



Published in final edited form as:

Arthritis Care Res (Hoboken). 2013 March ; 65(3): 372–381. doi:10.1002/acr.21835.

VALIDATION OF THE PEDIATRIC AUTOMATED NEUROPSYCHOLOGICAL ASSESSMENT METRICS IN CHILDHOOD-ONSET SYSTEMIC LUPUS ERYTHEMATOSUS

Hermine I. Brunner¹, Marisa S. Klein-Gitelman², Frank Zelko², Erin C. Thomas², Jessica Hummel¹, Shannen M. Nelson¹, Jennifer Huggins¹, Megan L. Curran², Tresa Roebuck-Spencer³, Dean W. Beebe¹, and Jun Ying⁴

¹ Cincinnati Children's Hospital Medical Center, University of Cincinnati, Chicago

² Ann & Robert H. Lurie Children's Hospital of Chicago; Feinberg School of Medicine, Northwestern University, Chicago

³ Cognitive Science Research Center, University of Oklahoma, Norman OK

⁴ Department of Environmental Health, University of Cincinnati, OH

Abstract

Objective—To evaluate the reproducibility and validity of the Pediatric Automated Neuropsychological Assessment Metrics (Ped-ANAM) when used in childhood-onset systemic lupus erythematosus (cSLE).

Methods—Forty children with cSLE and 40 matched controls were followed for up to 18 months. Formal neuropsychological testing at baseline was repeated after 18 months of follow-up; overall cognitive performance and domain-specific cognition (attention, working memory, processing speed and visuoconstructional ability) were measured and categorized as having normal cognition, mild/moderate or moderate/severe impairment. The 10 Ped-ANAM subtests were completed every 6 months and twice at baseline. Ped-ANAM performance was based on accuracy (AC), mean time to correct response (MNC), throughput, and coefficient of variation of the time required for a correct response (CVc) as a measure of response consistency.

Results—Particularly MNC scores demonstrated moderate to substantial reproducibility (intraclass correlation coefficients: 0.47-0.80). Means of select Ped-ANAM scores (MNC, AC, CVc) differed significantly between children with different levels of cognitive performance and allowed for the detection of moderate or severe cognitive impairment with 100% sensitivity and 86% specificity. Six Ped-ANAM subtests significantly correlated with the change in overall cognitive function in cSLE (baseline vs. 18 month; Spearman correlation coefficient $> \pm 0.4$; $p < 0.05$, $n = 24$).

Conclusions—The Ped-ANAM has moderate to substantial reproducibility, criterion and construct validity and may be responsive to change in cSLE. Additional research is required to confirm the Ped-ANAM's outstanding accuracy in identifying cognitive impairment and its usefulness in detecting clinically relevant changes in cognition over time.

Address reprint requests to: Hermine Brunner MD, M.Sc., Cincinnati Children's Hospital Medical Center, University of Cincinnati, William Rowe Division of Rheumatology, E 4010, 3333 Burnet Avenue, Cincinnati, OH 45229-3039, USA; hermine.brunner@cchmc.org..

The authors deny any conflicts of interest, real or perceived.

Keywords

SLE; cognitive dysfunction; SLE; children; Ped-ANAM; ANAM

Like adults, children with systemic lupus erythematosus (cSLE) frequently report cognitive problems, and several studies have documented significant cognitive deficits with traditional neuropsychological test batteries (1, 2). Most studies find neurocognitive dysfunction (NCD) on tests measuring attention or concentration, cognitive flexibility, free recall memory, visuoconstructional ability, and speed of information processing in a substantial subgroup of cSLE, suggesting the presence of a subcortical cognitive syndrome (3). NCD may represent active neuropsychiatric lupus. When present, NCD has detrimental effects on patient quality of life (4), and thus constitutes an important disease feature to consider in the medical management of cSLE.

The detection of NCD with cSLE is typically made by formal neuropsychological testing (1, 5), but this is costly, time intensive, not always readily available, and requires specialized advanced training. In recent years computer-administered tests have been explored in patients with various diseases as more cost-effective screening tools of NCD (6). One of these is the Automated Neuropsychological Assessment Metrics (ANAM™), designed as a library of automated tasks assessing various aspects of cognitive functioning (7).

The ANAM has been found to be well-suited to screen the cognitive abilities of adults with SLE (3, 8). Similarly, our group reported on the potential usefulness of the adaptation of the ANAM for pediatric use (Ped-ANAM™) (9) to screen for cSLE-associated NCD in a small cross-sectional pilot study (10). Building on our promising initial results, our objectives of this case-control study were to more thoroughly document the utility of the Ped-ANAM in this high-risk population; specifically, we aimed to (a) assess the feasibility of administration by non-neuropsychologist clinic staff, (b) document the short-term reproducibility of the instrument, (c) test its concurrent and criterion validity, as well as its (d) responsiveness to change, relative to more traditional neuropsychological tests.

PATIENTS AND METHODS

Forty patients with cSLE (11) were asked to each identify a friend who was within one year of his/her age, of the same gender, and in the same school grade. This “best-friend approach” to selecting control populations has been shown to result in good case-control matches on sociodemographic variables (12). Controls had to be healthy, without known structural brain abnormalities or known NCD. No potential controls needed to be excluded from participation by these criteria. The patients’ medical records were reviewed for cSLE relevant parameters, and additional information about the study population is provided elsewhere (13). To study the reproducibility (also known as test-retest reliability) of the Ped-ANAM, study participants completed the test twice during the first study visit. For use as external standard, participants also completed formal neuropsychological tests at the same visit. Finally, both the Ped-ANAM and neuropsychological tests were repeated 18 months later to address responsiveness to change.

Measures

Pediatric Automated Neuropsychological Assessment Metrics—The Ped-ANAM has been adapted from the traditional ANAM (14, 15) for use in children 10 years (16). Together, the Ped-ANAM subtests measure sustained concentration and attention, spatial processing, cognitive-processing efficiency, verbal reasoning, learning, recall, and

working memory. A full description of the Ped-ANAM subtests is published elsewhere (10) (also see supplemental material *online*).

Performance on each of the Ped-ANAM subtests can be gauged by four scores. They are (a) accuracy defined as the percentage of correct responses (AC); (b) mean reaction time for correct responses or speed (MNC; in seconds); (c) throughput (TP) which is considered a measure of effectiveness or cognitive efficiency, and is a combination of reaction time and accuracy (17); and (d) coefficient of variation of reaction time for correct responses (CVc= standard deviation MNC/ MNC) which reflects the consistency of a test-taker's response speed within a given subtest. For the *Simple Reaction Time* subtest only the MNC score was calculated as this subtest only allows for correct responses.

When taking the Ped-ANAM, the *Simple Reaction Time* subtest is completed at the beginning and again at the end of the testing session, with the MNC calculated as the average across both subtests. Differences in the performance of the *Simple Reaction Time* subtest over a given Ped-ANAM session (beginning vs. end of session), can be used to assess the change in speed of sensorimotor processing or participant fatigue when taking the Ped-ANAM (18). Higher scores at the end of the session than at the beginning indicate a decline in overall speed of sensorimotor processing. Higher AC and TP scores, and lower MNC and CVc scores, indicate better cognitive performance. The TP variable has been widely used in ANAM research given that it is sensitive to cognitive performance, incorporates speed and accuracy in one variable, and more closely conforms to a normal distribution than other variables (8, 18-20).

Formal Neuropsychological Testing Battery and Definition of Neurocognitive Dysfunction—Formal neuropsychological testing was performed by a trained psychometrician, using a standardized neuropsychological battery for cSLE, with details provided elsewhere (5). The battery assesses *working memory*, *psychomotor speed*, *attention*, and *visuoconstructional ability*, cognitive domains that have been found to be particularly affected by cSLE in prior research (**Table 1**). All measures included in this battery of neuropsychological tests are well-validated and provide age-normed scores compared to large, demographically diverse normative samples (5).

To categorize levels of cognitive function, participants' age-normed scores were converted to a common z-score metric (mean= 0, standard deviation = 1 for a normative healthy population), with higher scores reflecting better performance. The mean z-scores within each cognitive domain were then averaged to yield a composite domain score, with four composite domain scores overall. Finally, in the absence of a generally accepted definition for NCD (21), three levels of cognitive ability were defined: (a) *normal cognition/ no NCD*, if all z-scores ≥ -1 ; (b) *mild to moderate NCD*, if one or two domain z-scores < -1 or one domain z-score ≤ -2 ; and (c) *moderate to severe NCD*, if more than one domain z-score < -2 or more than two domain z-scores < -1 .

Statistics

Demographic and clinical characteristics were summarized by means and standard deviations (SD) for numerical variables and frequency (in %) for categorical variables. We calculated intraclass correlation coefficients (ICCs) to assess the reproducibility of the Ped-ANAM subtests. ICCs can be interpreted as follows: poor agreement: ICC < 0.4 ; fair to good agreement: ICC $0.4 - 0.75$; substantial to excellent agreement: ICC > 0.75 (22).

As previously suggested (18), performance fatigue was examined by comparing the difference of the MNC score on the *Simple Reaction Time* subtest when administered in the beginning as compared to the end of a Ped-ANAM session. Both Spearman's and Pearson's

correlation coefficients were calculated to assess relationships between Ped-ANAM scores and NCD domain and overall z-scores. Results from these two methods were similar, thus only Spearman's correlation coefficients (*r*'s) are reported. Absolute values of *r*'s can be interpreted as unrelated; weak, moderate or strongly correlated for values <0.2; 0.2- 0.39, 0.4- 0.59, or 0.6 (23) respectively. To establish criterion validity, fixed effect models were used to determine Ped-ANAM measures associations with cognitive function category (normal cognition, mild/moderate NCD, and moderate/severe NCD); *post-hoc*, means were compared between NCD groups under the fixed effect model framework and adjusted for multiple comparisons using a Tukey's method. Ped-ANAM's responsiveness to cognitive change (based on change as measured by the traditional cognitive tests) was assessed by determining the correlation between the change in the cognitive function (domain z-scores) and the change of the Ped-ANAM scores per subtest.

Both unadjusted and adjusted (after adding demographics and clinical characteristics as controlling covariate) fixed effect models were considered in computation. Results from adjusted models are not shown as they were similar to those of unadjusted models. In the prediction analyses, moderate/severe NCD was predicted by Ped-ANAM measures using multivariate logistical regression models. Stepwise selection methods were used in the logistical regression models to select Ped-ANAM scores. A propensity score was calculated from each of the logistical models and considered in receiver operating characteristics curve analysis to predict moderate/severe NCD. The area under the receiver operating characteristics curve (AUC) was calculated, and sensitivity and specificity determined under a preferred threshold approach. 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 (24).

As an alternative approach, classification and regression tree (CART) model analysis (25) was done and a CART-score developed from the final nodes of the CART model tested in receiver operating characteristics curve analysis. CART analysis was performed using SYSTAT version 11.0 software (Systat Software, Inc., Point Richmond, CA) and other statistical analyses using SAS 9.3 software (Cary, NC), with ROC curves plotted using Splus version 6.2 software (Insightful Corp, Durham, NC). P-values < 0.05 were considered statistically significant.

RESULTS

Study Participants & Formal Neuropsychological Testing

English was the native language for all study participants of the 40 cSLE and 40 controls matched for sex, school grade, and age (within 1 year of age of index patient with cSLE), with sociodemographic details and information about cSLE status provided elsewhere (13) and *online*. In brief, 45% of the cSLE patients were African American, and 85% of the study participants were female. Besides the mean age of the cSLE group being somewhat higher than that of the control group (14.8 ± 2.3 years vs. 13.98 ± 3.2 years; $p = 0.03$), groups were comparable on sociodemographic factors as measured by the maternal education level and family income. NCD was identified in both controls and the cSLE group, with a trend towards more pronounced NCD severity in the latter. At baseline, normal cognition, mild NCD or moderate/severe NCD were noted, respectively, in 60% ($n=24$), 35% ($n=14$) and 5% ($n=2$) of the controls as compared to 62.5% ($n=25$), 25% ($n=10$) and 12.5% ($n=5$) of the cSLE patients. Daily prednisone was prescribed to 78% (31/40) of the patients with cSLE. Disease activity as measured by the SLEDAI (mean \pm SD: 4.9 ± 4.4) and BILAG 2004 (mean \pm SD: 3.0 ± 3.8) indices (26) was in the mild to moderate range. 18-month follow-up data were available for 24 of the 40 cSLE participants.

Feasibility

None of the participants had difficulties in understanding the Ped-ANAM instructions. Each administration took 35 – 55 minutes. There were rare technical problems (non-responsiveness of the software due to repetitive triggering of the mouse in very short time intervals) in four patients, requiring the intervention of clinical research personnel.

Throughput Scores

We found the TP score to be problematic as a measure of cognitive function due to our observation of atypical response patterns in some participants. That is, some participants, especially those with NCD, responded to test items with unusually fast reaction times (i.e., substantially faster than the mean of the group) with correspondingly very low accuracy (often at chance levels), suggesting that these participants responded quickly without actually attending to the cognitive demands of the task. Historically, TP has been shown to be more strongly weighted toward reaction time than accuracy (given the inherent greater range of variability in the reaction time variable compared to accuracy). Therefore, in this study the TP score of more cognitively complex tests, tended to give too much ‘credit’ or to overestimate cognitive functioning in participants with a quick but inaccurate response style.

Reproducibility

The reproducibility of the Ped-ANAM subtests was highest for subtests of higher cognitive complexity and for the MNc score (**Table 2**). The reproducibility (MNc scores) of the Ped-ANAM ranged from fair to excellent but, for the most part was good to substantial. AC scores showed fair to good agreement for some subtests (Code Substitution, Matching to Sample, Matching Grids, and Sternberg Memory Search) but substantial to excellent consistency for the *Continuous Performance Test*. ICCs for the derived performance parameter CVc varied widely and ranged from poor (*Procedural Reaction Time*) to good (*Continuous Performance Test*).

Construct Validity

We expected participants with normal cognition to perform better on the Ped-ANAM compared to those with NCD, especially if NCD was more pronounced. There were statistically significant differences in scores (AC, MNc, CVc) on several Ped-ANAM subtests, particularly in contrasting the performance of the moderate/severe NCD group with the normal cognition and mild/moderate NCD groups (**Table 3**). Interestingly, when visually examining group means for the MNc score, the moderate/severe NCD group tended to respond the fastest. We believe this is due to the earlier described tendency of some participants to respond quickly without attending fully to the cognitive demands of the task. This possibility is supported by the observation of lower AC scores in this group compared to the other two groups and the finding that the moderate/severe NCD group performed significantly slower than the other two groups on *Simple Reaction Time* subtests which do not present significant cognitive demands (all: p (MNc) < 0.021).

Criterion Validity

In an effort to evaluate criterion validity and determine whether one can limit the number of Ped-ANAM subtests to be completed for the surveillance of NCD, we assessed which particular subtests were especially correlated with the classification of NCD (**Table 4**). Therefore, we calculated odds ratios using multivariate logistic models (outcome: normal cognition yes/no), while adjusting for age differences between groups. Our results, shown in **Figure 1**, suggest that the presence of NCD can be accurately predicted by a subset of Ped-ANAM subtests (Spatial Processing, Continuous Performance Test, Matching to Sample, Code Substitution Delayed).

As an alternative approach we also explored combinations of Ped-ANAM scores of the above subtests using CART analysis. This resulted in a CART score (range: 1 – 4), with a lower score indicating higher likelihood of moderate/severe NCD. Specifically, *Score = 1* if CVc of *Continuous Performance Test* > 0.6; *Score = 2* if CVc of *Continuous Performance Test* ≤ 0.6 but CVc of *Matching to Sample* > 1.4; *Score = 3* if CVc of *Continuous Performance Test* ≤ 0.6 and CVc of *Matching to Sample* ≤ 1.4 but MNC of *Continuous Performance Test* ≤ 375 milliseconds; and *Score = 4* if CVc of *Continuous Performance Test* ≤ 0.6 and CV of *Matching to Sample* ≤ 1.4 but MNC of *Continuous Performance Test* > 375 milliseconds. The CART-score of 3 and above had an AUC of 92%, a sensitivity of 83.3% and specificity of 100% in indicating whether the participant has no NCD vs. Moderate/ Severe NCD (**Figure 1, Model 6**).

Change in Speed of Sensorimotor Processing

Participants' MNC scores when performing the *Simple Reaction Time* subtest performed in the beginning of the Ped-ANAM session were subtracted from those when repeating this subtest in the end of the Ped-ANAM session, with larger increases considered to reflect a more pronounced decline in overall speed of sensorimotor processing. Our study showed increases in mean ± SE's of the MNC scores to be 25.6 ± 16.2 milliseconds (p=0.119), 50.8 ± 22.7 milliseconds (p=0.028), and 159.6 ± 46.4 milliseconds (p=0.0003) in the normal cognitive, mild/moderate NCD and moderate/severe NCD groups, respectively. However, these differences only reached statistical significance when comparing the moderate/severe NCD to the normal cognition group (p=0.021) while there was only a trend for between the moderate/severe NCD vs. the mild/moderate NCD group (p= 0.095).

Responsiveness to Change over an 18 Month Time Period

Eighteen month follow-up data were available for 24 of the 40 participants with cSLE. In an effort to identify those Ped-ANAM subtests that most closely capture changes in cognition over time, we examined Spearman correlations between changes in Ped-ANAM scores (AC, MNC, CVc) per subtest and the changes in cognitive performance (z-scores) as per repeat formal neuropsychological testing 18 months apart (**Table 5**). Differences in the Ped-ANAM scores were more closely associated with changes in visuoconstructional ability, followed by those in processing speed and attention, and to a much lesser degree of working memory

DISCUSSION

Neuropsychiatric SLE is thought to be more common in children as compared to adults, with up to 95% of children manifesting at least one symptom of neuropsychiatric SLE and NCD being present in up to 55% according to some studies (2, 27). NCD may be transitory (28) or persistent over years (29) and can occur in patients without other cSLE disease activity (30, 31). Because its symptoms are often subtle, the identification of NCD is sometimes difficult, requiring formal neuropsychological testing to be recognized. In an effort to allow efficient screening for NCD in a clinical setting, we explored the usefulness of the Ped-ANAM (16). In assessing the psychometric properties of the Ped-ANAM in cSLE, we found the administration of this computer test to be feasible, and also found that selected subtests and indexes had good reproducibility, criterion and concurrent validity and responsiveness to cognitive change, suggesting that these measures may be a useful screening tool for NCD.

Besides newly providing estimates of the reproducibility of the Ped-ANAM, our results support its construct validity for screening overall cognitive function in pediatric populations. Ped-ANAM scores with normal cognition significantly differed from those with the presence of mild/moderate and moderate/severe NCD, as indexed by traditional

neuropsychological testing. As would be expected based on the study of the Ped-ANAM in other pediatric diseases, these differences in Ped-ANAM scores were present irrespective of a diagnosis of cSLE. Our results are also in line with previous studies in adults and children with SLE that found a large number of the ANAM and Ped-ANAM scores to be associated with and predictive of formal cognitive test performance (3, 8). Likewise, responsiveness of Ped-ANAM to cognitive change is supported by associations of changes in the Ped-ANAM scores with the changes in formal neuropsychological test scores over time.

Low Ped-ANAM throughput (TP) scores have been reported in numerous clinical populations as demonstration of cognitive difficulties (8, 32-36). However, we did not find the TP scores to be useful to differentiate study participants according to the level of their cognitive function, most likely due to atypical performance patterns which led to this variable being an overestimate of actual cognitive abilities. Instead, we found other Ped-ANAM scores that measure accuracy (AC) and the speed of accurate responses in terms of overall mean (MNC) and variability (CVc) were better suited to estimate cognitive function than TP score (10). Likewise, Hanley et al. pointed out shortcomings of the TP scores on tasks that assess higher cognitive functions (e.g., working memory or executive function).

As noted in our previous study, AC scores for all Ped-ANAM subtests were generally high, irrespective of cognitive abilities of the participants. While this ensures that most individuals can complete the Ped-ANAM successfully (thereby supporting feasibility), the pronounced negative skew of AC scores represents a “ceiling effect” which limits their utility. It is likely that the ability to identify subtle NCD and possibly changes in NCD over time would be improved if the Ped-ANAM subtests were more difficult, resulting in more normally distributed AC scores. Given the versatility of the Ped-ANAM metrics, this hypothesis could be easily tested by modifying Ped-ANAM task parameters.

We found the Ped-ANAM to have moderate to substantial reproducibility when completed within one day, similar to what has been reported for the original (adult) ANAM (37). Furthermore, it has been suggested previously that the performance of test takers improves most between the first and second administrations, with marginal improvement upon a third administration (38) on the same day. This raises the possibility that our study provides conservative estimates of the Ped-ANAM's reproducibility.

Our data confirm the utility of the *Continuous Performance Test* and *Spatial Processing* subtests of the Ped-ANAM as screening tasks for NCD (10) and also indicate that the *Matching to Sample* and *Code Substitution Delayed* subtests are similarly useful. For reasons not well understood and different from our previous study, the *Mathematical Processing* subtest did not add to the identification of NCD. One reason may be that the participants of the current study were younger than those in our earlier study. Alternatively, as the previous study of only 27 cSLE patients was underpowered, the *Mathematical Processing* subtest association with NCD might constitute a type 2 error. Either way, these inconsistencies suggest that it is premature to use fewer Ped-ANAM subtests in future studies of cSLE-associated NCD.

We also provide initial evidence that the Ped-ANAM is responsive to change in cognition of children and adolescents with cSLE. Hence, even in the context of the ongoing brain development expected in this age group, our findings are in line with those reported for the ANAM in adults with other neuropsychiatric diseases unrelated to SLE (8). Nonetheless, a more detailed evaluation of the Ped-ANAM's ability to detect clinically relevant changes in cognition of children and adolescents with cSLE remains warranted.

A limitation of our study may be that the sampling strategy used does not allow for the provision of valid estimates of the prevalence of cSLE-associated NCD. This is because we used a convenience sample of patients that was not sampled in a strict consecutive fashion from clinic. Likewise, the best-friend matching strategy used to correct for sociodemographic variables affecting cognition resulted in a sizeable number of controls having NCD. None of the above, however, has affected the findings of the reliability and validity of the Ped-ANAM as is supported by exploratory analysis which assessed the measurement properties of the Ped-ANAM in cases and controls separately (data not shown).

In summary, we established that the Ped-ANAM has promising psychometric properties when used in cSLE and healthy pediatric controls, largely confirming the results of a previous pilot study. This is important because it suggests that this tool, which can be administered in busy clinics by non-specialist staff, can feasibly track functioning over time for “early warning signs” of cognitive decline, and can screen for patients who are in need for more specialized neuropsychological follow-up and interventions. However, prior to the use of the Ped-ANAM in clinical care, a more detailed assessment of the tool's discriminant validity appears warranted, including the delineation of minimally clinically important differences in Ped-ANAM scores. Ongoing studies are expected to provide a solid reference range of Ped-ANAM performance scores in healthy children with different socioeconomic backgrounds.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are indebted to Dr. Joseph Bleiberg, Dr. Dennis Reeves, Kerry Culligan, and Children's National Medical Center for their essential involvement in development of the pediatric version of the ANAM. Furthermore, this study would not have been possible without the dedicated clinical research personnel, namely Aimee Baker, Dina Blair, Adlin Cedeno, and Ashia Ali.

We thank Drs. Esi Morgan-DeWitt, Dan Lovell, Alexei Grom, Tracy Ting, Michael Henrickson and PNP Janalee Taylor for providing us with access to their patients with cSLE.

We would also like to thank Meredith Amaya, April German, Allison Clarke, Kate Dahl, Antoinette Dezzutti, Lev Gottlieb, Jennifer Heil, Jennifer Keller, Andrew Phillips, Michal Rischall, Rebecca Wasserman Lieb, Lisa Welcome, Donna Diedenhofer, and Cindy Scharf and Mariah Wells for their assistance with neuropsychological testing. A special thanks to Mrs. Elaine Holtkamp for her administrative support of the study and assistance with the manuscript.

This study is supported an NIAMS Center of Clinical Research Award, NIAMS P60 AR47784 and Institutional Clinical and Translational Science Awards, NIH/NCRR Grant Number 8 UL1 TR000077-04 and UL1 RR026314. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

LITERATURE

1. The American College of Rheumatology nomenclature and case definitions for neuropsychiatric lupus syndromes. *Arthritis Rheum.* 1999; 42(4):599–608. [PubMed: 10211873]
2. Sibbitt WL Jr, Brandt JR, Johnson CR, Maldonado ME, Patel SR, Ford CC, et al. The incidence and prevalence of neuropsychiatric syndromes in pediatric onset systemic lupus erythematosus. *J Rheumatol.* 2002; 29(7):1536–42. [PubMed: 12136916]
3. Holliday SL, Navarrete MG, Hermosillo-Romo D, Valdez CR, Saklad AR, Escalante A, et al. Validating a computerized neuropsychological test battery for mixed ethnic lupus patients. *Lupus.* 2003; 12(9):697–703. [PubMed: 14514133]

4. Sweet JJ, Doninger NA, Zee PC, Wagner LI. Factors influencing cognitive function, sleep, and quality of life in individuals with systemic lupus erythematosus: a review of the literature. *Clin Neuropsychol*. 2004; 18(1):132–47. [PubMed: 15595365]
5. Ross GS, Zelko F, Klein-Gitelman M, Levy DM, Muscal E, Schanberg LE, et al. A proposed framework to standardize the neurocognitive assessment of patients with pediatric systemic lupus erythematosus. *Arthritis Care Res (Hoboken)*. 2010; 62(7):1029–33. [PubMed: 20589693]
6. Levinson D, Reeves D, Watson J, Harrison M. Automated neuropsychological assessment metrics (ANAM) measures of cognitive effects of Alzheimer's disease. *Arch Clin Neuropsychol*. 2005; 20(3):403–8. [PubMed: 15797175]
7. C-Shop. Automated Neuropsychological Assessment Metrics (Version 4). Center for the Study of Human Operator Performance (C-SHOP; Norman, OK: 2007.
8. Roebuck-Spencer TM, Yarboro C, Nowak M, Takada K, Jacobs G, Lapteva L, et al. Use of computerized assessment to predict neuropsychological functioning and emotional distress in patients with systemic lupus erythematosus. *Arthritis Rheum*. 2006; 55(3):434–41. [PubMed: 16739211]
9. C-Shop. ANAM4 PED: User Manual. Center for the Study of Human Operator Performance; Norman, OK: 2007.
10. Brunner HI, Ruth NM, German A, Nelson S, Passo MH, Roebuck-Spencer T, et al. Initial validation of the Pediatric Automated Neuropsychological Assessment Metrics for childhood-onset systemic lupus erythematosus. *Arthritis Rheum*. 2007; 57(7):1174–82. [PubMed: 17907235]
11. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum*. 1997; 40(9):1725. [PubMed: 9324032]
12. Reiter-Purtill J, Gerhardt CA, Vannatta K, Passo MH, Noll RB. A controlled longitudinal study of the social functioning of children with juvenile rheumatoid arthritis. *J Pediatr Psychol*. 2003; 28(1):17–28. [PubMed: 12490627]
13. Zelko F, Beebe D, Baker A, Nelson S, Cedeno A, Klein-Gitelman MS, et al. Predictors of school competence in childhood-onset systemic lupus erythematosus. *Arthrit Care Res*. 2012; 65 (in press).
14. Kane RL, Kay GG. Computerized assessment in neuropsychology: a review of tests and test batteries. *Neuropsychol Rev*. 1992; 3(1):1–117. [PubMed: 1300218]
15. Kane RL, Roebuck-Spencer T, Short P, Kabat M, Wilken J. Identifying and monitoring cognitive deficits in clinical populations using Automated Neuropsychological Assessment Metrics (ANAM) tests. *Arch Clin Neuropsychol*. 2007; 22(Suppl 1):S115–26. [PubMed: 17101258]
16. Reeves, D.; Bleiberg, J.; Winter, K.; Roebuck-Spencer, TM. ANAM - Pediatric Version (Ped-ANAM) User's Manual. ; 2004. National Rehabilitation Hospital. Washington, D.C.; Washington: 2004.
17. Thorne DR. Throughput: a simple performance index with desirable characteristics. *Behav Res Methods*. 2006; 38(4):569–73. [PubMed: 17393825]
18. Hanly JG, Omisade A, Su L, Farewell V, Fisk JD. Assessment of cognitive function in systemic lupus erythematosus, rheumatoid arthritis, and multiple sclerosis by computerized neuropsychological tests. *Arthritis Rheum*. 2010; 62(5):1478–86. [PubMed: 20155829]
19. Antonchak MA, Saoudian M, Khan AR, Brunner HI, Luggen ME. Cognitive dysfunction in patients with systemic lupus erythematosus: a controlled study. *J Rheumatol*. 2011; 38(6):1020–5. [PubMed: 21459946]
20. Ivins BJ, Kane R, Schwab KA. Performance on the Automated Neuropsychological Assessment Metrics in a nonclinical sample of soldiers screened for mild TBI after returning from Iraq and Afghanistan: a descriptive analysis. *J Head Trauma Rehabil*. 2009; 24(1):24–31. [PubMed: 19158593]
21. Williams TS, Aranow C, Ross GS, Barsdorf A, Imundo LF, Eichenfield AH, et al. Neurocognitive impairment in childhood-onset systemic lupus erythematosus: measurement issues in diagnosis. *Arthritis Care Res (Hoboken)*. 2011; 63(8):1178–87. [PubMed: 21560254]
22. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979; 86(2):420–8. [PubMed: 18839484]

23. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Int J Nurs Stud.* 2010; 47(8):931–6.
24. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982; 143(1):29–36. [PubMed: 7063747]
25. Webb, AR. *Statistical pattern recognition.* 2nd ed.. West Sussex, England; New Jersey: Wiley: 2002.
26. Brunner HI, Feldman BM, Bombardier C, Silverman ED. Sensitivity of the Systemic Lupus Erythematosus Disease Activity Index, British Isles Lupus Assessment Group Index, and Systemic Lupus Activity Measure in the evaluation of clinical change in childhood-onset systemic lupus erythematosus. *Arthritis Rheum.* 1999; 42(7):1354–60. [PubMed: 10403262]
27. Yancey CL, Doughty RA, Athreya BH. Central nervous system involvement in childhood systemic lupus erythematosus. *Arthritis Rheum.* 1981; 24(11):1389–95. [PubMed: 7317116]
28. Hanly JG, Fisk JD, Sherwood G, Eastwood B. Clinical course of cognitive dysfunction in systemic lupus erythematosus. *J Rheumatol.* 1994; 21(10):1825–31. [PubMed: 7837145]
29. Carlomagno S, Migliaresi S, Ambrosone L, Sannino M, Sanges G, Di Iorio G. Cognitive impairment in systemic lupus erythematosus: a follow-up study. *J Neurol.* 2000; 247(4):273–9. [PubMed: 10836619]
30. Denburg SD, Carbotte RM, Denburg JA. Cognitive impairment in systemic lupus erythematosus: a neuropsychological study of individual and group deficits. *J Clin Exp Neuropsychol.* 1987; 9(4): 323–39. [PubMed: 3597726]
31. Hanly JG. Disease activity, cumulative damage and quality of life in systematic lupus erythematosus: results of a cross-sectional study. *Lupus.* 1997; 6(3):243–7. [PubMed: 9104730]
32. Kane RL, Roebuck-Spencer T, Short P, Kabat M, Wilken J. Identifying and monitoring cognitive deficits in clinical populations using automated neuropsychological assessment metrics (ANAM) tests. *Arch Clin Neuropsychol.* 2006
33. Eonta SE, Carr W, McArdle JJ, Kain JM, Tate C, Wesensten NJ, et al. Automated Neuropsychological Assessment Metrics: repeated assessment with two military samples. *Aviat Space Environ Med.* 2011; 82(1):34–9. [PubMed: 21235103]
34. Vincent AS, Bleiberg J, Yan S, Ivins B, Reeves DL, Schwab K, et al. Reference data from the automated Neuropsychological Assessment Metrics for use in traumatic brain injury in an active duty military sample. *Mil Med.* 2008; 173(9):836–52. [PubMed: 18816922]
35. Jones WP, Loe SA, Krach SK, Rager RY, Jones HM. Automated neuropsychological assessment metrics (ANAM) and Woodcock-Johnson III Tests of Cognitive Ability: a concurrent validity study. *Clin Neuropsychol.* 2008; 22(2):305–20. [PubMed: 17853133]
36. Cernich A, Reeves D, Sun W, Bleiberg J. Automated Neuropsychological Assessment Metrics sports medicine battery. *Arch Clin Neuropsychol.* 2007; 22(Suppl 1):S101–14. [PubMed: 17118625]
37. Segalowitz SJ, Mahaney P, Santesso DL, MacGregor L, Dywan J, Willer B. Retest reliability in adolescents of a computerized neuropsychological battery used to assess recovery from concussion. *NeuroRehabilitation.* 2007; 22(3):243–51. [PubMed: 17917174]
38. Bendetto, J.; Harris, W.; Goernert, P. *Automated neuropsychological assessment metrics (ANAM) performance stability during repeated cognitive assessments.* Mankato State University; Mankato, MN: 1995.

SIGNIFICANCE & INNOVATION

Innovation

This study supports the utility of the Ped-ANAM as a screening tool for neurocognitive dysfunction in childhood-onset systemic lupus erythematosus

Significance

- The Ped-ANAM is a feasible screening device with good reproducibility and criterion and concurrent validity in this population.
- Additionally, neurocognitive dysfunction can be determined with high accuracy
- Initial evidence of the Ped-ANAM's responsiveness to change in cognitive ability over time is provided.

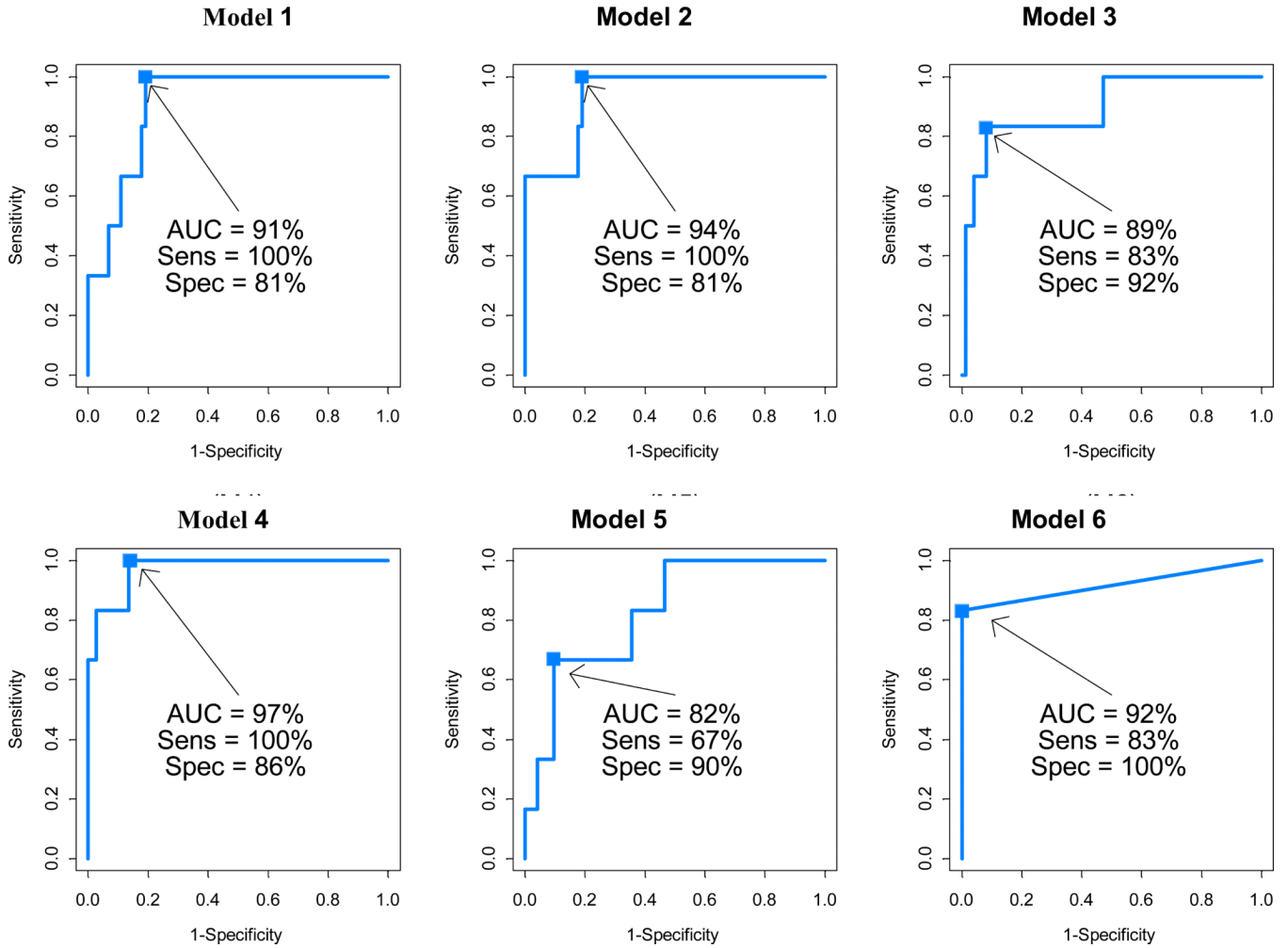


Figure 1. Receiver operating characteristic (curves of detecting moderate/severe NCD based on the performance scores of select Ped-ANAM subtests

AUC: area under the receiver operative characteristic curve (ROC); Sens: sensitivity; Spec: specificity

Model 1: ROC curve from Model1 using stepwise selected AC measures;

Model 2: ROC curve from Model2 using stepwise selected CVc measures;

Model 3: ROC curve from Model3 using stepwise selected MNc measures;

Model 4: ROC curve from Model4 using stepwise selected AC, CV and MNc measures;

Model 5: ROC curve using the model developed in pilot study; ref (11);

CART model: Score defined as 1 if CPT CV>0.6; 2 if CPT CVc 0.6 but M2S CVc>1.4; 3 if CPT CVc 0.6 and M2S CVc 1.4 but CPT MNc >375; and 4 if CPT CVc 0.6 and M2S CcV 1.4 but CPT MNc <375.

Table 1

Tests Used to Define Neurocognitive Dysfunction

Domain	Measure	Source	Description
<i>Working Memory</i>	Digit Span	Age-appropriate Wechsler Intelligence Scale ^{1,2}	Ability to repeat back in order, or in a re-sequenced order, increasingly difficult strings of numbers
	Letter-Number Sequencing	Age-appropriate Wechsler Intelligence Scale ^{1,2}	Ability to mentally resequence a series of letters and numbers before repeating them back
<i>Psychomotor Speed</i>	Coding	Age-appropriate Wechsler Intelligence Scale ^{1,2}	Test-takers “decode” and transcribe a series of symbols as quickly as possible
	Symbol Search	Age-appropriate Wechsler Intelligence Scale ^{1,2}	Score reflects speed and accuracy of test-takers’ visual search for matches in rows of symbols
<i>Attention</i>	Hit Reaction Time Standard Error	Conners Continuous Performance Test II ³	On a 15-minute-long boring task, the variability in reaction time to specific letters flashing on screen
	Inhibition vs. Color Naming Score	Delis-Kaplan Executive Functioning System ⁴	Relative ability to focus on the color of the ink in which a conflicting color word is printed (e.g., “blue” written in red ink).
<i>Visuoconstructive Abilities</i>	Block Design	Wechsler Abbreviated Scales of Intelligence ⁵	Ability to efficiently reproduce colored line drawings using blocks with sides that have varying patterns
	Block Counting	Kaufman Assessment Battery for Children ^{6,7}	Ability to mentally represent the volume of a 3-dimensional block construction printed in 2-dimensional space.

¹Wechsler, D. 2003. *Wechsler Intelligence Scale for Children – Fourth Edition*. Harcourt Assessment, Inc, San Antonio, TX.

²Wechsler, D. 2008. *Wechsler Adult Intelligence Scale – Fourth Edition*. NCS Pearson, San Antonio, TX.

³Conners, C. 2004. *Conners’ Continuous Performance Test II N*. Tonawanda, NY

⁴Delis, D. C., E. Kaplan, and J. H. Kramer. 2001. *The Delis-Kaplan Executive Function System: Technical Manual*. San Antonio

⁵Wechsler, D. 1999. *Wechsler abbreviated scale of intelligence : WASI*. Psychological Corporation, Harcourt Brace, San Antonio .

⁶Daleo, D. V., B. R. Lopez, J. C. Cole, A. S. Kaufman, N. L. Kaufman, B. L. Newcomer, and C. E. Murphy. 1999. K-ABC simultaneous processing, DAS nonverbal reasoning, and Horn’s expanded fluid-crystallized theory. *Psychol Rep* 84:563-574.

⁷Kaufman, A. S., M. R. O’Neal, A. H. Avant, and S. W. Long. 1987. Introduction to the Kaufman Assessment Battery for Children (K-ABC) for pediatric neuroclinicians. *J Child Neurol* 2:3-16.

Table 2Reproducibility of the Ped-ANAM Scores [†]

Ped-ANAM Subtests	Percentage of correct responses (AC)	Mean reaction time for correct response (MNe)	Consistency (CVc = SD of MNe/ MNe)
Code Substitution Delayed	0.38 (0.21, 0.55)	0.58 (0.44, 0.71)	0.21 (0.02, 0.40)
Code Substitution	0.44 (0.27, 0.61)	0.80 (0.73, 0.88)	0.43 (0.27, 0.60)
Continuous Performance Test	0.78 (0.70, 0.86)	0.71 (0.61, 0.81)	0.68 (0.57, 0.79)
Logical Relations	0.23 (0.03, 0.43)	0.77 (0.69, 0.86)	0.15 (0.00, 0.35)
Matching to Sample	0.52 (0.37, 0.67)	0.64 (0.52, 0.76)	0.57 (0.42, 0.72)
Matching Grids	0.43 (0.25, 0.60)	0.77 (0.68, 0.85)	0.38 (0.21, 0.56)
Mathematical Processing	0.25 (0.05, 0.45)	0.91 (0.87, 0.95)	0.62 (0.49, 0.75)
Procedural Reaction Time	0.10 (0.00, 0.31)	0.47 (0.32, 0.62)	0.16 (0.00, 0.38)
Spatial Processing	0.26 (0.03, 0.49)	0.71 (0.61, 0.82)	0.34 (0.17, 0.51)
Sternberg Memory Search	0.55 (0.41, 0.69)	0.52 (0.38, 0.67)	0.27 (0.10, 0.45)
Simple Reaction Time	not estimable	0.68 (0.57, 0.79)	not estimable

[†]Values are intraclass correlation coefficients and 95% confidence intervals (in brackets) adjusted for multiple comparisons using a Tukey's method.

Table 3

Differences in Ped-ANAM performance with differences in cognitive function as measured by formal neuropsychological assessment

Ped-ANAM scores #	Ped-ANAM Subtest	Mean ± SE				p-value [†]	
		(I) Normal cognition	(II) Mild NCD [‡]	(III) Moderate/severe NCD [‡]	(I) vs (III) (II) vs (III)		
AC	<i>Code Substitution Delayed</i>	79.61 ± 1.82	77.91 ± 2.54	64.27 ± 5.19	NS	0.018	0.050
	<i>Code Substitution</i>	96.16 ± 0.53	96.02 ± 0.74	89.86 ± 1.51	NS	0.001	0.001
	<i>Continuous Performance Test</i>	83.67 ± 2.79	78.20 ± 3.98	59.49 ± 7.97	NS	0.015	NS
	<i>Logical Relations</i>	95.56 ± 0.79	95.87 ± 1.11	87.06 ± 2.26	NS	0.002	0.002
	<i>Matching to Sample</i>	85.00 ± 2.29	80.52 ± 3.21	63.00 ± 6.56	NS	0.006	0.049
	Matching Grids	92.65 ± 1.27	92.90 ± 1.77	85.83 ± 3.62	NS	NS	NS
	Mathematical Processing	90.00 ± 0.83	90.00 ± 1.16	86.20 ± 2.37	NS	NS	NS
	Procedural Reaction Time	96.07 ± 0.73	94.24 ± 1.03	93.33 ± 2.10	NS	NS	NS
	<i>Spatial Processing</i>	91.99 ± 1.13	90.40 ± 1.58	74.81 ± 3.23	NS	0.000	0.000
	Sternberg Memory Search	92.08 ± 1.14	90.92 ± 1.59	87.78 ± 3.25	NS	NS	NS
	Code Substitution Delayed	0.55 ± 0.03	0.58 ± 0.04	0.76 ± 0.09	NS	NS	NS
	Code Substitution	0.34 ± 0.01	0.36 ± 0.02	0.42 ± 0.04	NS	NS	NS
	<i>Continuous Performance Test</i>	0.32 ± 0.01	0.36 ± 0.02	0.60 ± 0.04	NS	0.000	0.000
	<i>Logical Relations</i>	0.30 ± 0.01	0.32 ± 0.02	0.40 ± 0.04	NS	0.019	NS
<i>Matching to Sample</i>	0.46 ± 0.02	0.45 ± 0.03	0.71 ± 0.06	NS	0.001	0.001	
<i>Matching Grids</i>	0.28 ± 0.01	0.34 ± 0.02	0.38 ± 0.04	0.033	0.048	NS	
Mathematical Processing	0.36 ± 0.02	0.42 ± 0.03	0.48 ± 0.06	NS	NS	NS	
<i>Procedural Reaction Time</i>	0.29 ± 0.02	0.29 ± 0.02	0.41 ± 0.05	NS	0.050	NS	
<i>Spatial Processing</i>	0.36 ± 0.02	0.40 ± 0.02	0.48 ± 0.05	NS	0.030	NS	
Sternberg Memory Search	0.40 ± 0.03	0.51 ± 0.04	0.36 ± 0.08	NS	NS	NS	
<i>Code Substitution Delayed</i>	1.22 ± 0.05	1.29 ± 0.07	0.72 ± 0.15	NS	0.006	0.003	
Code Substitution	1.14 ± 0.04	1.25 ± 0.06	0.98 ± 0.12	NS	NS	NS	
Continuous Performance Test	0.59 ± 0.02	0.63 ± 0.03	0.55 ± 0.05	NS	NS	NS	
Logical Relations	1.54 ± 0.07	1.72 ± 0.10	1.46 ± 0.21	NS	NS	NS	
<i>Matching to Sample</i>	2.08 ± 0.08	2.49 ± 0.12	1.68 ± 0.24	NS	NS	0.009	

Ped-ANAM scores #	Ped-ANAM Subtest	Mean ± SE			p-value †		
		(I) Normal cognition	(II) Mild NCD ‡	(III) Moderate/severe NCD ‡	(I) vs (II)	(I) vs (III)	(II) vs (III)
	Matching Grids	1.71 ± 0.08	2.10 ± 0.11	1.67 ± 0.22	0.009	NS	NS
	Mathematical Processing	1.57 ± 0.10	1.78 ± 0.14	1.79 ± 0.28	NS	NS	NS
	Procedural Reaction Time	0.60 ± 0.02	0.61 ± 0.02	0.57 ± 0.05	NS	NS	NS
	Spatial Processing	2.09 ± 0.09	2.44 ± 0.13	1.71 ± 0.26	NS	NS	0.038
	Stemberg Memory Search	0.92 ± 0.04	1.03 ± 0.06	0.74 ± 0.11	NS	NS	NS
	Simple Reaction Time	0.29 ± 0.01	0.31 ± 0.02	0.43 ± 0.03	NS	0.001	0.003

AC = Percentage of correct responses per subtest; MNC = Mean reaction time for correct response in seconds; CVc = Consistency or SD of reaction time for a correct response / MNC

‡ Neurocognitive dysfunction (NCD) categories are defined as follows based on z-scores of the standardized tests completed for the formal neuropsychological testing in four functional domains (attention/executive function; processing speed; visuoconstructional memory; working memory). Normal if all z scores > -1; mild to moderate NCD if one or two domain z-scores < -1 or one domain z-score < -2; and severe NCD if more than one domain z-score < -2 or more than two domain z-scores < -1.

† p-values are adjusted for multiple comparison using a Tukey's method.

Table 4
 Stepwise logistic regression models of predicting moderate/ severe NCD using select Ped-ANAM performance scores

Ped-ANAM Score #	Ped-ANAM Subtest*	Model 1		Model 2		Model 3		Model 4		Model 5**			
		Slope ± SE	p-value	Slope ± SE	p-value	Slope ± SE	p-value	Slope ± SE	p-value	Slope ± SE	p-value		
AC	Continuous Performance Test	-0.04±0.02	0.043										
	Spatial Processing	-0.17± 0.07	0.020					-0.19±0.15	0.222				
CVc	Continuous Performance Test			14.42±4.80	0.003								
	Matching to Sample			5.21± 2.58	0.043								
	Mathematical Processing									0.96±2.30	0.677		
	Spatial Processing									6.17±3.79	0.103		
MNc	Code Substitution Delayed					-6.73±2.35	0.004			-0.01±0.01	0.064		
Intercept													
				15.42±6.80	0.023	-11.6±3.08	<0.0001	3.80±1.93	0.048	15.42±16.64	0.354	-2.40±1.83	0.188

Accuracy or AC = Percentage of correct responses per subtest; MNc = Mean reaction time for correct response in seconds; Consistency or CVc = SD of reaction time for a correct response / MNc

** Reference :Brunner HI, Ruth NM, German A, Nelson S, Passo MH, Roebuck-Spencer T et al: Initial validation of the Pediatric Automated Neuropsychological Assessment Metrics for childhood-onset systemic lupus erythematosus. Arthritis Rheum 2007, 57(7): 1174-1182.

Association of changes of Ped-ANAM scores and ranges cognitive ability based on formal neuropsychological testing over an 18 month period in 24 cSLE-Patients[§]

Table 5

Ped-ANAM scores [#]	Ped-ANAM Subtest*	Working Memory	Psychomotor Speed	Attention	Visuoconstructional Ability
	Code Substitution Delayed	-	0.30	-	-
	Code Substitution	0.25	-	-	-
	Continuous Performance Test	-	0.34	0.51	-
	Logical Relations	-	-0.23	-	0.21
	Matching to Sample	0.30	-	0.28	0.34
	Matching Grids	-	-	-0.24	-
	Mathematical Processing	-	-	0.51	-
	Procedural Reaction Time	0.39	-	0.32	0.51
	Spatial Processing	0.52	0.20	-	0.36
	Sternberg Memory Search	-	0.29	-	-
	Code Substitution Delayed	-	-	-	-
	Code Substitution	-	-	-0.43	-0.25
	Continuous Performance Test	-0.28	-	-0.51	-0.45
	Logical Relations	-0.26	-	-0.31	-0.22
	Matching to Sample	-0.41	-	-	-0.27
	Matching Grids	-	-	-	-
	Mathematical Processing	-	0.24	-0.26	-
	Procedural Reaction Time	-0.38	-	-	-0.49
	Spatial Processing	-0.22	-0.21	-	-
	Sternberg Memory Search	-	-0.25	-0.31	-
	Code Substitution Delayed	-	-	-	-
	Code Substitution	-	-	-	-
	Continuous Performance Test	-	0.25	-0.23	-
	Logical Relations	-	-	-	-
	Matching to Sample	-	-	-0.40	-0.40
	Matching Grids	-	-	-0.25	-

Ped-ANAM scores #	Ped-ANAM Subtest*	Working Memory	Psychomotor Speed	Attention	Visuoconstruction Ability
	Mathematical Processing	-	-	-	-
	Procedural Reaction Time	-	-	-	-
	Spatial Processing	-	-	-0.24	-
	Sternberg Memory Search	-	-0.23	-0.26	-
	Simple Reaction Time	-	-0.20	-	-

Accuracy or AC = Percentage of correct responses per subtest; Consistency CVc = SD of reaction time for a correct response / MNC; MNC = Mean reaction time for correct response

* Values are Spearman correlation coefficients. Ped-ANAM change score and changes in z-scores per cognitive domain during formal neuropsychological testing; *only* values of $r > |0.2|$ are shown and $r > |0.4|$ are statistically significant at $p < 0.05$.