Long-term full-scale intelligent quotient outcomes following pediatric and childhood epilepsy surgery: A systematic review and meta-analysis

Saman Arfaie a,1, Pouya Amin c,1, Angela T.H. Kwan d,1, Arad Solgi f,1, Ali Sarabi g,1, Benyamin Hakak-Zargar b, Tristan Brunette-Clément 1, Denys Pushenko h, j, Kamran Mir-Moghtadaei i, Mohammad Sadegh Mashayekhi k, Mohammad Mofatteh p, q, Faraz Honarvar l, Lily Yu Xi Ren m, Camille Noisieux-Lush n, Zahra Azizi o, p, Phillip Lawrence Pearl q, Torsten Baldeweg r, s, Alexander G Weil t, Aria Fallah u

a Department of Neurology and Neurosurgery, McGill University, Montreal, QC, Canada
b Faculty of Medicine, McGill University, Montreal, QC, Canada
c Faculty of Medical and Life Sciences, University of Warwick, Coventry, UK
d Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada
e Department of Chemistry and Physics, University of Toronto, Toronto, ON, Canada
f School of Kinesiology & Health Science, York University, Toronto, ON, Canada
g School of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ, United States of America
h School of Medicine, Faculty of Health, Deakin University, Geelong, VIC, Australia
i Department of Neurosurgery, Faculty of Medicine, University of Montreal, Montreal, QC, Canada
j Department of Biology, University of Toronto, Toronto, ON, Canada
k Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada
I Faculty of Medicine, Queen’s University, Kingston, ON, Canada
m Lane Medical Library and Knowledge Management Center, Stanford Medicine, Stanford University, Palo Alto, CA, United States of America
n Department of Neurology, University of Montreal, Montreal, QC, Canada
o Department of Pediatrics, University of Montreal, Montreal, QC, Canada
p School of Medicine, Dentistry and Biomedical Sciences, Queen’s University Belfast, United Kingdom
q Department of Developmental Neuroscience, UCL Great Ormond Street Institute for Child Health, London WC1N 1EH, United Kingdom
r Division of Pediatric Neurosurgery, Department of Surgery, Sainte Justine Hospital, University of Montreal, Montreal, QC, Canada
s Department of Neurosurgery and Pediatrics, David Geffen School of Medicine at University of California, Los Angeles, CA, United States of America

ARTICLE INFO

Keywords:
Cognitive outcome
Epilepsy
Full-scale intelligence quotient
Long-term
Pediatric
Surgery

ABSTRACT

Objective: Cognitive measures are an important primary outcome of pediatric, adolescents, and childhood epilepsy surgery. The purpose of this systematic review and meta-analysis is to assess whether there are long-term alterations (≥ 5 years) in the Full-Scale Intelligence Quotient (FSIQ) of pediatric patients undergoing epilepsy surgery.

Methods: Electronic databases (EMBASE, MEDLINE, and Scopus) were searched for English articles from inception to October 2022 that examined intelligence outcomes in pediatric epilepsy surgery patients. Inclusion criteria were defined as the patient sample size of ≥ 5, average follow-up of ≥ 5 years, and surgeries performed on individuals < 18 years old at the time of surgery. Exclusion criteria consisted of palliative surgery, animal studies, and studies not reporting surgical or FSIQ outcomes. Publication bias was assessed using a funnel plot and the Quality in Prognosis Studies (QUIPS) toolset was used for quality appraisal of the selected articles. A random-effects network meta-analysis was performed to compare FSIQ between surgical patients at baseline and follow-up and Mean Difference (MD) was used to calculate the effect size of each study. Point estimates for effects and 95% confidence intervals for moderation analysis were performed on variables putatively associated with the effect size.

* Corresponding author at: School of Medicine, Dentistry and Biomedical Sciences, Queen’s University Belfast, 97 Lisburn Rd, Belfast BT9 7BL, United Kingdom.
E-mail address: mnofatteh01@qub.ac.uk (M. Mofatteh).
1 Denotes equal contribution.

https://doi.org/10.1016/j.seizure.2023.01.020

Received 26 July 2022; Received in revised form 28 December 2022; Accepted 30 January 2023
Available online 1 February 2023
1059-1311/Crown Copyright © 2023 Published by Elsevier Ltd on behalf of British Epilepsy Association. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
1. Introduction

Epilepsy surgery is a treatment option for medically intractable epilepsy in pediatric populations [1,2]. The ultimate goals for operated patients are to obtain seizure freedom, cessation of anti-epileptic drugs (AEDs), and improvement of developmental capacities [3]. With early-onset epilepsy, cognitive impairment and mental retardation may result in infants less than two years of age [4]. Similarly, in late-onset epilepsy, children are at an increased risk for cognitive decline and behavioral deficits [4]. Determining seizure freedom is essential for evaluating post-operative cognitive outcomes [2,4]. Moreover, a pre-operative baseline is useful for evaluating cognitive functions for subsequent time points.

Predicting post-operative cognitive outcomes has proven difficult since epilepsy is a heterogeneous condition characterized by clinical, demographic, and etiological differences [5]. Additionally, maturational development, physiological, and functional plasticity predict post-operative outcomes challenging in pediatric patient populations [6–8]. Given that there are numerous between-study variations (i.e., cognitive domain studied using different psychometric tests, aetiologies, sample size, duration of follow-up) and methodological differences in determining the reliability of post-operative cognitive status, achieving a consensus about cognitive outcomes following epilepsy surgery has not been reached [5].

Intelligence quotient (IQ) is one of the reliable measures to assess cognitive outcomes following epilepsy surgery in pediatric patients. More specifically, the Full-Scale IQ (FSIQ) is often used as a standardized test that represents the global intellectual ability of an individual [9]. This test consists of scores from four domains including, perceptual reasoning (PRI), processing speed (PSI), verbal comprehension (VCI), and working memory (WMI) [10]. PRI evaluates non-verbal and fluid reasoning, PSI measures speed, economy, and accuracy of information processing, VCI examines verbal reasoning ability, and WMI that of storage and manipulation of information for short-term memory consolidation [11]. This multidimensionality provides a better outlook of intelligence than its core components assessed independently. FSIQ is consolidating [11]. This multidimensionality provides a better outlook of intelligence than its core components assessed independently. FSIQ remains an area of open investigation. Some studies have shown that children undergoing epilepsy surgery may subsequently present lower IQ scores, which predispose them to behavioral issues and psychosocial dysfunction later in life [13,14]. Therefore, it remains the cornerstone of measuring the cognitive ability of an individual and acts as a reliable tool in studies that examine intelligence.

While several studies support the efficacy of epilepsy surgery in achieving seizure freedom, the long-term impact of such surgeries on patients’ FSIQ remains an area of open investigation. Some studies have shown that children undergoing epilepsy surgery may subsequently present lower IQ scores, which predispose them to behavioral issues and psychosocial dysfunction later in life [13,14]. Therefore, the purpose of this systematic review and meta-analysis is to analyze the pre-operative cognitive function and long-term post-operative cognitive outcomes (≥ 5 years) in pediatric and early adolescent cohorts undergoing epilepsy surgery in the literature.

2. Methods

2.1. Search strategy

This systematic review and meta-analysis was conducted based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines. We conducted our electronic searches using the EMBASE, MEDLINE, and Scopus databases from inception to October 2022 for relevant articles. The following key terms were used in various combinations: “adolescent”, “childhood”, “cognitive”, “epilepsy”, “focal resection”, “full-scale intelligence quotient”, “hemispherectomy”, “intelligence quotient”, “laser interstitial thermal therapy”, “lesionectomy”, “lobectomy”, “outcomes”, “pediatric”, “psychosurgery”, “resection”, “seizure”, “surgery”, “tempotomy”, and “Wechsler”. Three reviewers (S.A., P.A., A.T.H.K) independently performed title and abstract screening. A full-text review was conducted by three reviewers (A. Solgi, P.A., B.H.Z.) and any discrepancies were resolved by S.A. To ensure that no appropriate articles were excluded, two reviewers (S.A., A. Solgi) manually searched from the selected databases until October 2022 and checked the references of relevant papers. Additionally, one reviewer (S.A.) contacted corresponding authors of articles meeting criteria for the papers missing data, and those who successfully responded with matching data were included [14–16].

2.2. Eligibility criteria

Inclusion criteria for the studies were as follows: English only, the patient sample size of ≥ 5, average follow-up of ≥ 5 years, individuals ≤ 18 years old at the time of surgery, and reports of pre-operative and post-operative FSIQ. If multiple studies from the same center examined overlapping patient populations, the study with the longest duration of follow-up was included. When studies with a pediatric and adolescent cohort of ≥ 5 patients did not report epilepsy surgery in the title or abstract, their eligibility was assessed through full-text review. Exclusion criteria consisted of palliative surgical procedures (i.e., corpus callosotomy, vagal nerve stimulator, deep brain stimulator insertion), studies not reporting on surgical or FSIQ outcomes, animal studies, gray literature, case reports, reviews, conference abstracts, and editorials.

2.3. Selection and coding of data

Our primary outcome measure was to assess any changes between preoperative and post-operative FSIQ in pediatric patients. Given that long-term FSIQ was assessed, studies with ≥ 5 years were selected, and their mean and standard deviation was obtained. Factors possibly associated with pre-and post-operative cognition reported in the literature consisted of the sample population, seizure outcome, sex, type of surgery, etiological characterization (acquired or progressive, congenital, and tumor or non-tumor), side of surgical focus, age at seizure onset, age at surgery, follow-up period, baseline (pre-operative) and follow-up (post-operative) FSIQ, and percentage of seizure freedom at...
follow-up. Individual participant data (IPD) were aggregated and added to study-level data to enable analysis based on our selection criteria. Independently abstracted data were managed on Microsoft Excel Spreadsheet (version 2016; Microsoft, Redmond, WA, USA).

2.4. Assessment of risk of bias and agreement

The risk of bias for each study was evaluated by two reviewers (A. Solgi, B.H.Z.) and verified by a third reviewer (S.A.) using the Quality in Prognosis Studies (QUIPS) tool [17]. In accordance with the QUIPS, each paper was assigned low, moderate, or high risk of bias for six domains: study participation, study attrition, prognostic factor measurement, outcome measurement, study confounding, and statistical analysis and reporting. Mean risk scores for each domain were calculated by associating the level of risk with numbers (low = 1, moderate = 2, high = 3). We calculated Cohen’s kappa score (k) to determine the strength of agreement, and inter-rater reliability, for the title and abstract, as well as full-text screening using the Covidence web application (www.covidence.org, Veritas Health Innovation Ltd, Melbourne, Australia) with the following thresholds for interpretation: < 0.20 as slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and > 0.81 as almost perfect agreement.

2.5. Assessment of publication bias and heterogeneity

We assessed publication bias in studies that included study-level data by visually assessing the symmetry of funnel plots. These papers reported stratified data for patients’ FSIQ five years after surgery. The I2 statistic was used to assess data heterogeneity. I2 levels of 0–30%, 30–60%, 60–90%, and 90–100% were considered low, moderate, substantial, and considerable, respectively. Moreover, the Baujat plot was used to identify studies that contribute to heterogeneity.

2.6. Statistical methods

All statistical analyses were performed using RStudio (version 3.3.3) and Metafor package. A random-effects network meta-analysis combining direct and indirect evidence was performed to compare overall intelligence (FSIQ) between surgical patients at baseline and follow-up. Mean differences were used as the primary summary measures for the analysis, and the results are presented as 95 percent confidence intervals. The null hypothesis of the test is that the effect size (the mean difference) is zero, while the alternative hypothesis is that it is not. Additionally, moderation analysis was performed to investigate the moderating effects of variables putatively associated with the mean difference in FSIQ for surgical patients. Etiology was subdivided into three major categories: acquired, congenital, and progressive. (1) Acquired included mesial temporal sclerosis, dysembryoplastic neuroepithelial tumors, cavernoma, ganglioglioma, atrophy, trauma, stroke, gliosis, and hippocampal sclerosis. (2) Congenital consisted of polymicrogyria, focal cortical dysplasia, tuberous sclerosis complex, and malformation of cortical development among others. Finally, (3) progressive comprised of Sturge-Weber Syndrome, Rasmussen encephalitis, and others.

3. Results

3.1. Individual study and overall estimates

After running the search strategy and removing duplicates, we identified 21,408 articles from the electronic databases MEDLINE, Embase, and Scopus with excellent agreement between the three reviewers (k = 0.854). We excluded 15,896 papers based on our inclusion criteria. Next, 797 articles that fit our inclusion and exclusion criteria were assessed as full text with an excellent agreement between our reviewers (k = 0.926). Overall, seven studies were identified and included in this study. Fig. 1 illustrates the search strategy and selection process in detail. Moreover, it should be noted that there were no patient duplications between Skirrow et al., 2011 and Skirrow et al., 2019, even though the patients share the same medical center [14,18]. This information was verified through correspondence with the senior author for data clarification.

3.2. Descriptive information

Studies were published between 2011 and 2021 and included data from hospital centers on three continents and six countries (Canada, USA, UK, China, France, and South Korea). Five studies reported data and factors possibly associated with pre- and post-epilepsy surgery at the study level [5,14–16,20]. In two studies, these measures were recorded at the individual participant data (IPD) level [18,21].

Most surgeries for the treatment of underlying epilepsies were functional hemispherectomies, lobectomies, or lesionectomies. The classification measures used to report seizure outcomes varied from one study to another. Almost all studies reported on the side of surgical focus, and seizure freedom showed significant clinical improvements. The underlying aetiologies behind the seizures are also reported in detail. Studies reported seizure status at follow-up either crudely as presence/absence or using the International League Against Epilepsy (ILAE) and Engel’s Classification (Tables 1 and 2).

From the seven included studies for the 330 surgical patients, comprising 168 females, 162 males, the average age at seizure onset ranged from 0.73 to 6.00 years, with a weighted mean of 5.01 years. Three studies also reported a control group of 57 non-surgical patients consisting of 38 females and 19 males, with an average age at seizure onset from 3.67 to 4.70 years, and a weighted mean of 4.34 years [14,16,18] (Table 1).

For the surgical cohort, the average age at surgery ranged from 6.16 to 14.00 years, with a weighted mean of 12.00 years, and the average time between surgery and follow-up ranged from 5.00 to 9.45 years, with a weighted mean of 6.60 years in these studies. Considering the side of focus, 146 surgeries were conducted on the left side and 147 on the right. In total 71% of patients had been seizure-free on average prior to the follow-up period (Table 2).

The age-appropriate Wechsler Intelligence Scales for FSIQ evaluation in each study have been included and further specified (Table 3). Baseline and FSIQ measurements for each respective study have been assessed and their mean, standard deviation, and sample size are reported. This meta-analysis included 330 surgical patients with baseline FSIQ recorded for 296 patients and follow-up FSIQ for 262 patients. Additionally, 57 non-surgical patients were reported from the studies with baseline FSIQ noted for 54 patients and a follow-up FSIQ for 55 patients. The observed discrepancy in these numbers is attributed to the follow-up attrition rate or the lack of reported FSIQ scores (Table 3).

3.3. Assessment of quality of studies

Overall, all seven studies had a low mean risk of bias with respect to the six domains of study participation, study attrition, outcome measure, study confounding, and statistical analysis and reporting (Figs 2 and 3). Papers assessed from 2019 and onwards have a lower risk of bias. Bias due to participation, outcome measurement, statistical analysis, and reporting was deemed low risk across all studies, with the remaining factors exhibiting moderate risk. One study had a high level of bias due to attrition [15].

3.4. Publication bias

Funnel plots are a visual tool to investigate publication bias for studies reporting study-level data. As Fig. 4 demonstrates, the funnel plot is symmetric and there was no significant evidence of funnel plot asymmetry according to Egger’s test (p-value = 0.82). Therefore, there
was no evidence of publication bias in our meta-analysis. Fig. 5 depicts a Baujat plot, in which each number represents a different study, and the studies on the top right have the most influence on the results and contribute the most to heterogeneity.

3.5. Primary outcome (Mean FSIQ, difference FSIQ)

Quantitative analysis was performed on the studies with surgical patients \((N = 330)\). According to the forest plot (Fig. 6) the range of FSIQ mean difference for surgical patients in follow-up showed a significant increase compared to baseline, with a mean difference of 3.36 and 95% confidence interval of \((0.14, 6.57)\) with a p-value of 0.04. The I² statistic for this analysis is 0%, indicating low heterogeneity. Therefore, we reject the null hypothesis because the 95 percent confidence interval for the random effect model does not contain zero. The results show that the FSIQ mean difference has slightly increased and this change is significant. Since only three studies reported on FSIQ in non-surgical individuals, quantitative analysis was performed on these studies \((N = 57)\) \([14,16,18]\). As the forest plot indicates, the range of FSIQ mean difference for non-surgical patients in follow-up showed a decrease compared to baseline, with a mean difference of \(-0.91\) and 95% confidence interval of \((-7.37, 5.54)\) with a p-value of 0.78.

3.6. Moderation analysis

According to Table 4, moderation analysis on sex, age at seizure epilepsy onset, age at surgery, side of surgical focus, time from epilepsy onset to surgery, follow-up period, type of surgery, etiological characterization, and post-surgical seizure freedom all revealed no correlation between these variables and the effect size. As a result, there was no evidence that these variables moderated the observed mean difference.

4. Discussion

4.1. Primary outcome

The natural progression of intractable epilepsy, if left untreated, leads to poor results in neurocognitive functions, language impairments, visuospatial skills, and other behavioral manifestations \([22–26]\). Seizure freedom is the main purpose behind pursuing neurosurgical intervention alongside improving quality of life by minimizing neurological sequelae as much as possible \([27]\). However, as with most surgical procedures requiring the removal of some brain tissue, a major concern raised by parents is whether there may be alterations to cognitive function \([28–32]\).

This meta-analysis aimed to evaluate the long-term changes in general intelligence, using FSIQ as a reliable estimate, among pediatric and adolescent patients following surgical procedures with curative intent,
corroborate those of recent large cohorts studying developmental and studies suggest a statistically, but not clinically significant increase in and FSIQ at a minimum of five years following surgery. Identified during childhood and adolescence [29], patient selection (such that but not clinically significant, and relatively stable many years after surgery. These findings demonstrated an increase in FSIQ [5,14,16,18,20,21], whereas only one study showed a relatively small decrease in FSIQ [15].

The calculated mean increase in FSIQ was statistically significant, indicating sustained improvement.

Other procedures include Multiple Subpial Transections. Operations involving two or more types of procedures on a single patient were included in this category.

The change in FSIQ observed at the long-term follow-up (average of 6.60 years) had a mean increase of 3.36 from baseline among the surgical cohort. As the forest plot (Fig. 6) indicates, six studies demonstrated an increase in FSIQ [5,14,16,18,20,21], whereas only one study showed a relatively small decrease in FSIQ [15].

The calculated mean increase in FSIQ was statistically significant, but not clinically significant, and relatively stable many years after surgery. This could be attributed to the observed variability of IQ scores during childhood and adolescence [29], patient selection (such that those who require urgent surgical attention are also more likely to benefit from the intervention and show minimal side effects), the extent of surgery (i.e., performing a smaller resection with the intention of avoiding functional brain regions) and other comorbidities such as autism. However, we suspect that it is rather a result of epilepsy surgery, consistent with recent reports in favor of surgical intervention for refractory epilepsy [30]. Lack thereof or delayed surgical intervention may translate to continuous epileptic episodes with permanent developmental damage and possible mental retardation [31]. In such instances, decreased FSIQ secondary to continued seizures would be an expected outcome. Moreover, epilepsy surgery yields relatively predictable functional outcomes, of which FSIQ is one, whereas continued seizures and side effects of anti-epileptic drugs leave significantly more unpredictable functional outcomes [27].

Table 1
Characteristics of Included Studies for Surgical/Non-Surgical Patients.

<table>
<thead>
<tr>
<th>ID</th>
<th>Author &amp; Year</th>
<th>Location</th>
<th>Sample Size (N)</th>
<th>Sex (Female / Male)</th>
<th>Age at Seizure Onset (Years)</th>
<th>Etiology (N) (Acquired/ Congenital/ Progressive/ Dual or Multiple/ Unspecified or Unknown)</th>
<th>Method Used to Report Seizure Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Laguitton et al., 2021</td>
<td>Marseille, France</td>
<td>81</td>
<td>39/42</td>
<td>6.0</td>
<td>(47 / 16 / 0 / 0 / 18)**</td>
<td>Reported as presence or absence of seizures</td>
</tr>
<tr>
<td>2</td>
<td>Qu et al., 2020</td>
<td>Xi’an, China</td>
<td>10</td>
<td>4/6</td>
<td>5.7</td>
<td>(4 / 4 / 2 / 0 / 0)</td>
<td>Reported as presence or absence of seizures</td>
</tr>
<tr>
<td>3</td>
<td>Ko et al., 2019</td>
<td>Seoul, South Korea</td>
<td>58</td>
<td>29/29</td>
<td>6.0</td>
<td>(58 / 0 / 0 / 0 / 0)</td>
<td>Reported as presence or absence of seizures</td>
</tr>
<tr>
<td>4</td>
<td>Skirrow et al., 2019</td>
<td>London, UK</td>
<td>31</td>
<td>13/18</td>
<td>5.67</td>
<td>(18 / 11 / 0 / 2 / 0)</td>
<td>Reported as presence or absence of seizures</td>
</tr>
<tr>
<td>5</td>
<td>Puka et al., 2016</td>
<td>Toronto, Canada</td>
<td>71</td>
<td>41/30</td>
<td>5.5</td>
<td>(33 / 21 / 1 / 13 / 3)</td>
<td>Reported as presence or absence of seizures</td>
</tr>
<tr>
<td>6</td>
<td>Arya et al., 2015</td>
<td>Ohio, USA</td>
<td>37</td>
<td>21/16</td>
<td>0.73</td>
<td>(0 / 37 / 0 / 0 / 0)</td>
<td>ILAE Classification</td>
</tr>
<tr>
<td>7</td>
<td>Skirrow et al., 2011</td>
<td>London, UK</td>
<td>42</td>
<td>21/21</td>
<td>4.01</td>
<td>(42 / 0 / 0 / 0 / 0)</td>
<td>Reported as presence or absence of seizures</td>
</tr>
<tr>
<td></td>
<td>Total (Weighted Average)</td>
<td></td>
<td>330</td>
<td>168 / 162</td>
<td>(5.01)</td>
<td>202 / 89 / 3 / 15 / 21</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Non-Surgical Patients</td>
<td></td>
<td>4</td>
<td>Skirrow et al., 2019</td>
<td>9</td>
<td>5/4</td>
<td>3.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>Puka et al., 2016</td>
<td>37</td>
<td>26/11</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td>Skirrow et al., 2011</td>
<td>11</td>
<td>7/4</td>
<td>3.68</td>
</tr>
<tr>
<td></td>
<td>Total (Weighted Average)</td>
<td></td>
<td>57</td>
<td>38/19</td>
<td>(4.34)</td>
<td>11 / 0 / 1 / 0 / 8</td>
<td>—</td>
</tr>
</tbody>
</table>

* Only patients who had a follow-up were included.
** Etiology is extracted based on abnormal MRI findings which were reported in 78% of the patients.

Table 2
Other Characteristics of Included Studies for Surgical Patients.

<table>
<thead>
<tr>
<th>ID</th>
<th>Author &amp; Year</th>
<th>Time from Epilepsy Onset to Surgery (Years)</th>
<th>Age at Surgery (Years)</th>
<th>Side of Focus in Surgery (Left/ Right)</th>
<th>Type of Surgery (N) (Lobectomy/ Hemispherectomy/ Lesionectomy/ Corticectomy/ Other)</th>
<th>Follow up Period (Years)</th>
<th>Post-surgical Seizure Freedom (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Laguitton et al., 2021</td>
<td>7</td>
<td>14.0</td>
<td>32/49</td>
<td>(69 / 0 / 0 / 12 / 0)</td>
<td>5</td>
<td>70%</td>
</tr>
<tr>
<td>2</td>
<td>Qu et al., 2020</td>
<td>6.12</td>
<td>11.8</td>
<td>4/6</td>
<td>(0 / 10 / 0 / 0 / 0)</td>
<td>7.5</td>
<td>70%</td>
</tr>
<tr>
<td>3</td>
<td>Ko et al., 2019</td>
<td>1.3</td>
<td>10.2</td>
<td>28/30</td>
<td>(0 / 0 / 58 / 0 / 0)</td>
<td>5.6</td>
<td>87.9%</td>
</tr>
<tr>
<td>4</td>
<td>Skirrow et al., 2019</td>
<td>6.72</td>
<td>13.19</td>
<td>16/15</td>
<td>(31 / 0 / 0 / 0 / 0)</td>
<td>9.45</td>
<td>71%</td>
</tr>
<tr>
<td>5</td>
<td>Puka et al., 2016</td>
<td>7.20</td>
<td>12.98</td>
<td>41/30</td>
<td>(41 / 0 / 13 / 17 / 0)</td>
<td>6.7</td>
<td>55%</td>
</tr>
<tr>
<td>6</td>
<td>Arya et al., 2015</td>
<td>5.52</td>
<td>6.16</td>
<td>N/A</td>
<td>(19 / 2 / 0 / 0 / 16)</td>
<td>5.68</td>
<td>57%</td>
</tr>
<tr>
<td>7</td>
<td>Skirrow et al., 2011</td>
<td>9.29</td>
<td>13.3</td>
<td>25/17</td>
<td>(42 / 0 / 0 / 0 / 0)</td>
<td>9.4</td>
<td>86%</td>
</tr>
<tr>
<td></td>
<td>Total (Weighted Mean)</td>
<td>(6.11)</td>
<td>(12.00)</td>
<td>146/147</td>
<td>202 / 12 / 71 / 29 / 16</td>
<td>(6.60)</td>
<td>(71%)</td>
</tr>
</tbody>
</table>

* Other procedures include Multiple Subpial Transections. Operations involving two or more types of procedures on a single patient were included in this category.
4.2. Variables associated with outcomes

Baseline cognitive scores and seizure status have been consistently identified as predictors of cognitive outcomes both with and without surgery [28,32,33]. We also sought to identify predictors of long-term post-operative cognitive outcomes with the reviewed literature. According to Table 4, moderation analysis on sex, age at seizure epilepsy onset, age at surgery, side of surgical focus, time from epilepsy onset to surgery, follow-up period, type of surgery, etiological characterization, and post-operative seizure freedom all revealed no correlation between these variables and the effect size. As a result, there was no evidence that these variables moderated the observed mean difference and none of them showed a statistically significant effect on FSIQ.

The absence of post-operative seizure freedom effect in our findings is at odds with some of the larger cohort studies, which have a shorter follow-up duration of 24 months or less [28,33]. However, it should also be noted that there are other studies that did not find a correlation between seizure freedom and cognitive improvement [34,35]. Given that we assessed a follow-up period of five or more years, such a longer period risks a greater degree of heterogeneity because of lower retention rates among patients with less successful surgical outcomes post-operation mostly due to seizures. Consequently, this minimizes the difference observed for seizure freedom in our patient population. However, this is not an issue for studies with short-term follow-up since they show significant correlation between seizure freedom and cognitive outcomes [2,36]. It is also possible that FSIQ may not be the most sensitive measure compared to more fluid measures such as PSI or WMI, which are impacted by seizure and anti-seizure medications.

Age is another usual predicting factor for postoperative outcomes. Despite this, we were unable to derive a meaningful relationship between age at epilepsy onset and surgery, and post-operative outcomes in our review. This is likely multifactorial and due to some extent to controversy in the literature, with some papers reporting that older age has been associated with better cognitive outcomes [37,38] while others report the opposite [39,40]. Even the most recent systematic review on this topic also reports heterogeneity in findings [26]. Regarding age at onset, early epilepsy, from focal cortical dysplasia for example, can negatively affect a child's neurological development, whereas later

<table>
<thead>
<tr>
<th>ID</th>
<th>Author &amp; Year</th>
<th>IQ Test Administered</th>
<th>Baseline FSIQ Mean</th>
<th>Standard Deviation</th>
<th>Sample Size</th>
<th>Follow-up FSIQ Mean</th>
<th>Standard Deviation</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Laguitton et al., 2021*</td>
<td>Wechsler Intelligence Scale (WISC-IV, WAIS-III)</td>
<td>81.89</td>
<td>17.67</td>
<td>81</td>
<td>85.74</td>
<td>20.02</td>
<td>68</td>
</tr>
<tr>
<td>2</td>
<td>Qu et al., 2020</td>
<td>Wechsler Intelligence Scale (WISC, WAIS)†</td>
<td>77.7</td>
<td>9.0</td>
<td>10</td>
<td>84.4</td>
<td>16.9</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Ko et al., 2019††</td>
<td>Wechsler Intelligence Scale (WPPSI, WISC, WAIS)</td>
<td>78.9</td>
<td>27.1</td>
<td>54</td>
<td>80.9</td>
<td>28.7</td>
<td>42</td>
</tr>
<tr>
<td>4</td>
<td>Skirrow et al., 2019</td>
<td>Wechsler Intelligence Scale (WPPSI-II/R, WISC-III/IV, WAIS-III/IV, WASI)</td>
<td>85.19</td>
<td>19.19</td>
<td>31</td>
<td>92.35</td>
<td>18.57</td>
<td>31</td>
</tr>
<tr>
<td>5</td>
<td>Puka et al., 2016**</td>
<td>Wechsler Intelligence Scale (WISC-IV, WAIS-III WMS-III)</td>
<td>80.3</td>
<td>18.7</td>
<td>59</td>
<td>80.7</td>
<td>18.8</td>
<td>61</td>
</tr>
<tr>
<td>6</td>
<td>Arya et al., 2015*</td>
<td>Wechsler Intelligence Scale (WISC, WAIS)</td>
<td>68.17</td>
<td>19.92</td>
<td>23</td>
<td>64.92</td>
<td>11.98</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>Skirrow et al., 2011**</td>
<td>Wechsler Intelligence Scale (WISC-III/R, WAIS-III/R)</td>
<td>82.5</td>
<td>19.82</td>
<td>38</td>
<td>90.24</td>
<td>19.73</td>
<td>38</td>
</tr>
</tbody>
</table>

WISC = Wechsler Intelligence Scale for Children; WAIS = Wechsler Intelligence Scale for Adults; WPPSI = Wechsler Preschool and Primary Scale of Intelligence; WAIS = Wechsler Abbreviated Scale of Intelligence; HAWIK = Hamburg-Wechsler Intelligence Scale for Children; HAWIE = Hamburg-Wechsler Intelligence Scale for Adults; R = Revised.

* Weighted statistics formulas are used to aggregate the values in the articles.
† The missing individual values for this article were obtained directly from the author via email correspondence.
†† FSIQ was measured and available in a subset of patients, which is reflected in the FSIQ sample size.

Fig. 2. Summary of the risk of bias assessment using the QUIPS tool.
epilepsy, from a tumor for example, can lead to regression, albeit in a patient with relatively superior baseline cognitive development, potentially leading to better overall post-operative cognitive function, as reported by Helmstaedter et al. and Cloppenborg et al. in their respective large cohorts [28,33]. Concerning age at surgery and its effect on cognitive function, younger patients might tolerate surgery better because of the neuroplasticity, which was also observed by Helmstaedter et al. [33]. Conversely, later surgery could mean less severe forms of epilepsy. Cloppenborg et al. found a positive correlation between post-operative IQ and older age at surgery [28]. Another important consideration derived from the age at onset and surgery is time to surgery because shorter delays imply less time for damaging seizures, and less neurodevelopmental delay or regression [28,33].

4.3. Strength and limitations

To our knowledge, this is the first study that assesses long-term cognitive outcomes in the pediatric population. The strengths of the study included screening all available literature, not limited to the date of publication, that adhered to our selection criteria, and ensuring that no key papers were missing. Furthermore, all variables putatively associated with the mean difference effect size have been analyzed and included to ensure a comprehensive conclusion. In cases of missing values, data was requested and included in the analysis. A significant change in FSIQ is rarely explained by a practice effect compared to studies with a follow-up in the first year alone. Since our interest resides in long-term FSIQ, the results of this study can be used as a surrogate of factors evaluating functional outcomes (i.e., employment, educational attainment, psychosocial factors) given that representative data from seven different countries have been included.

A limitation inherent to all meta-analyses is that the data are dependent on the available literature. It should be noted that the generalizability of FSIQ results may be limited to children who have the ability to perform the age appropriate FSIQ testing. In the neurosurgical literature, the primary source of data originates from case series as it would be deemed unethical to not offer surgical interventions for patients who need them. As such, there is often a lack of a control group for the non-surgical cohorts to compare the natural history of diseases without surgery to outcomes following surgery. Moreover, it is difficult to differentiate between purely non-surgical patients and patients who could have been offered surgery, but were not due to various reasons.
and complications. Therefore, natural disease history may be different between them. As it pertains to IQ measurements for epilepsy, non-surgical control patients may also show statistically significant test-retest longitudinal gains, so their absence adds uncertainty about the definitive statistical significance and clinically uncertain improvements in surgical patient groups.

Furthermore, there is a lack of IPD-level data across all studies and a small number of articles adhering to our selection criteria. The
aggregation of IPD, which consisted of two studies, was conducted to ensure homogeneity in our analysis [18,21]. Post-operative anti-seizure medication was not systematically reported in the studies. Considering the impacts these medications have on working memory, speed of processing, sustained attention, and coding, we are unable to deduce whether a decrease in anti-seizure medication usage following a successful operation could explain the improvement in FSIQ. With some studies reporting improvement in FSIQ upon reduction of antiseizure medication, this factor should be considered in future studies [15].

Depending on the sample sizes, some studies exerted a greater influence on our calculation of the FSIQ mean difference. Additionally, only English papers were evaluated. Despite these, it is reassuring to encounter consistent and homogenous findings across all papers, which lend validity to our results. Overall, our findings are significant but should be interpreted only in the context of pediatric epilepsy patients undergoing curative surgeries, and therefore it would be difficult to generalize these findings to any specific etiology or type of surgery. To improve this limitation, future studies should exclusively focus on individual aetiologies and surgeries and stratify the patient population.

4.4. Future directions

Each paper presented different criteria for patient selection and provided alternative definitions for their processes. For example, Laguittion et al. [5] relied on professional neuropsychologists as evaluators while Puka et al. [38] used trained research assistants. It is possible that heterogeneity in evaluator training may have influenced the quality of administration and the overall scores. This should be addressed in a future study establishing standardized neuropsychological evaluation criteria and reporting systems. Future research in the field would benefit from establishing guidelines to record consistent long-term outcomes systematically and rigorously, using FSIQ and other cognitive measurements, for patients undergoing epilepsy surgery. Information pertaining to pre- and post-surgery effects on different scales that comprise the FSIQ (i.e., VC, WM, PS, PR) would have also been helpful. Effects of surgery will differentially impact specific scales associated with FSIQ depending on the functional brain region being operated on. Moreover, if the intellectual profiles of the participants were heterogeneous (i.e., significant weakness in WM and PS), the FSIQ would not be considered the best metric to evaluate intellectual potential, and instead individual scales should be used to better address the strengths and weaknesses of each patient.

Another suggestion for future longitudinal studies, with a control group consisting of patients with medically refractory epilepsy, who are not candidates for surgery, is to match patients according to individual characteristics, IQ profiles, and epilepsy surgeries to better describe the cognitive changes observed following each surgery. Additionally, future investigations could benefit from multi-center studies, with a larger sample size observed over an extended period, that will account for any possible confounding variables. Specifically, it would be beneficial for prospective studies to consider social (i.e., SES, interpersonal support, parenting, nutrition) and psychological (i.e., resilience, psychiatric disorders, level of stress) aspects to gain a better understanding of potential factors that could influence cognitive outcomes.

5. Conclusion

Overall, this analysis showed a statistically significant increase in long-term FSIQ after epilepsy surgery in pediatrics and adolescents. Although this does not translate to a clinically significant change in FSIQ, it does provide evidence that most children do not show long-term deterioration. There are patients in some studies who show significant improvement while few demonstrate significant deterioration. Our findings emphasize the importance of discussing epilepsy surgery as a viable option for these patients. As such, longitudinal studies with larger patient cohorts should be used to evaluate cognitive outcomes before and after surgery utilizing standardized tests.

Funding

No funding has been obtained for this work.

Disclosures

This study adhered to ethical standards and guidelines. All authors contributed significantly to the production of the manuscript and reviewed it.

CRediT authorship contribution statement


Declaration of Competing Interest

The authors have no conflicts of interest to disclose.

Acknowledgements

Preliminary results of this work were presented at the 89th Annual Scientific Meeting of the American Association of Neurological Surgeons held in Orlando, Florida, USA, 21st-25th August 2021. We would also like to thank Professor Torsten Baldeweg for sharing the raw data to his papers, his insights on the manuscript, edits, and suggestions.

References
