited (an orthopedic control group, 16 with low back pain, 14 with knee

Reference	Ozyemisci-Taskiran 2019[24]
	osteoarthritis, 6 with meniscal degeneration, 9 with strain, 7 with sprains). The control groups age and gender were similar to the stroke subjects. **Inclusion criteria:** People with stroke who were over 18 years; medically stable at least 4 weeks after stroke onset; able to understand and speak Turkish; able to complete the questionnaire **Exclusion criteria:** Coexisting illnesses that might cause fatigue, such as cancer, thyroid disease, rheumatological disease, other neurological disorders; medication changes in the previous two weeks that may affect perception of fatigue; premature discharge; severe cognitive or communication deficits. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** **Subjects with stroke** • Mean age (SD): 57.9 (13.3) years • Male: 21 • HADS anxiety (SD): 7.3 (3.9) • HADS depression (SD): 7.3 (3.5) • VAS-F (SD): 44.4 (26.7) • SF-36 vitality (SD): 55.9 (22.5) • Mean duration after stroke (SD): 10.4 (16.7) months **Control subjects** • Mean age (SD): 53.0 (15.2) years • Male: 20 • HADS anxiety (SD): 6.4 (4.2) • HADS depression (SD): 6.2 (4.5) • VAS-F (SD): 47.4 (24.5) • SF-36 vitality (SD): 58.1 (20.5)
Intervention	**Fatigue Severity Scale – Turkish versions (FSS-T)**

Reference	Ozyemisci-Taskiran 2019[24]
	Hospital Anxiety and Depression Scale – Turkish version (HADS-T) Visual Analogue Scale – Fatigue (VAS-F) SF-36 vitality subscale (SF-36v)
Comparison	Itself, other scales, different populations
Length of follow-up	7 days
Outcome measures	Face/content/construct validity Criterion/Concurrent validity Discriminant/convergent validity Test-retest reliability Internal consistency (Cronbach's alpha)
Source of funding	No additional information
Outcomes	**Face/content/construct validity** Feasibility – Among 46 subjects, 93.5% responded to all items of the FSS. The remaining two subjects had left the 6th item incomplete, and one subject did not respond to the 8th item. Among the 52 control subjects, all completed each item, apart from one who missed the 7th item. Content validity – Among people with stroke, 89% found the questions understandable, and 72% viewed the scale as representative of fatigue. **Criterion/Concurrent validity** There is no gold standard to diagnose fatigue, hence criterion validity was not assessed with another measure. When FSS-T scores of stroke and control groups were evaluated, it was observed that FSS did not distinguish between two groups. Mean FSS scores and subjects with scores greater than 4 were similar in both groups. **Discriminant/convergent validity** FSS test 1: • Stroke survivors = 4.2 (1.7) • Control = 4.1 (1.4) • P value = 0.717

Reference	Ozyemisci-Taskiran 2019[24]

FSS test 2:

- Stroke survivors = 4.2 (1.7)
- Control = 3.7 (1.6)
- P value = 0.233

For people after stroke:
FSS-T correlated moderately with SF-36v: r = -0.531, p = 0.002.
No correlation between FSS-T and VAS-F: r = 0.197, p = 0.281.
Weak correlation between FSS-T and HADS-anxiety subscale: r = 0.310, p = 0.041.
Weak correlation between FSS-T and HADS-depression subscale: r = 0.334, p = 0.027.

For people after control group:
FSS-T correlated moderately with SF-36v: r = -0.485, p = 0.002.
No correlation between FSS-T and VAS-F: r = 0.236, p = 0.154.
No correlation between FSS-T and HADS-anxiety subscale: r = 0.050, p = 0.729.
No correlation between FSS-T and HADS-depression subscale: r = -0.034, p = 0.813.

Test-retest reliability

Item	ICC (95% CI)	P-value
1. My motivation is lower when I am fatigued	0.652 (0.328-0.820)	0.001
2. Exercise brings on my fatigue	0.573 (0.183-0.776)	0.005
3. I am easily fatigued	0.710 (0.447-0.847)	<0.001
4. Fatigue interferes with my physical functioning	0.575 (0.189-0.776)	0.005
5. Fatigue causes frequent problems for me	0.766 (0.556-0.877)	<0.001
6. My fatigue prevents sustained physical functioning	0.284 (-0.402-0.631)	0.162
7. Fatigue interferes with carrying out certain duties and responsibilities	0.741 (0.504-0.865)	<0.001
8. Fatigue is among my three most disabling symptoms	0.694 (0.415-0.839)	<0.001
9. Fatigue interferes with my work, family, or social life	0.725 (0.477-0.855)	<0.001
All items	0.742 (0.512-0.863)	<0.001

<table>
<tr><th>Reference</th><th>Ozyemisci-Taskiran 2019[24]</th></tr>
<tr><td></td><td>
Internal consistency
<table>
<tr><th>Item</th><th>Cronbach’s alpha if item deleted</th></tr>
<tr><td>1. My motivation is lower when I am fatigued</td><td>0.927</td></tr>
<tr><td>2. Exercise brings on my fatigue</td><td>0.917</td></tr>
<tr><td>3. I am easily fatigued</td><td>0.925</td></tr>
<tr><td>4. Fatigue interferes with my physical functioning</td><td>0.920</td></tr>
<tr><td>5. Fatigue causes frequent problems for me</td><td>0.921</td></tr>
<tr><td>6. My fatigue prevents sustained physical functioning</td><td>0.917</td></tr>
<tr><td>7. Fatigue interferes with carrying out certain duties and responsibilities</td><td>0.914</td></tr>
<tr><td>8. Fatigue is among my three most disabling symptoms</td><td>0.915</td></tr>
<tr><td>9. Fatigue interferes with my work, family, or social life</td><td>0.916</td></tr>
</table>
Cronbach’s alpha total (stroke survivor): 0.928
Cronbach’s alpha total (control): 0.874
</td></tr>
<tr><td>Risk of bias assessment</td><td>
Population
Stroke survivors. At least 4 weeks after stroke onset, but majority appear to be a long time after 4 weeks. Missing information on type of stroke, initial stroke treatment, physical activity prior to stroke and severity.

Control population
People with a range of orthopaedic conditions. People with these conditions could experience fatigue, which may explain why there is limited discriminant/convergent validity.

Recruitment/selection bias
People recruited from those admitted for rehabilitation. These people may experience more fatigue than the general population (if they are requiring additional support for rehabilitation). But equally this seems like an appropriate population otherwise. Baseline differences between groups are minimal.

Sample size
Sample size (N=98). <50 people after stroke. 3 did not consent to participate in the second test and three were prematurely discharged at their request, which means those people may not be included in the retest results.
</td></tr>
</table>

Reference	Ozyemisci-Taskiran 2019[24]
	Appropriateness of metrics Metrics used were appropriate.
Limitations	Author provided limitations: • Small sample size, which reduced the generalizability of the results to all Turkish stroke survivors • Control group was recruited among the individuals admitted to the outpatient clinic, which may explain the presence of fatigue being higher than the general population
Summary/author's conclusion	**Author's conclusion** The FSS-T is a valid and reliable scale to measure fatigue in stroke. The FSS-T is not sensitive to differentiate fatigue in stroke from the control subjects with orthopedic problems with similar age and gender. Reliability of items related to physical functioning needs further investigation to clarify the subject's perception about fatigue and physical impairments.

Reference	Saneii 2020[25]
Study type Setting/Location	Cross-sectional methodological study Iran, occupational therapy clinics between 8.00am and 1.00pm in a quiet room.
Number of participants and characteristics	N=280 (140 stroke survivors, 140 healthy adults) **Inclusion criteria:** Age range 45-70 years; ability to read and write; adequate cooperation; an MMSE score of greater than 21; Persian as their native language **Exclusion criteria:** History of substance abuse; comorbidity with psychiatric, orthopaedic, and neurological disorders (any lesions and anomalies in the central nervous system); sleep deprivation; chronic fatigue syndrome and other similar diseases; use of antifatigue medication. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** **Stroke survivors** • Male:Female = 71:69 • Mean age (SD): 58.85 (7.88) years

Reference	Saneii 2020[25]
	• Marital status (married:single) = 39:101 • Educational status (class) (at least 12:less than 12): 124:16 • Hemiplegia side (right:left): 91:49 • Mean time since stroke (SD): 20.27 (15.82) months **Healthy adults** • Male:Female = 77:63 • Mean age (SD): 58.16 (10.51) years • Marital status (married:single) = 40:100 • Educational status (class) (at least 12:less than 12): 108:32
Intervention	**Fatigue Impact Scale – Persian version (FIS-P)** Fatigue severity scale – Persian version (FSS-P) SF-36 questionnaire – Persian version (SF-36-P)
Comparison	Different scales, itself, different populations
Length of follow-up	1 week
Outcome measures	Face/content/construct validity Discriminant/convergent validity Test-retest reliability Internal consistency (Cronbach's alpha) Interrater reliability Responsiveness to change Dimensions of fatigue considered
Source of funding	No additional information.
Outcomes	**Face/content/construct validity** Face validity – "The relevance, suitability, clarity and simplicity of all questions were acceptable." Content validity: Ranged from 0.6-1, average of 0.85, universal agreement was 0.48. This was considered acceptable by the study (an agreement above 0.42 was considered acceptable based on Lawshe's method).

Reference	Saneii 2020[25]

Floor and ceiling effects – 2.1%, considered acceptable.

Discriminant/convergent validity

Convergent validity:

Significant negative correlations between FIS-P and all SF-36 subscales, with the only exception found for the cognitive subscale of FIS-P and the emotional domain of SF-36. The FIS-P and FSS scales had a significant positive correlation.

Scale/Subscale	FIS-P cognitive subscale	FIS-P physical subscale	FIS-P social subscale	FIS-P total
SF-36				
Physical function	-0.525	-0.480	-0.549	-0.546
Role physical	-0.249	-0.175	-0.228	-0.229
Role emotional	NS	-0.174	-0.188	-0.184
Vitality	-0.465	-0.455	-0.517	-0.508
Mental health	-0.332	-0.342	-0.384	-0.375
Social function	-0.400	-0.412	-0.423	-0.431
Physical pain	-0.420	-0.452	-0.466	-0.469
General health	-0.452	-0.463	-0.501	-0.498
FSS	0.697	0.731	0.771	0.772

Discriminant validity:

Significantly different results between stroke patients and healthy adults.

FIS-P	Stroke Patients (N=140) (mean [SD])	Healthy adults (N=140) (mean [SD])	Independent-Sampling T test (t [p value])
Total score	73.22 (34.19)	47.96 (33.39)	256.6 (<0.001)
Physical	22.75 (8.95)	14.81 (9.64)	7.14 (<0.001)
Cognitive	14.87 (8.97)	10.13 (8.64)	4.50 (<0.001)
Social	35.59 (17.63)	23.01 (16.91)	6.09 (<0.001)

Test-retest reliability

Intra-class correlation coefficient:

Reference	Saneii 2020[25]
	• FIS-P Total = 0.991 (0.983-0.995), p=<0.001 • FIS-P Physical = 0.961 (0.979-0.926), p=<0.001 • FIS-P Cognitive = 0.987 (0.993-0.976), p=<0.001 • FIS-P Social = 0.987 (0.976-0.993), p=<0.001 **Internal consistency** Cronbach's alpha: • FIS-P Total = 0.895 • FIS-P Physical = 0.87 • FIS-P Cognitive = 0.90 • FIS-P Social = 0.95 **Inter-rater reliability** Intra-class correlation coefficient: • FIS-P Total = 0.984 (0.848-0.848), p=0.001 (note: the confidence interval provided by the study is reported here, this is not a valid confidence interval) • FIS-P Physical = 0.911 (0.142-0.991), p=0.019 • FIS-P Cognitive = 0.987 (0.879-0.999), p=<0.001 • FIS-P Social = 0.987 (0.876-0.999), p=<0.001 **Responsiveness to change** Minimum detectable change • FIS-P Total = 8.26 • FIS-P Physical = 3.87 • FIS-P Cognitive = 2.76 • FIS-P Social = 5.27 **Dimensions of fatigue considered** Cognitive impact, physical impact, social impact.
Risk of bias assessment	**Population**

Reference	Saneii 2020[25]
	Stroke survivors. Chronic time horizon. Mixture of different genders. No information about the type of stroke, initial stroke treatment, physical activity prior to stroke and severity. **Control population** Healthy adults. No evidence of indirectness. **Recruitment/selection bias** People recruited from an occupational therapy clinic. May not capture a complete population. Unclear where the healthy adults were recruited. **Sample size** Larger sample size (N=280, with 140 people in each study arm). **Appropriateness of metrics** Metrics provided were appropriate
Limitations	Author provided limitations: • The use of convenience sampling method • The lack of a previous Persian version of a similar instrument
Summary/author's conclusion	**Author's conclusion** • Face validity: The FIS-P is suitable, easy to understand and unambiguous. • Content validity: The FIS was originally designed to be completed by MS patients. The results indicate that it can be used to assess the impact of fatigue in stroke victims' lives. • Convergent validity: The FSS-P and FIS-P were strongly positively correlated, meaning that both assessed a common concept. The relationship to SF-36 was low to moderate inverse relationships. These two tools are somewhat structurally different. • Test-retest reliability: Reasonable ICC levels, has satisfactory test-retest reliability. • Inter-rater reliability: Inter-rater reliability can be really important for this population due to people potentially needing support to complete the questionnaire. The ICC results showed strong agreement between two raters. • Internal consistency: The Cronbach's coefficient alpha values were high. The result is consistent to the Hungarian, Turkish and French versions. Therefore, it can be concluded that the tool is consistent with the original tool and can be used in stroke patients.

Reference	Smith 2008 [29]
Study type Setting/Location	Cross-sectional study The Netherlands, a stroke rehabilitation unit of a nursing home “De Hazelaar”, Tilburg. People with congestive heart failure were recruited from outpatient clinics and healthy controls were recruited from the general Dutch population (method unclear).
Number of participants and characteristics	N=377 (80 after stroke, 137 with end-stage congestive heart failure, 160 healthy controls from the general Dutch population without a history of stroke or cardiovascular disease). **Inclusion criteria:** People after a stroke, with congestive heart failure or healthy people. **Exclusion criteria:** Stroke: People with reduced level of consciousness, severe language deficits, multiple cognitive deficits reflecting dementia syndrome, or severe emotional problems; people with other life-threatening diseases. Heart failure: People with diastolic heart failure; aged 80 and older; myocardial infarction in the month before inclusion; other life threatening diseases; history of care. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** **People after stroke** • Mean age (SD): 74.1 (6.6) years • Male = 44 (55%) • Low educational level (primary school) = 38 (47.5%) • Right hemisphere stroke = 41 (51.3%) • Brainstem lesion = 19 (23.7%) • Subcortical lesion = 26 (32.5%) • Cortical lesion = 22 (27.5%) • No information on stroke location = 13 (16.3%) • Average physical dimension of Stroke-Adapted Sickness Impact Profile (SD) = 72.8 (31.5%) on body care and movement subscale, 77.9 (26.0%) on the mobility subscale, 82.1 (29.0%) on the ambulation subscale, 36.3 (30.6%) on the alertness behaviour subscale • Mean time between stroke and study inclusion (SD) = 7.6 (5.4) months

Reference	Smith 2008 [29]
	Congestive heart failure • Mean age (SD) = 67.6 (8.8) years • Male = 98 (71.5%) • Low educational level = 48 (35%) • Mean left ventricular ejection fraction (SD) = 28.9 (7.0%) **Healthy controls** • Mean age (SD) = 69.3 (6.0) years • Male = 74 (46.3%) • Low educational level = 34 (21.3%)
Intervention	**Fatigue Assessment Scale (FAS)** Beck Depression Inventory (BDI) Stroke-Adapted Sickness Impact Profile (SA-SIP30)
Comparison	Itself, different time periods, other scales
Length of follow-up	2 months (only 80 people did the test twice)
Outcome measures	Face/content/construct validity Convergent/discriminant validity Test-retest reliability Internal consistency (Cronbach's alpha)
Source of funding	This research was supported by a VICI grant (453-04-004) from the Netherlands Organisation for Scientific Research, The Hague, the Netherlands, and by a grant from the Dutch Heart Foundation (2003B038) to Johan Denollet.
Outcomes	**Face/content/construct validity** Principle component analysis with oblimin rotation revealed two factors. Two of the FAS items had higher values in the BDI, while the BDI item on fatigue had a higher score on the FAS. Overall, this indicates that the two scales measure different entities. However, the FAS may not be a unidimensional construct in people with stroke, as four of the 10 items had low component loadings (FAS-3 and FAS-7) or loaded on the BDI component (FAS-8 and FAS-6). Therefore, more research on the content validity is required. The association between the FAS and BDI was 0.44 ($p < 0.001$)

Reference	Smith 2008 [29]
	Convergent/discriminant validity Mean FAS scores by population groups: • People with stroke = 15.3 (7.6) • People with congestive heart failure = 16.5 (7.9) (when compared to people with stroke, p 0.44) • Healthy controls = 9.2 (5.6) (when compared to people with stroke, p <0.001). Therefore, fatigue scores do show a significant difference from healthy controls, but cannot distinguish between those with stroke and those with congestive heart failure. **Test-retest reliability** Intra-class correlation coefficient = 0.81 **Internal consistency (Cronbach's alpha)** Cronbach's alpha = 0.77
Risk of bias assessment	**Population** People after stroke. However, this reports the results for the population and control population together, which introduces potential indirectness to some results. Reports a range of genders and location of stroke. Does not report the type of stroke, initial stroke treatment, physical activity prior to stroke and severity. **Control population** People with congestive heart failure and healthy comparisons. People with congestive heart failure may also experience fatigue, but as the healthy comparison group is present this is not a source of bias. **Recruitment/selection bias** Baseline characteristics between groups were not balanced (partly by design of excluding older people with heart failure). This could have an effect on the fatigue levels experienced by a person. **Sample size** Sample size (n=377). However, only 80 participants had a stroke. **Appropriateness of metrics**

Reference	Smith 2008 [29]
	Metrics used were appropriate.
Limitations	Author provided limitations: • The relatively small sample size • The generalisability to the stroke population may be limited due to the exclusion of people suffering, among other things, from a reduced level of consciousness
Summary/author's conclusion	**Author's conclusion** Level of fatigue in people with stroke are equal to the levels of fatigue in people with congestive heart failure, who experience fatigue as one of their main complaints. Moreover, stroke and congestive heart failure each had a large effect on fatigue, emphasising its clinical important in people with stroke. The FAS is an adequate measure of fatigue in people with stroke, but the content validity of the FAS in these patients needs to be examined in future studies.

Reference	Taasen 2020[31]
Study type Setting/Location	Cross-sectional study Norway, stroke organisations, municipal healthcare and a hospital. The stroke organisation advertised information about the project on their websites, Facebook account and sent e-mails to their members. Physiotherapists at the mentioned institutions informed possible candidates about the project
Number of participants and characteristics	N=63 (82 eligible, 9 >4 on clock test [this may mean <4 as the study excluded people below this value, and included people above this 4], 5 no reason for leaving study, 2 aphasia). 3 did not participate in the re-test. The project is a sub-study of the Physical Activity after stroke capacity, activity, and life quality study (PASCAL, NCT00311025). **Inclusion criteria:** Age at least 18 years; score at least 4 at the Clock-Drawing Test; able to speak and write Norwegian. **Exclusion criteria:** People who could not communicate or who were cognitively reduced. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** • Mean age (SD): 60.25 (14.69) years

Reference	Taasen 2020[31]

	• Women (%) = 36 (54.5%)
	• >12 months post stroke (%) = 42 (63.6%)
	• <3 months post stroke (%) = 24 (36.4%)
	• More than one stroke (%) = 7 (10.5%)
	• Independent in activities of daily living:
	  ○ Dressing (%) = 63 (95.45%)
	  ○ Feeding (%) = 64 (96.96%)
	  ○ Mobility indoors (%) = 64 (96.96%)
	  ○ Stairs = 64 (96.96%)
	  ○ Toilet use = 65 (98.45%)
	• Medications, median (IQR): 2 (2)
	• No medication (%) = 6 (9.1%)
	• Mean multimorbidities (SD) = 1.90 (1.83)

Intervention	**Neurological Fatigue Index-Stroke Norwegian version (N-NFI-Stroke)**
Comparison	Itself, different times, between subacute and chronic stroke populations
Length of follow-up	Within 2 days up to a week apart (face-to-face, by phone or through written correspondences).
Outcome measures	Face/content/construct validity Test-retest reliability Internal consistency (Cronbach's alpha) Dimensions of fatigue considered
Source of funding	No additional information
Outcomes	**Face/content/construct validity** Floor and ceiling effects – There was a normal distribution of results with the average score being 17.88. Answered varied between totally disagree (1-17%), disagree (19-49%), agree (34-55%), totally agree (7-31%). **Test-retest reliability** Weighted Kappa (95% confidence intervals)

Reference | **Taasen 2020[31]**

Item	Total sample (n=63)	Chronic stroke (n=39)	Subacute stroke (n=24)
1. I get tired fast	0.59 (0.44-0.74)	0.61 (0.40-0.83)	0.35 (0.09-0.61)
2. Sometimes the body becomes weak	0.59 (0.41-0.77)	0.74 (0.58-0.89)	0.32 (-0.02-0.66)
3. My arms and legs can feel very heavy	0.70 (0.57-0.84)	0.78 (0.62-0.94)	0.54 (0.27-0.82)
4. My body cannot keep up with what I want to do	0.60 (0.43-0.76)	0.65 (0.49-0.81)	0.31 (-0.08-0.70)
5. The longer I am doing something, the more difficult it becomes	0.67 (0.50-0.83)	0.82 (0.65-0.98)	0.12 (-0.21-0.47)
6. Sometimes I cannot accomplish what I am doing and have to give up	0.78 (0.67-0.89)	0.88 (0.79-0.96)	0.49 (0.18-0.80)
7. I am tired almost everyday	0.70 (0.67-0.89)	0.76 (0.65-0.88)	0.40 (0.09-0.70)
8. I can become exhausted, even though I have not done anything	0.75 (0.64-0.87)	0.82 (0.70-0.93)	0.46 (0.15-0.77)
9. Sometimes I have to really concentrate on the tasks that usually are undemanding	0.55 (0.40-0.71)	0.68 (0.52-0.84)	0.02 (-0.34-0.38)
10. I have trouble to talk when I am tired	0.72 (0.54-0.89)	0.84 (0.72-0.96)	0.47 (0.13-0.82)
11. My coordination capacity deteriorates as the day is passing by	0.56 (0.36-0.77)	0.68 (0.47-0.90)	0.25 (-0.07-0.57)
12. Mental effort makes me totally exhausted	0.71 (0.58-0.85)	0.91 (0.82-0.99)	0.15 (-0.16-0.46)

Internal consistency

Cronbach's alpha

- Total = 0.90 (corrected correlation lowest and highest values: 0.50-0.78)
- Physical items = 0.89 (corrected correlation lowest and highest values: 0.55-0.79)
- Cognitive items = 0.74 (corrected correlation lowest and highest values: 0.46-0.60)

Dimensions of fatigue considered

Physical (first 8 questions) and cognitive (last 4 questions) aspects.

Risk of bias assessment

Population

Stroke survivors. Mixture of subacute and chronic population. Mixture of genders. No information on type of stroke, initial stroke treatment, physical activity prior to stroke and severity.

Reference	Taasen 2020[31]
	Control population No control population. **Recruitment/selection bias** Recruited from a range of different sources. No information about how many people came from each source, but seems like an appropriate method. **Sample size** Small sample size (N=63). **Appropriateness of metrics** Metrics used were appropriate.
Limitations	Author provided limitations: • The sample was of an adequate size according to the requirements for analysis of categorical data, but a larger sample would have been preferable to generalise results. • People with chronic stroke were predominant, with few people with subacute stroke. • The participants were slightly younger than the general stroke population and to a high degree independent with activities of daily living, which may limit generalisation of the results. • The test procedures were performed under different circumstances, which could affect how answers were given between participants. • The study used classical test theory rather than modern psychometric techniques. • It may have added to the results if invariance in item difficulty between language versions was evaluated.
Summary/author's conclusion	**Author's conclusion** N-NFI-Stroke is the first specific questionnaire for outcome measurements relating to the symptom of fatigue in post-stroke survivors. It is a valid and reliable measurement instrument that can be administered rapidly and is easily comprehended. It can play an important role in research, clinical practice and health assessment.

Reference	Tseng 2010[32]
Study type Setting/Location	Cross sectional study.

Reference	Tseng 2010[32]
	United States of America, people recruited from local stroke support groups and the ASTRA (Advancing Stroke Treatment through Research Alliances) participant database.
Number of participants and characteristics	N= 21 (28 people from local support groups and 72 from ASTRA contacted, 21 agreed to participate). Convenience sample. **Inclusion criteria:** Diagnosis of stroke at least 6 months and no more than 5 years ago; the ability to perform the exercise movement on a total-body recumbent stepper; receive medical clearance from their primary care physician to confirm they are medically stable and able to participate in exercise; score <2 on a dementia screening tool, the AD8. **Exclusion criteria:** Hospitalisation for myocardial infarction, heart surgery or congestive heart failure during the preceding 3 months; recent symptoms of chest discomfort; resting blood pressure of 160/100 or greater; currently using a pacemaker; currently smoking or significant pulmonary pathology; alcoholism or alcohol dependency; recreational drug use; medication change within the duration of the study (e.g. antidepressants, cardiac medications). **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** • Mean age (SD): 59.5 (10.3) years • Men:Women = 12:9 • Stroke lesion side (right:left:brain stem) = 15:4:2 • Stroke subtype (Ischaemic:Haemorrhagic) = 18:3 • Mean time post-stroke (SD) = 4.1 (3.5) years • Fugl-Meyer Total-Motor Score = 70.8 (28.8) • Geriatric Depression Scale = 10.2 (7.3)
Intervention	**Visual Analogue Fatigue Scale (VAFS)** Heart rate Systolic blood pressure increase Rate of perceived exertion
Comparison	To itself at different times (at rest, post-exercise, post recovery). Participants took part in a fatigue-inducing exercise as a part of the study (a 15-minute standardised exercise protocol on a total-body recumbent stepper at a rate of 75 steps per minute and an external power of 75-80 Watts).

<table>
<tr><th>Reference</th><th>Tseng 2010[32]</th></tr>
<tr><td>Length of follow-up</td><td>14 days</td></tr>
<tr><td>Outcome measures</td><td>Criterion/Concurrent validity

Test-retest reliability
Intratest reliability
Responsiveness to change</td></tr>
<tr><td>Source of funding</td><td>This project was made possible by the use of the General Clinical Research Center (GCRC) at the University of Kansas Medical Center, Grant number M01 RR023940 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH).</td></tr>
<tr><td>Outcomes</td><td>Criterion/Concurrent validity
A significant positive relationship was found using Pearson’s correlation coefficient for exertion fatigue and systolic blood pressure increase (r=0.630, P=0.02) and exertion fatigue and rate of perceived exertion (r=0.802, P=0.00). Exertion fatigue and heart rate (r=0.738, P=<0.01).

Test-retest reliability
Intraclass correlation coefficient
• VAFS at rest = 0.851
• VAFS post exercise = 0.846
• VAFS post recovery = 0.888
• Exertion fatigue (VAFS post exercise – VAFS at rest) = 0.829
• Recovery rate ([VAFS post exercise – VAFS post recovery]/[VAFS post exercise – VAFS at rest] x 100) = 0.893

Intratest reliability
<table>
<tr><th></th><th>Visit 1 (mean [SD])</th><th>Visit 2 (mean [SD])</th></tr>
<tr><td>VAFS at rest</td><td>7.2 (4.3)</td><td>8.3 (4.5)</td></tr>
<tr><td>VAFS post exercise</td><td>69.4 (30.5)</td><td>65.8 (31.9)</td></tr>
<tr><td>VAFS post recovery</td><td>48.5 (25.4)</td><td>47.5 (26.9)</td></tr>
<tr><td>Exertion fatigue</td><td>62.4 (29.3)</td><td>57.5 (30.9)</td></tr>
<tr><td>Recovery rate (%)</td><td>37.0 (17.3)</td><td>37.7 (15.9)</td></tr>
</table>
Conclusion: The values appear similar between the two visits, with some variation in exertion fatigue.</td></tr>
</table>

Reference	Tseng 2010[32]
Risk of bias assessment	**Population** Stroke survivors. Provides information on gender, stroke lesion side and stroke subtype. Chronic time horizon. Does not provide information on location of stroke, initial stroke treatment, physical activity prior to stroke and severity. **Control population** No control population. **Recruitment/selection bias** Convenience sampling. Unclear how many participants came from local support groups and the ASTRA database. Limited number of participants meaning that there are a small number of people filling specific subgroups which may affect results (for example: stroke subtype). **Sample size** Very small sample size (N=21). **Appropriateness of metrics** Metrics used were appropriate but limited.
Limitations	Author provided limitations: • Subgroup analysis was not possible to see if the reliability and validity would be different between gender, different types of stroke and different times post stroke. • Although baseline fatigue was detected, the study was not designed to distinguish between other types of fatigue.
Summary/author's conclusion	**Author's conclusion** Preliminary findings suggest that the reliability, responsiveness and validity of the Visual Analog Fatigue Scale appears to be promising to assess exertion fatigue in people with chronic stroke.

Reference	Valko 2008[34]
Study type Setting/Location	Cross sectional study Switzerland, Neurology and Pulmonary Departments of the University Hospital of Zurich, Switzerland.
Number of participants	N=1306 (454 healthy subjects, 188 with multiple sclerosis, 235 with previous ischaemic stroke, 429 with sleep-wake disorders).

Reference	Valko 2008[34]

and characteristics	**Inclusion criteria:**
	People with clinically definite multiple sclerosis and previous ischaemic stroke consecutively examined in the neurological clinic between January 2005 and January 2007. Consecutive people with sleep-wake disorders referred to the neurological or pulmonary sleep clinics since December 2005 (including people with narcolepsy with cataplexy [n=22], restless leg syndrome [n=79], sleep apnoea [n=108], insomnia [n=62], parasomnia [n=25], excessive daytime sleepiness/hypersomnia of other origin [n=84] and other sleep-wake disorders [n=49]). Healthy control subjects among relatives and friends.
	
	Exclusion criteria:
	People with a diagnosed sleep-wake disorder, previous sleep studies, or other diseases known to cause fatigue (e.g. advanced cancer, HIV infection, heart failure, rheumatic disorders, depression).
	
	Values listed below are presented as mean (SD) or number (%) unless stated otherwise
	
	Healthy subjects
	• Mean age (SD): 47 (18) years
	• Female (%): 60%
	• Education status:
	  ○ Primary school degree: 43 (10%)
	  ○ Secondary school degree: 128 (29%)
	  ○ College degree: 93 (21%)
	  ○ University degree: 183 (41%)
	
	Multiple sclerosis
	• Mean age (SD): 45 (13) years
	• Female (%): 67%
	• Education status:
	  ○ Primary school degree: 63 (36%)
	  ○ Secondary school degree: 57 (32%)
	  ○ College degree: 26 (15%)
	  ○ University degree: 30 (17%)
	• Mean duration since disease onset (SD): 11.07 (9.79) years

Reference	Valko 2008[34]

	Previous ischaemic stroke
	• Mean age (SD): 63 (14) years
	• Female (%): 31%
	• Education status:
	  ○ Primary school degree: 59 (28%)
	  ○ Secondary school degree: 105 (49%)
	  ○ College degree: 21 (10%)
	  ○ University degree: 29 (14%)
	• Mean duration since disease onset (SD): 1.21 (0.62) years
	
	Sleep-wake disorders
	• Mean age (SD): 52 (15) years
	• Female (%): 35%
	• Education status:
	  ○ Primary school degree: 80 (28%)
	  ○ Secondary school degree: 113 (40%)
	  ○ College degree: 34 (12%)
	  ○ University degree: 59 (21%)

Intervention	**Fatigue Severity Scale-German (FSS-G)** Visual analogue scale-fatigue (VAS-F) Epworth Sleepiness Scale-German (ESS-G)
Comparison	Between scales, different populations
Length of follow-up	21 days (this was only completed with 104 of the healthy subjects).
Outcome measures	Face/content/construct validity Discriminant/convergent validity Test-retest reliability Internal consistency (Cronbach's alpha)

<table>
<tr><th>Reference</th><th>Valko 2008[34]</th></tr>
<tr><td>Source of funding</td><td>This was not an industry supported study. Dr. Blochas received research support from Respironics, ResMed and Weinmann AG and has had the free use of monitoring equipment from VivoMetrics.</td></tr>
<tr><td>Outcomes</td><td>Face/content/construct validity
In people after stroke, no correlation was found between FSS-G scores and age, duration from disease onset, gender or educational status.

The results of a linear regression analysis showed a significantly higher FSS-G score for each of the 3 patient groups than healthy controls. The residual analysis revealed symmetrically distributed residuals around zero, and only a minor departure from the model assumptions, as the distribution of residuals showed a small ceiling and bottom effect. This means that slightly more observations lie in the far left and far right side of the histogram of the residuals then one would expect under a strict Gaussian distribution.

Discriminant/convergent validity

Discriminant validity
<table>
<tr><th>FSS-G item</th><th>Healthy subjects (N=454) (mean [SD], Cronbach alpha if item deleted)</th><th>Multiple sclerosis (N=188) (mean [SD], Cronbach alpha if item deleted)</th><th>Previous ischaemic stroke (N=235) (mean [SD], Cronbach alpha if item deleted)</th><th>Sleep-wake disorders (N=429) (mean [SD], Cronbach alpha if item deleted)</th></tr>
<tr><td>1</td><td>5.06 (1.60), 0.86</td><td>5.28 (1.58), 0.95</td><td>4.65 (1.98), 0.96</td><td>5.07 (1.80), 0.94</td></tr>
<tr><td>2</td><td>3.11 (1.66), 0.85</td><td>4.93 (1.91), 0.93</td><td>4.11 (2.00), 0.96</td><td>4.23 (1.90), 0.94</td></tr>
<tr><td>3</td><td>2.75 (1.44), 0.83</td><td>4.69 (2.02), 0.93</td><td>4.03 (2.12), 0.95</td><td>4.28 (1.99), 0.93</td></tr>
<tr><td>4</td><td>3.94 (1.71), 0.84</td><td>5.04 (1.91), 0.93</td><td>4.28 (2.13), 0.95</td><td>4.61 (1.92), 0.92</td></tr>
<tr><td>5</td><td>2.66 (1.58), 0.82</td><td>4.45 (1.98), 0.93</td><td>3.69 (2.21), 0.95</td><td>4.41 (2.10), 0.92</td></tr>
<tr><td>6</td><td>2.50 (1.57), 0.83</td><td>4.85 (2.06), 0.93</td><td>3.98 (2.26), 0.95</td><td>4.06 (2.05), 0.93</td></tr>
<tr><td>7</td><td>2.41 (1.55), 0.83</td><td>3.99 (2.05), 0.93</td><td>3.53 (2.19), 0.95</td><td>3.90 (2.06), 0.93</td></tr>
<tr><td>8</td><td>2.30 (1.73), 0.83</td><td>4.52 (2.20), 0.93</td><td>3.68 (2.34), 0.95</td><td>4.44 (2.21), 0.93</td></tr>
<tr><td>9</td><td>2.16 (1.55), 0.82</td><td>4.19 (2.21), 0.93</td><td>3.51 (2.23), 0.95</td><td>4.16 (2.22), 0.93</td></tr>
<tr><td>Total</td><td>3.00 (1.08), 0.85</td><td>4.66 (1.64), 0.94</td><td>3.90 (1.85), 0.96</td><td>4.34 (1.64), 0.94</td></tr>
</table>
Conclusion: Higher values are seen in the multiple sclerosis, previous ischaemic stroke and sleep-wake disorder groups, although the values in the previous ischaemic stroke and healthy subjects groups are in total within the same number. There may be sufficient discriminant validity.</td></tr>
</table>

Reference	Valko 2008[34]

Convergent validity:

	Healthy subjects (N=454) (mean [SD])	Multiple sclerosis (N=188) (mean [SD])	Previous ischaemic stroke (N=235) (mean [SD])	Sleep-wake disorders (N=429) (mean [SD])
FSS-G	3.00 (1.08)	4.66 (1.64)	3.90 (1.85)	4.34 (1.64)
VAS-F	3.47 (2.24)	4.83 (2.49)	4.65 (2.55)	5.12 (2.45)

A highly significant correlation was seen between FSS scores and fatigue as indicated on a VAS (r = 0.69, P <0.01). This was higher in patients (MS: r=0.79, stroke: r=0.70, sleep-wave disorders: r=0.71) than in healthy subjects (r=0.52).

Conclusion: The values appear to be broadly similar between the two scales with the VAS-F generally having higher values than the FSS. However, the scores would broadly appear to correlate. There is insufficient data provided to discuss the convergent validity further.

Test-retest reliability – this was only completed with 104 of the healthy subjects. Lin's concordance measure rho.
Healthy subjects (N=104) = 0.88 (0.84-0.92)

Internal consistency

- Cronbach's alpha
  - Healthy subjects = 0.85
  - Multiple sclerosis = 0.94
  - Previous ischaemic stroke = 0.96
  - Sleep-wake disorders = 0.94
  - Total = 0.93

Risk of bias assessment

Population
Only includes people with ischaemic stroke.

Control population
Includes people with multiple sclerosis and sleep-wave disorders, who may experience fatigue. However, also includes healthy subjects, so overall no indirectness.

Recruitment/selection bias

Reference	Valko 2008[34]
	Only a small number of healthy subjects were included in the evaluation of test-retest reliability, which does not tell us if this would be the same experience for people with the disorders listed. Baseline values were not comparable between populations (for example: age is higher in the stroke group than other groups, more people had university degrees in the healthy subject group). **Sample size** Larger sample size (N=1306, with 235 of those people having a previous ischaemic stroke). **Appropriateness of metrics** Metrics used were appropriate. There was no use of Pearson's correlations to explore convergent validity between FSS-G and VAS-F which would have helped interpretation.
Limitations	Author provided limitations: • The presence of depression was not assessed
Summary/author's conclusion	**Author's conclusion** The FSS constitutes a valid instrument to assess and quantify fatigue for clinical and research purposes.

Reference	Visser-Keizer 2015[35]
Study type Setting/Location	Cross-sectional study The Netherlands, academic rehabilitation center (performed between April 9, 2010 and November 15, 2012)
Number of participants and characteristics	N=134 (55 with ischaemic stroke, 22 with haemorrhagic stroke, 35 with traumatic brain injury, 22 with other acute brain injuries all at least 6 months after brain injury) **Inclusion criteria:** People with acute brain injury participating in outpatient neurorehabilitation **Exclusion criteria:** People in the acute phase of brain injury (<6 months); people with premorbid chronic fatigue that interfere with their daily life. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** **Ischaemic stroke (n=55)**

Reference	Visser-Keizer 2015[35]
	• Mean age (SD): 55.4 (9.7) years • Women/Men (%): 25:30 (46%:54%) • Mean time since injury (SD): 33.7 (34.6) months • Lesion location (left:right:bilateral:diffuse:no lesion visible on CT or MRI): 22:25:6:0:2 • Educational level (1:2:3:4:5:6:7): 0:1:3:8:22:17:3 **Haemorrhagic stroke (n=22)** • Mean age (SD): 53.3 (9.4) years • Women/Men (%): 15:7 • Mean time since injury (SD): 29.1 (27.9) months • Lesion location (left:right:bilateral:diffuse:no lesion visible on CT or MRI): 6:14:2:0:0 • Educational level (1:2:3:4:5:6:7): 0:0:1:0:8:8:5 **Traumatic brain injury (n=35)** • Mean age (SD): 43.4 (13.3) years • Women/Men (%): 20:15 • Mean time since injury (SD): 43.1 (59.2) months • Lesion location (left:right:bilateral:diffuse:no lesion visible on CT or MRI): 2:3:8:12:10 • Educational level (1:2:3:4:5:6:7): 0:1:2:4:7:5:3 **Other acute brain injuries (n=22)** • Mean age (SD): 49.8 (12.0) years • Women/Men (%): 9:13 • Mean time since injury (SD): 36.9 (43.1) months • Lesion location (left:right:bilateral:diffuse:no lesion visible on CT or MRI): 5:3:2:12:0 • Educational level (1:2:3:4:5:6:7): 0:1:2:4:7:5:3
Intervention	**Dutch Multifactor Fatigue Scale (DMFS)** Checklist Individual Strength (CIS) Hospital Anxiety and Depression Scale – Anxiety and Depression subscales (HADS-A and HADS-D)

Reference	Visser-Keizer 2015[35]
	Dutch Personality Questionnaire – Self-esteem (PDQ-Self-esteem)
Comparison	Other scales, different populations
Length of follow-up	No follow up
Outcome measures	Face/content/construct validity Convergent/discriminant validity Internal consistency (Cronbach's alpha) Domains of fatigue considered
Source of funding	Supported by the Foundation Beatrixoord North Netherlands, Center for Rehabilitation, University Medical Center Groningen, Groningen, The Netherlands (grant no. 210.101).
Outcomes	**Face/content/construct validity** 5 factors identified: Impact of fatigue, Mental fatigue, Signs and Direct consequences of fatigue, Physical fatigue, Coping with fatigue. The 5-factor solution contained a number of items with insufficient quality, and so a total of 19 items were excluded.

Scale/Item	Factor loading	Corrected Item-Scale Correlation
Impact of fatigue		
I am often tired	0.81	0.75
When I am too fatigued, all of a sudden, I can't go further	0.65	0.59
Fatigue hinders my doings	0.71	0.64
I can be overcome by fatigue	0.56	0.49
I can easily get over my fatigue	0.46	0.40
I don't need to have a rest to make it through the day	0.69	0.62
I suffer from severe fatigue	0.80	0.73
I suffer terribly from my fatigue	0.83	0.76
I am tired every day	0.79	0.72
Fatigue if my serious complaint	0.71	0.63
Fatigue affects my whole life	0.90	0.85

Reference	Visser-Keizer 2015[35]		
	Mental fatigue		
	I can follow conversations without getting tired	0.75	0.64
	A lot of impressions, such as bustle or noise, make me fatigued	0.73	0.62
	Thinking makes me fatigued	0.73	0.62
	When too fatigued, I suddenly cannot think anymore	0.77	0.67
	My complaints get worse when I am fatigued	0.70	0.59
	When fatigued, I make mistakes	0.69	0.57
	When fatigued, I have difficulty concentrating	0.83	0.74
	Signs and Direct consequences of fatigue		
	I get fatigued in the afternoon	0.64	0.52
	Even if I am very tired, I recovery easily	0.59	0.48
	When fatigued, I get a headache	0.63	0.51
	Things that move me emotionally make me tired	0.73	0.61
	When I am fatigued, I say things I regret afterwards	0.63	0.50
	After a lot of thinking, fatigue still bothers me the next day	0.70	0.59
	Fatigue makes me react emotionally	0.75	0.63
	Other people notice that I am fatigued before I do	0.52	0.40
	When fatigued, I have difficulty letting my thoughts go	0.65	0.53
	Physical fatigue		
	I feel physically fit	0.81	0.66
	My body ached when fatigued	0.58	0.42
	After a good night sleep, I wake up rested	0.54	0.39
	Physical exertion make me tired	0.72	0.56
	I have little energy	0.74	0.57
	I have a good physical condition	0.71	0.53
	Coping with fatigue		
	I consciously plan when I will rest	0.53	0.33
	I finish what I am doing, even if I am tired	0.57	0.36

Reference	Visser-Keizer 2015[35]

I avoid becoming overtired	0.65	0.43
I always let myself get tired out	0.82	0.61
I often let myself become overtired when circumstances demand it	0.77	0.53

Convergent/discriminant validity

Person correlations between the final DMFS subscales:

	Impact of Fatigue	Mental Fatigue	Signs and Direct Consequences of Fatigue	Physical Fatigue
Mental Fatigue	0.68			
Signs and Direct Consequences of Fatigue	0.62	0.64		
Physical Fatigue	0.51	0.43	0.41	
Coping with Fatigue	0.11	0.15	0.15	0.08

Pearson correlations between subscales and measures of fatigue, mood and self-esteem

	Impact	Mental	Signs and direct consequences	Physical	Coping
CIS	0.72	0.62	0.57	0.70	0.14
HADS-Anxiety	0.37	0.36	0.50	0.36	0.19
HADS-Depression	0.41	0.39	0.40	0.45	0.05
PDQ-Self-esteem	-0.21	-0.20	-0.23	-0.27	-0.15

Comparison between groups

	Ischaemic stroke	Haemorrhagic stroke	Traumatic brain injury	Other acute brain injuries
Impact of fatigue	38.7 (11.7)	37.0 (11.2)	39.7 (7.3)	39.5 (8.7)
Mental fatigue	25.9 (7.1)	26.4 (5.0	29.4 (4.6)	28.1 (5.2)

Reference	Visser-Keizer 2015[35]				
	Signs and direct consequences of fatigue	28.8 (8.4)	28.1 (9.0)	32.3 (5.6)	30.7 (5.3)
	Physical fatigue	19.0 (6.0)	16.4 (5.2)	17.9 (5.2)	19.5 (4.7)
	Coping with fatigue	14.8 (4.5)	15.9 (4.3)	16.1 (4.4)	15.5 (4.0)
	Internal consistency Cronbach's alpha • Impact of fatigue = 0.91 • Mental fatigue = 0.86 • Signs and Direct consequences of fatigue = 0.83 • Physical fatigue = 0.77 • Coping with fatigue = 0.69 **Domains of fatigue considered** Physical and mental				
Risk of bias assessment	**Population** People after stroke and people with traumatic brain injury/other acute brain injuries. Most outcomes merge these populations and so introduces indirectness. **Control population** Other types of acute brain injury. This is an appropriate comparison for looking for difference between them and so if it can be applied to all types of acute brain injury, but it does not provide information about its comparability to people without fatigue. **Recruitment/selection bias** Recruited from an academic rehabilitation centre. No obvious problems. **Sample size** Sample size (n=134). **Appropriateness of metrics** No obvious problems.				

Reference	Visser-Keizer 2015[35]
Limitations	Author provided limitations: • Only people from outpatient neurorehabilitation. This may have selected those who suffer from fatigue or do not cope with fatigue adequately • Excluded people with complaints of chronic and ongoing fatigue in life before brain injury, but still had people who experienced an episode of burnout before their brain injury which could affect results • Research in more people would be useful to increase the certainty in the results
Summary/author's conclusion	**Author's conclusion** The DMFS is the first brain injury-specific scale that measures the full spectrum of fatigue in the chronic phase after acute brain injury. The measurement of separate aspects of fatigue using the DMFS is believed to be valuable to tailor rehabilitation to individual needs.

Reference	Wu 2008[36]
Study type Setting/Location	Cross-sectional study China, Department of Rehabilitation, Beijing Friendship Hospital, Capital Medical University
Number of participants and characteristics	N=330 **Inclusion criteria:** All people met the diagnostic standard from the Fourth National Cerebrovascular Disease Academic Meeting of China Medical Association in 1995 and were diagnosed with CT or MRI examination; vital signs were stable; people did not have serious complications, such as acute heart failure, haemorrhage of upper digestive tract, respiratory failure, or serious pulmonary infection. **Exclusion criteria:** Patients, who were not willing to accept scale evaluation, because of communication or cognitive disorders; people who had cancer, systemic lupus erythematosus, or Parkinson's disease; people who had a transient ischaemic attack, but not cerebral infarction and vertebrobasilar arterial insufficient; patients whose modified Rankin Scale was at least 4 points; people who were not willing to accept the questionnaire. **Values listed below are presented as mean (SD) or number (%) unless stated otherwise** • Age ranges: ○ 35-50 years = 40 (18.7%)

Reference	Wu 2008[36]
	○ 51-65 years = 71 (33.2%) ○ 66-80 years = 95 (44.4%) ○ >80 years = 8 (3.7%) • Male:female = 139:75 • Marriage (single, divorced and loss of spouse:married) = 36:162 • Education level (primary school or below:middle school:university or above) = 61:131:22 • Duration from survey to cerebral infarction: ○ <1 month = 70 (32.7%) ○ 1-6 months = 52 (24.3%) ○ 7-13 months = 92 (43.0%) • Modified Rankin Scale ○ 0-1 = 132 (61.7%) ○ 2-3 = 82 (38.3%)
Intervention	**Fatigue Impact Scale – Chinese version (FIS-C)**
Comparison	Itself
Length of follow-up	No follow up
Outcome measures	Face/content/construct validity Internal consistency (Cronbach's alpha) Domains of fatigue considered
Source of funding	No additional information
Outcomes	**Face/content/construct validity** Content validity: Correlation coefficient was calculated between each item and total scores, and the three subscale were >0.5, which was significantly different. Structural validity: Kaiser-Meyer-Olkin measurement of sampling adequacy = 0.934. Suitable for factor analysis. Factor 1 – Physiological states was captured by items 10, 13, 14, 16, 17, 23, 24, 31, 32, 37 and 38.

Reference	Wu 2008[36]
	Factor 2 – Cognitive states was captured by items 1, 5, 6, 11, 18, 21, 26, 30, 34 and 35. Factor 3 – The effects of emotion and fatigue on living was captured by items 4, 12, 33, 20, 36, 39 and 40. Factor 4 – Social communication was captured by items 2, 15, 19 and 22. Factor 5 – Effect on work was captured by items 3, 7, 8, 9, 25 and 27. Factor 6 – Economic situation and sex life was captured by items 28 and 29. Correlation analysis was used to separate these items into the three subscales. **Internal consistency** Cronbach's alpha • Cognitive subscale = 0.937 • Physiological subscale = 0.918 • Social subscale = 0.940 **Domains of fatigue considered** Cognitive, physiological and social
Risk of bias assessment	**Population** People after stroke. Included a range of genders. No information on location of stroke, initial stroke treatment and prior activity level. **Control population** Not applicable **Recruitment/selection bias** Rehabilitation department. No extra information. No obvious problems. **Sample size** Larger sample size (n=330) **Appropriateness of metrics** Metrics were appropriate
Limitations	Author provided limitations:

Reference	Wu 2008[36]
	None provided
Summary/author's conclusion	**Author's conclusion** The FIS was suitable to evaluate the clinical effects of cerebral infarction.

Appendix E – Forest plots

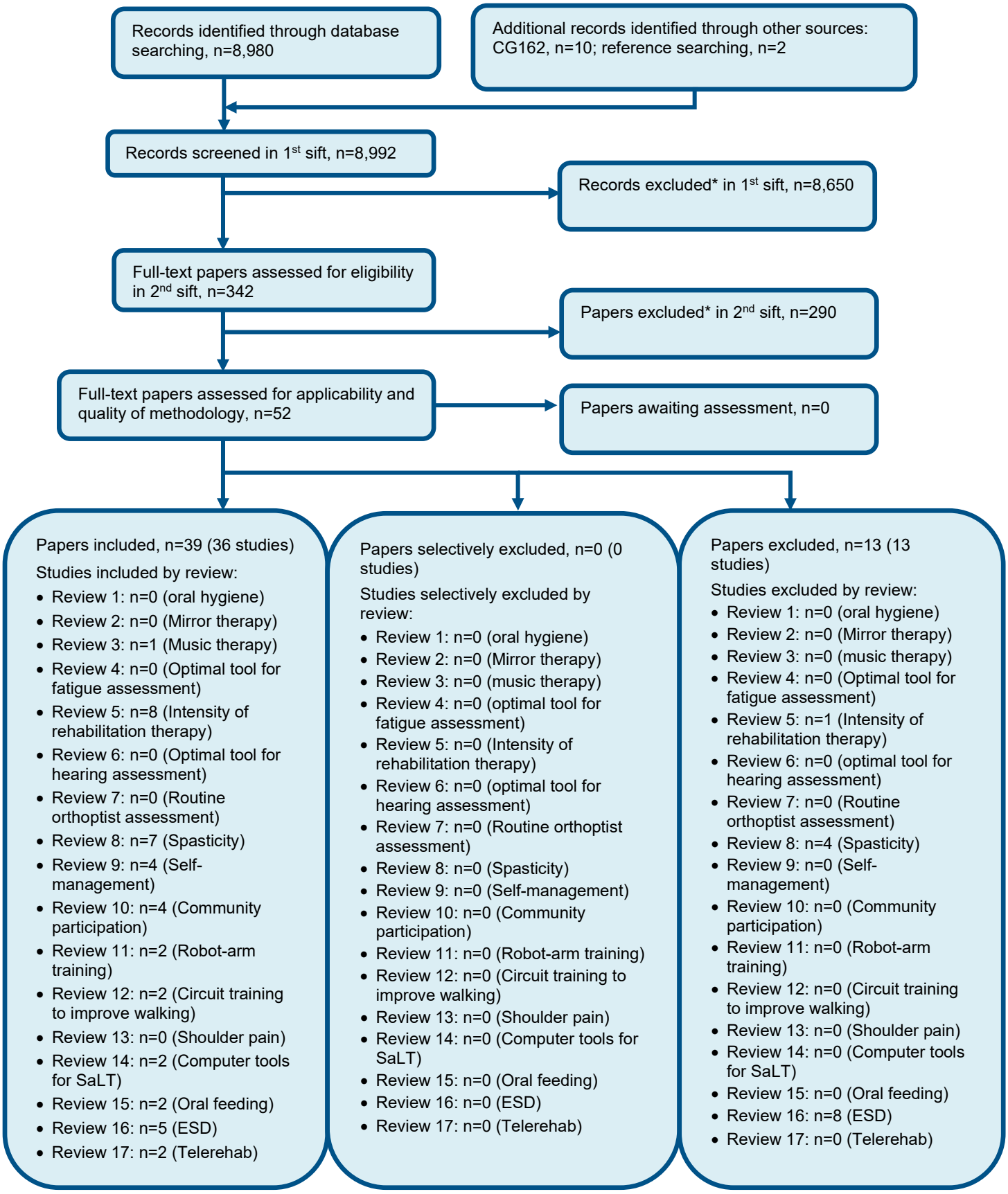
No assess-to-treat evidence was identified.

Appendix F – GRADE tables

No assess-to-treat evidence was identified.

Appendix G – Economic evidence study selection

Figure 22: Flow chart of health economic study selection for the guideline



* Non-relevant population, intervention, comparison, design or setting; non-English language

Appendix H – Economic evidence tables

There are no included health economic studies in this review.

Appendix I – Health economic model

New cost-effectiveness analysis was not conducted in this area.

Appendix J – Excluded studies

Clinical studies

Table 7: Studies excluded from the clinical review

Study	Code [Reason]
Almhdawi, K. A., Jaber, H. B., Khalil, H. W. et al. (2021) Post-stroke fatigue level is significantly associated with mental health component of health-related quality of life: a cross-sectional study. Quality of Life Research 30(4): 1165-1172	- No relevant outcomes
Borgaro, Susan R., Gierok, Susan, Caples, Heather et al. (2004) Fatigue after brain injury: Initial reliability study of the BNI Fatigue Scale. Brain Injury 18(7): 685-690	- Population not relevant to this review protocol *Only 45.2% of people had a cerebrovascular accident before the trial*
Buck, D., Jacoby, A., Massey, A. et al. (2004) Development and validation of NEWSQOL, the Newcastle Stroke-Specific Quality of Life Measure. Cerebrovascular Diseases 17(23): 143-52	- Quality of life scale that does not specifically focus on the use of the scale to measure fatigue
Cumming, T. B. and Mead, G. (2017) Classifying post-stroke fatigue: Optimal cut-off on the Fatigue Assessment Scale. Journal of Psychosomatic Research 103: 147-149	- No relevant outcomes
Dornonville de la Cour, F. L., Norup, A., Schow, T. et al. (2021) Evaluation of Response Processes to the Danish Version of the Dutch Multifactor Fatigue Scale in Stroke Using the Three-Step Test-Interview. Frontiers in Human Neuroscience 15: 642680	- Study design not relevant to this review protocol
Elbers, R. G., Rietberg, M. B., van Wegen, E. E. et al. (2012) Self-report fatigue questionnaires in multiple sclerosis, Parkinson's disease and stroke: a systematic review of measurement properties. Quality of Life Research 21(6): 925-44	- Systematic review used as source of primary studies
Hendriks, C., Drent, M., Elfferich, M. et al. (2018) The Fatigue Assessment Scale: quality and availability in sarcoidosis and other diseases. Current Opinion in Pulmonary Medicine 24(5): 495-503	- Study does not contain an intervention relevant to this review protocol

Study	Code [Reason]
Hubacher, M., Calabrese, P., Bassetti, C. et al. (2012) Assessment of post-stroke fatigue: the fatigue scale for motor and cognitive functions. European Neurology 67(6): 377-84	- No relevant outcomes
Jaracz, K.; Mielcarek, L.; Kozubski, W. (2007) Clinical and psychological correlates of poststroke fatigue. Preliminary results. Neurologia i Neurochirurgia Polska 41(1): 36-43	- No relevant outcomes
Johansson, S., Kottorp, A., Lee, K. A. et al. (2014) Can the Fatigue Severity Scale 7-item version be used across different patient populations as a generic fatigue measure--a comparative study using a Rasch model approach. Health & Quality of Life Outcomes 12: 24	- Study design not relevant to this review protocol
Kerber, K. A., Brown, D. L., Skolarus, L. E. et al. (2013) Validation of the 12-item stroke-specific quality of life scale in a biethnic stroke population. Journal of Stroke & Cerebrovascular Diseases 22(8): 1270-2	- Study does not contain an intervention relevant to this review protocol
LaChapelle, Diane L. and Finlayson, M. (1998) An evaluation of subjective and objective measures of fatigue in patients with brain injury and healthy controls. Brain Injury 12(8): 649-659	- Population not relevant to this review protocol
Legris, N., Devilliers, H., Daumas, A. et al. (2018) French validation of the Stroke Specific Quality of Life Scale (SS-QoL). Neurorehabilitation 42(1): 17-27	- Quality of life scale that does not specifically focus on the use of the scale to measure fatigue
Lenaert, B., van Kampen, N., van Heugten, C. et al. (2020) Real-time measurement of post-stroke fatigue in daily life and its relationship with the retrospective Fatigue Severity Scale. Neuropsychological Rehabilitation: 1-15	- Study design not relevant to this review protocol
Poulsen, M. B., Skovbolling, S. L., Kruuse, C. et al. (2020) How to identify fatigue in stroke patients: an investigation of the post-stroke fatigue case definition validity. Topics in Stroke Rehabilitation 27(5): 369-376	- Study design not relevant to this review protocol
Sallam, S. A., Al-Khamis, F. A., Muaidi, Q. I. et al. (2019) Translation and validation of the stroke specific quality of life scale into Arabic. Neurorehabilitation 44(2): 283-293	- Quality of life scale that does not specifically focus on the use of the scale to measure fatigue

Study	Code [Reason]
Stookey, A. D., Macko, R. F., Ivey, F. M. et al. (2021) Evaluating Test-Retest Reliability of Fatigability in Chronic Stroke. Journal of Stroke & Cerebrovascular Diseases 30(9): 105895	- Study does not contain an intervention relevant to this review protocol *On discussion with the topic expert it was decided that the 6-minute walk test would not be an appropriate comparison for this review as it could be confounded by so many factors and does not directly relate to fatigue. Therefore, this study was excluded.*
Tang, W. K., Lu, J. Y., Chen, Y. K. et al. (2010) Is fatigue associated with short-term health-related quality of life in stroke?. Archives of Physical Medicine & Rehabilitation 91(10): 1511-5	- Quality of life scale that does not specifically focus on the use of the scale to measure fatigue
Tyson, S. F. and Brown, P. (2014) How to measure fatigue in neurological conditions? A systematic review of psychometric properties and clinical utility of measures used so far. Clinical Rehabilitation 28(8): 804-816	- Systematic review used as source of primary studies
Vuletic, V.; Lezaic, Z.; Morovic, S. (2011) Post-stroke fatigue. Acta Clinica Croatica 50(3): 341-4	- No relevant outcomes
Wong, G. K., Lam, S. W., Ngai, K. et al. (2013) Development of a short form of Stroke-Specific Quality of Life Scale for patients after aneurysmal subarachnoid hemorrhage. Journal of the Neurological Sciences 335(12): 204-9	- Study does not contain an intervention relevant to this review protocol

Health Economic studies

Published health economic studies that met the inclusion criteria (relevant population, comparators, economic study design, published 2006 or later and not from non-OECD country or USA) but that were excluded following appraisal of applicability and methodological quality are listed below. See the health economic protocol for more details.

Table 8: Studies excluded from the health economic review

Reference	Reason for exclusion
None.	

Appendix K – Research recommendations – full details

K.1 Research recommendation

For people after stroke with communication difficulties, what is the optimal tool for assessing fatigue?

K.1.1 Why this is important

Fatigue after stroke is a common occurrence (estimated to occur in 51% of people 6 months after the stroke in one study[12]. Fatigue can have a significant effect on quality of life and the person's ability to engage with rehabilitation and so achieve success at the end of the program. If fatigue is identified early then it allows for strategies to be adapted and interventions which may help to reduce fatigue to be tried. The tools assessed in this review were primarily with populations who did not have communication difficulties, where the presentation of communication difficulties may be used as an exclusion criteria for the studies. Tools may need to be adapted so that they can be used by people with communication difficulties.

K.1.2 Rationale for research recommendation

Importance to 'patients' or the population	Fatigue is a common experience after stroke and can have a substantial effect on the person's quality of life and ability to engage with rehabilitation. Communication difficulties are common for people after stroke, however there can inequities in the support provided to people with communication difficulties when services are not able to provide adjustments to their needs. An effective tool for identifying fatigue for people with communication difficulties would help for this to be addressed earlier and help their recovery after stroke. Research into the best ways to recognise fatigue is a part of a priority identified in the James Lind Alliance Stroke Rehabilitation and Long-term Care Top 10 Priorities exercise (number 4), which takes into account feedback from people after stroke.
Relevance to NICE guidance	The evidence identified in this guideline either did not explicitly include people with communication difficulties or excluded them from the studies. This is a gap in the evidence that further research can be used to ensure equitable access to the guideline.
Relevance to the NHS	There are no studies that have been identified in this review investigating tools to assess fatigue for people with communication difficulties. Given this further evidence could allow NHS services to use an evidence-based led tool for fatigue that could be used across the service to allow for more consistent care across the country. If fatigue is identified earlier then this may help the person to engage more with rehabilitation, which

	can improve outcomes and reduce the chance of further rehabilitation needs.
National priorities	Developing high intensity care models for stroke rehabilitation is an aim in the NHS Long Term Plan. Identifying fatigue effectively may aid delivery of high intensity rehabilitation.
Current evidence base	The evidence identified in this review either does not explicity include people with communication difficulties or excludes people with communication difficulties. Therefore, new evidence investigating the use of tools for people with communication difficulties would be valued.
Equality considerations	People with communication difficulties often experience more difficulty accessing services than people without communication difficulties. Therefore, providing appropriate care is an equality consideration.

K.1.3 Modified PICO table

Population	Inclusion: • Adults (age ≥16 years) with communication difficulties who have had a first or recurrent stroke (including people after subarachnoid haemorrhage) Exclusion: • Children (age <16 years) • People who have had a transient ischaemic attack • People who do not have communication difficulties after stroke
Intervention	Tools for the assessment of fatigue after stroke, including adapted versions of fatigue scores used for people without communication difficulties (for example: the Fatigue Severity Scale, Fatigue Assessment Scale, Modified Fatigue Impact Scale). Followed by treatment for fatigue as assessed appropriate by the healthcare professional if fatigue is identified.
Comparator	Unadapted forms of tools to assess fatigue (for example: Fatigue Severity Scale, Fatigue Assessment Scale, Modified Fatigue Impact Scale) Clinical judgement of fatigue. Followed by treatment for fatigue as assessed appropriate by the healthcare professional if fatigue is identified.
Outcome	Health-related quality of life (EQ-5D-5L) Carer-related quality of life Activities of daily living

	Psychological distress Stroke-specific Patient Reported Outcome Measures Participation in leisure activities/social groups scores Withdrawal due to adverse events Resource use
Study design	Randomised controlled trial (test and treat)
Timeframe	6 months
Additional information	Subgroup analyses: • Type of stroke (subarachnoid haemorrhage, other type of haemorrhage, ischaemic) • Location of stroke (total anterior circulation stroke, partial anterior circulation stroke, lacunar stroke, posterior circulation stroke) • Initial stroke treatment (thrombolysis, thrombectomy) • Physical activity prior to stroke (low, moderate, high) • Gender (male, female, non-binary) • Severity of stroke (NIHSS scale, split into mild 1-5, moderate 5-14, severe 15-24, very severe >25) • Time after stroke on entry to the study (hyperacute <72 hours, acute 72 hours-7 days, subacute 7 days-6 months, chronic >6 months)

K.2 Research recommendation

What is the clinical and cost-effectiveness of the Fatigue Severity Scale and Fatigue Assessment Scale in informing the management of fatigue for people after stroke?

K.2.1 Why this is important

Fatigue after stroke is a common occurrence (estimated to occur in 51% of people 6 months after the stroke in one study[12]. Fatigue can have a significant effect on quality of life and the person's ability to engage with rehabilitation and so achieve success at the end of the program. If fatigue is identified early then it allows for strategies to be adapted and interventions which may help to reduce fatigue to be tried. The tools assessed for this review were assessed with tool validity and reliability outcomes, with no information being identified on the clinical and cost-effectiveness. Additional information about this can help to show if using tools to assess fatigue can lead to clinically important differences in care outcomes.

K.2.2 Rationale for research recommendation

Importance to 'patients' or the population	Fatigue is a common experience after stroke and can have a substantial effect on the person's quality of life and ability to engage with rehabilitation. Identifying a tool that can effectively identify fatigue may allow for more

	people to receive interventions and adaptations that can help them to reduce their fatigue and improve their quality of life.
Relevance to NICE guidance	The evidence identified in this guideline investigated tool validity and reliability but did not investigate the clinical and cost-effectiveness of tools to assess fatigue. Identifying this can help to show whether fatigue assessment tools are clinically and cost effective, helping to answer the initial question from the review.
Relevance to the NHS	There are no studies that have been identified in this review investigating the clinical and cost-effectiveness of tools to assess fatigue. Given that all people after stroke may experience fatigue, being able to understand if the assessment will be clinically and cost-effective is particularly important to ensure NHS resources are being used appropriately.
National priorities	Developing high intensity care models for stroke rehabilitation is an aim in the NHS Long Term Plan. Identifying fatigue effectively may aid delivery of high intensity rehabilitation.
Current evidence base	The evidence identified in this review investigated tool validity and reliability but did not investigate the clinical and cost-effectiveness of tools to assess fatigue.
Equality considerations	No specific equality considerations were identified. The committee noted that in general throughout the guideline, people with communication difficulties, older people and people who have had a previous stroke or transient ischaemic attack were excluded from trials but are people that the guideline is for. Therefore, research should aim to include these people where possible.

K.2.3 Modified PICO table

Population	Inclusion: • Adults (age ≥16 years) who have had a first or recurrent stroke (including people after subarachnoid haemorrhage) Exclusion: • Children (age <16 years) • People who have had a transient ischaemic attack
Intervention	Fatigue Severity Scale (cut off value 36) Fatigue Assessment Scale (cut off value 23) Combinations of the above

	The interventions provided for fatigue are not consistently prescribed in current practice, therefore any consistent management strategy used in this population would be accepted, including education, behaviour change interventions, nonpharmacological and pharmacological management.
Comparator	Each other Management based on self-reported fatigue (usual care)
Outcome	Health-related quality of life (EQ-5D-5L) Carer-related quality of life Activities of daily living Psychological distress Stroke-specific Patient Reported Outcome Measures Participation in leisure activities/social groups scores Withdrawal due to adverse events Resource use
Study design	Randomised controlled trial (test and treat)
Timeframe	6 months
Additional information	Subgroup analyses: • Type of stroke (subarachnoid haemorrhage, other type of haemorrhage, ischaemic) • Location of stroke (total anterior circulation stroke, partial anterior circulation stroke, lacunar stroke, posterior circulation stroke) • Initial stroke treatment (thrombolysis, thrombectomy) • Physical activity prior to stroke (low, moderate, high) • Gender (male, female, non-binary) • Severity of stroke (NIHSS scale, split into mild 1-5, moderate 5-14, severe 15-24, very severe >25) • Time after stroke on entry to the study (hyperacute <72 hours, acute 72 hours-7 days, subacute 7 days-6 months, chronic >6 months)