Journal of the International Neuropsychological Society (2003), **9**, 879–886.

DOI: 10.1017/S1355617703960085

Reexamining the effects of epilepsy surgery on IQ in children: Use of regression-based change scores

ELISABETH M.S. SHERMAN,[1,2] DANIEL J. SLICK,[1] MARY B. CONNOLLY,[3,4] PAUL STEINBOK,[5,6] ROY MARTIN,[7] ESTHER STRAUSS,[8] GORDON J. CHELUNE,[9] AND KEVIN FARRELL[3,4]

[1]Department of Psychology, British Columbia's Children's Hospital, Vancouver, British Columbia, Canada
[2]Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada
[3]Division of Neurology, British Columbia's Children's Hospital, Vancouver, British Columbia, Canada
[4]Division of Neurology, University of British Columbia, Vancouver, British Columbia, Canada
[5]Division of Neurosurgery, Section of Surgery, British Columbia's Children's Hospital, Vancouver, British Columbia, Canada
[6]Division of Neurosurgery, Department of Surgery, University of British Columbia, Vancouver, British Columbia, Canada
[7]University of Alabama at Birmingham Epilepsy Center, Birmingham, Alabama
[8]Department of Psychology, University of Victoria, Victoria, British Columbia, Canada
[9]Mellen Center, Cleveland Clinic, Cleveland, Ohio

(RECEIVED June 3, 2002; REVISED October 15, 2002; ACCEPTED November 23, 2002)

Abstract

Prior studies have found no adverse effects of pediatric epilepsy surgery on IQ. However, empirical techniques such as regression models, designed to account for confounding factors such as practice effects and test–retest reliability and able to provide a standardized method for evaluating outcome, have not been used in studying change after pediatric epilepsy. The goal of this study was to demonstrate the regression technique while empirically measuring the effect of epilepsy surgery on IQ in a group of pediatric patients. Predictors of retest IQ (e.g., baseline IQ, retest interval, demographics, epilepsy severity) were evaluated in a control group with intractable seizures ($N = 23$) assessed twice with the WISC–III. The resulting equation was used to evaluate IQ changes in a second group of children who underwent epilepsy surgery ($N = 22$). In controls, baseline IQ was a strong predictor of retest IQ. Number of AEDs was inversely related to retest IQ. Based on the control regression, four children (18%) in the surgical sample obtained significantly higher than expected postsurgical IQ scores and one child (5%) obtained a lower than expected IQ score. This study demonstrates that regression-based techniques yield informative estimates on outcome and may be an improvement over prior methods of measuring change after pediatric epilepsy surgery. (*JINS*, 2003, *9*, 879–886.)

Keywords: Epilepsy, IQ, Regression, Change, Surgery, Outcome, Pediatric

INTRODUCTION

It is generally accepted that, in most cases, pediatric epilepsy surgery does not adversely affect intelligence as measured by IQ tests (Adams et al., 1990; Caplan et al., 1993; Gilliam et al., 1997; Lindsay, 1990; Meyer et al., 1986; Miranda & Smith, 2001; Oxbury et al., 1995; Szabo et al., 1998; Tinuper et al., 1988; Westerveld et al., 2000). However, the studies upon which this assumption is based have assessed change by comparing pretreatment and posttreatment means or individual change scores (i.e., pretreatment score minus posttreatment score). In the latter case, these are then evaluated against a clinically significant cutoff such as 95% confidence bands defined by the standard error of measurement (*SEM*), the standard error of the difference (SE_{diff}), or by other clinical cutoffs deemed to reflect clinically significant change (i.e., 10 or 15 IQ points). These approaches do not allow researchers to disentangle "true" change attributable to surgery from the effects of other factors that also influence test scores. These factors include (1) practice effects, (2) test–retest reliability, (3) differing baseline levels, (4) regression towards the mean, and (5) sub-

Reprint requests to: E.M.S. Sherman, Ph.D., Department of Psychology, British Columbia's Children's Hospital, K3-122, 4480 Oak Street, Vancouver, British Columbia V6H 3V4, Canada. E-mail: esherman@cw.bc.ca

ject variables that operate independently or interact with the treatment in question to affect outcome (e.g., disease severity, maturation). Another methodological limitation of previous research is the practice of collapsing scores from different versions of IQ tests across children, despite psychometric differences between test versions.

Given the increasing availability of epilepsy surgery in pediatric centers, there is a need for outcome research that accurately measures the proportion of change attributable to surgery independent of confounding factors. This has been done in the adult literature. Empirical techniques for measuring change in adult epilepsy samples include the Reliable Change Index (RCI) and standardized regression-based change scores (Chelune et al., 1993; Hermann et al., 1991, 1996; McSweeney et al., 1993; Sawrie et al., 1996). The RCI method for deriving cutoff scores accounts for measurement error and practice effects and is based on the test–retest reliability coefficient and the *SEM* (Chelune et al., 1993). In addition to accounting for imperfect test–retest reliability as do RCIs, regression-based change (RBC) scores account for many other factors that may significantly affect outcome such as differential practice effects associated with variability in baseline level of function and test–retest interval, and subject variables that may be systematically associated with test–retest changes. In addition, RBC scores can be easily standardized, providing a single metric of effect size and directionality across measures (Sawrie et al., 1996). It is important to understand that RBC scores do not necessarily indicate whether a significant change from baseline level has occurred—for which use of RCIs may be more appropriate—but instead, assess the degree to which observed change conforms to established trends in a reference population. These trends may consist of increases or decreases in performance over time in association with combinations of influential predictor variables, such as type and severity of illness, treatment type, baseline cognitive level, gender, age, and test–retest interval (e.g., increased scores at retest from healthy children *vs.* decreased scores from children with progressive neurological disease).

The goal of this study was to use the regression-predicted change technique in order to provide empirically derived information on the effects of epilepsy surgery on IQ in children. The procedure entailed deriving the control regression equation using baseline and retest IQ data from a sample of children with intractable epilepsy and then deriving RBC scores using this equation with a second group of children who underwent epilepsy surgery.

METHOD

Participants and Procedure

Data for this study were obtained by archival chart review of pediatric patients seen at epilepsy surgery programs and seizure clinics at two tertiary care hospitals (British Columbia's Children's Hospital, Vancouver, Canada; and University of Alabama at Birmingham Epilepsy Center, Birmingham, Alabama). Inclusion criteria included (1) epilepsy diagnosis confirmed by a neurologist and defined as recurrent, afebrile epileptic seizures; (2) baseline and retest Full Scale Intelligence Quotient (FSIQ) scores using the Wechsler Intelligence Scale for Children–Third Edition (WISC–III; Wechsler, 1991); (3) age between 6 and 17 years; and (4) absence of any known progressive neurological condition. Because this was a retrospective clinical study, we could not ensure that all children were on the same medications at baseline and retest. However, to minimize a differing effect of antiepileptic drugs (AEDs) on baseline and retest, we excluded any children who were on different medications at baseline and retest if one of these medications included phenobarbital or topiramate, two AEDs with potentially major cognitive side effects (Aldenkamp, 2001).

The presurgical control group consisted of 23 patients who were specifically under evaluation for epilepsy surgery and who had been tested twice with the WISC–III prior to any surgery. Of these waitlist patients, approximately half ($N = 11$, 48%) had proceeded to surgery at the time of this writing. According to the International Classification of Epilepsy (ICE), most of the controls had localization-related epilepsy (Table 1). Means for medication status and seizure frequency in the month prior to the evaluation are shown in Table 2. Most children were on a single AED at the time of testing (range 0–4), but on average had been on three or more AEDs prior to their first evaluation (range 0–16). The range in seizure counts in the month prior to the baseline assessment across the group was extreme, ranging from 0 to 215 seizures. Seizure rating classifications are presented for descriptive purposes and ease of interpretation in Table 1.

The surgical group comprised 22 patients (Table 1). Like the controls, most surgical patients were on a single AED at the time of initial evaluation (range 0–4) but had had multiple AED trials in the past (Table 2). Surgical outcome for the group, according to Engel et al.'s (1993) classification system, is shown in Table 1.

RESULTS

Group Comparisons

The two groups were compared in terms of demographic and neurological characteristics (Table 2). Demographic and neurological data had a normal distribution with the exception of age at onset, number of AEDs, and seizure frequency (Shapiro–Wilk Test). Mann–Whitney tests were used to assess for group differences in these variables and *t* tests were used in all other cases. Given the particularly extreme range and positively skewed distribution of this variable, seizure frequency was log-transformed for purposes of all other statistical analyses.

Age and retest interval were significantly different between groups. Chi-square analyses indicated no significant group differences in sex or handedness. Following the ap-

https://doi.org/10.1017/S1355617703960085 Published online by Cambridge University Press

Table 1. Neurological characteristics of the presurgical and surgical groups

	Presurgical controls	Surgical group
ICE classification		
Localization-Related	19 (83%)	9 (90%)[a]
Generalized	3 (13%)	—
Undetermined	1 (4%)	1 (10%)[a]
Etiology[b]		
Tumor	1 (6%)	2 (10%)
Vascular	2 (11%)	—
Infection	4 (22%)	2 (10%)
Dysplasia	4 (22%)	4 (19%)
Cryptogenic	7 (39%)	15 (62%)
Seizure frequency rating		
None	6 (26%)	5 (23%)
1–10	9 (39%)	9 (43%)
Over 10 but under 100	6 (26%)	5 (23%)
100 or more seizures	2 (9%)	3 (14%)
Engel seizure outcome classification		
Class 1: Free of disabling seizures	—	11 (50%)
Class 2: Rare disabling seizures	—	2 (9%)
Class 3: Worthwhile improvement	—	3 (14%)
Class 4: No worthwhile improvement	—	6 (27%)

Note. Seizure frequency refers to the month preceding baseline IQ.
[a]ICE classifications were available for ten children in the surgical group.
[b]Etiologies were available for 18 patients in the presurgical control group.

proach used by Chelune et al. (1993) to assess for the effects of preexisting group differences on change scores, Pearson's correlations between age, retest interval, seizure frequency, and change in IQ were examined. These correlations were all small and nonsignificant. In addition, we examined whether subgroups within the control group were different with regards to IQ changes at retest. Specifically, we compared IQ changes between children in the control group who had proceeded to surgery (true controls) *versus* those who had not yet had surgery (waiting list only). Although the mean change in IQ was larger for the waitlist group ($M = -3.67$, $SD = 6.37$, $N = 12$), it was not significantly different from that of the true controls who had been assessed twice prior to proceeding to surgery ($M = -1.55$, $SD = 6.37$, $N = 11$).

Baseline and retest scores on the IQ measure are presented in Table 3. Baseline IQs were not significantly different between groups. There did not appear to be any group-level increase in IQ (i.e., a practice effect). Instead, the presurgical control mean decreased by 2.6 IQ points at retest; a paired samples *t* test indicated that this decrease approached significance ($p < .066$). The surgical group

Table 2. Demographic and neurological predictors: Differences between presurgical control and surgical groups

	Presurgical controls ($N = 23$)			Surgical group ($N = 22$)			
	Mean	SD	Range	Mean	SD	Range	*p*
Age	9.5	2.2	6.1–14.7	11.3	2.9	6.0–15.6	***.020***
Sex (male/female)[a]	8/15			11/11			.373
Handedness (right/left)[a]	21/2			21/1			.513
Age at onset[b]	3.8	3.0	0.2–9.0	4.7	4.1	0.5–14.7	.366
Duration of epilepsy	5.7	3.7	0–13	6.6	3.5	0.6–13.3	.429
Number of AEDs[b]	1.7	0.8	0–3	1.7	1.1	0–4	.799
Number of prior AEDs	5.1	3.7	0–16	4.0	2.4	0–9	.276
Seizure frequency[b]	27.7	52.7	0–215	25.0	44.6	0–180	.855
Test–retest interval	2.5	1.3	0.5–5.1	1.5	0.8	0.7–3.7	***.003***

Note. *p* values from *t* tests except for: [a]Chi-square Test and [b]Mann–Whitney *U* Test. Age, age at onset, duration of epilepsy, and test–retest interval are shown in years. Number of AEDs refers to AEDs at baseline. Seizure frequency refers to number of seizures in the month preceding baseline.

https://doi.org/10.1017/S1355617703960085 Published online by Cambridge University Press

Table 3. Baseline and retest IQ for the presurgical controls and surgical group

	Baseline IQ	Retest IQ	*t*	*p*
Presurgical controls	73.0 (17.3)	70.4 (19.7)	1.934	.066
Surgical group	78.2 (20.3)	77.5 (19.4)	0.427	.674

mean decreased by .7 IQ points, which was a nonsignificant change.

Predictors of Postsurgical IQ

To identify potential predictors to enter into the regression equation, simple correlations between demographic/neurological variables and retest scores within the control sample were examined (Table 4). As expected, baseline IQ was very highly correlated with retest IQ. Number of current AEDs was also strongly correlated. There was also a trend involving sex, with female gender associated with lower retest IQ. No other variable, including age or retest interval, was significantly correlated with retest IQ.

Regression Equation and Derivation of Change Scores

Data from the control group was used to construct the regression model. Because of the very high correlation with baseline FSIQ, a two-stage method was used to derive the regression equation for predicting FSIQ at retest. In the first stage, baseline FSIQ scores were regressed onto FSIQ scores at retest. The resulting R^2 value of .89 was significant at $p < .001$. In the second stage the other variable significantly correlated with FSIQ at retest, number of AEDs at baseline, was added to the regression equation as a second predictor. The resulting R^2 value of .92 was significantly larger than the R^2 value obtained when FSIQ at retest was predicted by baseline FSIQ alone ($p = .017$). The resulting regression equation is given below:

$$\text{FSIQ}_{\text{retest}} = (.965 \times \text{FSIQ}_{\text{baseline}}) + (-4.519 \times \text{AEDs}_{\text{baseline}}) + 7.358$$

It can be seen from inspection of this equation that predicted retest FSIQ values were inversely related to number of AEDs at baseline, and that for children who were not taking any AEDs at baseline, a modest increase in FSIQ at retest was expected. For example, given a baseline FSIQ of 100, the predicted FSIQs at retest for children taking 0, 1, 2, and 3 AEDs at baseline are 104, 99, 95, and 90, respectively.

The regression-based method for evaluating change is not limited to models that only include main effects. In our study, there was insufficient power to formally assess all the possible interactions between predictor variables. As an example of this approach, however, we compared the effect size when retest IQ was predicted by baseline IQ *versus* the prediction of retest IQ by the interaction of baseline IQ with number of AEDs. The former was associated with a much larger effect size, suggesting that IQ alone was the better predictor.

Table 4. Correlations between predictor variables and retest IQ for the presurgical controls

	Retest FSIQ	*p*
Baseline FSIQ	.945	***.000***
Number of AEDs	−.539	***.008***
Sex	−.357	.095
Seizure frequency[a]	−.308	.153
Handedness	.201	.355
Age at onset	.106	.629
Age	.095	.665
Test–retest interval	−.090	.684
Number of prior AEDs	−.026	.905
Duration of epilepsy	−.018	.933

Note. Number of AEDs refers to AEDs at baseline. Sex was coded Male = 1, Female = 2; handedness was coded as Right = 1, Left = 2; age at onset, age, test–retest interval, and duration of epilepsy are coded in years.
[a]Log-transformed; refers to number of seizures in the month preceding baseline.

Postsurgical Change Scores

Predicted retest FSIQ scores were derived using the equation shown above. Significance levels for differences between observed and predicted retest FSIQ scores were derived using the standard error for each individual's predicted score ($SE_{\hat{Y}}$), rather than the standard error of the regression (SE_{Reg}), as has been used in prior studies with adults.* This was done to obtain more accurate *t* and *p* values because, unlike the SE_{Reg}, the $SE_{\hat{Y}}$ is not constant across cases, but increases as individual values of independent variables deviate from the mean (Crawford & Howell, 1998). That is, persons who are outliers with respect to their scores on predictor variables will have larger margins of error associated with their predicted scores and thus larger changes in raw scores will be required to reach significance for these individuals.

RBC scores should only be derived when necessary assumptions for regression analyses are met (see Pedhazur, 1997, pp. 33–34). Evaluation of residuals from the control regression used to derive the RBC scores showed that these assumptions were met. In addition, the means and variances of the predictor variables in the new sample should approximate those of the original sample from which the

*This was accomplished using a computer program developed by Crawford and Howell (1998). The program (CLREGMUL) is available at the website of Professor John R. Crawford: *http://www.psyc.abdn.ac.uk/homedir/jcrawford/jcrawford.htm*

https://doi.org/10.1017/S1355617703960085 Published online by Cambridge University Press

regression model was derived. This condition was also met, as the control and surgery groups did not differ significantly in terms of mean or variance of FSIQ or number of AEDs at baseline testing.

Results of the analysis of change scores from the 22 children in the surgical group are presented in Table 5. Individuals are sorted in order of magnitude of difference between observed and predicted FSIQ at retest. Predicted retest IQs ranged from 14 points below to 5 points above baseline IQ, with the majority (15/22 or 68%) below baseline level, reflecting the fact that most of the children were taking two or more AEDs at baseline. The actual difference between baseline and retest IQ ranged from −16 to +16. Using a cutoff of $p \leq .05$, one child (5%) showed a larger than predicted decline in IQ of 16 points. This child underwent right hemispherectomy for Rasmussen's encephalitis. Four children (18%) did not show the predicted decline in IQ; two underwent frontal resections and two underwent temporal lobectomy. All four of these children instead showed an increase in IQ, from 2 to 16 points. Because this was a preliminary study, a less conservative cutoff of $p \leq .15$ was also applied. Two children (9%) showed a larger than expected decline while five (23%) did not show the predicted decline. Thus, using this criterion, approximately one-third of the sample did not conform to the control trend following surgery.

To determine whether RBC scores were related to outcome as defined by Engel et al.'s (1993) seizure outcome classification system, observed–predicted difference scores were correlated with numeric Engel outcome scores. Subclassifications of seizure outcome were collapsed across categories (e.g., scores of 4a and 4b were both coded as 4). There was no significant relationship between seizure outcome and raw RBC scores ($r = .23$, $p = .30$), or between seizure outcome and actual change in IQ from baseline to retest ($r = .14$, $p = .55$).

DISCUSSION

This study is the first demonstration of the use of the regression-based technique for measuring change after epilepsy surgery in a pediatric sample. Although preliminary given the small sample size, the results give credence to the hypothesis that most children encounter no adverse effects of epilepsy surgery on IQ. In fact, a subgroup of children (18% to 23%, depending on the *p* value) clearly appeared to benefit measurably from this intervention. Of course, accurately assessing the risks of a particular intervention such as surgery is as important as assessing any potential benefits. This study provided preliminary evidence for a relatively low prevalence of IQ decline associated with epilepsy surgery when other important variables are controlled. With

Table 5. FSIQ scores, change scores, and type of epilepsy surgery for the surgical group ($N = 22$)

Surgery type	Sex	Seizure frequency	Number of AEDs	Baseline	Retest	Predicted retest	Predicted change	Observed change	Observed–predicted	*p*	Engel outcome
RH	F	1	1	73	57	73	0	−16	−16	***.014***	2a
L ATL	M	2	1	104	92	103	−1	−12	−11	.093	4b
R ATL	F	0	0	104	99	108	4	−5	−9	.196	1a
LF	M	30	3	73	58	64	−9	−15	−6	.340	1a
LF	M	0	1	97	91	96	−1	−6	−5	.390	2a
R ATL	F	0	0	69	69	74	5	0	−5	.467	1a
L ATL	M	5	2	68	60	64	−4	−8	−4	.518	3a
L ATL	M	30	1	85	83	85	0	−2	−2	.761	1a
R ATL	F	0	0	119	120	122	3	1	−2	.752	1a
RF	M	5	3	83	77	74	−9	−6	3	.642	1a
RH	F	0	2	47	47	44	−3	0	3	.601	1a
CC	F	14	2	40	40	37	−3	0	3	.638	4b
L ATL	F	12	1	91	94	91	0	3	3	.590	3a
RP	F	8	2	73	73	69	−4	0	4	.488	1b
L ATL	F	10	2	72	72	68	−4	0	4	.491	1a
RP	M	180	2	107	106	102	−5	−1	4	.518	4b
R ATL	M	5	1	70	76	70	0	6	6	.368	4b
LF	F	100	2	72	78	68	−4	6	10	.103	1a
R ATL	M	10	2	53	63	49	−4	10	14	***.039***	1a
RF	M	23	4	93	95	79	−14	2	16	***.053***	4b
L ATL	M	8	2	56	72	52	−4	16	20	***.004***	3b
LF	F	7	4	72	83	59	−13	11	24	***.003***	4c

Note: Predicted scores are rounded to nearest integer; the *p* value applies to the *t* test of the difference between predicted and observed retest IQ. Surgery types are as follows: left hemispherectomy (LH), right hemispherectomy (RH), left anterior temporal lobectomy (L ATL), right anterior temporal lobectomy (R ATL), left frontal resection (LF), right frontal resection (RF), left parietal resection (LP), right parietal resection (RP), and corpus callosotomy (CC). Numbers of AEDs refer to AEDs at baseline. Seizure frequency refers to number of seizures in the month preceding baseline.

https://doi.org/10.1017/S1355617703960085 Published online by Cambridge University Press

a conservative *p* value, only one child in the surgical sample clearly experienced a greater than expected decline in IQ. This child underwent right hemispherectomy and was the only child in this series with Rasmussen's encephalitis, a condition that is associated with progressive decline in neurological and cognitive functioning if left untreated. Thus, it is possible that further decline in IQ had occurred between preoperative testing and surgery in this child. Interestingly, on a group level, outcome with regards to IQ was unrelated to seizure frequency postsurgery.

In evaluating surgical outcome on an individual basis, RBC results were quite different than those that would have been obtained using the clinical cutoffs used in previous research. Specifically, using $2 \times SEM$ (WISC–III $SEM = 3.2$; Wechsler, 1991), four cases of significant decrease (18%) and three cases of significant increase (14%) would have been detected. In comparison, the RBC technique identified fewer children with retest IQs that were lower than expected, while identifying a similar number of children who experienced better than expected outcomes in terms of IQ. Further, because only a relatively large change is considered significant with this technique, at least one of the children who obtained a significant, positive RBC score would not have been identified as benefiting from surgery using RCIs. In a population such as epilepsy, where disease course may combine with other factors such as medications to produce changes in IQ levels independent of surgical treatment, the regression-based approach to measuring change confers a distinct advantage over other techniques and may be an essential asset to researchers interested in accurately measuring the effects of surgical treatment on outcome.

On a group level, the regression-based technique helped identify predictors with substantial influence on retest scores and their degree of influence when other variables were accounted for. In this study, baseline scores were a major predictor of retest scores, which was expected (Dikmen et al., 1999; Heaton et al., 2001). In contrast, the strong relation between AEDs and retest IQs and the utility of AEDs as a predictor of retest IQ was somewhat surprising given previous research indicating smaller associations between AEDs and IQs in adult patients (Dodrill & Matthews, 1992). In our sample, for each AED taken at baseline, retest IQ values were expected to drop by approximately 4.5 points compared to baseline scores. One explanation for this finding is that the number of AEDs may be a marker for epilepsy severity, an interpretation that has been made by others (e.g., Dodrill & Matthews, 1992). The association between this putative marker of epilepsy severity and outcome, along with a trend for decreasing IQ in our controls over time, is consistent with the theory that some types of intractable epilepsy may be associated with progressive cognitive decline (Bjornaes et al., 2001; Hermann et al., 2001; but see Camfield, 1997 for an opposing point of view). From this perspective, our study provides some preliminary support for this theory and raises the possibility that surgery may prevent this decline in a subgroup of patients. For example, the four children in the surgical sample with significantly higher than expected postsurgical scores were all on multiple AEDs (2 to 4) and thus should have experienced a decrease in IQ as predicted by the regression equation; following surgery, all showed IQ increases. Replication with larger samples of children with intractable epilepsy would be helpful in supporting or countering this important observation, as would studies using the RBC methodology to examine outcome in other important cognitive domains such as memory and language. Also of note is the fact that age at onset was not significantly related to IQ in this study, which could have been due to sample characteristics such as the large number of children with early-onset epilepsy and intellectual impairment which resulted in a restriction of range for these variables. We specifically included children with IQs below 70 because this group is frequently excluded in research on postsurgical outcome (Miranda & Smith, 2001). Children with intellectual impairments are frequently represented in clinical samples evaluated for pediatric epilepsy surgery, and we felt that to exclude this group would distort the nature of our true clinical population. Visual inspection of the change score table suggests that children with intellectual delays at baseline did not fare any worse than other children in terms of outcome.

Some limitations of the study deserve mention. First, there were preexisting differences between the presurgical control group and the surgical group. This included a younger age and longer retest interval in the presurgical control group. These differences were likely secondary to the fact that this was a retrospective study that did not use randomization to groups. As such, surgery was likely undertaken when children were older and all other avenues for seizure relief had been explored unsuccessfully. The postsurgical evaluations also followed a standard protocol of a one-year postsurgery time interval, which tended to impose limits on the time between baseline and retest in the surgical group. Second, the control group, although all considered potential surgical candidates, comprised a heterogeneous group of children, only half of whom had had surgery at the time of this writing. Although the difference was nonsignificant, there was a smaller drop in IQ in the true control subgroup (i.e., those who went on to have surgery) compared to those children who were still on the waiting list. While small, this difference was not trivial and would have reached significance with a larger sample. Children who are found not to be appropriate surgical candidates are often those with multifocal epilepsy; in contrast, ideal surgical candidates are those with a single, well-defined epileptogenic focus. These two groups of children may show a different pattern of cognitive change over time, with children with multifocal epilepsies likely more at risk for cognitive decline. Additionally, unlike in the adult literature, the control children were not restricted to children with a specific type of epilepsy (i.e., temporal lobe epilepsy), nor were surgical candidates composed of children who received the same operation (i.e., temporal lobectomy). However, the heterogeneity of the controls mirrors the nature of pediatric epilepsy surgery candidates, who, unlike adult patients, present with a wide

https://doi.org/10.1017/S1355617703960085 Published online by Cambridge University Press

variety of epilepsy syndromes and etiologies. More homogenous control data and more homogenous surgical groups would be preferable in future to better delineate the effects of treatment and expected outcome in particular surgical subpopulations in children. Lastly, the small sample size in conjunction with group nonhomogeneity likely resulted in relatively large standard errors of prediction for the regression model. Thus, the proportion of surgical patients who showed significantly greater than expected change may well have been underestimated. Cross-validation of regression models for predicting change with larger and better-defined surgical and presurgical samples is needed, as are studies with sufficient power to assess the impact of potential interactions between predictor variables.

In sum, this study represents a first attempt at using empirically based methods for detecting change after epilepsy surgery in children. There are many avenues for further refinement of this approach, including deriving change scores for specific surgical groups, the use of regression models based on surgical data to identify good and poor responders to surgery, replication of these preliminary results with different outcome measures, and the use of interaction terms and other statistical techniques such as hierarchical linear models and nonlinear models to model change. It is hoped that our preliminary study is a first step in contributing positively to the field of outcome measurement after pediatric epilepsy surgery.

ACKNOWLEDGMENTS

This project was supported by a grant to the first author from the British Columbia's Children's Hospital Foundation Telethon Projects New Research Fund and from the British Columbia Medical Services Foundation/Vancouver Foundation. Portions of this paper were presented at the Thirtieth International Neuropsychological Society Meeting, February 2002, Toronto, Canada. The authors thank Kim Eyrl for assistance with data collection and the anonymous reviewers for their helpful comments on an earlier draft of this paper.

REFERENCES

Adams, C.B.T., Beardsworth, E.D., Oxbury, S.M., Oxbury, J.M., & Fenwick, P.B.C. (1990). Temporal lobectomy in 44 children: Outcome and neuropsychological follow-up. *Journal of Epilepsy*, *3* (*Suppl.*), 157–168.

Aldenkamp, A.P. (2001). Effects of antiepileptic drugs on cognition. *Epilepsia*, *42* (*Suppl. 1*), 46–49.

Bjornaes, H., Stabell, K., Heriksen, O., & Loyning, Y. (2001). Intelligence in pediatric and adult patients with refractory epileptic seizures: A longitudinal study [Abstract]. *Epilepsia*, *42* (*Suppl. 7*), 159.

Camfield, P.R. (1997). Recurrent seizures in the developing brain are not harmful. *Epilepsia*, *38*, 735–737.

Caplan, R., Guthrie, D., Shields, W.D., Peacock, W.J., Vinters, H.V., & Yudovin, S. (1993). Communication deficits in children undergoing temporal lobectomy. *Journal of the American Academy of Child and Adolescent Psychiatry*, *32*, 604–611.

Chelune, G.J., Naugle, R.I., Luders, H., Sedlak, J., & Awad, I.A. (1993). Individual change after epilepsy surgery: Practice effects and base rate information. *Neuropsychology*, *1*, 41–52.

Crawford, J.R. & Howell, D.C. (1998). Regression equations in clinical neuropsychology: An evaluation of statistical methods for comparing predicted and obtained scores. *Journal of Clinical and Experimental Neuropsychology*, *20*, 755–762.

Dikmen, S.S., Heaton, R.K., Grant, I., & Temkin, N.R. (1999). Test–retest reliability and practice effects of expanded Halstead–Reitan Neuropsychological Test Battery. *Journal of the International Neuropsychological Society*, *5*, 346–356.

Dodrill, C.B. & Matthews, C.G. (1992). The role of neuropsychology in the assessment and treatment of persons with epilepsy. *American Psychologist*, *47*, 1139–1142.

Engel, J., Jr., Van Ness, P.C., Rasmussen, T.B., & Ojemann, L.M. (1993). Outcome with respect to epileptic seizures. In J. Engel (Ed.), *Surgical treatment of the epilepsies* (2nd ed.), pp. 609–621. New York: Raven Press.

Gilliam, F., Wyllie, E., Kashden, J., Faught, E., Kotagal, P., Bebin, M., Wise, M., Comair, Y., Morawetz, R., & Kuzniecky, R. (1997). Epilepsy surgery outcome: Comprehensive assessment in children. *Neurology*, *48*, 1368–1374.

Heaton, R.K., Temkin, N., Dikmen, S., Avitable, N., Taylor, M.J., Marcotte, T.D., & Grant, I. (2001). Detecting change: A comparison of three neuropsychological methods, using normal and clinical samples. *Archives of Clinical Neuropsychology*, *16*, 75–91.

Hermann, B.P., Wyler, A.R., VanderZwagg, R., LeBailly, R.K., Whitman, S., Somes, G., & Ward, J. (1991). Predictors of neuropsychological change following anterior temporal lobectomy: Role of regression toward the mean. *Journal of Epilepsy*, *4*, 139–148.

Hermann, B.P., Seidenberg, M., Schoenfeld, J., Peterson, J., Leveroni, C., & Wyler, A.R. (1996). Empirical techniques for determining the reliability, magnitude, and pattern of neuropsychological change following epilepsy surgery. *Epilepsia*, *37*, 942–950.

Hermann, B.P., Seidenberg, M., Bell, B., Rutecki, P., Sheth, R., Jones, J., & Wendt, G. (2001). The neurodevelopmental impact of childhood onset temporal lobe epilepsy on brain structure and cognition. [Abstract]. *Epilepsia*, *42* (*Suppl. 7*), 151–152.

Lindsay, J. (1990). Hemispherectomy outcome. *Journal of Epilepsy*, *3* (*Suppl. 1*), 253–256.

McSweeney, A.J., Naugle, R.I., Chelune, G.J., & Luders, H. (1993). "T-scores for change": An illustration of a regression approach to depicting change in clinical neuropsychology. *The Clinical Neuropsychologist*, *6*, 32–42.

Meyer, F.B., Marsh, R., Law, S.E., & Sharborough, F.W. (1986). Temporal lobectomy in children with epilepsy. *Journal of Neurosurgery*, *64*, 371–376.

Miranda, C. & Smith, M.L. (2001). Predictors of intelligence after temporal lobectomy in children with epilepsy. *Epilepsy and Behavior*, *2*, 13–19.

Oxbury, S., Zwailla, Z., Adams, C., Middleton, J., & Oxbury, J. (1995). Hemispherectomy in childhood: Serial neuropsychological follow-up and prolonged improvement. *Epilepsia*, *36* (*Suppl. 3*), S24.

Pedhazur, E. (1997). *Multiple regression in behavioral research*. New York: Harcourt Brace.

Sawrie, S.M., Chelune, G.J., Naugle, R.I., & Luders, H.O. (1996). Empirical methods for assessing meaningful neuropsycholog-

ical change following epilepsy surgery. *Journal of the International Neuropsychological Society*, *2*, 556–564.

Szabo, C.A., Wyllie, E., Stanford, L.D., Geckler, C., Kotagal, P., Comair, Y.G., & Thornton, A. (1998). Neuropsychological effect of temporal lobe resection in preadolescent children with epilepsy. *Epilepsia*, *39*, 814–819.

Tinuper, P., Andermann, F., Villemure, J., Rasmussen, T., & Quesney, L. (1988). Functional hemispherectomy for treatment of epilepsy associated with hemiplegia: Rationale, indications, results and comparison with callosotomy. *Annals of Neurology*, *24*, 27–34.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children–Third Edition*. San Antonio, Texas: The Psychological Corporation.

Westerveld, M., Sass, K.J., Chelune, G., Hermann, B., Barr, W., Loring, D., Strauss, E., Trennery, M., Perrine, K., & Spencer, D.D. (2000). Temporal lobectomy in children: Cognitive outcome. *Journal of Neurosurgery*, *92*, 24–30.

https://doi.org/10.1017/S1355617703960085 Published online by Cambridge University Press