I, Patricia Vega-Fernandez, hereby submit this original work as part of the requirements for the degree of Master of Science in Clinical and Translational Research.

It is entitled: Cognitive Performance Scores for the Pediatric Automated Neuropsychological Assessment Metrics in Childhood-Onset Systemic Lupus Erythematosus

Student's name: Patricia Vega-Fernandez

This work and its defense approved by:

Committee chair: Erin Haynes, Dr.P.H.
Committee member: Dean Beebe, Ph.D.
Committee member: Hermine Brunner, M.D.
Committee member: Jun Ying, Ph.D.
Cognitive Performance Scores for the Pediatric Automated Neuropsychological Assessment Metrics in Childhood-Onset Systemic Lupus Erythematosus

A thesis submitted to the

Graduate School

of the University of Cincinnati

in partial fulfillment of the

requirements for the degree of

Master of Science

in Clinical & Translational Research

In the Department of Environmental Health
Division of Epidemiology & Biostatistics
of the College of Medicine

June, 2014

by

Patricia Vega-Fernandez

MD, Universidad del Cauca, April 2004

Committee Chair: Erin N. Haynes, DrPH
Abstract

Background/Objective: Patients with childhood-onset SLE (cSLE) may experience neuropsychiatric SLE (NPSLE) manifested as neurocognitive dysfunction (NCD). Formal neurocognitive testing (FNCT) is the most accepted method for diagnosing NCD. However, access is limited and it is costly and time-consuming. The Pediatric Automated Neuropsychological Assessment Metrics (PedANAM) is a computerized test battery that assesses multiple domains of cognitive performance. However, it is unclear how PedANAM-generated variables can be interpreted in a clinical setting as measures of NCD.

Our objective was to explore and initially test approaches to the calculation of a summary score (PedANAM Cognitive Performance Score (PedANAM-CPS)) to assess cognitive performance for the screening of NCD in cSLE with high sensitivity.

Methods: Two cohorts of subjects were analyzed. The development cohort included cSLE patients and controls that completed the PedANAM and FNCT at two research study visits 18 months apart. The validation cohort consisted of cSLE patients and controls recruited in a clinical setting who completed the PedANAM and Pediatric Perceived Cognitive Function-43 questionnaire (PedsPCF-43). Candidate PedANAM-CPSs were explored in the development cohort’s first visit via 3 statistical methods: 1) Simple-summary score: mean of all PedANAM subtest’s accuracy scores; 2) Logit-based score developed by logistic regression modeling; 3) PCA-based score derived from Principal Component Analysis (PCA). The latter 2 methods assigned in a different way a statistical weight to each subtest accuracy score. Receiver
operating characteristic curve analysis was used to assess the accuracy of candidate scores as predictors of NCD as suggested by FNCT or by PedsPCF-43 in the study cohorts. A previously proposed performance score (Inclusive-score) was also analyzed.

**Results:** A total of 166 subjects were studied, including 108 cSLE patients. All candidate PedANAM-CPSs were closely related to one another and significantly differentiated subjects based on the presence of NCD as measured by FNCT or PedsPCF-43. Logit-based and Inclusive scores exhibited good performance detecting NCD when applied to the second visit of the development and the validation cohorts.

**Conclusion:** Candidate PedANAM-CPS showed good construct and criterion validity, with a Logit-based score and Inclusive score performing somewhat better for discriminating subjects based on the presence or absence of NCD. The PedANAM-CPS may be a useful tool to summarize cognitive performance in an effort to assist with the screening for NCD in cSLE. Additional studies will confirm its overall accuracy and clinical implications.
# Table of Contents

Abstract ......................................................................................................................................................... ii

Table of Contents .......................................................................................................................................... v

Introduction .................................................................................................................................................. 1

Patients and Methods ................................................................................................................................... 3
  
  Measurements .......................................................................................................................................... 4
  
  Statistical Analysis ..................................................................................................................................... 6

Results ........................................................................................................................................................... 9

Discussion .................................................................................................................................................... 12

Bibliography ................................................................................................................................................ 15

Appendices .................................................................................................................................................. 18

  Table 1 .................................................................................................................................................... 19

  Table 3 .................................................................................................................................................... 22

  Table 4 .................................................................................................................................................... 23

  Table 5 .................................................................................................................................................... 24

  Figure 1 .................................................................................................................................................. 25
**Introduction**

Systemic Lupus Erythematosus (SLE) is a complex multisystem autoimmune disease with many manifestations. SLE is a relatively common disease in the United States with an estimated prevalence of childhood-onset SLE (cSLE) of 1 in 5,000 children. Children with cSLE present with an increased frequency of organ involvement including neuropsychiatric disease (NPSLE) [1, 2]. The prevalence of NPSLE in cSLE has been estimated to be 43-95% [3, 4]. Children with cSLE report school failure, concentration problems, and learning difficulties [5]. NPSLE has been associated with poor outcomes in all age groups [6-11]. The presence of NPSLE in a maturing brain may be of greater consequence than in adults, and it may interfere with the child’s acquisition of skills and cause a loss of function [12].

The standard assessment of neurocognitive function is done through formal neurocognitive testing (FNCT) [13], a battery of standardized neuropsychological tests that measure a variety of cognitive skills. However, access to it is limited, its performance is time-consuming, costly, and often not readily available. Screening has been defined as the systematic application of a test, for early identification of individuals at risk of a specific disorder that will benefit from further investigation [14]. Within a general pediatric rheumatology practice context the emphasis should be on prompt identification of neurocognitive dysfunction. Comprehensive assessment of the performance on each cognition domain will assist with in the diagnostic and therapeutic processes. A high-quality, low cost screening test that facilitates the clinical diagnosis and follow up of neurocognitive deficit (NCD) appears highly desirable.
Although the usefulness of questionnaires, computerized tests and imaging techniques to diagnose NPSLE have been studied [15-20], no tool is used regularly in daily clinical practice, likely because of lack of reliability and difficulties in implementation into clinical settings [14, 20, 21]. The *Pediatric Automated Neurophysiological Assessments Metrics (PedANAM)* is a computerized library of tests designed to measure cognitive ability, mental processing speed, memory, and efficiency in children age 10 years and older [22]. Significant advantages of the PedANAM are that it is time efficient, uses a standard computer with no need for specialized equipment, and has minimal practice effects. In a published study [17], specific PedANAM performance scores have been moderately correlated with performance on FNCT as measured by Z-scores; indicating good concurrent validity with traditional neuropsychological testing. Furthermore, this study started to explore the use of prediction models for the screening of NCD in cSLE.

A present short coming of the PedANAM is that a multitude of test statistics are offered rather than a performance summary score that can be easily tracked in a clinical setting. Such a summary measure, i.e. *PedANAM Cognitive Performance Score* (PedANAM-CPS) would be useful when employing the software as a screening tool for NCD and to monitor cognitive performance in cSLE patients.

The objective of this study is to explore and initially test approaches to the calculation of a PedANAM-CPS to summarize cognitive performance estimates provided by this software in an effort to assist with the screening for NCD in cSLE.
Patients and Methods

All of the cSLE participants fulfilled the updated American College of Rheumatology classification criteria prior to age 17 years [23], and were 9 to 17 years old at the time of study enrollment. English needed to be the first language spoken by participants and their caregivers. Routine sociodemographic and clinical data including SLE disease activity (SLEDAI), laboratory data, and medications were collected throughout the study course. All the studies were approved by the institutional review boards, with written consent and assent obtained as appropriate.

Participants were divided in a development and a validation cohort to develop and initially validate the PedANAM-CPS.

The development cohort

As part of a larger investigation at two tertiary pediatric rheumatology centers, 40 pairs of cSLE patients and their age plus sex-matched healthy controls enrolled in a study of cognitive functioning [24] were study at two study visits 18 months apart. During both visits participants completed the PedANAM and FNCT. Children were excluded if they had a history of comorbid conditions affecting their neurocognitive functioning prior to cSLE diagnosis, or known structural brain abnormalities, neuropathies, or movement disorders [13].

The validation cohort

At a general pediatric rheumatology clinical setting cSLE and non-cSLE subjects were recruited and asked to complete the PedANAM. Their parents were asked to complete the pediatric perceived cognitive function questionnaire-43 (PedsPCF-43)[25]. The requirements to
be part of the study were participant and parent agreement of participation, and completion of the assessments.

**Measurements**

The *Pediatric Automated Neuropsychological Assessment Metrics (PedANAM)*, completed by all participants, is a series of computer administered and scored subtests of cognitive processing efficiency which are designed for repeated measures (administrations over time) for a single subject. PedANAM subtests have been adapted to display age-appropriate stimuli as well to allow sufficient time for a response in a pediatric population [22]. The PedANAM subtests included in this study were: Simple Reaction Time, Procedural Reaction Time, Code Substitution Learning, Spatial Processing, Mathematical Processing, Matching to Sample Test, Sternberg Memory Search, Code Substitution Delayed Memory Test, and Continuous Performance Test. Each subtest produces three scores: *accuracy score* or percent of correct responses, *mean reaction time for correct responses*, and *throughput* or correct responses/minute a measurement of effectiveness.

The accuracy score has been linked to NCD [17]. Hence, the accuracy score of each subtest was considered as the raw score and the measurement of the subject’s performance for that subtest. All the subtest’s accuracy scores were normalized using Z-scores, i.e. $Z\text{-score} = \frac{\text{raw score} - \text{mean}}{\text{standard deviation}}$.

*Formal Neuropsychological Testing (FNCT)* was applied to the development cohort by a trained psychometrician using a standardized neuropsychological battery for cSLE [26]. Using published norms, participants’ performance on each of the neuropsychological tests were expressed as Z-scores with a normative mean of 0 and standard deviation of 1. The tests were
clustered into four cognitive domains: Working Memory, Psychomotor Speed, Attention/Executive Functioning, and Visuoconstructional Ability. Functioning in each cognitive domain was estimated by calculating a mean Z-score for the tests clustered in that domain.

The Pediatric perceived cognitive function questionnaire-43 (PedsPCF-43) was completed by the parents of participants of the validation cohort. It samples the caregiver’s perceptions of child’s cognitive functioning as observed in children’s everyday lives [25]. The questions were developed via IRT (Item Response Theory) analyses with input from parents, teachers, and clinicians experienced at working with pediatric cancer survivors. The focus of the PedsPCF-43 is upon fluid cognitive abilities sensitive to changes in mental status secondary to neurologic and systemic medical events. It measures facets of cognitive functioning such as attention, memory retrieval, and working memory. Its reliability, validity, and clinical utility has been investigated in the U.S. general population [27]. However it has not been validated for use in cSLE. A full description of the PedsPCF-43 questionnaire is published elsewhere [25, 27].

Ratings of the PedsPCF-43 are expressed in sex-standardized scores (T scores), with lower scores indicating greater cognitive dysfunction.

Definition of NCD or important reduction of cognitive performance

Cognitive performance was assessed in the development and validation cohorts in two different ways. For the development cohort, the results of FNCT were available to assess cognitive performance and define NCD. Following the definition used in previous studies [15, 17, 26], cognitive dysfunction was operationally defined as having at least one domain Z-score below -2 or at least two domain Z-scores below -1. For the validation cohort FNCT was unavailable. Rather perceived cognitive performance as rated by the caregiver via the PedsPCF-
43 was measured. In this study, a subject was considered to be at risk for NCD if the PedsPCF-43 T-score was <50.

**Statistical Analysis**

Information collected on both cohorts included baseline disease characteristics, ongoing disease activity, physical examination, standard clinical and laboratory information, and the PedANAM test. FNCT or the PedsPCF-43 questionnaire were collected on the development and validation cohorts, respectively. Demographic and clinical data were summarized by mean and standard deviation percentage values. Of notice, as cognition is closely related to brain development, and given the influence of age on PedANAM scores[28], our analysis was adjusted for age. However, unadjusted data did not differ significantly from adjusted ones. Hence, we presented only results of the adjusted analyses in the result section.

Development cohort’s first visit data was used to construct the PedANAM-CPS candidates. First, each of the PedANAM-CPS subtests accuracy score was normalized using Z-scores. Then, PedANAM-CPS candidates were created using three statistical methods. Subsequently for external and cross-validation purposes, the performance of those candidate scores was evaluated using data from the second visit of the development cohort and the validation cohort. The analytical plans used were the following.

**Construction of PedANAM-CPS candidates**

The PedANAM-CPS candidates were developed using the following statistical methods.  
1) **Simple-summary score** – using equal weighted average of all subtest’s accuracy scores.  
2) **Logit-based score** – this score uses all the subtests accuracy scores to predict the likelihood of
the NCD using a logistic regression model. NCD outcome from FNCT was used as the dependent variable. The intercept and the slope coefficients from such a model were used as the parameters to combine individual accuracy scores and construct the predicted Logit score for each subject. The Wald statistics were used to assess the importance of the slopes coefficients of the Logit-based score. 3) PCA based-score – uses a Principal Component Analysis (PCA). This method is the simplest of the true eigenvector-based multivariate analyses. Often, its operation can be thought of as revealing the internal structure of the data in a way that best explains the variance in the data. PCA is mostly used as a tool in exploratory data analysis and for making predictive models. The PCA was used to decompose the variance–covariance matrix of all PedANAM subtest accuracy scores and its first eigenvector was considered the way to combine the subtest scores based upon their contributions to the total variation.

The stated statistical methods were chosen for the analysis as they treat the information differently: the PCA approach, an unsupervised method, analyzes the data without taking into account presence or absence of the outcome, while the logistic regression model, a supervised approach, does consider the outcome when building the model. Also, the PCA approach is considered helpful in reducing the number of observed variables to a smaller number of principal components which accounts for most of the variance of the observed variables.

By definition, the higher the Simple summary score or PCA-based score, the more accuracy obtained in the PedANAM and hence less likely to have NCD. The Logit-based score is constructed differently: the higher the score the more likely to have NCD.

*Validation of PedANAM-CPS candidates*
To establish *construct validity*, the PedANAM-CPS candidates and their component scores, i.e. the individual subtest scores, were compared between NCD and non-NCD subjects using t-tests. In addition, fixed effect models were used to assess the associations of the PedANAM-CPS to cSLE groups. Groups were defined as cSLE with NCD (SLE NCD), cSLE with no NCD (cSLE Non-NCD), and controls.

To determine *convergent validity*, Pearson’s correlation coefficients were calculated to assess relationships between Simple-summary, PCA-based and Logit-based scores.

To assess the accuracy of the PedANAM-CPS candidates predicting and discriminating subjects with NCD (*concurrent validity*), a receiver operating characteristics (ROC) curve analysis was used. The area under the ROC curve (AUC) was calculated, and the scores’ sensitivity and specificity were determined under a preferred threshold approach. PedANAM-CPS candidates’ performance was evaluated in the last visit of the developmental cohort and the only visit of the validation cohort. Values of the AUC can be interpreted as outstanding, excellent, good, fair, and poor performance in predicting NCD, for values of 1.0–0.91, 0.81–0.90, 0.71–0.8, 0.61–0.7, and 0.6, respectively [29].

In addition, we analyzed the performance of a previously proposed prediction model of NCD, i.e. *Inclusive-score*. This score encompasses the accuracy score of spatial processing subtest as well as the coefficient of variation of reaction time for correct responses score of continuous performance and matching to sample subtest, and the mean reaction time to correct responses score of code substitution delayed subtest. Stepwise selection methods were used to select PedANAM scores included in this final logistic regression model. The Inclusive-score has a reported sensitivity and specificity to detect NCD of 100 and 86% respectively [15].
The performance of the Inclusive-score was evaluated on the second visit of the development cohort and on the validation cohort.

**Results**

*Study subjects*

Demographics of the study subjects and disease information are presented in Table 1. The age mean of cSLE patients was 14.8 ± 2.3 in development and 15.3 ± 3.3 in the validation cohort. Over 85% of the study subjects were female and at least a third or more of the cSLE patients were African American. cSLE and controls of both cohorts were comparable on sociodemographic variables as measured by subject and maternal education level, and family income. Most of the cSLE subjects were prescribed prednisone. Disease activity as measured by the Systemic Lupus Erythematosus Activity Index (SLEDAI) was in the mild to moderate range (mean ± SD, 4.9 ± 4.4 and 4.3 ± 4.7) in the development and validation cohorts, respectively. NCD was present in 22.5% of the cSLE participants of the development cohort. Interestingly, the validation cohort SLE NCD group had a lower prevalence of NCD when compared to the control group. However, the T-score median (min, max) of the PCF-43 questionnaire was significantly different among groups: (35 (33, 53) in the control group, 36 (33, 49) in the SLE Non-NCD group, and 53.5 (53,78) in the SLE NCD group (p-value=<0.001).

*PedANAM-CPS candidates*

Development of the PCA-based score revealed that all eigenvectors were consistent across the studies visits. This methodology suggested that the Code Substitution, Logical
Relations, Matching Grids, and Spatial Processing subtests are the ones contributing the most to the total variation. The analysis of the slope coefficients obtained for the Logit-based score agreed with the PCA in that the Spatial Processing subtest is an important contributor. Nonetheless, under this method the Code Substitution Delay and Stenberg Memory Search subtests were also found to be significant contributors.

Construct validity

The ability of the PedANAM-CPS candidates to recognize presence or absence of NCD was assessed by comparing the scores between NCD and non-NCD subjects of the development cohort while adjusting for age differences between groups. All the three score methods (simple-summary, PCA-based, and Logit-based scores) significantly differentiated subjects with and without NCD (Table 2). Not surprisingly, we found that none of the individual accuracy subtests scores differentiated consistently the subjects with NCD. When comparing the three PedANAM-CPS candidates among controls and cSLE subjects with and without NCD, the Logit-based score significantly differentiated cSLE NCD from SLE Non-NCD patients (p-value=0.031).

Convergent validity

The relationship between the three PedANAM-CPS candidates was examined (Table 3). The PCA-based score and simple-summary score were found to be highly related among themselves. Logit-based scores related moderately to the PCA-based and simple-summary scores.
Concurrent validity

The performance of the PedANAM-CPS candidate scores at discriminating presence or absence of NCD on the first visit of the development cohort was studied using a ROC curve analysis. Figure 1 shows the AUC for each score. Of noticed, the Inclusive-score and the Logit-based score have an excellent to good performance predicting NCD in this cohort.

We then calculated the sensitivity and specificity of the PedANAM-CPS candidates’ cut-off points. Cut-off points were selected when the sensitivity was at least above 80% and the overall accuracy (the average of sensitivity and specificity) was among the highest among all possible cut off points (Table 4). Two of the PedANAM-CPS candidates, the Simple-summary and the PCA-based scores performed at the same level. The logit-based score cut-off have better accuracy than the other two scores’ cut-off points.

The predictive power of the PedANAM-CPS candidates was examined on the second visit of the development cohort and the validation cohort (Table 5). The sensitivity of all the scores detecting NCD was near 90% when tested on the second visit of the development cohort. On the validation cohort, the sensitivity to detect NCD was around 80% for all the scores but the Inclusive-score. However, the Inclusive-score exhibited excellent to good performance detecting NCD when applied to development (second visit) and validation cohorts (AUC= 0.89 and 0.74 respectively).
Discussion

Although NPSLE can be clinically evident, SLE patients without history of overt neuropsychiatric disorder can have decreased cognitive performance when compared to non-SLE subjects [30]. Hence, a screening test that allows early identification should be adopted [31]. Screening for changes in neurocognitive performance and the presence of NCD is difficult in a clinical setting using FNCT. Based on recent research the PedANAM software may offer a cost effective approach for the screening of cSLE patients in daily clinical practice. To enhance the usefulness of the PedANAM we set out to develop and initially validate PedANAM-CPS candidates, using various statistical approaches, to evaluate cognitive performance for the screening of NCD in cSLE.

The construct and convergent validity of the PedANAM-CPS candidates confirmed their ability to recognize the presence or absence of NCD. The Inclusive-score and Logit-based score exhibited excellent to good performance in predicting the presence of NCD; even when the Inclusive-score was designed under more a strict definition of NCD [15]. According with our results the Inclusive-score and Logit-based score both seem reasonable tools to evaluate cognitive performance for the screening of NCD.

The Logit-based and PCA-based scores confirmed the utility of assessing visual perception and working memory via the Spatial Processing Subtest [15, 17]. Additionally, the contribution of learning and memory processes as assessed through the Code Substitution Delayed Memory subtest was found to be significant, especially under the Logit-based score. Nevertheless the Logit-based and PCA-based scores weighted differently the contribution of most PedANAM subtests. As the cognitive domains that are impaired in children with cSLE are
diverse, seems that the PedANAM-CPS should be obtained from the full PedANAM battery rather than from an abbreviated form [32]. However, the performance of the Inclusive-score, a method based on few subtests, was equivalent to the others PedANAM-CPS candidates performance. Ultimately, the PedANAM-CPS needs to be a comprehensive screening test that allows the detection of typical and atypical NCD.

To be congruent with the high sensitivity required for the desirable neurocognitive screening test [14], the cut-off points calculated for the scores needed to have greater than 80% sensibility in detecting NCD. Since the damage caused by untreated central nervous system disease can be irreversible and have long term unbearable consequences we favored a test with high likelihood of detecting cognitive dysfunction. The PedANAM is a child-friendly test, inexpensive and can be widely available. PedANAM benefits surpass the possible risk associated with a false positive result: close evaluation and follow-up of a given individual.

Interestingly, the performance of the Logit-based and PCA-based scores predicting NCD in the validation cohort wasn’t similar to the one observed in the development cohort. This phenomenon can have several explanations. First, subjects from the validation cohort were recruited in a clinical setting without the close observation and advantages offered by a research environment. Feasibility of the PedANAM as a point of care test should also be considered. However, the average time of the PedANAM completion was similar among cohorts (30-40 min). Only four subjects from the validation cohort were unable to complete the battery without an interruption; reasons for interruption were: software malfunction, inability to understand the test and/or fatigue secondary to the PedANAM. Furthermore, several studies have found that the PedANAM is a feasible tool [15, 17, 33, 34]. Participant’s age was
not found to impact the ability of finish the battery. Another factor that influences the performance of the PedANAM-CPS candidates is the different methods used to define NCD. Indeed, recently published data suggests that questionnaire-based screening for NCD complements the information provided by FNCT rather than supplements it [35].

A limitation of our study may be the sample size. As the performance of a measurement relates to the prevalence the disease that is intended to measure, a larger study population can produce different findings. Even though the PedANAM performance was not appreciated to be associated to the patient’s ethnic background we were unable to test for such [34].

Early detection of NCD and cognitive decline are the first steps to improve the prognosis and decrease the morbidity in affected cSLE patients. The development of the PedANAM-CPS for cSLE is the first step to integrate a neurocognitive screening tool into clinical care. Based on our findings, we considered that the Logit-based score and the Inclusive-score are suitable scores to become the PedANAM-CPS. However, larger studies to explore the performance of the proposed scores will elucidate a robust and reliable PedANAM-CPS that will allow systematic evaluation of cognitive performance for the screening of NCD in cSLE. Subsequently, evaluation of the ability of the performance score to detect clinically significant changes and its role in other disease processes are granted. The establishment of PedANAM-CPS will ultimately impact the clinical care and quality of life of patients with cSLE.
Bibliography


Table 1.
Demographics of Development and Validation Dataset at Enrollment *

<table>
<thead>
<tr>
<th>Variable</th>
<th>Development Dataset</th>
<th>Validation Dataset</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cSLE (n=40)</td>
<td>Controls (n=40)</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>14.8 ± 2.3</td>
<td>13.9 ± 3.2</td>
<td>0.03</td>
</tr>
<tr>
<td>Age (years) in median (range)</td>
<td>14 (9, 17)</td>
<td>14 (9, 17)</td>
<td>0.767</td>
</tr>
<tr>
<td></td>
<td>15 (10, 20)</td>
<td>14 (11, 18)</td>
<td>0.239</td>
</tr>
<tr>
<td>Female</td>
<td>Yes 85</td>
<td>85</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>91.2</td>
<td>72.2</td>
<td>0.032</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>30</td>
<td>32.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30.9</td>
<td>94.4</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>45</td>
<td>47.5</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>30.9</td>
<td>5.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Hispanic</td>
<td>17.5</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11.8</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Asian and other</td>
<td>7.5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>26.5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Grade Level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary School (4-6)</td>
<td>20</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>18.6</td>
<td>22.2</td>
<td></td>
</tr>
<tr>
<td>Middle School (7-8)</td>
<td>17.5</td>
<td>17.5</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>15.3</td>
<td>33.3</td>
<td>0.321</td>
</tr>
<tr>
<td>High School (9-12)</td>
<td>62.5</td>
<td>62.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>55.9</td>
<td>38.9</td>
<td></td>
</tr>
<tr>
<td>College</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.2</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>Maternal education level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No High School Diploma</td>
<td>7.5</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.5</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>Complete High School Diploma</td>
<td>30</td>
<td>37.5</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>16.7</td>
<td>11.1</td>
<td>0.523</td>
</tr>
<tr>
<td>Education Beyond High School</td>
<td>62.5</td>
<td>52.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>78.8</td>
<td>83.3</td>
<td></td>
</tr>
<tr>
<td>Annual Family Income</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$ 25K</td>
<td>20</td>
<td>15.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14.3</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>$26K-$50K</td>
<td>35</td>
<td>34.2</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>31.7</td>
<td>5.6</td>
<td>0.057</td>
</tr>
<tr>
<td>$51K-$75K</td>
<td>20</td>
<td>28.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.7</td>
<td>16.7</td>
<td></td>
</tr>
<tr>
<td>&gt;$75K</td>
<td>25</td>
<td>21.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>41.3</td>
<td>72.1</td>
<td></td>
</tr>
<tr>
<td>On Prednisone therapy</td>
<td>77.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prednisone dose (mg/day)</td>
<td>19.8 ± 17.4</td>
<td>17.1 ± 16.1</td>
<td></td>
</tr>
<tr>
<td>Disease activity, SLEDAI (mean ± SD)‡</td>
<td>4.9 ± 4.4</td>
<td>4.3 ± 4.7</td>
<td></td>
</tr>
<tr>
<td>PCF-43† T-score (mean ± SD)</td>
<td>60.5 ± 7.9</td>
<td>63.2 ± 5.8</td>
<td>0.167</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------</td>
<td>------------</td>
<td>-------</td>
</tr>
<tr>
<td>Neurocognitive dysfunction¥</td>
<td>22.5</td>
<td>7.5</td>
<td>8.8</td>
</tr>
</tbody>
</table>

*Except where indicated otherwise, values are percentages; cSLE = childhood-onset systemic lupus erythematosus

‡ Systemic Lupus Disease Activity Index 2k version; range 0 – 104; 0 = inactive SLE

† PCF-43 questionnaire: Perceived Cognition Functioning -43 questionnaire

¥ Neurocognitive dysfunction categories are defined based on z-scores of the standardized tests completed for the formal neurocognitive testing (FNCT) on the research cohort, and on T-scores of the pediatric perceived cognitive function questionnaire-43 (PedsPCF-43) on the clinical cohort.

FNCT measures following cognitive domains: working memory, psychomotor speed, attention and executive functioning, visuoconstructional ability.
Table 2.

Ability of the candidate PedANAM-CPS to identify neurocognitive deficit

<table>
<thead>
<tr>
<th>Candidate PedANAM-CPS</th>
<th>Development Cohort</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-NCD (n=68)*</td>
<td>NCD (n=12)*</td>
<td>p-value†</td>
<td></td>
</tr>
<tr>
<td>Simple Summary Score</td>
<td>0.08 ± 0.07</td>
<td>-0.39 ± 0.18</td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td>PCA-based Score</td>
<td>0.09 ± 0.08</td>
<td>-0.42 ± 0.19</td>
<td>0.027</td>
<td></td>
</tr>
<tr>
<td>Logit-based Score</td>
<td>-2.26 ± 0.16</td>
<td>-0.60 ± 0.37</td>
<td>0.001</td>
<td></td>
</tr>
</tbody>
</table>

^ Neurocognitive deficit as measured by formal neuropsychological assessment in the development cohort and by the Pediatric Perceived Cognitive Function questionnaire-43 (PedsPCF-43) in the validation cohort; NCD = neurocognitive deficit; PCA = Principal Component Analysis; Logit = logistic regression model

* Values are mean ± SD

† P values are adjusted for age
Table 3.
Pearson correlation coefficients of the candidate PedANAM-CPS

<table>
<thead>
<tr>
<th>Candidate PedANAM-CPS</th>
<th>Development Cohort (n=80)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simple-summary score</td>
<td>PCA-based score</td>
<td>Logit-based score</td>
<td></td>
</tr>
<tr>
<td>Simple-Summary Score</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCA-based Score</td>
<td>0.990</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logit-based Score</td>
<td>-0.579</td>
<td>-0.601</td>
<td>1.000</td>
<td></td>
</tr>
</tbody>
</table>

PCA= Principal Component Analysis; Logit= logistic regression model
Table 4.
Cut-off point sensitivity and specificity of the PedANAM-CPS candidates development cohort†

<table>
<thead>
<tr>
<th>Candidate PedANAM-CPS</th>
<th>AUC ¥</th>
<th>Cut-off*</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple-Summary Score</td>
<td>0.60</td>
<td>0.25</td>
<td>83.3 %</td>
<td>37.3 %</td>
</tr>
<tr>
<td>PCA-based Score</td>
<td>0.60</td>
<td>0.25</td>
<td>83.3 %</td>
<td>41.8 %</td>
</tr>
<tr>
<td>Logit-based Score</td>
<td>0.77</td>
<td>-2.70</td>
<td>91.7%</td>
<td>31.3 %</td>
</tr>
</tbody>
</table>

PedANAM-CPS = Pediatric Automated Neuropsychological Assessment Metrics – Cognitive Performance Score; PCA = Principal Component Analysis; Logit = logistic regression model

† First visit of the development cohort, n=80 subjects.

¥ Interpretation of AUC values: 1.0–0.91: outstanding, 0.81–0.90: excellent, 0.71–0.8: good, 0.61–0.7: fair, and <0.6: poor

*Simple-summary or PCA-based scores <0.25 indicate higher likelihood of NCD. Logit-based scores > -2.70 indicate higher likelihood of NCD
Table 5.
Summary of sensitivity and specificity from the external validation of the candidate PedANAM performance scores

<table>
<thead>
<tr>
<th>Candidate PedANAM-CPS</th>
<th>Cut-off*</th>
<th>Development Cohort Second Visit (n=61)</th>
<th>Validation Cohort (n=86)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AUC¥</td>
<td>Sensitivity</td>
</tr>
<tr>
<td>Simple-Summary Score</td>
<td>0.25</td>
<td>0.78</td>
<td>80 %</td>
</tr>
<tr>
<td>PCA-based Score</td>
<td>0.25</td>
<td>0.80</td>
<td>90 %</td>
</tr>
<tr>
<td>Logit-based Score</td>
<td>-2.70</td>
<td>0.80</td>
<td>90 %</td>
</tr>
<tr>
<td>Inclusive-score</td>
<td>0.09</td>
<td>0.89</td>
<td>89 %</td>
</tr>
</tbody>
</table>

PedANAM= Pediatric Automated Neuropsychological Assessment Metrics; PCA = Principal Component Analysis; Logit = logistic regression model.

¥ Interpretation of AUC values: 1.0–0.91: outstanding, 0.81–0.90: excellent, 0.71–0.8: good, 0.61–0.7: fair, and <0.6: poor

*Simple-summary or PCA-based scores <0.25 indicate higher likelihood of NCD. Logit-based scores > -2.70 indicate higher likelihood of NCD. Inclusive-score >0.09 indicate higher likelihood of NCD
Figure 1.
Simple-summary score

PCA-based score
Logit-based score

![Logit-based Score Diagram](image)

- AUC = 77%
- Sens = 91.7%
- Spec = 31.3%
- Logit Score cut = -2.7

Inclusive-score

![Inclusive-score Diagram](image)

- AUC = 89%
- Sens = 89%
- Spec = 55%
- Propensity Score cut = 0.09

**Figure 1.** Area under (AUC) the receiver operative characteristic curve (ROC) calculated at the first visit of the development cohort. Sens = sensitivity; Spec = specificity; PCA = principal component analysis; Logit = logistic regression model.