Epilepsy & Behavior 29 (2013) 443–448

Contents lists available at ScienceDirect

Epilepsy & Behavior

journal homepage: www.elsevier.com/locate/yebeh

An evidence-based checklist to assess neuropsychological outcomes of epilepsy surgery: How good is the evidence?

M. Hrabok [a,b,c,d,*], J. Dykeman [e], E.M.S. Sherman [a,b,e,f], S. Wiebe [a,b,g,h,i]

[a] Alberta Health Services, Canada
[b] The Alberta Children's Hospital Research Institute, Canada
[c] Copeman Healthcare Centre, Edmonton, Canada
[d] University of Alberta Hospital, Edmonton, Canada
[e] Faculty of Medicine, University of Calgary, Canada
[f] Copeman Healthcare Centre, Calgary, Canada
[g] Department of Clinical Neurosciences, Faculty of Medicine, University of Calgary, Canada
[h] Department of Community Health Sciences, Faculty of Medicine, University of Calgary, Canada
[i] Faculty-wide Clinical Research Unit, Faculty of Medicine, University of Calgary, Canada

ARTICLE INFO

Article history:
Received 7 August 2013
Accepted 30 August 2013
Available online 11 October 2013

Keywords:
Epilepsy
Neuropsychology
Evidence-based
Surgery
Outcome
Methodology

ABSTRACT

Purpose: We aimed to assess the quality of evidence on neuropsychological outcomes after epilepsy surgery (ES). Accordingly, we created an evidence-based neuropsychology (EBNP) checklist to assess neuropsychological outcomes and applied this tool to studies from a systematic review.
Methods: The EBNP checklist was created using clinical expert input, scale development methodology for item generation and reduction and inter-rater reliability, and critical appraisal guidelines for studies about treatment. The checklist was applied to articles obtained through a systematic review of resective ES neuropsychological outcomes. The proportion of studies fulfilling the quality criteria and the total quality score were used to assess the quality of the evidence.
Results: An initial 45-item checklist was applied to 147 articles, with excellent inter-rater agreement (kappa = 0.80). The mean quality score was 23 (SD: 4, range: 12–33). There was substantial variability in the percentage of studies meeting the criteria for specific items (0–99%). The median proportion of papers fulfilling various quality criteria was 1.4% for items related to group comparisons, 37% for clinical applicability, 67% for patient description, 78% for outcome assessment, and 91% for interventions. Higher quality correlated with longitudinal design, reporting presurgical IQ, seizure frequency and antiepileptic drugs, and using validated measures of change in individual patients. The final EBNP checklist consisted of 19 items.
Discussion: The EBNP checklist reliably identified quality strengths and threats to validity of neuropsychological outcome studies in ES. Studies would be most improved by the inclusion of random allocation to interventions or at minimum blinded outcome assessment, empirically based measures of reliable change and completeness of reporting of follow-up.

1. Introduction

Assessment of neuropsychological functioning is essential in epilepsy. Not only are cognitive deficits present in up to 75% of adults [1] and 25% of children [2] with new onset epilepsy, but deficits also relate to the etiology and type of epilepsy, interictal epileptiform activity, seizure frequency and severity, psychiatric comorbidities, and the effects of medical and surgical treatment [3]. The effectiveness of temporal lobe epilepsy surgery (ES) in achieving seizure remission has been demonstrated in randomized controlled trials [4,5] and is supported by Grade "A" recommendations from the American Academy of Neurology [6]. However, assessment of cognitive outcomes in temporal lobe ES has not been subjected to the same rigor as seizure outcomes, despite data indicating that reliable declines in cognitive function occur in approximately 40% of patients after dominant temporal lobe resections [7].

Rating the quality of evidence has become standard practice in many clinical disciplines. Using principles of evidence-based medicine [8] (i.e., critically appraising the evidence for its validity and usefulness to assist in managing individual patients), authors of clinical practice guidelines rate the quality of individual studies to determine the strength of their statements and recommendations [9,10]. Scientific journals like Neurology require authors to explicitly grade their reports based on methodological quality (http://www.neurology.org/site/misc/auth2.xhtml).

* Corresponding author at: Brain Health Program, Copeman Healthcare Centre, 700-10216 124 Street, Edmonton AB T5N 4A3, Canada.
E-mail addresses: Marianne.Hrabok@gmail.com (M. Hrabok), jdykeman@ucalgary.ca (J. Dykeman), ESherman@copemanhealthcare.com (E.M.S. Sherman), swiebe@ucalgary.ca (S. Wiebe).

1525-5050/$ – see front matter © 2013 Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.yebeh.2013.08.035

Critical appraisal of the quality of the evidence regarding epilepsy treatment has been applied almost exclusively to assessing seizure outcomes following antiepileptic drug (AED) treatment [11–13] or ES [6]. The quality of evidence pertaining to cognitive outcomes following treatments for epilepsy has received little attention. In general, EBM principles have not been systematically or explicitly applied to research studies in clinical neuropsychology [14], despite application in other areas of clinical psychology [15].

Determining the quality of the evidence pertaining to neuropsychological outcomes after surgery is relevant to assist in lateralization and localization of the seizure focus and to help identify patients at risk for postoperative decline in cognition [16]. Postoperatively, neuropsychological evaluation is used to detect cognitive and psychological strengths and challenges, assist in treatment planning [17], and assess the impact of cognitive rehabilitation [18].

Our primary aims were to (1) assess the quality of the evidence regarding neuropsychological function after ES and (2) develop and apply a checklist to assess the quality and clinical usefulness of primary studies describing neuropsychological outcomes after ES.

2. Methods

Development of the evidence-based neuropsychology (EBNP) checklist included a conceptual framework and item generation, tryouts and refinement to assess face validity and consistency, application of the items to relevant articles to determine frequency of endorsement, and item reduction to produce the final checklist.

2.1. Development sample — systematic review

The articles used for checklist development and testing were taken from a previously published systematic review of neuropsychological outcomes after resective ES. Details of the exhaustive literature search are described previously [7]. PubMed, EmBase, PsycInfo, and the Cochrane databases were searched, and we included articles up to 2010 describing neuropsychological outcomes of ES. The references of a sample of studies were reviewed for additional relevant studies. Two reviewers independently applied study inclusion criteria and abstracted data for analyses.

Neuropsychological outcomes were defined as IQ, memory, language, executive functioning, attention, and subjective cognitive changes. We restricted our sample to studies reporting original research that reported on a cognitive outcome in surgical patients. Only articles with original data were included.

2.2. Conceptual design and item generation

The checklist was derived by incorporating elements from published, general critical appraisal checklists based on EBM principles [8,19,20] as well as methodological and clinical aspects uniquely important to ES and neuropsychology [14,16,17,21]. Each coauthor had the specific expertise necessary for conceptual design and item generation of the checklist [neuropsychology (EMS, MH), epileptology (SW), and clinical epidemiology (SW, JD)]. Our goal was not to create a tool to derive a quantitative quality score or a threshold for high- and low-quality studies. The goal was to generate a framework for critically appraising ES articles that described neuropsychological outcomes with regard to (a) scientific validity and (b) clinical usefulness in managing individual patients.

2.3. Tryouts, refinement, and application of the EBNP checklist

Content validity of a comprehensive checklist was evaluated during two iterations of item tryout in a random sample of 10% of the articles. Two raters (MH/JD) independently applied the checklist items to the articles. Following each tryout, the checklist was reevaluated for item clarity and consistency and revised by all authors. Two raters (MH/JD) independently applied the final checklist to all the articles included in the study. Discordant ratings were discussed, and consensus was reached for each article. The kappa statistic was used to assess inter-rater agreement.

2.4. Item reduction and final checklist

To enhance practical application, we created an abbreviated checklist. Clinical and statistical approaches to item reduction were employed to derive the final EBNP checklist [22]. The clinical approach involved identifying themes across questions that related to each other or that provided information relevant to applicability of the study results to patient care. These themes included patient characteristics, comparison groups, randomization, sample descriptors, study design/outcomes, and clinical applicability. The statistical approach involved recursive partitioning and split sample validation to identify items that best predicted other items within clinically identified groups of questions. Analyses were completed using the *rpart* package in *R* version 2.14. The final checklist items were grouped into the four categories of the PICO system, a common heuristic in EBM which stands for (P) patients, (I) interventions, (C) comparison, and (O) outcomes. The PICO framework, widely used to ask focused research questions [23] and to identify gaps in clinical research [24], allows for a clinically relevant structure to enhance the checklist's ease of use.

2.5. Assessment of quality of studies

We assessed areas of methodological strength and weakness in the literature by calculating the proportion of studies fulfilling each criterion in the checklist. To examine factors associated with methodological quality, a total quality score was obtained by unweighted addition of the number of checklist items fulfilled for each article. A higher score reflected higher quality. Linear regression analysis assessed the association between year of publication and journal and total quality score. We used the median score as a threshold for classifying studies as having higher or lower quality. Recursive partitioning was used to identify items associated with studies that had quality scores above or below the median.

3. Results

Of 187 articles identified in the systematic review, 147 articles met the eligibility criteria and were used for checklist development. Of the excluded articles, 30 were not original research, six did not include surgical patients, and five did not report on cognitive outcomes.

3.1. Checklist development

After two iterations of tryouts and refinement, a 45-item checklist (44 items with yes/no answers and 1 item with a numerical response for sample size) was applied to each of the 147 articles. Consensus was reached in 6615 ratings, with excellent inter-rater agreement (kappa = 0.80). The percentage of studies possessing each characteristic is shown in Table 1.

3.2. Quality of the evidence

The total quality scores were normally distributed, with a mean of 23.2 (SD: 4.1), a median of 24 (interquartile range: 21 to 26), and a range of 12 to 33 (possible minimum: 0, maximum: 44). The journal of publication (n = 41) was not related to the total quality score ($r = -0.03, p = 0.68$). Although the date of publication was associated with total quality score, with increasing quality over time ($r = 0.18$, $p = 0.026$) (Fig. 1), the association was not significant after removing the two studies published before 1988. Longitudinal studies and those