Cognitive assessment scale for stroke patients (CASP): A multicentric validation study

Charles Benaim, Grégoire Wauquiez, Dominic Pérennou, Céline Piscicelli, Brigitte Lucas-Pineau, Huei-Yune Bonnin-Koang, Philippe Vuadens, Christine Binquet, Abderrahmane Bourredjem, Hervé Devilliers

Annals of Physical and Rehabilitation Medicine 65 (2022) 101594

Available online at ScienceDirect www.sciencedirect.com

Elsevier Masson France EM|consulte www.em-consulte.com

Original Article

Cognitive assessment scale for stroke patients (CASP): A multicentric validation study

Charles Benaim[a,b,*], Grégoire Wauquiez[c], Dominic Pérennou[d,e], Céline Piscicelli[d,e], Brigitte Lucas-Pineau[f], Huei-Yune Bonnin-Koang[g], Philippe Vuadens[b], Christine Binquet[h], Abderrahmane Bourredjem[h], Hervé Devilliers[h]

[a] *Physical Medicine and Rehabilitation, Lausanne University Hospital and University of Lausanne, Avenue Pierre Decker 4, CH-1011, Switzerland*
[b] *Clinique Romande de Réadaptation - SuvaCare, Sion, Switzerland*
[c] *Physical Medicine and Rehabilitation, Dijon University Hospital, France*
[d] *Neurorehabilitation Department, Institute of Rehabilitation, Grenoble Alpes University Hospital, Echirolles 38434, France*
[e] *Laboratoire de Psychologie et NeuroCognition, UMR CNRS 5105, Université Grenoble Alpes, Grenoble, France*
[f] *Centre de Rééducation Fonctionnelle Divio, Dijon, France*
[g] *Neurorehabilitation, Nimes University Hospital, France*
[h] *Centre d'Investigation Clinique CIC 1432 and University of Bourgogne-Franche-Comté, France*

ARTICLE INFO

Article History:
Received 16 December 2020
Accepted 19 September 2021
Available online xxx

Keywords:
Validation studies
Stroke
Cognitive
Assessment

ABSTRACT

Background: The Mini Mental State Examination and Montreal Cognitive Assessment are commonly used as short screening batteries for assessing cognitive impairment after stroke. However, aphasia or hemispatial neglect may interfere with the results. For this reason, we developed the Cognitive Assessment scale for Stroke Patients (CASP), which takes these conditions into consideration and previously demonstrated its superiority over these scales in terms of feasibility.
Objectives: Our goal was to verify the psychometric properties of the (original) French version of the CASP.
Methods: We included 201 patients with a recent first hemispheric stroke and 50 controls. Stroke patients were examined 4 times (visit 1 [V1] to visit 4 [V4]) in the subacute post-stroke phase. The structural validity of the CASP was studied by principal factorial analysis, convergent validity by comparison with several variables including a comprehensive neuropsychological assessment, divergent validity by comparison with the total score between stroke patients and controls, and sub-scores between right and left stroke. Internal consistency, reproducibility and sensitivity to change were assessed. We propose the Minimal Clinically Important Difference (MCID) value and a pathological threshold as well as a threshold to predict cognitive change between V1 and V4.
Results: Of the 201 participants included (63% male; mean [SD] age 63 [13] years), CASP data were available for 199/150/133/93 at V1/V2/V3/V4, respectively. CASP has a one-dimensional structure. The hypotheses of convergent/divergent validities were confirmed. Internal consistency was good and reliability excellent. Responsiveness was small to moderate, but the MCID could still be estimated. We discuss the choice of a pathological threshold and a predictive threshold of V1 over V4.
Conclusions: CASP has good psychometric properties for screening cognitive impairment in the subacute post-stroke phase, which is consistent with its Italian and Korean versions. It can be used for patients with severe motor aphasia or left hemispatial neglect but not in case of severe oral comprehension or visual impairment.

Abbreviations: ASRS, aphasia severity rating scale; AUC, area under the receiver operating characteristic curve; BDAE, boston diagnostic aphasia examination; CASP, cognitive assessment scale for stroke patients; CFI, comparative fit index; CNA, comprehensive neuropsychological assessment; MCID, minimal clinically important difference; MMSE, mini mental state examination; MoCA, montreal cognitive assessment; PFA, principal factor analysis; RMSEA, root mean square error of approximation; SEM, standard error of measurement; SRM, standardized response mean; SRMR, standardized root mean square residual; TLI, tucker lewis index

* Corresponding author at: Physical Medicine and Rehabilitation, Lausanne University Hospital and University of Lausanne, Avenue Pierre Decker 4, CH-1011, Switzerland.
E-mail address: Charles.benaim@chuv.ch (H. Devilliers).

https://doi.org/10.1016/j.rehab.2021.101594

1. Introduction

The Mini Mental State Examination (MMSE) [1] and Montreal Cognitive Assessment (MoCA) [2] are 2 short screening batteries assessing cognitive impairment that are widely used in neurology but are not suitable for many post-stroke patients. For example, a patient with motor aphasia cannot answer the MMSE and MoCA questions assessing anterograde amnesia or time orientation. Similarly, a patient with severe spatial neglect may be uncomfortable with the MoCA naming item because it is visual (drawing of a lion on the left side of the page, and so should be compensated by positioning the sheet to the right of the patient).

For this reason, we developed the Cognitive Assessment scale for Stroke Patients (CASP) [3,4] to minimize these disadvantages (Supplemental Material 1 and 2): all answers can be given without using language (with the exception of the naming test, of course), and several items are arranged on the test page so that that they are not hindered by left spatial neglect (side most often affected). The CASP contains 9 items grouped in 6 dimensions: communication, spatial/visuo-construction, executive functions, short-term memory, praxis and time orientation (6×6 points). Its main features are that it focuses specifically on post-stroke cognitive impairments, it is more feasible to administer than the MMSE and MoCA in these patients, and the severity of aphasia has much less influence on the overall CASP score than the other 2 tests [3–5]. For example, it can be used to establish that a mute patient is well oriented to time or does not have anterograde amnesia, which is impossible to reveal with the MMSE and MoCA. In Barnay et al., the CASP was impossible to administer in 18% of 44 unselected aphasic stroke patients as compared with 36% for the MMSE and 30% for the MoCA [3]. In Benaim et al., the CASP was impossible to administer in 0% of 50 unselected non-aphasic patients versus 0% for the MMSE and 6% for the MoCA [4]. A cultural adaptation of the drawings was suggested by Park et al., who cross-culturally validated the CASP in Korean [6], then by an Iranian team that is currently validating the instrument in Persian (unpublished data). Its main limitation is that it cannot be used in case of severe oral comprehension impairment.

This work was the final validation of CASP (French version) including structural validity, reliability and responsiveness. We hypothesized that the CASP has a predominantly one-dimensional structure, which corresponds to overall cognitive performance, and possibly 2 additional dimensions supported by language and hemispatial neglect, usually found in left and right hemispheric strokes, respectively.

2. Material and methods

2.1. Study design and setting

Five rehabilitation services in France and Switzerland participated in this multicenter cohort prospective study. The protocol was approved by the ethics committees in France (2013-A00913-42, East-France CCP, 2015-02-20) and Switzerland (CCVEM 048/14, 2014-12.01).

2.2. Patients

In total, 201 consecutive patients aged 18 to 90 years were recruited between January 2015 and February 2019. The inclusion criteria were first-ever unilateral hemispheric stroke, time since stroke < 2 months, and informed consent given. Non-inclusion criteria were known cognitive deficiencies, psychosis or severe visual impairment before the stroke, non-French-speaking, and severe stroke-induced impairment of oral comprehension. Regarding the latter and as explained previously [3,4], the assessment of oral comprehension by precise clinical tools such as the Boston Diagnostic Aphasia Examination (BDAE) [7] items is not compatible with the objective of the CASP, which must be able to be administered quickly and by non-specialists. For this reason, we use the Aphasia Severity Rating Scale (ASRS) of the BDAE to detect patients unable to complete the CASP. However, the examiner is instructed to consider only oral comprehension because severe speaking impairment does not preclude taking the CASP. Under these conditions, the ASRS score should not be ≤2/5.

2.3. Control groups

We included 50 age-matched patients recruited from orthopaedic and geriatric rehabilitation services. Apart from stroke, the criteria for inclusion and non-inclusion were the same as for stroke patients. Data for these patients were used for assessing divergent validity. In daily practice, the primary goal of the CASP is to detect cognitive impairment in stroke patients. Therefore, we set up a second control group of stroke patients without any cognitive impairment detected by the neuropsychologist at the fourth visit. Data for these patients were used for assessing pathological and predictive thresholds.

2.4. Variables and follow-up

After the inclusion visit (V0), patients were assessed for an initial assessment (V1), then for 3 more visits (V2 to V4), according to the schedule described in the Fig. 1.

Data collected included general socio-demographic data; brain lesion description: side, mechanism (ischemia/hemorrhage), US National Institute of Health Stroke Score[7] (NIHSS); and neurological deficiencies and impairments: CASP, MMSE, ASRS [8], modified Rankin Scale [9] (mRankin), and the "sadness" item of the visual analogue mood scale [10]. As a clinical anchor to define cognitive impairment, we let the neuropsychologists choose among valid clinical scales available in their respective rehabilitation centers the best-fitting tests for their patients. However, we encouraged them to use the BDAE or the Montreal-Toulouse 86 test [11] for aphasia and the Line bisection [12] or Bell's test [13] for spatial neglect. Then, they had to rate each of the following domains on a 5-point Likert scale (0: no deficiency to 5: severe deficiency): language production and comprehension, praxis, amnesia, spatial unilateral neglect, executive functioning and timed-orientation (same domains as for the CASP). We called it the Comprehensive Neuropsychological Assessment (CNA). All the neuropsychologists had at least 5 years of experience with stroke.

On visits V1 to V4, the CASP was administered by experts (neuropsychologists, senior neurologists or physiatrists) or non-expert clinicians (physicians in training or clinical psychologists), depending on the local organization. In all cases, the neuropsychologist who administered the CNA was blinded to CASP results. The sequence of examiners for visits 1 to 3 (reliability) was "1st rater–2nd rater–1st rater" in all but one center (1st rater–1st rater–2nd rater). All other clinical scales were administered by physicians (physiatrists or neurologists). V1 to V4 always occurred during the rehabilitation stay; patients who left the hospital before V4 were not contacted after discharge.

2.5. Statistical methods

The statistical methods used follow the COSMIN guidelines, and the reporting is in accordance with the STARD 2015 checklist.

Structural validity: The aim was to determine the underlying dimensions of the CASP (i.e., 1, 2 or more concepts measured by the items). First, we conducted principal factor analysis (PFA), which builds up uncorrelated variables (factors). Coefficients defining these linear combinations, called factor loadings, may be interpreted as correlation coefficients. To determine the number of retained factors, we used the Kaiser criterion [14]: eigenvalues (proxy of the explained

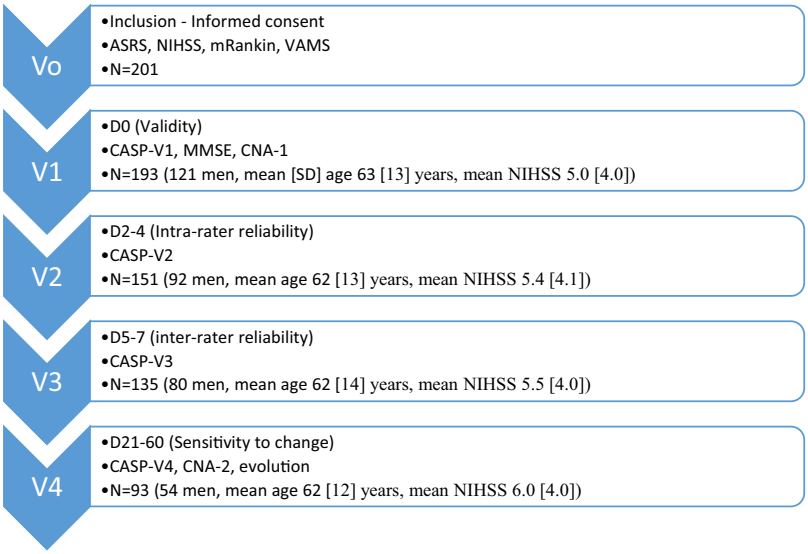

Fig. 1. Scheduling and content of visits from the inclusion visit (V0) to visit 4 (V4), number of patients per visit. ASRS, Aphasia Severity Rating Scale; CASP, Cognitive Assessment scale for Stroke Patients; CNA, Comprehensive Neuropsychological Assessment; D, day; MMSE, Mini Mental State Examination; mRankin, modified Rankin Scale; NIHSS, US National Institute of Health Stroke Score; VAMS, Visual Analogous Mood Scale.

variance) >1. We next studied item loading on each factor, considering 0.3 as significant loading [15]. Finally we performed a confirmatory factor analysis [16] by using the retained PFA structure of the CASP. In this analysis, each item was defined to represent only one domain, but the domain scores were allowed to correlate with each other. To assess the model quality of fit, we report the following: (1) ratio of the chi-square statistic and degrees of freedom (chisq/df ≤3 indicates acceptable fit) [17]; (2) standardized root mean square residual (SRMR ≤.08 indicates acceptable fit) [18]; (3) root mean square error of approximation (RMSEA<0.05 or even 0.08 is generally considered good fit and >0.1 corresponds to poor fit) [19,20]; and (4) indexes to describe incremental fit, the Bentler Comparative Fit Index (CFI) and Tucker Lewis Index (TLI) (CFI and TLI >0.9 indicate good fit) [21].

External validity: (1) convergent validity was determined by comparing the results of the CASP with those of other clinical scores by Spearman's rank correlation. We expected at least moderate correlation (>0.35) with the CNA, MMSE, and mRankin scores and NIHSS. (2) Divergent validity: we compared CASP domain scores between stroke groups defined by lesion side (left vs right) and between cases and controls by Student *t* test [22]. Left stroke patients were expected to have lower scores on language items and higher scores on spatial/visuo-construction items. Controls were expected to have higher total scores than stroke patients.

Reliability: (1) for internal consistency, Cronbach's alpha coefficient was computed for CASP items, with ≥ 0.7 considered satisfactory [23]; (2) inter- and intra-rater concordance were assessed by computing intra-class coefficients (ICCs). Agreement was considered excellent at ICC >0.75, good at 0.4−0.75 and poor otherwise [24].

Responsiveness: The sensitivity to change observed by neuropsychologists between V1 and V4 was quantified on a 9-point Likert scale (from -4 to 4). Effect size (ES=Δscore/6initial score) and the standardized response mean (SRM=Δscore/6 [Δscore]) were calculated for patients with improved condition (improvement of at least 2 points on the Likert scale). Values of 0.20, 0.50, and 0.80 for these statistics correspond to a small, moderate or important change, respectively [25]. The mean difference in scores between patients with improvement of at least 2 points on the CNA and those with no improvement was considered a first estimate of the CASP Minimal Clinically Important Difference (MCID), according to the Anchor-based method [26]. A second estimate of the MCID was obtained with the distribution approach by calculating the standard error of measurement (SEM=6*Square root [1−*r*], where 6=baseline standard deviation and *r*=Cronbach's alpha) [27].

Pathological thresholds of the CASP were determined by comparing patients without cognitive deficiencies (CNA=0) to the others. We used V4 data to ensure sufficient patients without impairment. The area under the receiver operating characteristic curve (AUC) was calculated for the CASP. Because the purpose of the CASP was for screening cognitive deficiencies by non-expert clinicians before referral to neuropsychologists, the thresholds were determined by prioritizing sensitivity to specificity.

Predictive thresholds were determined in the same manner: we estimated the extent to which the CASP score at V1 could predict the CNA score at V4.

2.6. Study size

The number of patients required for the PFA was 20 times the number of CASP items, so 180. In total, 120 patients were required to ensure an ICC of 0.7 with accuracy of ±0.1. Fifty control participants were needed to ensure a difference of 1.3 for divergent validity, with precision ±1.5.

3. Results

3.1. Participant characteristics (Table 1, Fig. 1)

We included 201 patients. Because the number of participants needed for reliability (V2 and V3) was reached before the end of the inclusions, the last patients had only V0 (inclusion), V1 (validity) and V4 (sensitivity to change). The number of patients assessed at the 5 visits was 201, 193, 151, 135 and 93. The number of missing CASP scores at visits V1 to V4 was 2, 1, 2 and 0, or 5/572 (1%). In all cases, only one item was not completed (4, "Reproducing a copy of a cube" and 1, "Inhibition/Flexibility"). The proportion of incomplete MMSE scores was 10/152 (7%). For participants who required a long stay in rehabilitation services and who were still present at V4, the mean (SD) CASP score improved by 1.87 (3) points ($p < 0.001$). At initial assessment, they had more severe global impairments and disabilities than the others (mean NIHSS: 5.84 [3.89] vs 4.01 [3.90], $p = 0.0017$; mRankin score: median [interquartile range] 4 [3, 4] vs 2

Table 1
Main baseline characteristics of stroke patients (n = 201).

	Overall n = 201	Patients attending V4 n = 93
Age, years	63 (13)	62 (12)
Sex (M)	127 (63%)	54 (58%)
Lesion side (n = 199)		
Right	106 (53%)	56 (60%)
Left	92 (46%)	35 (38%)
Bilateral	1	5 (1%)
Lesion type		
Ischemia	152 (76%)	64 (68%)
Hemorrhage	49 (24%)	29 (32%)
Days post-stroke (V0)	31.1 (12.7)	33.7 (13.7)
NIHSS (V0, n = 196)	5.0 (4.0)	6.0 (4.0)
ASRS (V0, n = 200)	5.0 (1.4)	5.1 (1.4)
MMSE (V1, n = 151)	25.2 (4.2)	24.5 (5.0)
CNA (V1, n = 188)	5.4 (3.9)	5.9 (4.2)
mRankin		
- No disability (0, 1)	47 (23%)	7 (8%)
- Light-mild disability (2, 3)	70 (35%)	31 (33%)
- Severe disability (4, 5)	79 (39%)	53 (57%)
VAMS sadness (n = 178)	71.4 (26.8)	74.9 (26.7)
CASP		
V1 (n = 199)	30.6 (4.9)	30.1 (5.6)
V2 (n = 150)	31.1 (4.9)	30.8 (5.3)
V3 (n = 133)	31.6 (4.5)	31.1 (5.0)
V4 (n = 93)	31.8 (4.6)	31.8 (4.6)

Data are mean (SD) unless indicated. ASRS, Aphasia Severity Rating Sacle; CASP, Cognitive Assessment scale for Stroke Patients; CNA, Comprehensive Neuropsychological Assessment; MMSE, Mini Mental State Examination; mRankin, modified Rankin Scale; NIHSS, US National Institute of Health Stroke Score; V, visit; VAMS, Visual Analogous Mood Scale.

[1–3], $p < 0.0001$) but equivalent cognitive impairments (mean CNA score: 5.9 [4.2] vs 4.9 [3.6], p = 0.1194; mean CASP score: 30.1 [5.6] vs 31.1 [3.91], p = 0.4024). After checking and validation, data for 201 participants (193 at V1 and 8 at V2) were available for assessing structural and external divergent validity (initial CASP data), 193 for external convergent validity (V1), 132 for inter-rater reliability (mainly V1 and V2), 129 for intra-rater reliability (mainly V2 and V3) and 93 for sensitivity to change (V4). The mean age of the 50 non-stroke participants was comparable to that of stroke patients (64 [17] vs 63 [13], p = 0.68) but with fewer men (42% vs 63%, p = 0.0097).

3.2. Structural validity

PFA: Only one factor was selected by the Kaiser Criterion, which suggests that CASP is rather unidimensional. Factor loadings of all items were higher for Factor 1 than other factors (Table 2), which therefore can be considered the "overall cognitive performance factor". Factors 2 and 3 should not be taken into account, but of note, language items (left lesions) were best correlated with Factor 2 and the hemispatial neglect item (right lesions) with Factor 3.

Confirmatory factor analysis: Chisq/df and SRMR values were acceptable (2.79 and 0.0643, respectively), the RMSEA was adequate (0.095, 95% confidence interval [CI] 0.069–0.121), and CFI and TLI were fair (0.864 and 0.812).

3.3. External validity

Convergent validity: the CASP score was strongly correlated with the CNA score (Rho = -0.788, $p < 10^{-4}$) and MMSE score (Rho = 0.640, $p < 10^{-4}$), moderately with the mRankin score (Rho = -0.388, $p < 10^{-4}$) and weakly with the NIHSS (Rho = -0.288, $p < 10^{-4}$); the latter did not support the convergent validity.

Divergent validity: Mean CASP scores were similar for left and right hemispheric stroke patients (30.8 [5.3] and 30.3 [4.6], p = 0.5001). As expected, the mean language score (Naming +Comprehension) was significantly lower and spatial/visuo-construction score (Cube+Line bisection) higher with left than right hemispheric stroke (5.5 [1.1] vs 5.8 [0.4], p = 0.0053 and 4.9 [1.4] vs 4.1 [1.7], p = 0.0006). The mean CASP score was significantly lower for stroke patients than controls (30.6 [4.9] vs. 32.9 [3.2], $p < 10^{-4}$). Scores for all 6 CASP domains were lower for stroke patients than controls, but the difference was not significant for praxis and time orientation (p = 0.1911 and p = 0.4556, respectively).

Table 2
Factor loadings of the principal factor analysis. All items were best represented on the first factor.

Factor pattern CASP items	Factor1 (81%)	Factor2 (13%)	Factor3 (6%)
Naming	0.45154	0.39273	-0.00212
Comprehension	0.53658	0.33151	0.12017
Cube	0.55542	-0.29031	0.11372
Line bissection	0.33905	-0.19757	0.26231
Graphic series	0.62828	-0.03099	-0.03026
Inhibition/flexibility	0.59698	-0.12880	-0.25295
Short-term memory	0.52860	0.06323	0.07957
Praxis	0.58015	-0.09607	-0.06064
Temporal orientation	0.54706	-0.01520	-0.09595

3.4. Reliability

Internal consistency: Cronbach's alpha was 0.78 (95% CI 0.69–0.83), which indicates good internal consistency.

Inter-rater reliability: the mean time between the 2 administrations was 2.8 (1.0) days. The ICC was excellent for the total CASP score (0.78, 95% CI 0.73–0.83]) and the language domain (0.80, 95% CI 0.74–0.84]) and good for the other 5 domains (0.42–0.65). Among the 9 individual items, only Line bisection had a poor ICC (0.37, 95% CI 0.27–0.48]).

Intra-rater reliability: The mean time between the 2 administrations was 3.7 (1.4) days. The ICC was excellent for the total CASP score (0.85, 95% CI 0.79–0.89) and language (0.89, 95% CI 0.85–0.92) and spatial/visuo-construction domains (0.77, 95% CI 0.69–0.83) and good for the other 4 domains (0.51–0.73). For the 9 individual items, the ICCs were good or excellent (0.51–0.96).

3.5. Responsiveness

If we consider as significantly improved individuals showing improvement by at least 2 points on the CNA (78/93 stroke patients), then the responsiveness was small (ES=0.41) to moderate (SRM=0.76). For these patients, CASP scores improved by a mean of 2.3 (3.0) points (p = 0.002), which can be considered a first estimate of the MCID (anchor-based method). The second estimate (distribution approach) was very close: 2.3.

3.6. Pathological threshold (Table 3)

At V4, the mean CASP score for the 11 patients without cognitive impairment (CNA score =0 at V4) was 35.6 (0.5) versus 31.2 (4.6) for the other 82. The AUC was excellent: 0.90 (95% CI 0.84–0.96). The threshold was 34.5 to 35 depending on whether sensitivity or specificity was prioritized. A priori, we retained a threshold of 35, which gave a sensitivity of 89% and specificity 64%. Positive and negative predictive values to predict cognitive impairment were 95% and 56%, respectively.

3.7. Predictive threshold (Table 4)

At V1, the mean CASP score for the 11 patients who fully recovered cognitive functions at V4 (CNA=0 at V4) was 34.2 (1.7) versus

Table 3
Pathological thresholds. Sensitivity, specificity and Youden index for the CASP at visit 4 (CASP V4) as compared with the Comprehensive Neuropsychological Assessment at visit 4 (CNA V4).

CASP V4: Potential cut-off	Sensitivity	Specificity	Youden Index
33.0	0.54	1.00	0.54
33.5	0.56	1.00	0.56
34.0	0.74	1.00	0.74
34.5	0.76	1.00	0.76
35.0	**0.89**	**0.63**	**0.52**

29.5 (5.7) for the other 82. The AUC was very good: 0.81 (95% CI 0.70 −0.92). The threshold was 32.5 to 34.5 depending on whether sensitivity or specificity was prioritized. A priori, we retained a threshold of 34.5, which gave a sensitivity of 84% and specificity of 55%. Positive and negative predictive values to predict full recovery were 93% and 32%, respectively.

4. Discussion

Our study suggests that the CASP could be an interesting tool for screening patients with cognitive impairment in the subacute phase of stroke. This finding is consistent with results from its Italian and Korean versions [5,6]. It could also give an idea of the cognitive evolution in the medium term.

In the present study, 7% of the MMSE scores could not be calculated because of missing items versus only 1% for the CASP, which shows the good feasibility of this test.

The CASP is essentially unidimensional in structure and measures cognitive impairment overall. However, the language items also contributed somewhat to the formation of a second factor and the neglect item contributed somewhat to the formation of a third factor. This result was not surprising because patients with severe aphasia and those with severe unilateral spatial neglect are known to have a very different clinical profile.

We verified the convergent and divergent validity, reproducibility and responsiveness of the CASP. The reproducibility of some items could be perfected, especially Line bisection between 2 different examiners. This observation could be due in part to the heterogeneity of our examiners because in 52% of cases, one of the 2 examiners was not an experienced clinician (physician in training). However, we would need to improve the instructions for completing the tests for better reproducibility because this tool is intended for rapid screening "at the bedside", and young doctors are often on the front line. Consequently, the following sentence has been added to the instruction for the bisection test: "*. . .the form must be placed exactly in front of the patient and on a table without distractors*."

The determination of the MCID was beyond our expectations because both estimation methods yielded almost identical values, which is relatively rare. The pathological threshold below which cognitive impairment is likely was quite high, 35/36, which indicates that the CASP is an "easy" scale. However, in this part of the calculations, the number of stroke patients without cognitive impairment was quite small as compared with other patients, and this value should be confirmed with a larger population. In contrast, we prioritized the sensitivity of this screening test; otherwise the threshold would have been lower. Several control participants had a score below this value, which may be a surprise. We had included these participants for the divergent validity check, and for this, they did not need to undergo a full cognitive assessment before enrolment. We simply ensured no known cognitive impairments in medical records. Thus, the CASP likely detected some participants as having early-onset cognitive impairment that had not been previously explored. The threshold of 34/36 in the first month post-stroke allows us to differentiate patients who will have an excellent medium-term evolution (between V1 to V4) from other patients. This can be useful in establishing the prognosis and for planning management, as was pointed out with the Birmingham Cognitive Screen [28].

Table 4
Predictive thresholds. Sensitivity, specificity and Youden index of the CASP at visit 1 (CASP V1) as compared with the CNA at visit 4 (CNA V4).

CASP V1 Potential cut-off	Sensitivity	Specificity	Youden Index
31.0	0.56	0.90	0.47
31.5	0.59	0.90	0.50
32.0	0.65	0.81	0.47
32.5	0.69	0.81	0.50
33.0	0.71	0.63	0.35
33.5	0.72	0.63	0.36
34.0	0.80	0.54	0.34
34.5	**0.83**	**0.54**	**0.38**

4.1. Interpretation

The CASP is not free of flaws. The main one is certainly that it cannot be administered in the presence of severe comprehension deficiencies because the patient may not understand instructions, but this limitation is common to most cognitive assessment tests. Furthermore, although to our knowledge, the CASP is the only test that offers a temporal orientation item that does not rely on language, we did not find a non-verbal equivalent of the spatial orientation test ("What country/city are we in?", "What floor of the building?" etc.). However, patients without aphasia could possibly be invited to give an oral response to that item.

The CASP is not the only clinical scale specifically designed for post-stroke. To the best of our knowledge, one of the oldest (2009) is the Brief Neuropsychological Screening test [29], which is very rapid (5−10 min) but does not explore temporal orientation and is only available in Italian. The Birmingham Cognitive Screen [30] was proposed in 2012 for assessing apraxia and determining functional prognosis after stroke. Its usefulness for global cognitive assessment and prognosis was demonstrated in 2015 [30]. Thus, the tool is interesting and comprehensive, but it takes about 1 h to complete, which we felt was too long for our purposes. Moreover, it is not free of charge. We found 5 other stroke-specific batteries in the literature, all published in 2015 or later but containing a variable number of verbal items to assess functions other than language: the Brief Memory and Executive Test [31], the Oxford Cognitive Screen [32], the Mild Vascular Cognitive Impairment assessment tool [33] and the Northwick Park Examination of Cognition [34].

In our opinion, the most interesting result of the previous CASP studies was that the severity of language impairment affected the CASP scores for items (other than language) much less than did the MMSE or MoCA (this last result having been established for the French and Italian versions of the CASP) [3,5]. However, the CASP cannot be used with severe oral comprehension disorders, the limit of feasibility being set to a score of <3/5 for the ASRS [8]. The face and content validity of the French version of the CASP have been established [3,4]. The Italian and Korean versions of the CASP showed good psychometric properties [5,6]. Persian and Chinese versions are being validated, after minor cultural adaptation (data not yet published).

4.2. Study limitations

The first limitation of this study was the absence of a standardized CNA across the 5 centers. However, we thought that imposing a single assessment battery such as that proposed by Hachinski et al. [35] on 5 university services would be less effective than letting experienced neuropsychologists choose tests best suited to the patient's

impairments, from among those they were used to administering. There were some deviations from the protocol in our study. For example, some patients were included beyond the planned 2 months post-stroke (up to 90 days) and some assessments for reproducibility were > 4 days apart (up to 6 days). These protocol deviations may have underestimated the responsiveness and reproducibility of the CASP, but the results were still satisfactory. Another limitation of the study is the sample size, which did not guarantee the validation or the determination of pathological thresholds in different age and educational groups or in left/right stroke patients. This would justify the continued inclusion of control and stroke patients. Finally, we had 201 patients at V1 and 93 at V4, which could indicate attrition bias. Actually, those who left the units before V4 were comparable to others in terms of cognitive impairments but less affected physically and therefore required a shorter stay.

4.3. Generalisability

From our work and clinical experience, the CASP would also be useful in the chronic post-stroke phase, but we do not know if it could also be used in the early phase. In an intensive care unit, patients are typically lying down, which is already a feasibility issue and requires adaptation. In addition, some patients would probably not be able to be assessed by the CASP at one time because of fatigability.

Finally, this test should not be given more prominence than it was designed for: rapid screening for post-stroke cognitive impairment by less experienced examiners (e.g., young residents). For this reason, it should be sensitive rather than specific. If the screening is positive, the patient must be assessed more closely by a neuropsychologist or speech-language pathologist to organize management. After a few years of using this test, within this framework, the CASP would certainly have its place as an alternative to other rapid cognitive assessment batteries.

5. Conclusions

These results are encouraging and show that the CASP is a useful screening tool in daily practice in the subacute phase after stroke. Patients with a pathological score should be referred to a neuropsychologist or speech-language therapist for full assessment and treatment.

Supplemental material:
- English CASP (on-line)
- French CASP (on-line)

Declaration of Competing Interest

None declared.

Acknowledgments

We thank all the young and experienced physicians, therapists and patients for their kind participation in this study.

Funding

French National “Projet Hospitalier de Recherche Clinique” (PHRC) 2012.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.rehab.2021.101594.

References

[1] Folstein MF, Folstein SE, McHugh PR. Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–98. doi: 10.1016/0022-3956(75)90026-6.
[2] Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005;53:695–9. doi: 10.1111/j.1532-5415.2005.53221.x.
[3] Barnay JL, Wauquiez G, Bonnin-Koang HY, Anquetil C, Pérennou D, Piscicelli C, et al. Feasibility of the cognitive assessment scale for stroke patients (CASP) vs. MMSE and MoCA in aphasic left hemispheric stroke patients. Ann Phys Rehabil Med 2014;57:422–35. doi: 10.1016/j.rehab.2014.05.010.
[4] Benaim C, Barnay JL, Wauquiez G, Bonnin-Koang HY, Anquetil C, Pérennou D, et al. The cognitive assessment scale for stroke patients (CASP) vs. MMSE and MoCA in non-aphasic hemispheric stroke patients. Ann Phys Rehabil Med 2015;58:78–85. doi: 10.1016/j.rehab.2014.12.001.
[5] Crivelli D, Angelillo MT, Grippa E, Colucci A, Nardulli R, Balconi M. When is a novel psychometric measure needed? A preliminary analysis regarding the cognitive assessment for stroke patients (CASP) battery compared with MMSE and MoCA. Appl Neuropsychol Adult 2018;25:410–6. doi: 10.1080/23279095.2017.1320556.
[6] Park KH, Lee HW, Park KB, Lee JY, Cho AR, Oh HM, et al. The Korean version of the cognitive assessment scale for stroke patients (K-CASP): a reliability and validity study. Ann Rehabil Med 2017;41:362–75. doi: 10.5535/arm.2017.41.3.362.
[7] Mazaux JM, Orgogozo JM. Échelle d'évaluation de l'aphasie (BDAE). Editions Scientifiques et Psychologiques, Issy-les-Moulineau 1982.
[8] Reinvang I, Graves R. A basic aphasia examination: description with discussion of first results. Scand J Rehabil Med 1975;7:129–35.
[9] van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. Stroke 1988;19:604–7. doi: 10.1161/01.str.19.5.604.
[10] Stern RA, Arruda JE, Hooper CR, Wolfner GD, Morey CE. Visual analogue mood scales to measure internal mood state in neurologically impaired patients: description and initial validity evidence. Aphasiology 1997;11:59–71. doi: 10.1080/02687039708248455.
[11] Nespoulos J., Lecours A., Lafond D., Lemay M., Puel M., Joanette Y., et al. Protocole Montréal-Toulouse d'examen linguistique de l'aphasie: MT-86 module standard initial, M1b(2e édition révisée par Renée Béland et Francine Giroux). 1992.
[12] Schenkenberg T, Bradford DC, Ajax ET. Line bisection and unilateral visual neglect in patients with neurologic impairment. Neurology 1980;30:509–17. doi: 10.1212/wnl.30.5.509.
[13] Gauthier L, Dehaut F, Joanette Y. The bells test: a quantitative and qualitative test for visual neglect. Int J Clin Neuropsychol 1989;11:49–54.
[14] Kaiser HF. The application of electronic computers to factor analysis. Educ Psychol Meas 2016. doi: 10.1177/001316446002000116.
[15] McElhone K, Abbott J, Teh LS. A review of health related quality of life in systemic lupus erythematosus. Lupus 2006;15:633–43. doi: 10.1177/0961203306071710.
[16] Thompson B. Exploratory and confirmatory factor analysis: understanding concepts and applications. Washington, DC: American Psychological Association; 2004.
[17] Jackson DL, Gillaspy JA, Purc-Stephenson R. Reporting practices in confirmatory factor analysis: an overview and some recommendations. Psychol Methods 2009;14:6–23. doi: 10.1037/a0014694.
[18] Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: a review. J Educ Res 2006;99:323–38. doi: 10.3200/JOER.99.6.323-338.
[19] MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods 1996;1:130–49. doi: 10.1037/1082-989X.1.2.130.
[20] Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model Multidiscip J 1999;6:1–55. doi: 10.1080/10705519909540118.
[21] Kline RB. Principles and practice of structural equation modeling. Guilford Publications; 2011.
[22] Rudner LM, Getson PR, Knight DL. Biased item detection techniques. J Educ Stat 1980;5:213–33. doi: 10.2307/1164965.
[23] Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16:297–334. doi: 10.1007/BF02310555.
[24] Enderlein G, Fleiss JL. The design and analysis of clinical experiments Wiley, New York – Chichester – Brislane – Toronto – Singapore 1986, 432 S., £38.35 Biom J 1988;30 304–304. doi: 10.1002/bimj.4710300308.
[25] Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol 2003;56:395–407. doi: 10.1016/s0895-4356(03)00044-1.
[26] Make B. How can we assess outcomes of clinical trials: the MCID approach. COPD 2007;4:191–4. doi: 10.1080/15412550701471231.
[27] Wells G, Beaton D, Shea B, Boers M, Simon L, Strand V, et al. Minimal clinically important differences: review of methods. J Rheumatol 2001;28:406–12.
[28] Bickerton WL, Riddoch MJ, Samson D, Balani AB, Mistry B, Humphreys GW. Systematic assessment of apraxia and functional predictions from the Birmingham cognitive screen. J Neurol Neurosurg Psychiatry 2012;83:513–21. doi: 10.1136/jnnp-2011-300968.
[29] Lunardelli A, Mengotti P, Pesavento A, Sverzut A, Zadini A. The brief neuropsychological screening (BNS): valuation of its clinical validity. Eur J Phys Rehabil Med 2009;45:85–91.

[30] Bickerton WL, Demeyere N, Francis D, Kumar V, Remoundou M, Balani A, et al. The BCoS cognitive profile screen: utility and predictive value for stroke. Neuropsychology 2015;29:638–48. doi: 10.1037/neu0000160.
[31] Brookes RL, Hollocks MJ, Khan U, Morris RG, Markus HS. The brief memory and executive test (BMET) for detecting vascular cognitive impairment in small vessel disease: a validation study. BMC Med 2015;13:51. doi: 10.1186/s12916-015-0290-y.
[32] Demeyere N, Riddoch MJ, Slavkova ED, Bickerton WL, Humphreys GW. The oxford cognitive screen (OCS): validation of a stroke-specific short cognitive screening tool. Psychol Assess 2015;27:883–94. doi: 10.1037/pas0000082.
[33] Oh HS, Kim JS, Shim EB, Seo WS. Development and clinical validity of a mild vascular cognitive impairment assessment tool for Korean stroke patients. Asian Nurs Res 2015;9:226–34. doi: 10.1016/j.anr.2015.04.005.
[34] Williams PM, Johnson C, Swan S, Barber C, Murphy P, Devine J, et al. The northwick park examination of cognition: a brief cognitive assessment tool for use in acute stroke services. Int J Ther Rehabil 2016;23:314–22.
[35] Hachinski V., et al. National institute of neurological disorders and stroke–Canadian stroke network vascular cognitive impairment harmonization standards n.d. 10.1161/01.STR.0000237236.88823.47. 2020