Epileptic Disord 2017; 19 (2): 166-77

ILAE survey of neuropsychology practice in pediatric epilepsy surgery evaluation

Madison M. Berl [1,2], Mary Lou Smith [3,4], Christine Bulteau [5,6] and the Task Force for Pediatric Epilepsy Surgery for the ILAE Commissions of Pediatrics, Surgical Therapies

[1] Children's National Health System, Division of Pediatric Neuropsychology, Washington, DC
[2] George Washington University, Washington, DC, USA
[3] Department of Psychology, University of Toronto Mississauga, Mississauga
[4] Neurosciences and Mental Health Program, the Hospital for Sick Children, Toronto, Canada
[5] INSERM U1129 "Infantile Epilepsies and Brain Plasticity", Paris ; Université Paris Descartes, Sorbonne Paris Cité; CEA, Gif sur Yvette
[6] Rothschild Foundation Hospital, Pediatric Neurosurgery Department, Paris, France

Received September 15, 2016; Accepted March 05, 2017

ABSTRACT – *Aims.* To determine the extent to which specific neuropsychological measures are in common use around the world for the assessment of children who are candidates for epilepsy surgery.
Materials and methods. As part of the work of the International League Against Epilepsy Pediatric Surgical Task Force, a survey was developed and distributed online. The survey consisted of questions related to demographics, training experience, general practice, and specific measures used and at what frequency.
Results. Seventy-eight clinicians with an average of 13.5 years of experience from 19 countries responded to the survey; 69% were English-speaking. Pre- and post-neuropsychological evaluations were conducted with a majority of children undergoing surgical resection for epilepsy. There was high consistency (>90%) among the domains evaluated, while consistency rate among specific measures was more variable (range: 0-100%). Consistency rates were also lower among respondents in non-English-speaking countries. For English-speaking respondents, at least one measure within each domain was used by a majority (>75%) of clinicians; 19 specific measures met this criterion.
Conclusion. There is consensus of measures used in neuropsychological studies of pediatric epilepsy patients which provides a basis for determining which measures to include in establishing a collaborative data repository to study surgical outcomes of pediatric epilepsy. Challenges include selecting measures that promote collaboration with centers in non-English-speaking countries and providing data from children under age 5.

Key words: measure, assessment, presurgical, outcome, neurodevelopment

Correspondence:
Madison M. Berl
Children's National Health System,
111 Michigan Avenue NW,
Washington DC, 20010, USA
<mberl@childrensnational.org>

doi:10.1684/epd.2017.0908

The role of neuropsychological assessment in pediatric epilepsy surgery

Neuropsychological assessment is accepted as a key component of the evaluation of children being considered for epilepsy surgery. The goal is to evaluate cognition and behavior (Baker, 2001; Baker and Goldstein, 2004; Cross *et al.*, 2006; Jones-Gotman *et al.*, 2010; Helmstaedter and Witt, 2012; Jayakar *et al.*, 2014; Smith and Berl, 2017), with the results providing an understanding of the child's cerebral organization and the site and laterality of the seizure focus. The findings provide an indication of how the epileptic process and any underlying brain dysfunction have impacted the child's development, and potential risks of the planned surgery. The results delineate areas of intact and delayed performance, yielding information on the functional integrity of the epileptogenic area and on the non-epileptogenic regions of the brain. Neuropsychological assessment differs from other diagnostic modalities by its capability to detect the cognitive and psychosocial functional consequences of epilepsy or epilepsy surgery (Jones-Gotman *et al.*, 2010; Sherman *et al.*, 2011; Smith and Berl, 2017). The pre-operative assessment also provides a baseline against which to track the child's development after surgery, and critical information for planning and implementing interventions that allow for optimal cognitive, academic, behavioral, and psychosocial outcomes across the life span.

Neuropsychology, development, and optimizing outcomes of epilepsy surgery

Although epilepsy surgery is widely used for children around the world, and has been for several decades (Cross *et al.*, 2006; Harvey *et al.*, 2008; Jayakar *et al.*, 2014), there are important gaps in knowledge where neuropsychological function could potentially inform on optimal timing and outcomes of surgery. We have limited knowledge of the ideal point in development when to conduct surgery; although there are some data suggesting that early surgery may lead to better cognitive outcomes (Freitag and Tuxhorn, 2005; Honda *et al.*, 2013), other studies have found that children who are older at the time of surgery have superior cognitive outcome (Puka and Smith, 2016). It is unknown whether there may be a key age at which to intervene surgically; is it in the child's best interest to wait until certain cognitive skills are established or is it better to remove the epileptogenic tissue early in the hope that the skill will develop more optimally? This point is worth considering given that functional tissue may be removed during the resection of the epileptogenic area. Development is not a steadily progressive process, particularly in children with epilepsy, and there may be important markers such as plateaus or declines that signal the need for surgical intervention in order to prevent or minimize lasting deficits.

Other questions for which the best evidence entails knowledge of neuropsychological outcomes include whether time to surgery should be shortened if there is an identifiable lesion, regardless if seizure control is obtained using medication, and whether surgical timing should be influenced by the number of failed medications. Although the child may obtain adequate seizure control on medication, this outcome has to be weighed against the potential side effects of medication (Loring *et al.*, 2007; Meador and Loring, 2016). It is unknown whether the location of the epileptogenic region should influence the optimal timing for surgery, and whether different decisions should be made depending on the maturational time frame of the functional area involved. For example, one can ask whether it is advantageous to intervene earlier when the seizure focus is in an area with a shorter developmental window (*i.e.* language with a superior temporal focus) as compared with one with a longer developmental period (*e.g.* executive functioning with a frontal focus). Lateralization, and perhaps localization of function, follows a protracted timespan (Holland *et al.*, 2007; Berl *et al.*, 2010, 2014), such that the mapping of function onto brain regions can be challenging in younger children (Anderson *et al.*, 2011; Smith *et al.*, 2011). What are the functions that become lateralized and at what point in development may they also inform on the timing of surgery? It has been suggested that the maturational stage of a cognitive ability at the time of seizure onset influences whether or not deficits in that function are manifested (Gonzalez *et al.*, 2014), and thus it is possible that optimizing cognitive outcomes may reside critically with such maturational issues.

The field of pediatric neuropsychology itself has needs for guidance in the selection and development of optimal test instruments for children. Systematic research is required to identify which measures are the most accurate for mapping function on to brain regions, and which are the most sensitive in detecting change after surgery (Smith and Berl, 2017). Answers to these questions will lead to the development of evidence-based test batteries.

Impetus for the present survey

The questions raised above are complicated and require large samples of patients to be addressed properly. Pediatric epilepsy surgery centers operate on anywhere from 7 to 120 patients per year, with most centers between 20 and 40 (Cukiert *et al*., 2016), and there is considerable heterogeneity in the age of children, etiologies, location of resection, and the surgical procedures carried out (Harvey *et al*., 2008; Cukiert *et al*., 2016). Thus, no one center has the sample size to answer these complex questions within a reasonable time span. Multicenter studies are needed to collect a sufficient number of cases, and such collaboration requires consistency in data collection. Knowledge of current practice with respect to neuropsychological tests is an important first step in establishing the groundwork for such studies. Therefore, we designed a survey to gather information on the practices of pediatric neuropsychologists with respect to test use. The goal of the survey was to determine the extent to which measures are already in common use which is an important first step to achieve the long-term goal of determining which neuropsychological measures will be meaningful in determining surgical outcomes. An additional and important goal was to determine barriers to collaborative studies.

Neuropsychologists in pediatric epilepsy surgery centers around the world were surveyed to identify current practice and tests commonly used. There are no current sources of such information. A much earlier survey of neuropsychological test use, performed in 1993, yielded responses from 43 centers where children were evaluated, however, on average, pediatric cases constituted a relatively small percent (17%) of the patients seen, and very few neuropsychologists tested children younger than adolescence (Jones-Gotman *et al*., 1993). A more recent survey in 2011 published on responses from 75 epilepsy centers, but included data on test for use with adults only because there were too few responses from pediatric centers (Djordjevic, 2011).

Materials and methods

As part of the work of the International League Against Epilepsy (ILAE) Pediatric Surgical Task Force, a survey was developed and distributed online. The survey is available online at https://cri-datacap.org/surveys/?s=q4xLY4FDUj.

Participants and procedures

A web-based survey was open from June 2, 2014 to September 30, 2015. An invitation explaining and providing a link to the survey was distributed through three sources:

– individual emails of pediatric neuropsychologists provided by members of the ILAE Pediatric Surgical Task Force, a list that was informed by other surveys of the task force (Cukiert *et al*., 2016; Harvey *et al*., 2008);
– posting on the International Mail List for Pediatric Neuropsychology [PED-NPSY@LISTS.UMN.EDU];
– announcement at the end of professional presentations of two ILAE meetings (The European Congress on Epilepsy, Stockholm 2014 and The International Epilepsy Congress, Istanbul 2015).

To complete the survey, email reminders were sent twice to those invited, when an email address was known.

The survey was adapted from previous work (Djordjevic, 2011) and customized for pediatric practice. The survey was in English. A range of question formats was used including open-ended and multiple-choice with the primary goal of maximizing response rate by maintaining time to complete the survey to 15-20 minutes. There was also an option for respondents to simply upload a document that delineated what measures they used for a presurgical battery.

The survey consisted of 30 questions related to demographics, training experience, general practice questions (*i.e.* rationale for test selection, barriers to clinical and research work), and other activities conducted related to presurgical evaluation (*i.e.* IAT/WADA, fMRI, *etc*.). The remainder of the survey focused on specific measures used and at what frequency. If respondents endorsed evaluating a specific domain, they were then asked if they used specific measures. Eighty-three specific measures were listed across 10 domains (Intellectual Functioning, Language, Visual/Motor, Attention/Executive Function [EF], Memory/Learning, Academic Achievement, Social/Emotional, Adaptive Behavior, Quality of Life, and Symptom Validity). Within each domain, "Other" was also an option to enter measures not mentioned. How often a measure was used was ranked with descriptive anchors: standardly used (>85% of the time); often used (50-85% of the time); occasionally (20-50% of the time); rarely (<20% of the time); or never (0% of the time). At the end of the survey, an open text box was available for any other comments. For respondents who uploaded a document, their measures were entered and counted as "standardly used." Descriptive statistics including central tendency measures, range, and frequencies were the primary analyses. Results for test use were reported for the sample as a whole, as well as separately for English (USA, Canada, UK, Australia, and South Africa) and non-English-speaking countries.This project was undertaken as a Quality Improvement Initiative at the Children's National Health System; it did not constitute research on human subjects and as such was not under the oversight of the Institutional Review Board.

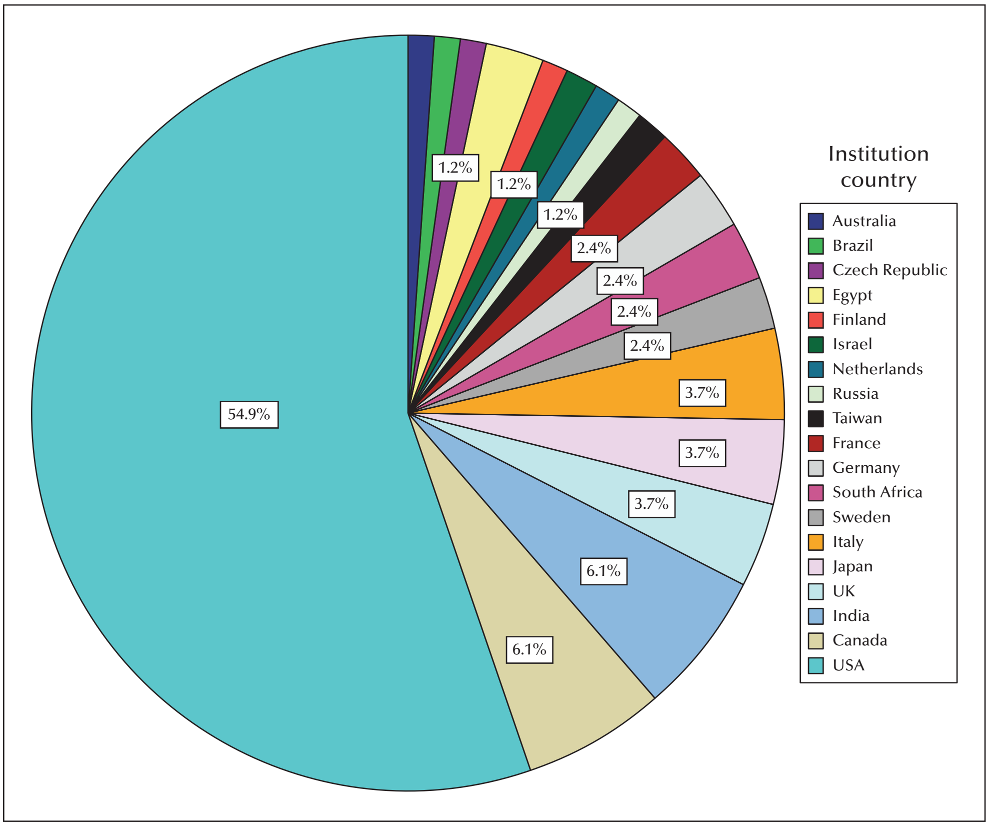

Figure 1. Geographic distribution of respondents.

Results

Eighty-four respondents started the survey, which was completed by 78 respondents and left incomplete by six respondents. Response rate is difficult to calculate, however, the invitation was sent to 216 email addresses, with <5% of emails corresponding to different addresses for the same person. The professional listserve has over 1,900 subscribers, however, not all work in the field of epilepsy, nor conduct presurgical evaluations in children. The majority of responses were from specific email invitations.

Demographics of respondents

Respondents represented 19 countries with 69% of respondents from English-speaking countries (*figure 1*). Seventy different institutions or private practices were represented (*supplementary table 1*). Fifty-eight percent of the respondents were female. Educational background included doctoral (PhD [72%]; PsyD [9%]; EdD [1%]; MD [4%]; MD, PhD [1%]) and master's degrees (11%). Years of experience ranged from 1 to 35 with an average of 13.5. The majority (95%) of respondents identified themselves as neuropsychologists; the remaining four respondents were physicians. A majority (83%) had specific training with epilepsy populations.

Description of typical surgical evaluations

Respondents conducted a neuropsychological evaluation for the majority (91%) of patients undergoing consideration for epilepsy surgery (69%: "almost always"; 22%: "for 75% of cases"). A majority (90%) also conducted post-surgical evaluations (58%: "almost always"; 32%: "for 75% of cases"). Many neuropsychologists (63%) evaluated children aged 12 months or younger while 37% reported a lower age limit ranging from 2 to 6 years old.

The time interval between surgery and follow-up evaluation was not consistent, with only 22% of respondents conducting a postsurgical evaluation at a specific time interval; 62% reported six months postoperatively as the minimum amount of time before conducting an evaluation. Approximately half of the respondents conducted a 12-month follow-up, while others assessed at six months (27%) or two years (21%). Barriers to conducting evaluations were fewer for presurgical compared to postsurgical evaluations. Eighteen percent reported insurance authorization as a barrier to conducting a presurgical evaluation; the only other barrier was not enough time between the request and date of surgery. Barriers to post-surgical follow-up were more varied, including: lack of insurance authorization, loss to follow-up when patients were seizure-free, clinical demands not allowing sufficient time to see follow-up patients, or referrals being made elsewhere.

The majority of respondents (92%) conducted comprehensive evaluations that required at least three hours of testing. Approximately 60% of evaluations were conducted in one session, with <8% occurring over three sessions or more. Of those, 91% used a systematic battery with alternative batteries because of age, functioning level, location of seizure focus, or limited access to measures at the time of evaluation.

There was high consistency among respondents for domains assessed; 93-95% of respondents assessed Intellectual Functioning, Language, Attention/EF, and Memory. Visual/Motor was evaluated by 90% of respondents. Other areas commonly evaluated included: Academic Achievement (72%); Social/Emotional (76%); and Adaptive Behavior (74%). A minority of respondents conducted Quality of Life (37%) and Symptom Validity testing (19%).

Specific measures used across domains

While there was strong consistency at the domain level, there was less consistency among specific measures and this variability was accentuated when examined by English-speaking versus non-English-speaking countries (*table 1*). The range for whether a specific measure was used was 0-100%. Consistency was generally lower for respondents from non-English-speaking countries as the highest rate of consistency was 63%, compared to 100% for English-speaking countries. Although lower, the relative frequencies were generally similar among English and non-English-speaking groups. Only three measures were used more often in non-English-speaking countries: the Trail Making Test (non-DKEFS), Handedness inventory, SNAP-IV (a questionnaire for identifying symptoms of ADHD), and HRQOL.

There was at least one measure in each domain that was clearly used most frequently. The highest consistency across the entire sample was among the Wechsler measures within the Intellectual Functioning domain. As such, within specific domains, any Wechsler subtest or index that also fitted under a certain domain was also assumed to be consistently used. For example, for the Visual domain, the Perceptual Reasoning Index is commonly used. Therefore, these tests are not reiterated below.

Among respondents from non-English-speaking countries, only two other measures were used by a majority of respondents: the Rey-Osterreith Complex Figure Test (RCFT) and Trail Making Test. In contrast, among respondents from English-speaking countries, there were several measures that were used by the majority. Within the Language domain, verbal fluency (NEPSY-II or DKEFS), Comprehension of Instructions (NEPSY-II), Boston Naming Test, and Peabody Picture Vocabulary Test were in common usage. Even though only used by a third of respondents, the Boston Naming Test, NEPSY-II, and Token test were the most common language tasks reported by respondents from non-English-speaking countries.

Within the Visual/Motor domains, the RCFT was used often among all respondents. Respondents from English-speaking countries also often used Beery VMI, Grooved Pegboard, Judgment of Line Orientation (JLO), finger tapping, and grip strength. Within Attention/EF measures, the Trail Making (DKEFS and non-DKEFS) and the Wisconsin Card Sorting Test (WCST) were used often among all respondents. Respondents from English-speaking countries also often used the Conners' CPT and Behavior Rating Inventory of Executive Functioning (BRIEF).

Within Memory measures, the RCFT was often used among all respondents. Respondents from English-speaking countries also often used the California Verbal Learning Test for Children (CVLT-C), NEPSY-II, Children's Memory Scale (CMS), and WRAML-2. Specific subtests for the CMS included Stories (59.3%), Dots (50%), Faces (46.3%), and Word Pairs (31.5%); for the WRAML-2, Story Memory (55.6%), Verbal Learning (42.6%), Design Memory (31.5%), and Picture Memory (25.9%) were included.

For Academic Achievement, the Woodcock-Johnson was used by a majority of respondents from English-speaking countries. Among Social/Emotional and Adaptive Behavior measures, the Behavior Assessment System for Children, Second Edition (BASC-2); Children's Depression Inventory, Second Edition (CDI-2); Child Behavior Checklist (CBCL); Adaptive Behavior Assessment System, Second Edition (ABAS-II); and Vineland Adaptive Behavior Scales, Second Edition (VABS-II) were used by a majority of respondents from English-speaking countries. There were no measures within Quality of Life or Symptom Validity domains that were commonly used.

Over 45 measures across all domains were named in the "other" option, however, there was no measure endorsed by more than five respondents. Among the measures included in the option to write-in tests that were not specifically included within the survey, several were tests administered to low functioning/non-verbal children (including the Griffith Mental Development Scales, Leiter, and Test of Nonverbal Intelligence), while others were different or older versions of commonly used tests (*e.g.* WISC-III, Bayley-II, Purdue Pegboard, Rey Auditory Verbal Learning Test, *etc.*).

Two thirds of the respondents endorsed two primary reasons for using specific measures: the use of historical data and in their opinion, it was the best measure. A minority of respondents also used measures based on recommendations from the National Institute of Neurological Disorders and Stroke (NINDS) Epilepsy Common Data Elements (CDE) or based their use simply on what was available to them. Notably, a majority of respondents (96.3% of English-speakers/83.3% of non-English speakers) reported their willingness to adopt alternative measures to what they currently used.

Other surgical team activities

The majority (82%) of respondents presented neuropsychological results during surgical team meetings. Respondents were also involved in mapping of cognitive functions including IAT/Wada (54%), bedside mapping (40%), fMRI (25%), DTI (6%), intraoperative mapping (3%), and TMS (3%). Other techniques endorsed by one respondent included MEG, PET, and SPECT.

With respect to IAT/Wada, 11% of respondents noted that they conducted them at least on a monthly basis, but most respondents conducted them only 1-4 times per year. Seventy percent of respondents endorsed that there has been a decrease in use of IAT/Wada due to the increased use of other mapping techniques. Of the IAT/Wada protocols used, 62% used a protocol from the literature or elsewhere, but 17% did not know the basis of their protocol. The protocols used most often were from the Medical College of Georgia (MCG) (Loring *et al.*, 2000, 2009; Lee *et al.*, 2002) and the Montreal Neurological Institute (MNI) (Dade and Jones-Gotman, 1997; Jones-Gotman *et al.*, 2005). The drug used in the IAT/Wada procedure was most often sodium amobarbital but availability of the drug was sometimes an issue. Other drugs used included etomidate, methohexital, and propofol. Forty percent of respondents conducted IAT/Wada in children as young as 4-6 years old, and the remainder had a minimum age of 7 or older.

For fMRI, 94% mapped language functions, 74% mapped motor functions, and 42% mapped memory functions. Of the 18 respondents who used language fMRI, tasks used included noun-verb generation (79%), verbal fluency (58%), listening to stories (58%), and auditory description decision/sentence judgment (42%). One respondent used sedated fMRI and during which passive listening to a story and passive movement were conducted. Of the eight respondents who used memory fMRI, tasks used included encoding scenes with post-task testing (50%), word encoding with post-task testing (38%), hometown/school walking (25%), and paired word association (25%). Of the 14 respondents who used motor fMRI, tasks used included finger tapping (93%), foot tapping (57%), and tongue wiggling (21%).

Discussion

Our report of current clinical neuropsychological practice within pediatric epilepsy surgical evaluation is based on a strong international response to a web-based survey among an experienced group of clinicians with specific training in epilepsy. A goal of the survey was to determine what measures are commonly in use among experts in the field. For English-speaking sites, there is at least one measure used by a majority of clinicians across domains. We highlight 19 measures that are used by a majority of English-speaking sites. We determined these 19 by selecting the measures with the highest percentages of use within each domain and any other tests commonly used by at least 75% of the sample. There was not a commonly used measure for quality of life or symptom validity testing, however, these domains are relatively newer areas of assessment. In fact, the original NINDS Epilepsy CDE did not list a quality of life measure because no scale was sufficiently studied to be recommended in epilepsy populations (Loring *et al.*, 2011). An update of the CDE again noted that no one scale can be recommended, but seven measures were discussed that included four epilepsy-specific scales and three generic scales (Austin *et al.*, 2012). Despite only a minority of respondents indicating that they considered the NINDS Epilepsy CDE, 11 of the 19 measures that were the most commonly used were also recommended by the NINDS Epilepsy CDE Working Group. In line with the NINDS Epilepsy CDE recommendations, most of the commonly used measures were for children aged 5-16. Thus, an area of need is to identify sensitive measures for children under the age of 5 given that many children in this age range undergo surgery (Wyllie *et al.*, 1998; Harvey *et al.*, 2008).

The drawback of the consensus approach used in this survey is that perhaps none of the measures listed are effective and efforts should be spent developing new measures. However, the approach here is pragmatic by leveraging what is already being done to conduct more definitive studies on the utility of specific measures. By pooling existing data, neuropsychologists can begin to address the utility of the commonly used tests. For example, the RCFT is used by a

Table 1. Use of neuropsychological measures[1].

				Frequency (%), *n*=78			
	Use overall (%), *n*=78	**English-speaking (%), *n*=54**	**non-English-speaking (%), *n*=24**	**Rarely**	**Occasionally**	**Often**	**Standardly**
INTELLECTUAL FUNCTIONING TESTS							
WISC-IV*	88.4	100.0	62.5	5.1	15.4	19.2	48.7
WAIS-IV*	74.3	92.7	33.3	38.5	11.5	6.4	17.9
WPPSI-IV**	61.5	72.2	37.5	20.5	15.4	7.7	17.9
WASI-2*	43.6	63.0	0.0	20.5	10.3	6.4	6.4
Bayley-III**	39.7	50.1	16.6	32.1	3.8	0.0	3.8
DAS	32.1	46.3	0.0	21.8	5.1	2.6	2.6
Mullen**	20.6	29.7	0.0	10.3	6.4	2.6	1.3
SB-V	20.5	26.0	8.4	15.4	3.8	0.0	1.3
K-ABC	12.8	14.9	8.3	11.5	1.3	0.0	0.0
RIAS	7.7	11.2	0.0	6.4	0.0	1.3	0.0
LANGUAGE TESTS							
NEPSY-II	70.5	87.0	33.3	7.7	20.5	23.1	19.2
BNT*	65.4	77.9	37.5	12.8	14.1	14.1	24.4
DKEFS VF	62.8	83.3	16.7	12.8	5.1	23.1	21.8
PPVT-4**	57.7	76.0	16.7	20.5	14.1	10.3	12.8
EOWPVT	42.3	61.1	0.0	12.8	6.4	10.3	12.8
CELF 4*	39.7	55.6	4.2	15.4	12.8	5.1	6.4
TokenTest	34.6	35.2	33.4	10.3	6.4	5.1	12.8
COWA*	30.8	40.8	8.4	7.7	2.6	7.7	12.8
CTOPP 2	25.7	37.1	0.0	14.1	10.3	0.0	1.3
PLS-5	11.6	16.8	0.0	9.0	1.3	1.3	0.0
TLC	3.9	5.6	0.0	2.6	1.3	0.0	0.0
Menyuk	2.6	3.7	0.0	2.6	0.0	0.0	0.0
TOLD4	1.3	1.9	0.0	1.3	0.0	0.0	0.0

Table 1. Use of neuropsychological measures (*continued*).

VISUAL/VISUAL-MOTOR/FINE MOTOR TESTS							
RCFT	74.4	79.7	62.6	10.3	15.4	29.5	19.2
Beery VMI	69.2	87.1	29.2	6.4	9.0	14.1	39.7
Grooved Pegboard*	62.8	81.6	20.8	6.4	3.8	9.0	43.6
JLO	51.3	64.8	20.9	19.2	14.1	9.0	9.0
Finger Tapping	48.7	59.3	25.1	16.7	9.0	5.1	17.9
NEPSY-IIVisuomotor	47.5	55.6	29.2	24.4	12.8	7.7	2.6
Grip Strength	37.2	50.1	8.4	14.1	5.1	2.6	15.4
Handedness Inventory	35.9	33.4	41.7	6.4	3.8	10.3	15.4
Hooper VOT	26.8	33.4	12.5	11.5	7.7	3.8	3.8
WravmaPegboard	19.3	27.9	0.0	14.1	1.3	2.6	1.3
MFVPT	19.3	22.3	12.5	7.7	6.4	2.6	2.6
TVP	14.2	20.4	0.0	9.0	2.6	2.6	0.0
FMS-2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ATTENTION/EXECUTIVE FUNCTIONING TESTS							
BRIEF*	69.3	88.8	25.0	7.7	6.4	9.0	46.2
WCST*	64.0	72.3	45.9	17.9	14.1	12.8	19.2
DKEFS	61.5	81.5	16.7	9.0	11.5	14.1	26.9
Conners CPT*	57.7	70.5	29.2	10.3	12.8	12.8	21.8
TrailMaking (non-DKEFS)*	52.6	50.0	58.3	9.0	7.7	16.7	19.2
TEA-Ch	43.6	51.9	25.0	19.2	10.3	9.0	5.1
TOL-DX	39.7	42.6	33.3	11.5	9.0	6.4	12.8
TOVA	12.9	14.9	8.4	6.4	2.6	2.6	1.3
ACT	6.4	7.5	4.2	5.1	1.3	0.0	0.0
TEC	3.9	3.7	4.2	2.6	1.3	0.0	0.0
MEMORY/LEARNING TESTS							
RCFT	73.1	83.4	50.0	10.3	11.5	26.9	24.4
CVLT-C*	59.0	74.1	25.0	10.3	5.1	7.7	35.9
NEPSY II	52.5	66.7	20.9	14.1	19.2	11.5	7.7
CMS	51.3	68.6	12.5	9.0	1.3	17.9	23.1
WRAML-2	43.6	63.0	0.0	7.7	7.7	11.5	16.7
WMS IV	26.9	31.5	16.7	7.7	2.6	5.1	11.5
TOMAL2	15.4	18.6	8.3	1.3	5.1	1.3	7.7

Table 1. Use of neuropsychological measures (*continued*).

ACADEMIC ACHIEVEMENT TESTS							
WJ-III	39.8	57.4	0.0	7.7	5.1	10.3	16.7
WIAT-3	32.1	46.4	0.0	3.8	2.6	9.0	16.7
Bracken 1-5 SRC	30.8	44.5	0.0	11.5	15.4	2.6	1.3
WRAT 4	30.7	42.6	4.2	12.8	5.1	3.8	9.0
GORT-5	21.8	31.6	0.0	12.8	6.4	1.3	1.3
TOWL-4	18.0	26.0	0.0	14.1	0.0	1.3	2.6
TOWRE	6.4	9.4	0.0	3.8	1.3	1.3	0.0
SOCIAL/EMOTIONAL TESTS							
BASC-2	48.8	68.6	4.2	11.5	2.6	2.6	32.1
CBCL*	43.6	50.1	29.2	5.1	1.3	1.3	35.9
CDI-2	43.6	57.4	12.5	15.4	10.3	5.1	12.8
Conners Rating Scales	30.8	35.2	20.8	14.1	9.0	2.6	5.1
RCMAS-2	24.3	33.4	4.2	7.7	5.1	3.8	7.7
MMPI-A	21.8	27.8	8.4	17.9	2.6	0.0	1.3
SCQ	20.6	25.9	8.3	10.3	7.7	0.0	2.6
MASC	18.0	24.2	4.2	10.3	3.8	2.6	1.3
DuPaul ADHD Rating Scale	15.5	18.7	8.4	2.6	1.3	1.3	10.3
SRS-2	15.4	18.6	8.3	6.4	6.4	0.0	2.6
M-CHAT	11.6	13.0	8.4	10.3	0.0	1.3	0.0
SNAP-IV	7.7	5.6	12.6	3.8	0.0	1.3	2.6
SCARED	6.4	7.5	4.2	3.8	1.3	0.0	1.3
ADAPTIVE BEHAVIOR TESTS							
ABAS-II*	48.8	68.6	4.2	9.0	16.7	9.0	14.1
VABS-2*	42.3	51.9	20.8	12.8	16.7	9.0	3.8
SIB-R*	15.4	22.3	0.0	1.3	2.6	5.1	6.4
QUALITY OF LIFE TESTS							
QOLCE	20.5	22.3	16.7	5.1	1.3	1.3	12.8
QOLIE	19.2	20.5	16.7	3.8	7.7	1.3	6.4
PedsQL	16.6	20.5	8.3	2.6	3.8	3.8	6.4
HRQOL	6.5	5.7	8.4	1.3	1.3	1.3	2.6
ICND	3.9	3.8	4.2	0	1.3	0	2.6
PROMIS	1.3	0	4.2	0	0	0	1.3
SYMPTOM VALIDITY TESTING							
TOMM	16.7	24.1	0.0	2.6	6.4	2.6	5.1
AST	1.3	1.9	0.0	1.3	0.0	0.0	0.0

[1]See supplementary *table 2* for complete test names associated with the listed acronyms.
Measures highlighted in gray were most commonly used. Data are presented in order of highest to lowest percentages of Overall Use.
*NINDS Epilepsy Common Data Element (CDE) recommendation; **NINDS Epilepsy CDE for ages 0-5.

majority of neuropsychologists worldwide, yet in adults, most studies show that the RCFT is poor for lateralizing and localizing function (Barr *et al.*, 1997; Frank and Landeira-Fernandez, 2008; McConley *et al.*, 2008; Schouten *et al.*, 2009; Wisniewski *et al.*, 2012). There is a need to conduct studies in children as one study suggests that the RCFT may be localizing in children (Schouten *et al.*, 2009). In addition to study regarding the utility of certain measures, a large database may also allow us to consider other sophisticated analyses. For example, it may be more informative to analyze a set of measures to generate a profile rather than a single measure to predict outcome. In addition, other aspects of our approach to evaluation might be investigated, including determining when is the optimal time to conduct follow-up testing and what are the most important ages to target for follow-up.

The lower consistency of measures between English and non-English-speaking neuropsychologists is not unexpected. The non-English-speaking respondents were not a homogeneous group and neuropsychologists were reluctant to use culturally different neuropsychological tools without studying their validity in their own cultural context and having an appropriate cross-cultural standardization sample. However, transcultural adaptation and equivalence studies are often not conducted by test publishers because it is cost prohibitive for smaller or resource-limited countries, for less common languages, and in children. In addition, even when performed, the complexity of how to properly norm a measure presents challenges (Shuttleworth-Edwards, 2016). As such, this important task is often taken on by individual research groups that need to develop a culturally sensitive outcome measure. With the advent of larger research networks, the power to develop more cross-cultural, multilingual tools may increase. Our results indicate that linguistic and cultural considerations are barriers for collaborative international studies. There are significant challenges with the availability of common measures in other languages that are culturally adapted.

A limitation of our study is that there was a bias regarding who received our survey and thus, there are regions under-represented in our survey. The under-representation parallels the geographic map of where there are disparities in basic epilepsy treatment (Meyer *et al.*, 2010), which include many areas within Africa, Asia, the Middle East, and Central and South America. These countries are resource-limited and do not have a major epilepsy center that we were aware of to contact. In fact, none of our respondents were from low-income economy countries. The majority of respondents to our survey were from high-income economies and the remaining 11.5% of our sample were from higher or lower-middle-income economy countries.

A challenge when understanding outcomes is that post-operative evaluations are made less frequently for all patients (<60% of the time), which likely contributes to the lack of outcome data available. Nonetheless, with over 50 clinicians from English-speaking countries, the ability to amass data quickly becomes a possibility. Another challenge is developing the procedure and incentive to input data into a common repository. The procedural aspects would need funding at some level to organize data and ensure confidentiality and integrity of data, however, with the advent of several web-based databases and programs that make exporting and importing different datasets achievable, this challenge could be overcome. In terms of incentive for many clinicians to participate, an open access policy would be important such that any person who contributes data would be able to access the wider database. Moreover, a scientific team approach would be appropriate such that working groups might be formed according to area of interest. For example, specific questions within different domains might be established and once a plan for a manuscript is outlined and submitted, an oversight committee could approve the topic and provide a timeline for completion; if the project is not completed on time, the topic would return to the general group as available for publication.

Conclusions

Our results confirm that there are several measures that are consistently used by a majority of clinicians when conducting neuropsychological assessment with children undergoing pre- and post-surgical evaluations. The consensus from neuropsychologists conducting presurgical evaluations in children with epilepsy provides the basis for recommendations for measures to be used in developing a multi-site collaboration through a common data repository. Participation is anticipated to be high if practitioners are already using the measure. Moreover, there was much interest in the results as indicated by open-text comments and informal conversations, and respondents endorsed a willingness to use alternative measures, both of which make the likelihood of participation even higher. Last, there is undoubtedly a need for more robust outcome studies in those children who undergo epilepsy surgery, yet there are no clear recommendations for the use of specific measures in children (Jones-Gotman *et al.*, 2010; Loring *et al.*, 2011).

This survey provides a rationale for how to proceed without getting mired in debate over which measures to include. The set of 19 most commonly used measures narrows down the field of possible tests. With minor refinement, a neuropsychological protocol, that is brief enough to allow for individual clinicians to still have room to customize to their individual needs, is possible. Future goals include inviting clinicians (including, but not limited to, all survey respondents) to be part of the development of a common data repository. With a committed group, meetings would take place at annual meetings (*e.g.* the American Epilepsy Society, International Neuropsychological Society, European Congress on Epileptology) and/or through teleconferencing to refine recommendations for the set of common measures and pursue funding mechanisms to create the collaborative database. Any other interested parties may contact any of the authors of this article. □

Supplementary data.
Summary didactic slides and supplementary tables are available on the www.epilepticdisorders.com website.

Acknowledgements and disclosures.
The authors and the ILAE Task Force would like to thank all the respondents who took time out of their busy schedules to complete the survey. The authors would also like to thank Virginia Terwilliger and Joy Facella-Ervolini for their assistance in manuscript preparation.
This report was written by experts selected by the International League Against Epilepsy (ILAE) and was approved for publication by the ILAE including formal presentation to the Neuropsychology Task Force of the Diagnostic Methods Commission and the Commission on Surgical Therapies. Opinions expressed by the authors, however, do not necessarily represent the policy or position of the ILAE.

References

Anderson V, Spencer-Smith M, Wood A. Do children really recover better? Neurobehavioural plasticity after early brain insult. *Brain* 2011; 134(8): 2197-221.

Austin J, Cnaan A, O'Dell C. *Epilepsy common data elements: quality of life scales pediatric measures*. 2012. https://www.commondataelements.ninds.nih.gov/Epilepsy.aspx#tab=Data_Standards.

Baker GA. Psychological and neuropsychological assessment before and after surgery for epilepsy: implications for the management of learning-disabled people. *Epilepsia* 2001; 42: 41-3.

Baker GA, Goldstein LH. The dos and don'ts of neuropsychological assessment in epilepsy. *Epilepsy & Behav* 2004; 5(1): S77-80.

Barr WB, Chelune GJ, Hermann BP, *et al*. The use of figural reproduction tests as measures of nonverbal memory in epilepsy surgery candidates. *J Int Neuropsychol Soc* 1997; 3(5): 435-43.

Berl MM, Duke ES, Mayo J, *et al*. Functional anatomy of listening and reading comprehension during development. *Brain Lang* 2010; 114(2): 115-25.

Berl MM, Mayo J, Parks EN, *et al*. Regional differences in the developmental trajectory of lateralization of the language network: developmental trajectories of lateralization of language. *Human Brain Mapping* 2014; 35(1): 270-84.

Cross JH, Jayakar P, Nordli D, *et al*. Proposed criteria for referral and evaluation of children for epilepsy surgery: recommendations of the subcommission for pediatric epilepsy surgery. *Epilepsia* 2006; 47(6): 952-9.

Cukiert A, Rydenhag B, Harkness W, *et al*. Technical aspects of pediatric epilepsy surgery: report of a multicenter, multinational web-based survey by the ILAE Task Force on Pediatric Epilepsy Surgery. *Epilepsia* 2016; 57(2): 194-200.

Dade LA, Jones-Gotman M. Sodium amobarbital memory tests: what do they predict? *Brain Cogn* 1997; 33(2): 189-209.

Djordjevic J. Inquiry on assessments across epilepsy centers in different countries. In: Helmstaedter C, Hermann B, Lassonde M, et al, eds. *Neuropsychology in the care of people with epilepsy. Progress in epileptic disorders*. John Libbey Eurotext, 2011, 13-26.

Frank J, Landeira-Fernandez J. Comparison between two scoring systems of the Rey-Osterrieth Complex Figure in left and right temporal lobe epileptic patients. *Arch Clin Neuropsychol* 2008; 23(7–8): 839-45.

Freitag H, Tuxhorn I. Cognitive function in preschool children after epilepsy surgery: rationale for early intervention. *Epilepsia* 2005; 46(4): 561-7.

Gonzalez LM, Embuldeniya US, Harvey AS, *et al*. Developmental stage affects cognition in children with recently-diagnosed symptomatic focal epilepsy. *Epilepsy & Behav* 2014; 39: 97-104.

Harvey AS, Cross JH, Shinnar S, *et al*. Defining the spectrum of international practice in pediatric epilepsy surgery patients. *Epilepsia* 2008; 49(1): 146-55.

Helmstaedter C, Witt J-A. Clinical neuropsychology in epilepsy: theoretical and practical issues. *Handbook Clin Neurol* 2012; 107: 437-59.

Holland SK, Vannest J, Mecoli M, *et al*. Functional MRI of language lateralization during development in children. *Int J Audiol* 2007; 46(9): 533-51.

Honda R, Kaido T, Sugai K, *et al*. Long-term developmental outcome after early hemispherotomy for hemimegalencephaly in infants with epileptic encephalopathy. *Epilepsy & Behav* 2013; 29(1): 30-5.

Jayakar P, Gaillard WD, Tripathi M, *et al*. Diagnostic test utilization in evaluation for resective epilepsy surgery in children. *Epilepsia* 2014; 55(4): 507-18.

Jones-Gotman M, Smith ML, Zatorre RJ. Neuropsychological testing for localizing and lateralizing the epileptogenic region. In: Engel J, ed. *Surgical treatment of the epilepsies*. New York: Raven Press, 1993, 245-61.

Jones-Gotman M, Sziklas V, Djordjevic J, *et al*. Etomidate speech and memory test (eSAM): a new drug and improved intracarotid procedure. *Neurology* 2005; 65(11): 1723-9.

Jones-Gotman M, Smith ML, Risse GL, *et al*. The contribution of neuropsychology to diagnostic assessment in epilepsy. *Epilepsy & Behav* 2010; 18(1–2): 3-12.

Lee GP, Park YD, Westerveld M, *et al*. Effect of Wada methodology in predicting lateralized memory impairment in pediatric epilepsy surgery candidates. *Epilepsy & Behav* 2002; 3(5): 439-47.

Loring DW, Meador KJ, Westerveld M. The Wada test in the evaluation for epilepsy surgery. *Neurosciences (Riyadh, Saudi Arabia)* 2000; 5(4): 203-8.

Loring DW, Marino S, Meador KJ. Neuropsychological and behavioral effects of antiepilepsy drugs. *Neuropsychology Rev* 2007; 17(4): 413-25.

Loring DW, Bowden SC, Lee GP, *et al*. Diagnostic utility of Wada Memory Asymmetries: sensitivity, specificity, and likelihood ratio characterization. *Neuropsychology* 2009; 23(6): 687-93.

Loring DW, Lowenstein DH, Barbaro NM, *et al*. Common data elements in epilepsy research: development and implementation of the NINDS epilepsy CDE project: NINDS Epilepsy CDE Project. *Epilepsia* 2011; 52(6): 1186-91.

McConley R, Martin R, Palmer CA, *et al*. Rey Osterrieth complex figure test spatial and figural scoring: relations to seizure focus and hippocampal pathology in patients with temporal lobe epilepsy. *Epilepsy & Behav* 2008; 13(1): 174-7.

Meador KJ, Loring DW. Developmental effects of antiepileptic drugs and the need for improved regulations. *Neurology* 2016; 86(3): 297-306.

Meyer A-C, Dua T, Ma J, *et al*. Global disparities in the epilepsy treatment gap: a systematic review. *Bulletin of the World Health Organization* 2010; 88(4): 260-6.

Puka K, Smith ML. Academic skills in the long term after epilepsy surgery in childhood. *Epilepsy & Behav* 2016; 62: 97-103.

Schouten D, Hendriksen JGM, Aldenkamp AP. Performance of children with epilepsy on the Rey-Osterrieth complex figure test: is there an effect of localization or lateralization? *Epilepsy Res* 2009; 83(2–3): 184-9.

Sherman EMS, Wiebe S, Fay-McClymont TB, *et al*. Neuropsychological outcomes after epilepsy surgery: systematic review and pooled estimates. *Epilepsia* 2011; 52(5): 857-69.

Shuttleworth-Edwards AB. Generally representative is representative of none: commentary on the pitfalls of IQ test standardization in multicultural settings. *Clinical Neuropsychologist* 2016: 1-24.

Smith ML, Berl MM. The role and limitations of neuropsychological evaluation. In: Arzimanoglou A, Cross JH, Gaillard WD, et al, eds. *Pediatric epilepsy surgery*. Montrouge: John Libbey Eurotext, 2017.

Smith ML, Lah S, Elliot I. Outcomes of pediatric epilepsy surgery. In: Helmstaedter C, Hermann B, Lassonde M, et al, eds. *Neuropsychology in the care of people with epilepsy. Progress in epileptic disorders*. Montrouge: John Libbey Eurotext, 2011, 239-47.

Wisniewski I, Wendling A-S, Manning L, *et al*. Visuospatial memory tests in right temporal lobe epilepsy foci: clinical validity. *Epilepsy & Behav* 2012; 23(3): 254-60.

Wyllie E, Comair YG, Kotagal P, *et al*. Seizure outcome after epilepsy surgery in children and adolescents. *Annal Neurol* 1998; 44(5): 740-8.

TEST YOURSELF

(1) Why is neuropsychological evaluation highly recommended in the presurgical work-up of pediatric epilepsy surgery?

(2) What are the common domains included in neuropsychological assessment for pediatric epilepsy surgery?

(3) What are the challenges in neuropsychological evaluation within pediatric epilepsy?

Note: Reading the manuscript provides an answer to all questions. Correct answers may be accessed on the website, www.epilepticdisorders.com, under the section "The EpiCentre".