



HAL
open science

Additional validation study and French cross-cultural adaptation of the Pediatric Stroke Outcome Measure–Summary of Impressions (PSOM-SOI)

Marie-Doriane Morard, Mickaël Dinomais, Kim Bull, Pascal Rippert, Mathilde Chevignard, Gabrielle Deveber, Stéphane Chabrier, Carole Vuillerot

► To cite this version:

Marie-Doriane Morard, Mickaël Dinomais, Kim Bull, Pascal Rippert, Mathilde Chevignard, et al.. Additional validation study and French cross-cultural adaptation of the Pediatric Stroke Outcome Measure–Summary of Impressions (PSOM-SOI). *Annals of Physical and Rehabilitation Medicine*, 2019, pp.S1877-0657(19)30189-7. 10.1016/j.rehab.2019.10.010 . hal-02510176

HAL Id: hal-02510176

<https://univ-angers.hal.science/hal-02510176>

Submitted on 24 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Additional validation study and French cross-cultural adaptation of the Pediatric Stroke Outcome Measure–Summary of Impressions (PSOM-SOI)

Marie-Doriane Morard, MD^{1,2}, Mickael Dinomais, MD, PhD^{3,4}, Kim Bull, BSc, MSc, PhD⁵, Pascal Rippert, PhD⁶, Mathilde Chevignard, MD, PhD,^{7,8} Gabrielle deVeber, MD⁹, Stéphane Chabrier, MD, PhD^{2,10}, Carole Vuillerot, MD, PhD^{1,11}; AVCnn Study Group

¹ Hospices Civils de Lyon, Hôpital Femme-Mère-Enfant, L'Escale, Département de Médecine Physique et de Réadaptation Pédiatrique, F-69500 Bron, France

² INSERM et Univ Lyon, UMR 1059 SAINBIOSE, F-42023 Saint-Étienne, France.

³ CHU d'Angers, Département de Médecine Physique et de Réadaptation, Université Angers F-49933 Angers, France

⁴ Laboratoire Angevin de Recherche en Ingénierie des systèmes (LARIS)-EA7315, Université Angers, F-49933 Angers, France.

⁵ Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, United Kingdom

⁶ Hospices Civils de Lyon, Pôle de Santé Publique, F-69003 Lyon, France

⁷ Hôpitaux de Saint Maurice, Département de Rééducation des Pathologies Neurologiques Acquises de l'enfant, F-94410, Saint Maurice, France

⁸ Sorbonne Université, UPMC Univ Paris 06, UMR 7371, UMR_S 1146, LIB, F-75005, Paris, France

⁹ Hospital for Sick Children, Child Health Evaluative Sciences Program, Division of Neurology, 555 University Ave, Toronto, Ontario M5G1X8, Canada

¹⁰ CHU Saint-Étienne, Département de Médecine Physique et de Réadaptation Pédiatrique, Centre National de Référence de l'AVC de l'enfant, INSERM CIC1408, F-42055 Saint-Étienne, France

¹¹ CNRS UMR 5558, Laboratoire de Biométrie et Biologie Evolutive, Equipe Bio statistique Santé, F-69310 Pierre-Bénite, France; Université Lyon I, F-69100, Villeurbanne, France; Université de Lyon, F-69000 Lyon, France.

The AVCnn Group: Stéphane Darteyre, MD, PhD (Univ Saint-Étienne, data collection and analysis); Céline Dégano, MSc (Univ Saint-Étienne, data collection and analysis); Matthieu Delion, MD, PhD (Univ Angers, data collection and analysis); Johanna Deron, MSc (CHU Saint-Étienne, language assessment); Gérard Dray, PhD (Mines Alès, statistical analysis supervision); Laure Drutel, MSc (CHU Saint-Étienne, language assessment); Samuel Groeschel, MD (University Children's Hospital Tübingen, data collection and analysis); Lucie Hertz-Pannier, MD, PhD (CEA, imaging study); Béatrice Husson, MD (CHU Bicêtre, imaging study); Manoëlle Kossorotoff, MD, PhD (Hôpital Universitaire Necker-Enfants Malades, data collection and analysis); Leila Lazaro, MD (CH Côte Basque, data collection); Jérémie Lefranc, MD (CHU Brest, data collection); Sylvie Nguyen The Tich, MD, PhD (CHU Lille, study design and data collection and analysis); Emeline Peyric, MSc (CHU Saint-Etienne, cognitive assessment); Emilie Presles, MSc (CHU Saint-Etienne, statistical plan); Magaly Ravel, MD (Univ Saint-Étienne, data collection and analysis); Cyrille Renaud, PhD (CHU Saint-Étienne, study coordinator and data collection and analysis); and Carole Vuillerot, MD, PhD (CHU Lyon, data collection and analysis).

Corresponding author:

Dr. Marie-Doriane Morard

L'Escale, Service central de rééducation, Hôpital Femme Mère Enfant, F-69677, Bron,
France.

Phone: +33 (0)472129491 fax: +33 (0)427856790, e-mail: marie-doriane.morard@chu-lyon.fr

Additional validation study and French cross-cultural adaptation of the Pediatric Stroke Outcome Measure–Summary of Impressions (PSOM-SOI)

Abstract

Background. The Pediatric Stroke Outcome Measure–Summary of Impressions (PSOM-SOI) measures neurological function across right and left sensorimotor domains (Item A), language production (Item B), language comprehension (Item C), and cognition/behavior (Item D).

Objective. This study was a cross-cultural adaptation into French of the PSOM-SOI and an assessment of its reliability and limitations of use.

Material and Methods. The translation and adaptation of the PSOM-SOI was followed by the assessment of its reliability in a cohort of 69 children with diagnosed acute neonatal arterial ischemic stroke. Three independent raters retrospectively scored the PSOM-SOI based on data from in-person neurological examination and results of standardized tests performed at age 7 in the cohort database. Comparison 1 (C1) involved a less experienced rater and an experienced rater and comparison 2 (C2) involved 2 experienced raters. Inter-rater reliability (IRR) was measured with Kappa coefficients.

Results. The cross-cultural adaptation was easily performed, and no rater had difficulties using the French PSOM-SOI. The IRR was better in C1 than C2. For Item A, the agreement in C1 ($\kappa=0.47$) and C2 ($\kappa=0.44$) was moderate. The C1 agreement was substantial for Items B ($\kappa=0.71$) and C ($\kappa=0.70$); the C2 agreement was fair for Item B ($\kappa=0.23$) and slight for Item C ($\kappa=0.16$). For Item D, the agreement was moderate in C1 ($\kappa=0.52$) and fair in C2 ($\kappa=0.35$). In all but one comparison, agreement or minor disagreement (≤ 0.5 points) was obtained for more than 90% of the item scores. Regarding the total score, agreement for normal function (≤ 0.5) versus abnormal function (> 0.5) was achieved for 90% in C1 and 67% in C2.

Conclusion. The IRR of the French PSOM-SOI gave variable results depending on the item and rater's experience, but the extent of disagreements was minor for individual items and total score. Additional prospective validation studies using the French PSOM-Short Neurological Exam to score the PSOM-SOI are needed. A dichotomized total score (cutoff ≤ 0.5) could be used to define normal function versus poor outcome.

Keywords: outcome measure; ischemic stroke; pediatric; psychometrics, reliability

Introduction

With a birth prevalence of 6 to 17/100,000 in term/near-term newborns, neonatal arterial ischemic stroke is one of the most frequent forms of pediatric stroke [1–3]. Persisting impairment resulting from stroke in neonates and children includes motor, language, cognitive, and behavioral disorders, with subsequent limitations in activity and participation [4–8].

To date, pediatric management of stroke is mainly based on low levels of evidence. Furthermore, to estimate and compare outcomes, there is a need for a standardized tool that is valid, reliable, sensitive to change, and easy to use in children. Many standardized scales are used for adults after stroke, but they do not take into account children's progressive abilities with maturation and neurodevelopment. In the pediatric field, some reliable classification systems have been developed specifically for cerebral palsy to classify manual abilities (Manual Ability Classification System [9]) or hand function (Bimanual Fine Motor Function [BFMF] [10,11]) and more generally gross motor function (Gross Motor Function–Classification System [GMF-CS]) [12]. However, other clinical tools were developed to evaluate change in gross motor function (Gross Motor Function Measure [12]) or measure how the affected hand and arm are used in bimanual performance (Assisting Hand Assessment [13]). The other available outcome measures used after pediatric stroke concern

the neuropsychological field, such as the Wechsler Intelligence Scale for Children (WISC) [14].

The Pediatric Stroke Outcome Measure (PSOM) was developed by deVeber *et al.* [5]. Recently entitled the PSOM–Short Neurological Examination (PSOM-SNE), it includes a final scoring section, the PSOM–Summary of Impression (PSOM-SOI). According to the International Classification of Functioning, Disability and Health–Children and Youth Version, the PSOM-SNE assesses body functions and structures with some items that assess activity [15] and is currently the only scale to assess neurological function in children after stroke. It is widely used in clinical research because of its good construct validity and reliability [16]. Thus, its translation and cross-cultural adaptation and reliability research are mandatory to enable its widespread use in international clinical trials and in clinical practice in French patients. However, in clinical practice one can assume that the PSOM-SOI, which comprises 5 subscale scores that reflect relevant functional impairments, is most often scored directly from the neurological examination or from medical dictation rather than using (as recommended) the PSOM-SNE.

Currently the PSOM (SNE and SOI) have been translated into Portuguese and Dutch. The aim of the present study was the cross-cultural adaptation of the PSOM-SOI into French and to assess its reliability in order to evaluate its possible limitation of use in clinical practice.

Methods

This study was conducted in 2 stages. First, the PSOM-SOI was translated and adapted into French (from France) by several multi-lingual physicians according to published methodology for cross-cultural adaptation [17,18]. Second, a reliability study was performed by 3 independent raters with data for children enrolled in the AVCnn cohort (*Accident Vasculaire*

Cérébral du nouveau-né, i.e., neonatal stroke; EudraCT no. 2010-A00329-30; ClinicalTrials.gov NCT02511249).

The PSOM-SOI

The PSOM-SNE consists of 115 developmentally ordered test items in the areas of behavior, mental status, cranial nerves, and motor and sensory, cerebellar, and gait functions [5,16,19]. After completing the PSOM examination, a deficit severity score is summarized in the PSOM-SOI, which consists of 5 domains summarizing the different items tested previously in the PSOM-SNE: right sensorimotor domain (labeled Item A in the current study), left sensorimotor domain (Item A), language production domain (Item B), language comprehension domain (Item C), and cognitive and behavioral performance domain (Item D). Each of these 5 domains is evaluated and a score is assigned: each child is evaluated as having no deficit (scored 0), mild deficit and normal function (scored 0.5), moderate deficit and decreased function (scored 1), or severe deficit with at least one missing function (scored 2). The PSOM-SOI total score is the sum of the 5 subscale scores and ranges from 0 (no deficit) to 10 points (maximum deficit) [5,16]. In the seminal paper, “good outcome/normal function” was defined by a total score ≤ 0.5 (5).

Cross-cultural adaptation of the PSOM-SOI

Additional material 1 describes the translation-adaptation process following recommendations proposed by Guillemin et al. [18] to obtain semantic, idiomatic, experiential, and conceptual equivalence in translation-adaptation by using back-translation techniques and committee review. First, 2 independent English-to-French translations were completed. The 2 versions were compared and reconciliated, leading to a first French version.

During this process, the cultural sensitivity of the terms and the phraseology were carefully considered. Then a back-translation of the first French version was completed and compared with the initial English version by a third independent English bilingual translator. The French version of the PSOM-SOI was finalized after the 2 native French-speaking investigators identified and resolved misinterpretations of the English language by comparing their consensus translation with the original English PSOM-SOI, followed by phone call discussions with the native English-speaking French back translator.

Assessment of the French PSOM-SOI in children with neonatal stroke

Patients and data collection

The AVCnn cohort includes 100 term newborns with neonatal arterial ischemic stroke (i.e., acutely diagnosed, excluding presumed perinatal ischemic stroke) included between November 2003 and October 2006 [20,21]. In 2010, families were asked to participate in a cross-sectional assessment at age 7 (AVCnn^{7years}) [21]. This study was conducted in accordance with international ethical standards and the Declaration of Helsinki and was approved by the regional ethics committee in May 2010. Informed consent was obtained from each participant [21].

Several validated outcome measures were used for each child [21]. The AVCnn^{7years} assessment took place over a full day in a medical setting close to the family residence and consisted of 1) a face-to-face clinical neurologic examination completed by one physician experienced in childhood neurodevelopment (specialized in neuropsychiatry or in pediatric neurorehabilitation, representing the AVCnn study principal investigator [SC]) and including the GMF-CS and BFMF; 2) a formalized language assessment completed by a speech therapist including the standardized French assessment of oral language *Nouvelles Épreuves pour l'Examen du Langage* (N-EEL; New Test for Assessment of Language) [22]; and 3)

psychometric testing completed by a neuropsychologist (WISC-IV). The assessment results were entered into the AVCnn database, which included a free-text field containing additional clinical observation during the assessment.

Reliability testing of the French PSOM-SOI

For each child, 3 physician raters (one specialized [SC] in neuropediatrics and 2 [MD, MDM] specialized in pediatric neurorehabilitation) independently scored the subscales of the French PSOM-SOI based on the results of the above assessments recorded in the AVCnn^{7years} database. Two raters (MD, SC) were physicians with more than 10 years of experience in neurological assessment of children with stroke and good knowledge of the PSOM. The third rater (MDM) was a junior physician with 1 year of experience in childhood neurological assessment. Inspired by the methodology used by the original authors, we conducted a reliability study of the PSOM-SOI scores abstracted from medical dictations. We used the outcome measures of the AVCnn^{7years} database for scoring the PSOM-SOI for each domain. Item A, “sensorimotor deficit,” was scored based on the results of the GMF-CS and the BFMF derived from the neurological examination and encompassed both right and left sensorimotor scores (scored 0-4). Item B, “deficit in language production,” was scored based on the sub-tests of the N-EEL: “*Vocabulaire 1*” (Vocabulary 1), “*Vocabulaire 2*”, “*Vocabulaire 1 avec ébauche*” (Beginning of the vocabulary), “*Vocabulaire 2 avec ébauche*”. Item C, “deficit in language–comprehension” was scored based on the subtests of the N-EEL “*Lexique*” (lexicon), “*Lexique 2*”, “*Abstrait–Différence*” (abstract-difference), “*Abstrait–Position*” (abstract-position). Item D, “deficit in cognition or behavior,” was scored based on the WISC-IV. Items B, C and D were scored from 0 to 2. The total score was calculated by summing the 4 individual item scores and ranged from 0 to 10.

Statistical analysis

Quantitative variables are described with mean (SD). Categorical variables are described with frequencies (%). The inter-rater reliability (IRR) analysis was first conducted on the ratings of the less experienced rater as compared with the experienced rater [MDM vs MD; Comparison 1 (C1)] and the second on the ratings of the 2 experienced raters [MD vs SC; Comparison 2 (C2)].

We conducted reliability testing in 3 ways. First, we examined IRR for each of the 4 items and the total score. In all, 10 comparisons were calculated between the 3 raters: IRR across the 4 items (8 comparisons) and IRR for the total score (2 comparisons). Second, we assessed the percentage of agreement between raters, across items, and the total score. Third, we assessed the agreement between raters for the original dichotomized total score classified as “good outcome/normal function” (i.e., total score 0 or 0.5) versus “abnormal function” (total score > 0.5) (5). The strength of agreement between the different scores, two-by-two, was assessed by Cohen’s Kappa coefficient (κ), a measure of chance agreement, interpreted according to Landis and Koch as slight (< 0.20), fair (0.21-0.40), moderate (0.41-0.60), substantial (0.61-0.80), or almost perfect (> 0.81) [23]. The κ coefficients are provided with their 95% confidence intervals (CIs). The range of disagreement and the agreement of the results were also evaluated. Statistical analyses were performed with IBM SPSS v20.0.

Results

Cross-cultural adaptation of the PSOM-SOI (additional material 2)

Only minor and easily correctable disagreements were observed during the translation-adaptation process. For example, the expression “missing function” was initially translated as “*perte de fonction*”. Because the back translation was “loss of function”, the reconciliation work enabled correction of this discrepancy by using “*fonction manquante*”. Another

example: “mild but no impact on function” was initially translated as “*léger sans déficit fonctionnel*” and the back translation was “slight without affecting function”. During the reconciliation step, this back translation could be confusing with the lack of impairment despite a mild deficit. It was eventually translated as “*léger sans repercussion sur la fonction*”. “Language deficit – production” was initially translated as “*langage déficit – expression*” but the reconciliation work allowed for keeping the French term “*production*”. None of the raters had difficulties understanding or using the French PSOM-SOI.

Assessment of the French PSOM-SOI in children with neonatal stroke

Patients and data collection (Fig. 1)

Overall, 73 of the 100 children initially enrolled underwent the 7-year assessment; 4 had incomplete N-EEL data. Thus, data for 69 children (mean age 6.8 years) were analyzed for the reliability study. Their main characteristics are in Table 1. Additional clinical information about the neurological examination was available in the free-text field for 24 children (35%).

Reliability testing of the French PSOM-SOI

Percentage of agreement (Tables 2 and 3)

For Item A, we found disagreement of 2.5 points for only 1 patient (in C2), disagreement of 1.5 points for 4 patients (6%) in C1 and in 2 patients (3%) in C2, disagreement of 1 point for 11 patients (16%) in C1 and 15 patients (22%) in C2. A disagreement of 0.5 points was most frequent for the other items, especially in C2 (Item B: 13% in C1 vs 45% in C2; Item C: 19% in C1 vs 39% in C2; Item D: 30% in C1 vs 44% in C2).

We observed agreement or minor disagreement (i.e., ≤ 0.5 points) for most patients: Item A, 78% in C1 and 92% in C2; Item B, 91% in C1 and 90% in C2; Item C, 97% in C1 and 81% in C2; and Item D, 96% in C1 and 97% in C2. For the total score, agreement was

reached for only 26% of patients in C1 and 17% in C2; however, we observed agreement or minor disagreement (≤ 0.5 points) for 69% in C1 and 38% in C2.

Agreement for dichotomized total score

Considering the dichotomization between “good outcome/normal function” (i.e., total score 0 or 0.5) versus “abnormal function” (total score > 0.5), 62/69 (90%) children were in the same category in C1 and 47/69 (67%) in C2.

IRR for items and total score (Table 4)

For Item A, IRR was moderate in both C1 and C2 ($\kappa=0.47$, 95% CI 0.34-0.60; $\kappa=0.44$, 0.31-0.58). For Items B and C, IRR was substantial in C1 ($\kappa=0.71$, 0.57-0.83; $\kappa=0.70$, 0.55-0.82). For the total score, IRR was slight in both C1 and C2 ($\kappa=0.19$, 0.08-0.30; $\kappa=0.09$, 0.01-0.17).

Discussion

In the present study, the original English version of the PSOM-SOI was cross-culturally adapted into French. The reliability of the French version was then evaluated by using data for children at age 7 from the database of the AVCnn cohort. We found agreement or minor disagreement across most individual item scores and the total score for the French PSOM-SOI.

Cross-cultural adaptation of the PSOM-SOI

The rigorous process of cross-cultural adaptation is mandatory to enable the widespread use of a tool in international clinical trials and clinical practice [24]. Even though a recent review of guidelines for cross-cultural adaptation of questionnaires could not arrive at a consensus [17], the methodology used in the present study is rigorous and in accordance with Guillemin

et al. [18]. In this process, the back-translation step and the reconciliation are indispensable to correct misinterpretations. To our knowledge, there is no tool to objectively measure the quality of the translation, but the committee review chosen for the present study involved an independent English translator as well as numerous multilingual health-professionals (additional material 1). We assume that this strengthened the quality of the cross-cultural adaptation.

Reliability testing of the French PSOM-SOI

The IRR across the 4 items of the French PSOM-SOI gave variable results depending on the item and the experience of the raters. Surprisingly, agreement between inexperienced versus experienced raters (C1) was better than between experienced raters (C2). The reasons for this may be multiple but may first indicate that the experienced rater (SC), who knew the cohort, also used knowledge of the patients for scoring.

Other hypotheses may be postulated when examining the scores in more detail. First, the choice to rely on existing validated scales to score the PSOM-SOI may have affected our IRR values. Indeed, for scoring the PSOM SOI, as near to the reality of clinical practice, the raters did not reach consensus on how to score from the validated scales used, which may explain the moderate IRR for C1 and C2 for Item A (sensorimotor). Indeed, the clinical examination of the sensorimotor domain of the PSOM-SOI measures additional motor details and sensory deficits that are not being captured by the GMF-CS and BFMF [10,25].

Otherwise, the percentage of agreement between C1 and C2 was close for Item A. For C1, disagreement was found for 4 children who were attributed a score of 0.5 by the experienced rater (MD) but a score of 2 by the other rater (MDM). This difference may be explained by the differential use of information contained in the free-text field.

The IRR was better for Items B and C (language), but discordance was observed between C1 and C2. As reported for the English PSOM, this discordance can likely be encountered in the language and cognitive areas, for which standard psychometric testing would be more sensitive and reliable than the neurological examination items of the PSOM-SNE [5]. In the prior PSOM study, the intraclass correlation coefficient was strong (> 0.7) for Item B (language production) but poor for Item C (language comprehension; < 0.2) [16]. For Item D (cognition/behavior), scoring was based on only] the WISC-IV psychometric measure; thus behavior was not considered, although this is an important aspect in such patients [16,26].

Despite the variable IRR and diversity in level of experience of the raters, it is reassuring to observe that the extent of disagreements was low. Indeed, for most children, we found agreement or minor disagreement (≤ 0.5 points) across PSOM-SOI items including the dichotomized total PSOM-SOI score. This makes sense because a total score ≤ 0.5 means normal function and thus no activity limitation according to the International Classification of Functioning, Disability and Health [15], even if the neurological examination shows a mild deficit. Hence, this classification of the total score is significant for the patient because a total score ≤ 0.5 in the PSOM-SOI was originally defined to dichotomize children with good neurological outcome/normal function versus poor neurological outcome [5].

Limitations

In the original PSOM paper, the neurological examinations performed by the 3 study neurologists were based on the standardized detailed PSOM-SNE, and the retrospective PSOM-SOI was completed based on their clinical dictation letters [5]. Therefore, the dictated neurological exams were more likely to be consistent. Here, in-person neurological examination was mostly performed by one rater (SC), with the completion of the PSOM-SOI

mainly based on the retrospective analysis of results of standardized tests pooled in the AVCnn database for raters MDM and MD.

Three comments can be made from these limitations. First, our results may be representative of the diversity of clinicians with different levels of experience. Second, the extent of disagreement for individual items and the total score may also indicate that the French version of the PSOM-SOI is not reliable in daily practice when it is directly scored from medical dictation. Finally, it strengthens the fact that using the total score in a dichotomic manner (with a cutoff ≤ 0.5 , as is planned in upcoming clinical trials, i.e., PASTA study NCT03249844) is a preferential approach to the use of the French PSOM-SOI.

The original PSOM-SOI also has inherent limitations related to its construction, including the lack of exploratory factor analysis that could affect its construct validity, although this is likely minimal in the dichotomization of normal versus abnormal function (5). The use of a mathematical sum to reflect the total score is simple but not likely valid because identical summary scores can reflect different clinical pictures. For example, a score of 0.5 in 2, 3 or 4 individual items would be dichotomized into an SOI score > 0.5 , that is, poor outcome, but classifying children without any activity limitation as poor outcome is not clinically meaningful according to the International Classification of Functioning, Disability and Health. An increasing number of international publications on stroke in children have used the total score of the PSOM-SOI in a dichotomized manner with the cutoff of 0.5 and to our knowledge, our paper is the first work on the metrologic justification of such a cutoff [27–29]. Thus, new schemes based on sub-scale scores that would accurately reflect and distinguish the full range of function and ability of children with stroke are under development (personal data [30]). This development will enable interpretation of the PSOM-SOI using the sub-scale scores rather than the total score.

Conclusion

After cross-cultural adaptation of the PSOM-SOI into French, we found variable results in terms of reliability as assessed by retrospective scoring. However, according to the aim of our work (i.e., reflecting the clinical practice of the PSOM-SOI sometimes directly scored from medical dictations), this does not restrict its use as an objective disease-specific measure of neurological recovery after pediatric stroke, notably when using dichotomized cutoff values (≤ 0.5 vs > 0.5) to define normal function versus poor outcome [14]. However, to ensure a sufficient level of reliability, the PSOM-SOI must be scored based on the standardized details of the PSOM-SNE.

This is a first step to advocate the cross-cultural translation into multiple languages of the full PSOM-SNE, which remains the only validated tool to assess outcome in the pediatric stroke population. Furthermore, international studies require this adaptation for inclusion of non-English-speaking patients. Thus, additional prospective validation studies are needed, and especially a study of reliability, using the French PSOM-SNE to score the French PSOM-SOI.

Figure legend

Figure 1. Flow chart of children included in the reliability study for the French Pediatric Stroke Outcome Measure–Summary of Impressions. N-EEL, *Nouvelles Épreuves pour l'examen du Langage* (standardized assessment of oral French); AVCnn cohort, *Accident Vasculaire Cérébral du nouveau-né*; AVCnn^{7years}, AVC cohort at 7 years follow-up

Acknowledgments. We thank Philip Robinson for help in manuscript preparation including critical review of study proposal; Emilie Presles, MSc, for reviewing the manuscript

and assistance with the references; and Eleni Panagiotakaki MD, PhD, for assisting in translating of the original PSOM into French.

Funding. The AVCnn study was funded by the *Ministère des Affaires Sociales et de la Santé* (PHRC nos. 0308052 and 1008026), CHU Saint-Étienne (AOL no. 085), and the following supporting organizations: Fondation Paralysie Cérébrale, Fondation Garches, Association des Paralysés de France, and Fondation de l’Avenir.

References

- [1] Grunt S, Mazenauer L, Buerki SE, Boltshauser E, Mori AC, Datta AN, et al. Incidence and outcomes of symptomatic neonatal arterial ischemic stroke. *Pediatrics* 2015;135:e1220-1228. doi:10.1542/peds.2014-1520.
- [2] Kirton A, deVeber G. Paediatric stroke: pressing issues and promising directions. *Lancet Neurol* 2015;14:92–102. doi:10.1016/S1474-4422(14)70227-3.
- [3] Darmency-Stamboul V, Cordier AG, Chabrier S. [Neonatal arterial ischemic stroke in term or near-term newborns: prevalence and risk factors]. *Arch Pediatr Organe Off Soc Francaise Pediatr* 2017;24:9S3–11. doi:10.1016/S0929-693X(17)30325-1.
- [4] Galvin J, Hewish S, Rice J, Mackay MT. Functional outcome following paediatric stroke. *Dev Neurorehabilitation* 2011;14:67–71. doi:10.3109/17518423.2010.547241.
- [5] deVeber GA, MacGregor D, Curtis R, Mayank S. Neurologic outcome in survivors of childhood arterial ischemic stroke and sinovenous thrombosis. *J Child Neurol* 2000;15:316–24.
- [6] Kolk A, Ennok M, Laugesaar R, Kaldoja M-L, Talvik T. Long-term cognitive outcomes after pediatric stroke. *Pediatr Neurol* 2011;44:101–9. doi:10.1016/j.pediatrneurol.2010.08.012.
- [7] Studer M, Boltshauser E, Capone Mori A, Datta A, Fluss J, Mercati D, et al. Factors affecting cognitive outcome in early pediatric stroke. *Neurology* 2014;82:784–92. doi:10.1212/WNL.0000000000000162.
- [8] Fuentes A, Deotto A, Desrocher M, deVeber G, Westmacott R. Determinants of cognitive outcomes of perinatal and childhood stroke: A review. *Child Neuropsychol J Norm Abnorm Dev Child Adolesc* 2016;22:1–38. doi:10.1080/09297049.2014.969694.
- [9] Eliasson A-C, Krumlinde-Sundholm L, Rösblad B, Beckung E, Arner M, Öhrvall A-M, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Dev Med Child Neurol* 2006;48:549–54. doi:10.1017/S0012162206001162.
- [10] Elvrum A-KG, Andersen GL, Himmelmann K, Beckung E, Öhrvall A-M, Lydersen S, et al. Bimanual Fine Motor Function (BFMF) Classification in Children with Cerebral Palsy: Aspects of Construct and Content Validity. *Phys Occup Ther Pediatr* 2016;36:1–16. doi:10.3109/01942638.2014.975314.

- [11] Beckung E, Hagberg G. Neuroimpairments, activity limitations, and participation restrictions in children with cerebral palsy. *Dev Med Child Neurol* 2002;44:309–16.
- [12] Russell DJ, Rosenbaum PL, Cadman DT, Gowland C, Hardy S, Jarvis S. The gross motor function measure: a means to evaluate the effects of physical therapy. *Dev Med Child Neurol* 1989;31:341–52.
- [13] Krumlinde-Sundholm L, Holmefur M, Kottorp A, Eliasson A-C. The Assisting Hand Assessment: current evidence of validity, reliability, and responsiveness to change. *Dev Med Child Neurol* 2007;49:259–64. doi:10.1111/j.1469-8749.2007.00259.x.
- [14] Engelmann KA, Jordan LC. Outcome measures used in pediatric stroke studies: a systematic review. *Arch Neurol* 2012;69:23–7. doi:10.1001/archneurol.2011.1015.
- [15] World Health Organization. International Classification of Functioning, Disability and Health - Children and Youth Version (ICF-CY). Geneva : World Health Organization , 2007 n.d.
- [16] Kitchen L, Westmacott R, Friefeld S, MacGregor D, Curtis R, Allen A, et al. The pediatric stroke outcome measure: a validation and reliability study. *Stroke J Cereb Circ* 2012;43:1602–8. doi:10.1161/STROKEAHA.111.639583.
- [17] Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol* 2015;68:435–41. doi:10.1016/j.jclinepi.2014.11.021.
- [18] Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993;46:1417–32.
- [19] Gordon AL, Ganesan V, Towell A, Kirkham FJ. Functional outcome following stroke in children. *J Child Neurol* 2002;17:429–34.
- [20] Chabrier S, Saliba E, Nguyen The Tich S, Charollais A, Varlet M-N, Tardy B, et al. Obstetrical and neonatal characteristics vary with birthweight in a cohort of 100 term newborns with symptomatic arterial ischemic stroke. *Eur J Paediatr Neurol EJPJN Off J Eur Paediatr Neurol Soc* 2010;14:206–13. doi:10.1016/j.ejpn.2009.05.004.
- [21] Chabrier S, Peyric E, Drutel L, Deron J, Kossorotoff M, Dinomais M, et al. Multimodal Outcome at 7 Years of Age after Neonatal Arterial Ischemic Stroke. *J Pediatr* 2016;172:156-161.e3. doi:10.1016/j.jpeds.2016.01.069.
- [22] Chevrier-Muller C, Plaza M. N-EEL : Nouvelles Epreuves pour l'Examen du Langage. Bayeux: ECPA Editions; 2001 n.d.
- [23] Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
- [24] Hunt SM, Alonso J, Bucquet D, Niero M, Wiklund I, McKenna S. Cross-cultural adaptation of health measures. European Group for Health Management and Quality of Life Assessment. *Health Policy Amst Neth* 1991;19:33–44.
- [25] Palisano RJ, Hanna SE, Rosenbaum PL, Russell DJ, Walter SD, Wood EP, et al. Validation of a model of gross motor function for children with cerebral palsy. *Phys Ther* 2000;80:974–85.
- [26] Greenham M, Anderson V, Hearps S, Ditchfield M, Coleman L, Mackay MT, et al. Psychosocial function in the first year after childhood stroke. *Dev Med Child Neurol* 2017. doi:10.1111/dmcn.13387.
- [27] Kirton A, Williams E, Dowling M, Mah S, Hodge J, Carlson H, et al. Diffusion imaging of cerebral diaschisis in childhood arterial ischemic stroke. *Int J Stroke Off J Int Stroke Soc* 2016;11:1028–35. doi:10.1177/1747493016666089.
- [28] Kirton A, Shroff M, Visvanathan T, deVeber G. Quantified corticospinal tract diffusion restriction predicts neonatal stroke outcome. *Stroke* 2007;38:974–80. doi:10.1161/01.STR.0000258101.67119.72.

- [29] Cooper AN, Anderson V, Greenham M, Hearps S, Hunt RW, Mackay MT, et al. Motor function daily living skills 5 years after paediatric arterial ischaemic stroke: a prospective longitudinal study. *Dev Med Child Neurol* 2019;61:161–7. doi:10.1111/dmcn.13915.
- [30] Slim M, Dlamini N, Fox C, Friefeld S, Westmacott R, Mohairir M, et al. Outcome Severity Assessment with the Pediatric Stroke Outcome Measure (PSOM): Validation of a Novel Classification System (abstract). *Ann Neurol* 2018;84(suppl 22):S406.

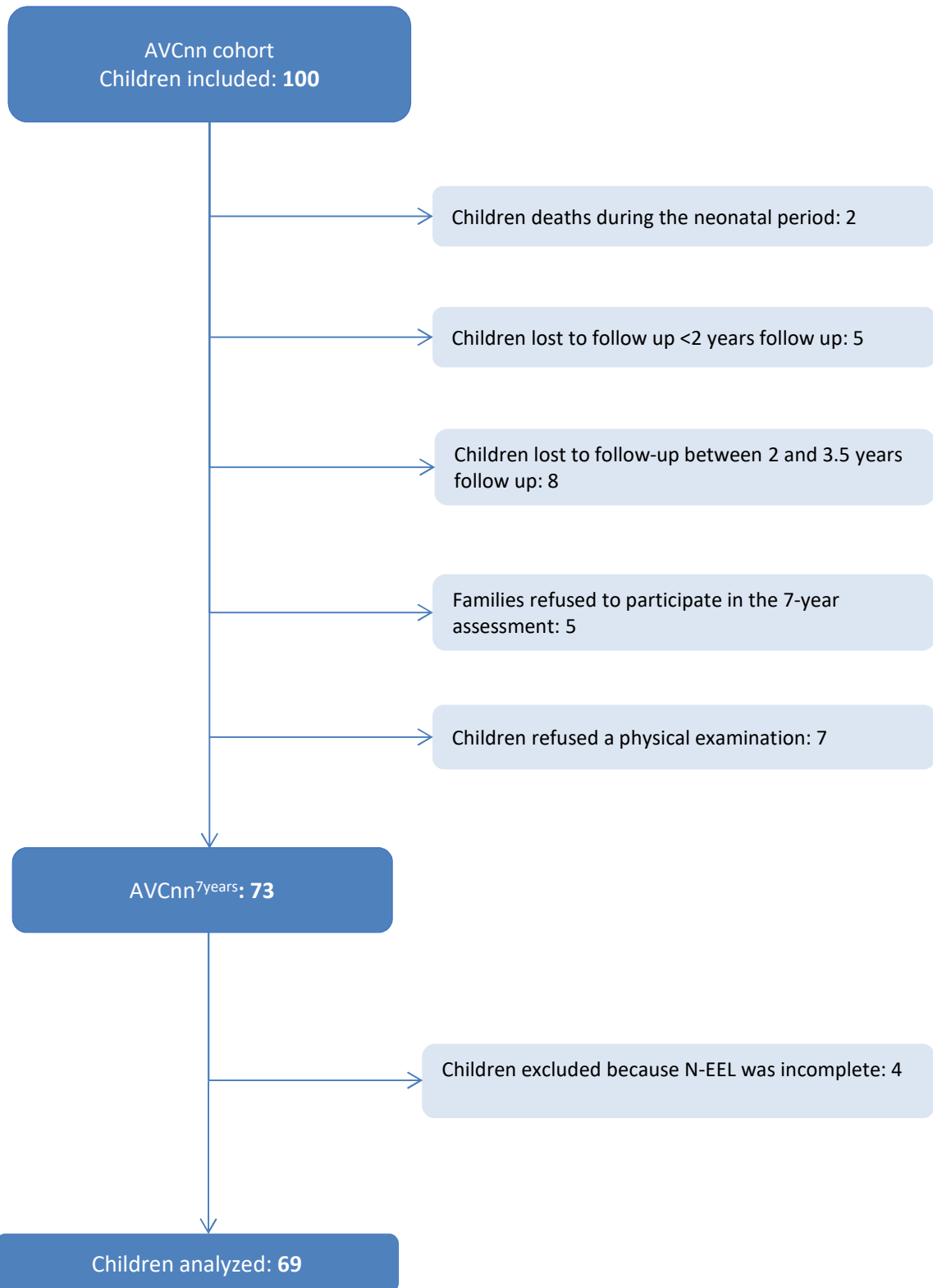


Table 1. Characteristics of children with neonatal stroke (n=69).

Mean age, years (SD)	6.8 (0.5)
Sex, male, n (%)	41 (59)
Side of stroke, n (%)	
Left	46 (67)
Right	20 (29)
Bilateral	3 (4)
School, n (%)	
Mainstream school	64 (93)
Inclusion in a mainstream school	2 (3)
Specialized school	3 (4)
Epilepsy, n (%)	9 (13)
GMF-CS score, n (%)	
I	62 (90)
II	7 (10)
BFMF score, n (%)	
0	44 (64)
1	10 (15)
2	11 (16)
3	2 (3)
No score	2 (3)
WISC-IV, median (range)	
Verbal Comprehension Index	99 (49-130)
Perceptual Reasoning Index	97.5 (49-119)
Working Memory Index	88 (49-133)
Processing Speed Index	100 (49-134)
N-EEL, mean (SD)	
Vocabulaire 1	47.9 (12.0)
Vocabulaire 2	33.2 (6.0)
Vocabulaire 1 avec ébauche	56.2 (11.1)
Vocabulaire 2 avec ébauche	36.5 (4.5)
Lexique 1	34.5 (2.1)
Lexique 2	19.2 (1.7)
Abstrait-Différence	6.7 (2.1)
Abstrait-Position	3.1 (1.3)

GMF-CS, Gross Motor Function Classification System; BFMF, Bimanual Fine Motor Function Classification; WISC-IV, Wechsler Intelligence Scale for Children-Fourth Edition; N-EEL, *Nouvelles épreuves pour l'examen du langage*; Vocabulaire 1, Vocabulary 1;

Vocabulaire 2, Vocabulary 2;Vocabulaire 1 avec ébauche, Beginning of vocabulary 1;
Vocabulaire 2 avec ébauche; Beginning of vocabulary 2; Lexique 1, Lexicon 1; Lexique 2,
Lexicon 2; Abstrait-Difference, Abstract-Difference; Abstrait-Position, Abstract-Position;
PSOM, Pediatric Stroke Outcome Measure

Table 2. Distribution of agreement and disagreement between the scores for each item of the French Pediatric Stroke Outcome Measure–Summary of Impressions (PSOM-SOI).

	Comparison 1						Comparison 2					
	Agreement			Disagreement			Agreement			Disagreement		
	Score	n	%	Score	n	%	Score	n	%	Score	n	%
Item A	0	34	49	-	-	-	0	38	55	-	-	-
	0.5	5	7	-	-	-	0.5	4	6	-	-	-
	1	4	6	-	-	-	1	2	3	-	-	-
	2	2	3	-	-	-	2	2	3	-	-	-
	-	-	-	-	-	-	2.5	0	0	-	-	-
	-	-	-	0 / 0.5	3	4	-	-	-	0 / 0.5	1	4
	-	-	-	0 / 1	2	4	-	-	-	0 / 1	6	26
	-	-	-	0.5 / 1	6	9	-	-	-	0 / 2.5	0	0
	-	-	-	0.5 / 2	4	6	-	-	-	0.5 / 1	5	22
	-	-	-	1 / 2	9	13	-	-	-	0.5 / 2	1	4
	-	-	-	-	-	-	-	-	-	1 / 2	9	39
-	-	-	-	-	-	-	-	-	1 / 2.5	1	4	
Total		45	65		24	35		46	67		23	33
Item B	0	21	30	-	-	-	0	21	30	-	-	-
	0.5	14	20	-	-	-	0.5	1	1	-	-	-
	1	10	15	-	-	-	1	5	7	-	-	-
	2	9	13	-	-	-	2	4	6	-	-	-
	-	-	-	0 / 0.5	5	7	-	-	-	0 / 0.5	21	30
	-	-	-	0 / 1	0	0	-	-	-	0 / 1	1	1
	-	-	-	0.5 / 1	4	6	-	-	-	0.5 / 1	10	15
	-	-	-	1 / 2	6	9	-	-	-	1 / 2	6	9
Total		54	78		15	22		31	45		38	55
Item C	0	20	29	-	-	-	0	22	32	-	-	-
	0.5	12	17	-	-	-	0.5	1	1	-	-	-
	1	16	23	-	-	-	1	4	6	-	-	-
	2	6	9	-	-	-	2	2	3	-	-	-
	-	-	-	0 / 0.5	7	10	-	-	-	0 / 0.5	18	26
	-	-	-	0 / 1	0	0	-	-	-	0 / 1	9	13
	-	-	-	0.5 / 1	6	9	-	-	-	0.5 / 1	9	13
-	-	-	1 / 2	2	4	-	-	-	1 / 2	4	6	
Total		54	78		15	22		29	42		40	58
Item D	0	21	30	-	-	-	0	21	30	-	-	-
	0.5	9	13	-	-	-	0.5	2	3	-	-	-
	1	9	13	-	-	-	1	9	13	-	-	-
	2	6	9	-	-	-	2	5	7	-	-	-
	-	-	-	0 / 0.5	16	23	-	-	-	0 / 0.5	25	36
	-	-	-	0.5 / 1	5	7	-	-	-	0.5 / 1	5	7
-	-	-	1 / 2	3	4	-	-	-	1 / 2	2	3	
Total		45	65		24	35		37	54		32	46

Comparison 1, between the experienced practitioner rater (MD) and less experienced practitioner rater (MDM); Comparison 2, between the 2 experienced raters (rater MD and rater SC)

Table 3. Distribution of agreements and disagreements with the score differences (between 2 scores) for the total score of the French PSOM-SOI.

	Comparison 1						Comparison 2					
	Agreement			Disagreement			Agreement			Disagreement		
	Score difference*	n	%	Score difference*	n	%	Score difference*	n	%	Score difference*	n	%
	0	18	26	-	-	-	0	12	17	-	-	-
	-	-	-	0.5	30	43	-	-	-	0.5	14	20
	-	-	-	1	17	25	-	-	-	1	22	32
	-	-	-	1.5	4	6	-	-	-	1.5	12	17
	-	-	-	2	0	0	-	-	-	2	6	9
	-	-	-	-	-	-	-	-	-	2.5	2	3
	-	-	-	-	-	-	-	-	-	3	1	1
Total score		18	26		51	74	-	12	17	-	57	83

* Difference between two scores; Comparison 1, between the experienced practitioner rater (MD) and less experienced practitioner rater (MDM); Comparison 2, between the 2 experienced raters (rater MD and rater SC)

Table 4. Agreement between raters and inter-rater reliability as judged by Cohen's kappa (κ) for each item and the total score of the PSOM-SOI.

Item	Comparison 1		Comparison 2	
	% Agreement	κ [95% CI]	% Agreement	κ [95% CI]
A	65	0.47 [0.34–0.60]	67	0.44 [0.31–0.58]
B	78	0.71 [0.57–0.83]	45	0.23 [0.09–0.37]
C	78	0.70 [0.55–0.82]	42	0.16 [0.04–0.29]
D	65	0.52 [0.36–0.67]	55	0.35 [0.21–0.50]
Total score	26	0.19 [0.08–0.30]	17	0.09 [0.01–0.17]

95% CI, 95% confidence interval