Chapter 3. Results

Definitions of Treatment-Resistant Epilepsy

In this section of the Evidence Report, we addressed Key Question #1: *What are the definitions of treatment-resistant epilepsy used in the literature?*

The purpose of this question is to catalogue the definitions of treatment-resistant epilepsy that appear in the published literature. To address this question, we abstracted the phrase or sentence used to describe treatment-resistant patients with epilepsy in clinical studies, in clinical practice guidelines, and in reviews that met the inclusion criteria listed below. To tally the number of publications defining treatment resistance, we considered even the least specific of definitions (Evidence Tables 1-3). However, a synonym was not considered a definition. If patients were described with any of the following terms, and those terms were not further defined, we considered the definition to be "Not Reported":

- Medically intractable
- Medication-resistant
- Medically refractory
- Medically resistant
- Medically uncontrolled
- Drug-resistant
- Refractory
- Intractable
- Pharmaco-resistant
- Chronic treatment-resistant
- Inadequately controlled
- Uncontrolled
- Poorly controlled
- Therapy-resistant

Because the majority of studies and reviews did not report a definition, we also examined the patient inclusion criteria that were used in published studies. Although these criteria do not comprise a formal *explicit* definition of treatment-resistant epilepsy, they can be used to determine whether there is a consistently applied *implicit* definition of this term. Such implicit definitions, however, are less informative than explicit definitions. This is because inclusion criteria are constructed to meet the specific demands of the study rather than to address the general concept of what constitutes treatment-resistant epilepsy.

Question specific inclusion criteria

As noted in the Methodology section, the general inclusion criteria listed in that section do not apply to this question. Rather, we included:

1. Any clinical study that was evaluated in Questions 2 to 6, that enrolled at least 50 patients, and that was published in 1996 or later. All such studies meeting the initial inclusion criteria for each question were included, regardless of whether they were later excluded from the analysis of that question. We abstracted definitions from clinical studies in an effort to obtain a broad sample of definitions. We did not include articles retrieved for Key Questions 7 – 9, because the nature of these studies made them less likely to include definitions of treatment-resistant epilepsy.
2. A random sample of 100 review articles on treatment-resistant epilepsy published between 1996 and 2001, inclusive. We chose this random sample by using a random number generator to assign a random number to each of the 298 review articles identified in our searches for this Evidence Report. We chose to use a random sample rather than a comprehensive dataset out of consideration of the time and budget for this project.
3. Any evidence-based clinical practice guidelines identified during our searches. We termed a guideline as "evidence-based" if it was included in the National Guidelines Clearinghouse (NGC).[a]

Evidence base

For this question, we included 82 published clinical studies, 100 randomly selected review articles and 3 clinical practice guidelines. Thus, we examined 185 publications. The number of publications reporting a definition are listed in Table 3.

Design and conduct of included studies

This question addresses definitions, not an intervention or diagnostic. As such, an evaluation of the quality of the literature is not relevant.

Definitions in Included Articles

Of the 82 clinical studies that met our inclusion criteria for this question, only 24 (29 percent) reported an explicit definition of "intractable", "refractory", "treatment-resistant," or any similar term. The remainder merely stated that the patients they enrolled had treatment-resistant epilepsy (or some equivalent term) without defining that term. Of the 24 articles reporting a definition, five definitions did not include any specific information (e.g. "incompletely controlled by existing therapy").[30] One study defined treatment resistance in terms of seizure frequency with no mention of treatment.[31]

Of the remaining 19 studies, 15 reported the number of AEDs patients tried before being considered treatment-resistant. Two studies required at least one AED, four required at least two, and four required three. Five were nonspecific (e.g. "multiple"). Six of the studies named the AEDs that they required patients to have tried before being considered treatment-resistant.

Three definitions mentioned intolerable side effects or ineffectiveness at maximum tolerated dose as a reason to consider drug treatment unsuccessful, four included seizure frequency as part of the definition, six included duration of symptoms, and one mentioned monitoring serum drug levels (Evidence Table 1). None of the studies mentioned auras. Because only 29 percent of the initial 82 articles reported definitions, a common definition of treatment-resistant epilepsy does not seem to be used in the literature. Even among the studies reporting a definition, no consensus can be discerned.

[a] For further information on the National Guidelines Clearinghouse go to www.guideline.gov

Definitions in Clinical Practice Guidelines

Of the three guidelines identified by our searches, only one reported a definition of treatment-resistant epilepsy. This guideline defined a patient with treatment-resistant epilepsy as having "inadequately controlled seizures or significant side effects for whom no options had been available.[32]" The reported definitions from guidelines are listed in Evidence Table 2.

Definitions in Review Articles

Of the 100 review articles surveyed, 79 articles defined treatment-resistant epilepsy as the presence of uncontrolled seizures or a similar term. Two of the remaining definitions were not specific.[33,34] Of the remaining 19 reviews, eight reported the number of AEDs patients tried before being declared treatment-resistant. Three of the eight reviews required at least two AEDs and three required three. Two of the eight reviews were not specific (e.g. "several"). Only one of the reviews named the AEDs that they required patients to have tried before being considered treatment-resistant.

Twelve reviews mentioned intolerable side effects or ineffectiveness at maximum tolerated dose as a reason to consider drug treatment unsuccessful. Four reviews mentioned frequent seizures as part of their definition, but none of these quantified what was meant by "frequent." Rather, their effect on the ability of the patient to lead a normal life was considered the proper criterion in three of these four reviews.

Four definitions included duration of symptoms, with one simply stating that duration was not a criterion.[35] Three mentioned monitoring serum AED levels, with one stating that dosage of AEDs should be increased to the maximum tolerated regardless of serum concentrations.[36] None of the reviews mentioned auras as part of their definitions. Reported definitions are listed in Evidence Table 3. No consensus definition of treatment-resistant epilepsy can be inferred from the available information.

Definitions Implied by Inclusion Criteria and Patient Characteristics in Clinical Studies

Because definitions were infrequently reported, we examined the inclusion/exclusion criteria and the characteristics of patients in clinical studies to determine the characteristics of patients deemed to have treatment-resistant epilepsy. These characteristics may *imply* a definition. However, the requirements of a trial are not necessarily the same as the requirements of a patient seeking treatment. A patient experiencing one seizure a year may be considered treatment-resistant but is unlikely to be included in a clinical trial. Thus, patient inclusion criteria may be biased toward enrolling more severely ill patients.

In addition to listing inclusion/exclusion criteria of studies (Evidence Table 4), we examined, at the request of the Expert Panel and Technical Experts, whether these criteria differed depending on the purpose of the trial or the target population of the intervention being studied (Evidence Table 5).

Of the 82 clinical studies included, eight specifically examined pediatric patients, two focused on Lennox-Gastaut syndrome, while two examined mesial temporal sclerosis (MTS), and one examined non-MTS focal lesions.

There were 19 drug trials for US Food and Drug Administration (FDA) approval and 17 additional drug trials that were not performed for this purpose. There were seven nonsurgical studies of nondrug treatments and 39 surgery trials. The surgery trials can be further broken down into control patients (2), temporal lobe surgery (27), hemispherectomy (2), frontal lobe resection (3), multiple subpial transection (2), and corpus callosotomy (3). As can be seen in Evidence Table 5, only nine of 82 studies (11 percent) reported whether AEDs were given until the maximum tolerated dose was reached before treatment was considered a failure. The majority of these studies (six) were drug studies that were not performed in order to obtain FDA approval. However, six studies is still a minority (35 percent) of the 17 non-FDA drug studies meeting the inclusion criteria. Only 13 studies (16 percent) required a minimum duration of illness before patients were considered treatment-resistant. We considered this number too small for a meaningful analysis of whether different types of studies required different durations of illness.

In contrast, 44 studies (54 percent) required patients to have tried a minimum number of AEDs before being considered treatment-resistant. In Figures 2 and 3, we examine whether a study of a particular type of treatment, or patient has a different requirement for the minimum number of AEDs compared to other studies. To be included in this summary, a subgroup of studies had to include at least five studies reporting such a requirement.

In those studies that reported a minimum number of AEDs, the majority required at least one AED. This proportion did not differ dramatically from the proportion in studies of pediatric patients or studies in which there was no special patient group.

All FDA drug studies and most non-FDA drug studies reported that a minimum number of AEDs must have been tried without success before a patient was considered treatment-resistant. In both cases, the minimum number was nearly always one. Most studies of surgery (80 percent) did not report a minimum number of AEDs that had been tried. However, when a number was reported, it was always greater than one. This difference between drug and surgical trials probably reflects differences in trial qualifications rather than differences in definitions of treatment resistance.

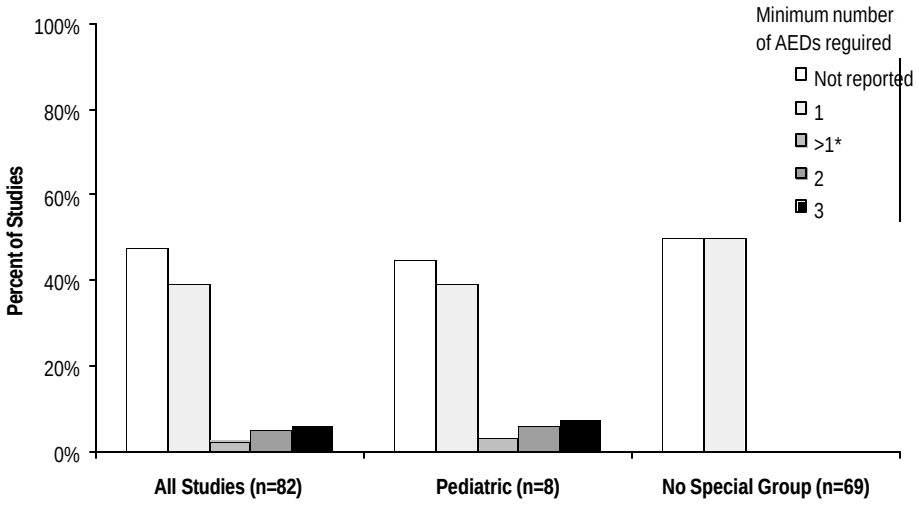
Nearly half (49 percent) of the studies reported a minimum seizure frequency before patients were considered treatment-resistant. This number ranged from less than 1 per month to 60 per month. Some studies (2.4 percent) were not specific about the precise number required, reporting only that seizures were “frequent” or some equivalent term. Pediatric studies differed from studies in which no special group was examined (Figure 4) in that a higher proportion of studies required a minimum seizure frequency (75 percent, as opposed to 49 percent) and the required seizure frequency tended to be lower. Among pediatric studies, 38 percent required a minimum seizure frequency of less than two per month, while among studies of no special group, 36 percent required a minimum of two to five.

When studies of different treatments are compared, studies of surgery seldom (8 percent) reported a minimum seizure frequency (Figure 5), while drug studies conducted to obtain FDA approval always reported these data.

Table 3. Definitions of treatment resistance

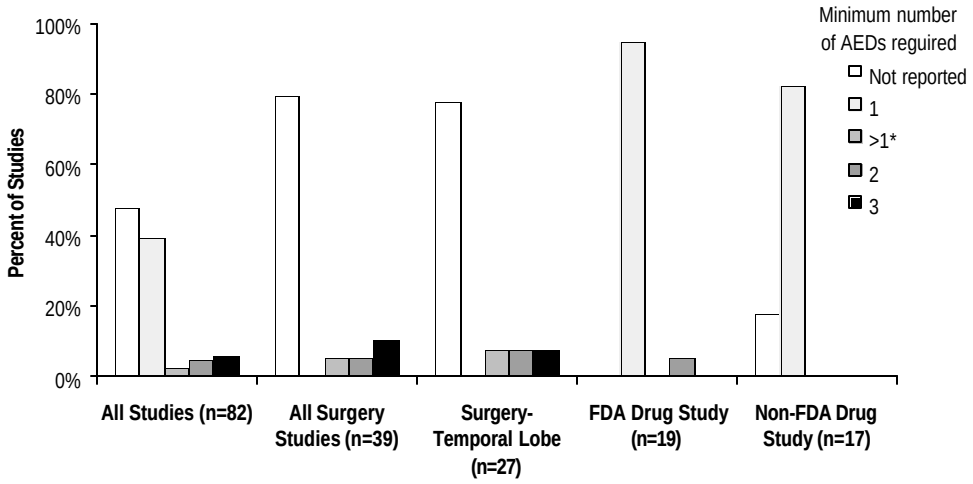
Source	Number of Publications Selected for Question	Number of Publications Reporting Definitions	Percentage of Publications Reporting Definitions
Question 2	6	3	50%
Question 3	10	4	40%
Question 4	26	7	27%
Question 5	39	9	23%
Question 6	1	1	100%
Research Articles Total	**82**	**24**	**29%**
Treatment Guidelines	3	1	33%
Review Articles	100	21	21%
Grand Total	**185**	**46**	**25%**

Figure 2. Minimum number of AEDs: different patient types



* ">1" indicates that the study used a nonspecific term such as "several"

Figure 3. Minimum number of AEDs: different treatments



* ">1" indicates that the study used a nonspecific term such as "several"

Figure 4. Minimum baseline seizure frequency: different patient types

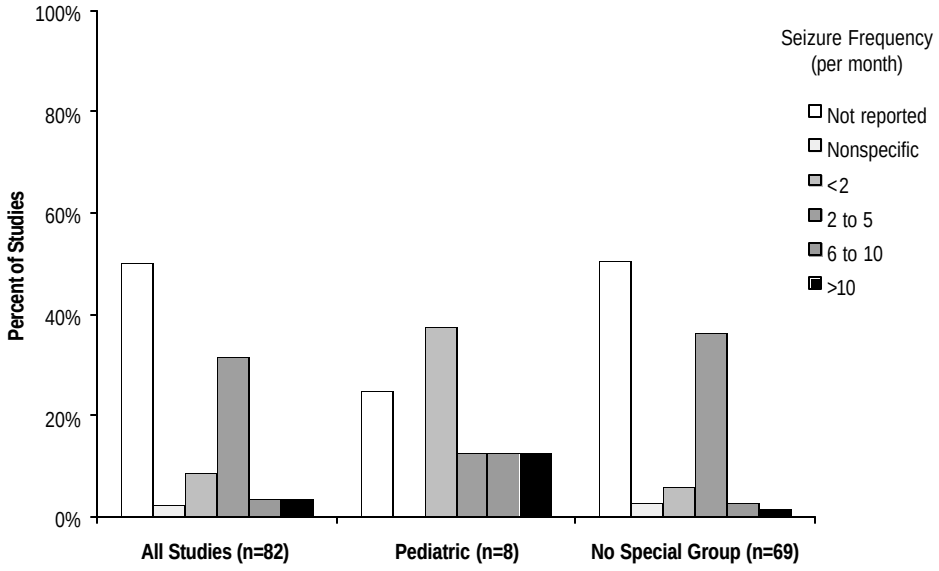
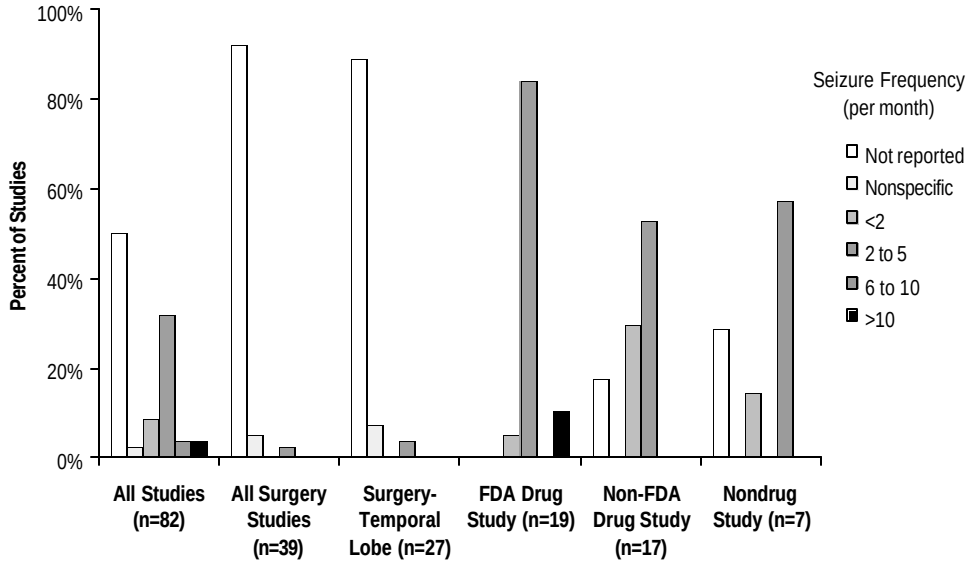


Figure 5. Minimum baseline se izure frequency: different treatments



Rediagnosing and Reevaluating Treatment-Resistant Epilepsy

In this section of the Evidence Report, we addressed Key Question #2: *Which methods of rediagnosing or reevaluating treatment-resistant epilepsy lead to, or can be expected to lead to improved patient outcomes?*

There are three primary roles for diagnostics in the management of patients with epilepsy. The first is to determine whether the patient is experiencing epileptic or nonepileptic seizures. Once a firm diagnosis of epilepsy has been established, the second role is to aid in the classification of epileptic seizures into seizure type. The third role of a diagnostic in the management of patients with epilepsy is to aid in the lateralization and localization of epileptic foci prior to epilepsy surgery. In this section of the report, we address the first two of these roles and how they apply to the subpopulation of patients with treatment-resistant epilepsy. We did not address the third role, which was in keeping with the desires of the Technical Experts and the Expert Panel.

We partitioned Question 2 into four subquestions (A – D). The first two subquestions address the differential diagnosis of epileptic seizures from nonepileptic seizures. The remaining two subquestions address the differential diagnosis of different seizures types. Whether we addressed some questions depended on the findings from previous questions.

We examined the peer-reviewed literature to determine whether there was evidence to suggest that some patients with a diagnosis of treatment-resistant epilepsy were misdiagnosed and their seizures were either not epileptic or they consisted of a combination of epileptic and nonepileptic seizures (Question 2A). If such evidence was found, we then examined the literature to determine which diagnostic technologies were likely to aid in the differential diagnosis of epileptic seizures from nonepileptic seizures (Question 2B).

Similarly, we examined the peer-reviewed literature to determine whether there was evidence to suggest that some patients with treatment-resistant epilepsy were diagnosed with an incorrect seizure type (Question 2C). If such evidence was found, we then examined the literature to determine which diagnostic technologies were likely to aid in the differential diagnosis of one seizure type from another (Question 2D).

Do all patients diagnosed with epilepsy that is deemed to be treatment-resistant truly have epilepsy?

To address Question 2A, we looked for studies that attempted to estimate the prevalence of patients with nonepileptic seizures among populations of patients with a diagnosis of treatment-resistant epilepsy. These nonepileptic seizures may have been the sole seizure type experienced by a patient (in which case the patient was misdiagnosed), or they may have occurred in addition to true epileptic seizures (in which case the patient was correctly diagnosed with epilepsy but the additional diagnosis describing the nonepileptic seizures was missed). In the former case, patients would not be expected to respond satisfactorily to treatment with AEDs. In the latter case, the epileptic seizures may be well controlled by AEDs, and the seizures experienced by the patient are nonepileptic in nature. In either case, such patients would, unless given a new diagnosis, remain incorrectly labeled as exclusively having treatment-resistant epilepsy.

Question specific inclusion criteria

Articles were included for Question 2A if they met the general criteria for inclusion presented in the Methodology section and the article reported that patients originally diagnosed as having epilepsy at the time of enrollment into the study were considered treatment-resistant. Studies that enrolled patients with known nonepileptic seizures (either alone or in combination with epileptic seizures), in addition to patients considered to have treatment-resistant epilepsy alone, cannot be used to answer Question 2A. Consequently, such studies were not considered for inclusion in this section of the report unless data from these patients were presented separately. We did not exclude studies that enrolled patients with a diagnosis of treatment-resistant epilepsy but who were suspected of having nonepileptic seizures. This is because all patients who were enrolled in such studies did have a diagnosis of treatment-resistant epilepsy on entry into the study and such studies do contain information on the accuracy of the original diagnosis of epilepsy.

Excluded studies

We did not exclude any of the articles that met both the general criteria for inclusion in this report and the question-specific inclusion criterion for reasons related to poor quality.

Evidence base

Five articles met both the general inclusion and the question-specific inclusion criterion presented above. These five articles are listed in Table 4. Details of these studies are presented in Evidence Tables 6 through 9.

All five articles in the evidence base were cross-sectional, case series. In these studies, a series of patients (total N = 744) were given diagnostic reassessment in order to determine the prevalence of patients with nonepileptic seizures among specific subgroups of patients, all of who were considered, prior to reassessment, to have treatment-resistant epilepsy.

Four of the five articles included in Table 4 described studies that were carried out at a single center. The remaining article described a study in which patients were recruited at two different centers. However, all patients in this latter study had their diagnosis reassessed at a single study center by a single diagnostic team.

Design and conduct of included studies

The following section presents the findings of our systematic assessment of the quality of the evidence base on the prevalence of patients with nonepileptic seizures (alone or in combination with epileptic seizures) among patients with a diagnosis of treatment-resistant epilepsy. This systematic assessment consists of an appraisal of each study's internal and external validity.

Internal validity

The internal validity of a study designed to measure the prevalence of some disease in a population of interest can be weakened by a number of potential biases. Sampling bias is not a concern in these studies because patients were consecutively enrolled during a fixed period. Reference standard bias is a concern in all of the studies because at present no stand alone "gold-standard" for diagnosing epilepsy is available for routine use in clinical practice. Thus, in practice, the differential diagnosis of epileptic seizures is based on a clinical judgment made by one or more specialists. This judgment is based on information from many sources. These sources include medical history, routine-EEG, ambulatory EEG, video EEG, imaging data, cardiac monitoring data, etc.

The potential biases in each study included in the evidence base for this question are discussed in greater detail in Appendix B.

External validity

The generalizability of a study's results were evaluated by examining the study's inclusion and exclusion criteria, and by evaluating the characteristics of the patients actually enrolled in the study. Details of the inclusion/exclusion criteria used by each of the relevant studies, along with the characteristics of the patients actually recruited by these studies, are presented in Evidence Tables 7 and 8.

The ability to draw conclusions about the generalizability of the studies addressing this question is limited because details on patient characteristics were incompletely reported. All five of the studies included in the present evidence base were carried out at specialist referral centers (three were specialist electrophysiology centers, one was a specialist neurosurgery center and one was a specialist epilepsy center). Such patients are unlikely to be representative of the general population of patients with treatment-resistant epilepsy. In addition, none of the five studies included children less than sixteen years of age. Therefore, the prevalence data extracted from the studies included in the present evidence base may not be generalizable to pediatric populations.

In four of the studies, some of the patients were referred to the specialist center for a diagnostic reassessment because their original seizure diagnosis was deemed questionable. Estimates of the prevalence of patients with an incorrect diagnosis based on data collected from these studies are likely to lead to an overestimate of the true extent of the misdiagnosis problem as it occurs in the more general population of patients with treatment-resistant epilepsy. In the remaining study, the study sample consisted of patients who were all considered candidates for epilepsy surgery. Not all patients with treatment-resistant epilepsy are surgical candidates and, thus, the findings of this study can only be generalized to a very select population of patients.

Synthesis of study results

The prevalence of nonepileptic seizures among the patients in each of the five studies used to address Question 2A are presented in Evidence Table 9 and are summarized in Figure 6. This figure demonstrates that between 8.3 percent and 37.6 percent of patients believed to have treatment-resistant epilepsy turned out to either not have epilepsy or to suffer from a combination of both epileptic and nonepileptic seizures. With the exception of the study of Henry and Drury,[37] the majority of these patients suffered from nonepileptic seizures alone. Only a small proportion of patients had a combination of both epileptic and nonepileptic seizures (Range: 0 percent to 1.0 percent).

The patients examined in the study of Henry and Drury[37] were undergoing presurgical evaluation. Such patients were probably assessed more often and/or more completely by clinicians who specialize in epilepsy compared to most other patients with a diagnosis of treatment-resistant epilepsy. Consequently, that no patients in their sample suffered from nonepileptic seizures alone is not a surprise. However, that 8.3 percent of the patients in this study experienced a combination of epileptic and nonepileptic seizures is surprising. Followup of these patients revealed that the changes in patient management that resulted from the reassessment led to a complete cessation of seizures in three patients (25 percent). Whether these three patients would have been identified prior to surgery had the study not been performed is unknown. However, they may possibly have undergone unnecessary surgery.

To determine an overall estimate of the prevalence of patients with nonepileptic seizures among patients who, prior to re-evaluation, were considered to have treatment-resistant epilepsy, we performed a meta-analysis. This meta-analysis did not include prevalence data abstracted from the study of Henry and Drury[37] for reasons explained above. The results of this homogenous (Q = 0.28; p = 0.96439) fixed-effects meta-analysis are presented in Evidence Table 10.

This meta-analysis shows that the proportion of patients who were misdiagnosed as having treatment-resistant epilepsy was substantial (35 percent; CI: 29 percent to 41 percent). Therefore, a problem of misdiagnosis clearly exists in clinical practice. However, these findings do not accurately represent the proportion of misdiagnosed patients in the overall population of patients with treatment-resistant epilepsy because none of the studies included in the meta-analysis were population-based. Also, the patients included in the four studies that we did meta-analyze represent a subpopulation of patients referred for specialist evaluation of their seizures.

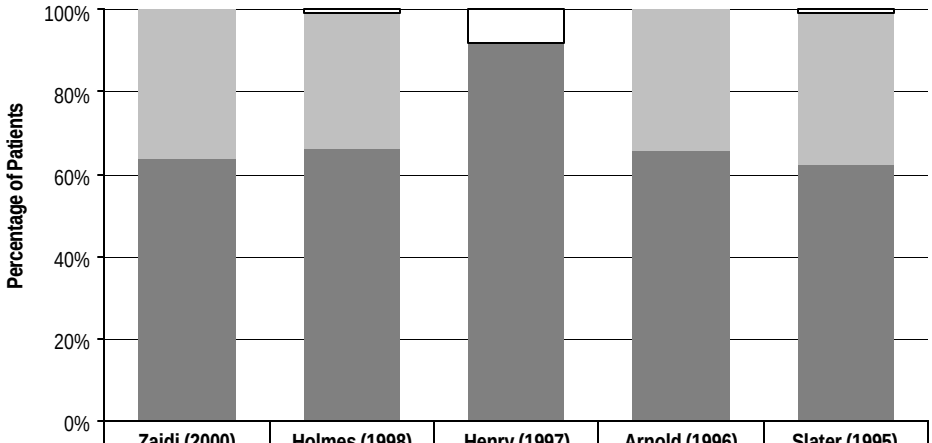
Table 4. Evidence base for determining if patients diagnosed with treatment-resistant epilepsy actually have epilepsy

Reference	Study Design	Country in Which Study Performed	Size (N)	Multicenter	Number of Centers
Zaidi (2000)[38]	Cross-sectional case series	United Kingdom	74	Yes	2[a]
Holmes (1998)[39]	Cross-sectional case series	United States	379	No	1
Henry (1997)[37]	Cross-sectional case series	United States	145	No	1
Arnold (1996)[40]	Cross-sectional case series	United States	45	No	1
Slater (1995)[41]	Cross-sectional case series	United States	101	No	1

[a] Patients enrolled at two centers but diagnostic reassessment was performed at a single study center

Figure 6. Prevalence of nonepileptic seizures

Prevalence of nonepileptic seizures among patients diagnosed with treatment-resistant epilepsy



	Zaidi (2000)	Holmes (1998)	Henry (1997)	Arnold (1996)	Slater (1995)
NES plus ES	0	3	12	0	1
NES Only	13	111	0	14	37
ES Only	23	224	133	27	63

NES: Non-epileptic seizure
ES: Epileptic seizure

Which diagnostic modalities are useful in differentiating seizure types commonly mistaken for epilepsy from true epileptic seizures?

Based on the results of Question 2A, we addressed Question 2B by evaluating the evidence on the diagnostic technologies most commonly used to differentiate epileptic seizures from nonepileptic seizures. As stated above, in clinical practice, the differential diagnosis of epileptic seizures is usually based on information from many sources (medical history, etc.). The clinical diagnosis is seldom based on one diagnostic technology alone. Ultimately then, to answer this question, diagnostic performance data from each analysis of each individual diagnostic technology must be combined into a single decision model which better describes the true clinical picture. As will be seen, a paucity of available evidence precluded the construction of such a model. Thus, we were limited to an analysis of the clinical utility of individual "stand alone" diagnostic technologies.

Question specific inclusion criteria

In addition to employing the general inclusion criteria, we included articles if they met the following criteria:

1. The study must have evaluated the effectiveness of a diagnostic technology used for the differential diagnosis of epileptic seizures from nonepileptic seizures.
2. The patients enrolled in the study were not restricted to only those with treatment-resistant epilepsy. Because the intent of Question 2B is to determine the utility of those diagnostic technologies that have been used to differentiate epileptic seizures from nonepileptic seizures, addressing this question requires a study that enrolls both kinds of patients.
3. The study must have either reported diagnostic test performance characteristics (e.g., sensitivity and specificity) or presented data in a format that allows calculation of test performance characteristics based on a comparison with some "reference" standard.[b] Alternatively, the study must have included followup data that allow conclusions about the effects of using a diagnostic on patient outcomes.

<u>Number of articles addressing each diagnostic</u>

Forty-three articles met the inclusion criteria for Question 2B. The numbers of articles that address each of the diagnostics meeting the inclusion criteria are presented in Table 5. A full list of articles and the diagnostics that they addressed are presented in Evidence Table 11.

The most common type of excluded article reported on a case-series study in which a group of patients was diagnosed with a given modality, and this diagnosis was then used to influence medical management. However, none of these studies reported whether these management changes led to improvements in patient outcomes. Although this study design is seen by some as being a legitimate design for the assessment of a diagnostic,[42] this assumes that the diagnostic test was accurate. Requiring the assumption that a diagnostic be accurate in order to assess the accuracy of that same diagnostic is circular reasoning. It also assumes perfect sensitivity and specificity of the test, which is not possible. This sort of study design is particularly common in the literature on EEG technologies (i.e. routine EEG, ambulatory-EEG, and video-EEG).

[b] There is no practical "gold" standard for the diagnosis of epilepsy, which remains a clinical diagnosis. Thus, we use the term "reference" standard instead of gold standard.

Of the fourteen diagnostics considered, only four (blood prolactin levels, the MMPI, video-EEG, and ambulatory-EEG) were addressed by five or more studies. As per the general inclusion criteria specified in the Methodology section, data about diagnostics not addressed by at least five studies are not considered further in this report. Consequently, we do not include further information about provocation techniques, routine EEG, creatinine kinase levels, tilt tables, auditory evoked potentials, hypnotic recall, MRI, SPECT, tongue biting, or CT in this report.

Blood Prolactin Level Monitoring

Interest in the use of blood prolactin levels as a diagnostic tool for the differentiation of epileptic seizures from nonepileptic seizures began in 1976 when Ohman, Walinder, Balldin et al.[43] reported that blood prolactin levels rose following epileptic seizures induced by electroconvulsive therapy. Blood levels of prolactin were found to peak approximately 30 to 40 minutes after the seizure occurred and then decline to normal preseizure levels. These findings were confirmed by Trimble[44] who demonstrated that blood prolactin levels increased following spontaneous epileptic seizures and noted that blood prolactin levels did not rise following a nonepileptic, psychogenic seizure (also known as a pseudoseizure, hysterical seizure, or psychological seizure). Since then, a number of reports have been published that have assessed the relationship between blood or plasma prolactin levels in patients with epilepsy and a number of different nonepileptic seizure types.[44-57] In the following section of the report, we assess the evidence related to the value of measuring blood prolactin levels in differentiating epileptic seizures from nonepileptic seizures.

Excluded articles

We excluded three studies for reasons of quality. These studies, and the reasons for which they were excluded, are listed in Evidence Table 12.

Evidence base

Following the exclusion of the articles, five articles describing five separate studies that enrolled 305 patients remained. Details on each study (study design characteristics, patient characteristics, and study results) are presented in Evidence Tables 13 through 21.

Design and conduct of included studies

The following section presents the findings of our systematic assessment of the quality of the evidence on the diagnostic utility of blood prolactin level measurements in the differentiation of epileptic seizures from nonepileptic seizures.

Internal validity

All five studies included in the present evidence base utilized a diagnostic case-control study design. The case-control study design is commonly used in the early stages of the evaluation of a diagnostic and is particularly susceptible to a number of biases that lead to overestimation of a test's true diagnostic performance.[24-28] No studies presented patient outcome data. Thus, no direct determination is possible about whether the use of blood prolactin level measurements will lead to improvements in patient outcome. However, a reasonable assumption is that a good diagnostic test will allow patients' nonepileptic seizures to be identified and treated more appropriately, thus leading to improved patient outcomes.

Imperfect reference standard bias (all five studies), prevalence bias, and spectrum bias (four studies each) were the most common potential biases in the studies of blood prolactin measurements. Patient bias, diagnostic yield bias, and verification bias were not present in these studies. These potential biases with respect to this question are discussed in detail in Appendix B.

External validity

Complete details of the inclusion/exclusion criteria used by each of the studies that comprise the present evidence base, along with the characteristics of the patients actually recruited by these studies are presented in Evidence Tables 17 and 18.

Details of both study inclusion/exclusion criteria and the patient characteristics included in the relevant studies were incompletely reported. Four of the five articles described the inclusion/exclusion criteria. Four articles reported on age, no article reported on sex distribution, one article reported on the duration of disease, no article reported on seizure frequency, only one article reported the number of patients who had cognitive or developmental deficits, and only one article reported the number of AEDs used by patients in the study.

Synthesis of study results

The assessment of study quality presented above indicates that, given the present evidence, definitive conclusions cannot be drawn about whether blood prolactin level measurements have a useful role in differentiating epileptic seizures from nonepileptic seizures. Acknowledging this, we have instead evaluated the available data with the aim of determining the plausibility of blood prolactin measurements having a role in differentiating epileptic seizures from nonepileptic seizures.

Not all of the studies in the present evidence base evaluated the ability of blood prolactin level measurements to differentiate epileptic seizures from the same type of nonepileptic seizure. The specific differential diagnoses assessed by each of the studies included in this section of the report are presented in Table 6.

These studies primarily assessed the ability of blood prolactin level measurements to differentiate several epileptic seizure types (mixed seizures, generalized tonic-clonic seizures, complex partial seizures, and simple partial seizures) from two paroxysmal seizure disorders that are often misdiagnosed as epileptic. These two nonepileptic seizure types were syncopal seizures and psychogenic seizures. Thus, the findings of this assessment are not applicable to the differentiation of epileptic seizures from any other nonepileptic seizure type.

Differentiating epileptic seizures from syncopal seizures

This section summarizes the findings of the studies that reported on the diagnostic utility of blood prolactin level measurement in differentiating epileptic seizures from syncopal seizures. Three of the five studies (Anzola,[45] Lusic, Pintaric, Hozo, et al.,[53] and Zelnik, Kahana, Rafael, et al.[56]) presented data on this differentiation. Of these, two (Lusic, Pintaric, Hozo, et al.[53] and Anzola[45]) presented dichotomous diagnostic performance data that allowed sensitivity, specificity, PPV and NPV of the test at a predetermined threshold to be directly determined. These terms are defined in Evidence Table 19. Data from these studies are presented in Evidence Table 20.

Typically, diagnostic performance data is captured by Receiver Operator Characteristic (ROC) curves that describe the trade off between sensitivity and specificity. If enough studies report appropriate data, the data can be meta-analytically combined into a single summary ROC

(SROC) curve. Because the present data set consisted of only three studies, we have not attempted such a meta-analysis. As mentioned earlier, Zelnik, Kahana, Rafael, et al.[56] did not present their diagnostic performance data in a typical 2 by 2 format. Instead, they summarized their data in the form of mean (and standard deviation) blood prolactin levels.

Because the present data could not be meta-analyzed, and because ROC curves synthesized from the available continuous data sets may be misleading, we have instead summarized these data as the effect size, Hedges' d. Although this effect size cannot be used to describe the diagnostic performance of blood prolactin measurements, it does allow diagnostic data to be compared and contrasted among studies that reported this data in different formats.

The data from the three studies reporting the diagnostic utility of blood prolactin level measurement in differentiating epileptic seizures from syncopal seizures is summarized in Figure 7. The effect sizes with confidence intervals that overlap zero indicate that the diagnostic test did not discriminate epileptic seizures from syncopal seizures any better than chance. Thus, data from Lusic, Pintaric, Hozo, et al.[53] did not show that blood prolactin level measurements were useful in differentiating epileptic seizures from nonepileptic seizures. The studies by Anzola[45] and Zelnik, Kahana, Rafael, et al.,[56] however, both found that the test did discriminate these two seizure types from one another statistically significantly better than chance.

Exploration suggests that Lusic, Pintaric, Hozo, et al.[53] did not find that the test was significantly better than chance because the performance characteristics data were collected at a threshold that was not optimal for the test. This is illustrated in Table 5, which shows three point estimates plotted in ROC space. These point estimates, along with their confidence intervals, came from the dichotomous data presented by Lusic, Pintaric, Hozo, et al.[53] and Anzola[45] (Evidence Table 20). All three point estimates can conceivably originate from a single underlying ROC curve (Figure 8).[c] However, because Lusic, Pintaric, Hozo, et al.[53] chose to use a lower threshold compared to any of the thresholds used by Anzola[45], the point estimate falls nearer the chance line. Thus, given the available data, blood prolactin measurements may plausibly provide information that aids in differentiating epileptic seizures from syncopal seizures. Further data are required, however, before stating that this test performs well enough to be used in actual clinical practice.

Differentiating epileptic seizures from psychogenic seizures

Two of the five included studies attempted to use blood prolactin levels to differentiate epileptic seizures from psychogenic seizures. These data are presented in Evidence Table 20 and 21. Wroe, Henlet, John et al.[50] presented dichotomous diagnostic performance data that allowed the sensitivity, specificity, PPV and NPV of the test at a predetermined threshold to be directly determined. Mishra, Gahlaut, and Kumar[57] presented summary statistics (means, standard deviations, etc.) that describe the distributions of blood prolactin levels in the two diagnostic groups.

We summarized the available data from these two studies in terms of Hedges' d. This summary, shown in Figure 9, suggests that blood prolactin measures can plausibly provide information that aids in differentiating epileptic seizures from psychogenic seizures.

In addition to the data presented in Figure 9, Mishra, Gahlaut, and Kumar[57] also presented data about the effectiveness of blood prolactin level measurement in differentiating three types of

[c] The ROC curve plotted in Figure 8 should not be confused with a summary ROC curve. Construction of a summary ROC curve requires that each data point contained within it be independent. Because two of the data points contained within Figure 8 originated from a single group of patients in a single study (Anzola[45]), the data contained within the ROC are not independent. Consequently, this ROC cannot be considered to be a summary ROC.

epileptic seizures (generalized tonic-clonic seizures; complex partial seizures; and simple partial seizures) from psychogenic seizures. These data, presented in Figure 10, suggest that while blood prolactin level measurements may be of some use in differentiating generalized tonic-clonic seizures and complex partial seizures from psychogenic seizures, the test appears to have little or no value in differentiating patients with simple partial seizures from those with psychogenic seizures. The data from Mishra, Gahlaut, and Kumar[57] show that blood levels of prolactin do not increase following simple partial seizures. Thus, blood prolactin levels will probably have little value in differentiating simple partial epileptic seizures from syncopal seizures as well.

Minnesota Multiphasic Personality Inventory

The Minnesota Multiphasic Personality Inventory-2 (MMPI-2) and its predecessor (MMPI) instruments are among the most widely used and widely researched tests of adult psychopathology. These instruments provide a broad psychological profile across a number of domains. Some investigators believe that the psychological profile of a patient with psychogenic seizures (either alone or in combination with epileptic seizures) may be different from that of patients with epileptic seizures alone.

Excluded studies

Not all of the articles that met the general and subquestion specific inclusion criteria were included in the evidence base for this diagnostic. We list the studies that were excluded for reasons of quality in Evidence Table 22, along with an explanation as to why they were excluded.

After the exclusion of the two articles listed in the table, four studies remained. As per the general inclusion criteria specified in the Methodology section, data about treatments not addressed by at least five included studies (or at least one large RCT with 50 or more patients in each study arm) are not considered further. Consequently, we do not further assess the MMPI-2 or MMPI in this report.

Video-EEG

Video-EEG monitoring is used in clinical practice to verify the seizure type, to localize the area of seizure onset if surgery is being considered, and to verify the diagnosis of epilepsy if the diagnosis is in doubt. Video-EEG monitoring consists of the simultaneous recording of EEG brain wave activity combined with time synchronized video recording of the patient. This diagnostic procedure is performed on an in-patient basis, and requires highly specialized equipment and dedicated space. Patients are monitored for extended periods in order to capture typical seizure events on video and simultaneously capture EEG activity during that event. In some centers, patients' medications are withdrawn in order to increase the chance of recording a seizure.

Excluded studies

We list the studies that were excluded for reasons of quality in Evidence Table 23, along with an explanation as to why they were excluded.

After the exclusion of the two articles listed in the table, four studies remained. As per the general inclusion criteria specified in the Methodology section, data about diagnostics not addressed by at least five included studies (or at least one large RCT with 50 or more patients in

each study arm) are not considered. Consequently, we do not further consider video-EEG in this report.

Ambulatory-EEG

Ambulatory-EEG monitoring is the recording of EEG brain wave activity remotely from the hospital environment. Ambulatory-EEG, like video-EEG, has the advantage over routine EEG of allowing EEG traces to be recorded continuously over long periods. This increases the chance of recording an ictal event. Unlike video-EEG, no video record of seizure events is available, and patients or their caregivers must accurately record the occurrence of a typical seizure event in order to temporally compare the occurrence of a seizure with the EEG trace.

Early clinical investigations documented the ability of ambulatory-EEG to record identifiable focal and generalized epileptiform activity. In 1983, a cassette tape ambulatory-EEG system was introduced. This system had continuous 8-channel recording capability, real-time identification, gain adjustment, and filter adjustments. Since then, improvements in computer technology have led to the development of instruments that can perform portable continuous recording of more than 16 channels with sampling rates of over 200 Hz.

Excluded studies

Not all of the articles that met the general and subquestion specific inclusion criteria were included in the evidence base for this diagnostic. We list the studies that were excluded for reasons of quality in Evidence Table 24 along with an explanation as to why they were excluded.

After the exclusion of the two articles listed in the table, three studies remained. As per the general inclusion criteria specified in the Methodology section, data about treatments not addressed by at least five included studies (or at least one RCT with more than 50 patients in each study arm) are not considered. Consequently, we do not further consider ambulatory-EEG in this report.

Comments on EEG Technologies

That the evidence base for the three EEG technologies (routine ictal and interictal EEG, ambulatory-EEG and video-EEG) did not reach the required minimum of five acceptable studies is surprising. Even with the expansion of the inclusion criteria so that we included articles published between 1980 and 1985 and articles describing retrospective studies, we did not reach five studies. This may be because all three of these EEG technologies are commonly used as aids in the diagnosis of epilepsy, and video-EEG is considered by many to be the "gold standard" for the differentiation of epileptic seizures from nonepileptic seizures.[58] Some reviews of the use of video-EEG in the differential diagnosis of epileptic seizures from nonepileptic seizures include studies in which provocation was used in an attempt to induce a seizure that was then captured by video-EEG, citing such studies as evidence of the effectiveness of video-EEG in combination with provocation. Without exception, however, these studies used video-EEG as a "reference standard" against which the effectiveness of provocation was measured. Thus, such studies cannot be considered as part of the evidence base for the diagnostic utility of video-EEG and instead form the evidence base for seizure provocation techniques.[59]

Two previous technology assessments looked at the clinical utility of video-EEG and addressed much the same issue being addressed in this report by subquestion 2B.[60,61] These technology assessments and the relevant references mentioned in each assessment are presented

in Evidence Table 25. Also in this Evidence Table is an indication as to whether each article was included in the current report and, if the article was not included, an explanation as to why.

Evidence Table 25 shows that none of the articles included in the previous two technology assessments met the inclusion criteria for the current report. The primary reason for not being included in the present report was that the studies utilized a case series design in which a group of patients were evaluated with video-EEG and a diagnosis or change in diagnosis was made based on the information gained from the assessment. No reference standards were used against which to compare the effectiveness of video-EEG, nor were patients followed up in order to verify the accuracy of the diagnosis. Thus, the investigators in these studies made the implicit assumption that video-EEG did accurately differentiate epileptic seizures from nonepileptic seizures. In other words, the investigators assumed that false-negative (making an incorrect diagnosis of non-epileptic seizure) and false-positive decisions (making an incorrect diagnosis of epileptic seizures) will not occur when video-EEG is used. Such assumptions, though they may be true for some seizure types[d], do not always hold true. Both assumptions rely on the supposition that an abnormal EEG always accompanies a true epileptic seizure. While this may be true for many seizure manifestations, this is not always the case.

For example, a number of studies of patients with implanted electrodes have demonstrated that epileptic seizures originating in the medial or orbital surface of the frontal lobe, the parietal lobe, or the temporal lobe, often occur in the absence of a measurable EEG abnormality when the EEG is performed using scalp electrodes.[62-65] These types of seizures may arguably be relatively rare. However, given that the appearance of a nonepileptic seizure is often very similar to epileptic seizures originating in the medial or orbital surface of the frontal lobe, the parietal lobe, or the temporal lobe,[66] these are the very patients who are the most likely to be misdiagnosed as having epileptic seizures. Thus, some false-negative decisions must be assumed to occur when video-EEG is used.

The fact that evidence-based conclusions were not drawn in the present report regarding the ability of vEEG to differentiate epileptic seizures from nonepileptic seizures should not be interpreted as evidence that this technology is not effective or useful. Indeed, vEEG may very well have an important role in diagnostic algorithms that are designed to make such a differential diagnosis. Until more high quality studies become available, however, the diagnostic performance characteristics of vEEG and its place in such diagnostic algorithms cannot be determined.

[d] Seizures resembling tonic-clonic convulsions, absence seizures, or complex partial seizures with automatism that are not accompanied by an ictal EEG abnormality can confidently be classified as nonepileptic seizures.

Table 5. Articles addressing each diagnostic

Diagnostic	Number of Articles
Blood Prolactin Levels	8
Minnesota Multiphasic Personality Inventory	6
Video-EEG	6
Ambulatory -EEG	5
Provocation techniques	4
Routine EEG	4
Creatinine kinase levels	3
Tilt table	2
Auditory evoked potentials	1
Hypnotic recall	1
Magnetic Resonance Imaging	1
Single Photon Emission Computed Tomography	1
Tongue biting	1
Computed Tomography	0

Table 6. Differential diagnoses of seizures

	Epileptic Seizure Type					Nonepileptic Seizure Type		
Reference	Mixed ES	GTCS	CPS	SPS		Mixed NES	PsyS	SynS
Lusic (1999)[53]			✓		From			✓
Anzola (1993)[45]		✓			From			✓
Zelnik (1991)[56]		✓			From			✓
Mishra (1990)[57]	✓				From			✓
Wroe (1989)[50]	✓	✓	✓	✓	From		✓	
Lusic (1999)[53]	✓				From		✓	

CPS Complex partial seizure
ES Epileptic seizure
FS Febrile seizure
GTCS Generalized tonic-clonic seizure
NES Nonepileptic seizure
PsyS Psychogenic seizure
SPS Simple partial seizure
SynS Syncopal seizure

Figure 7. Blood prolactin: discrimination between epileptic and syncopal seizures

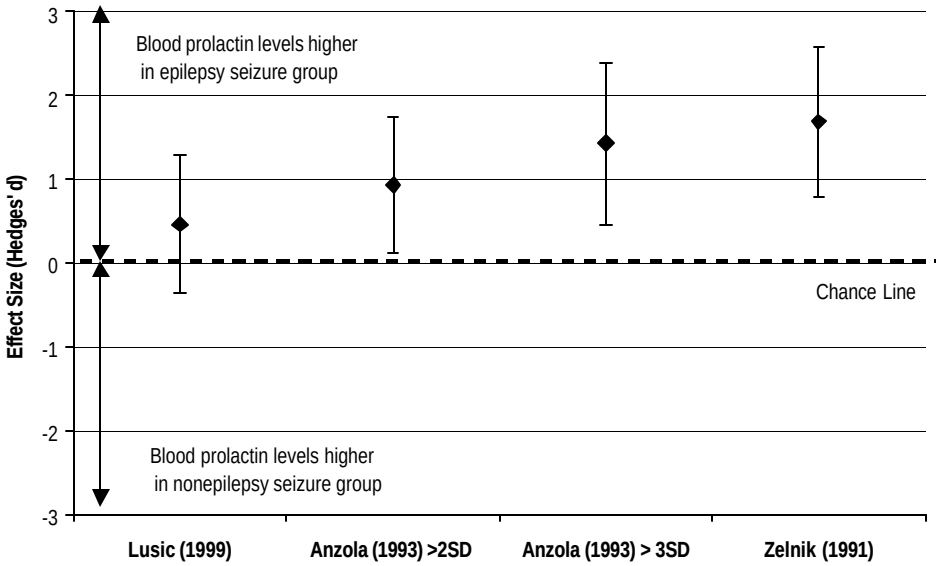


Figure 8. Differences in threshold when evaluating test performance in studies of blood prolactin measurement

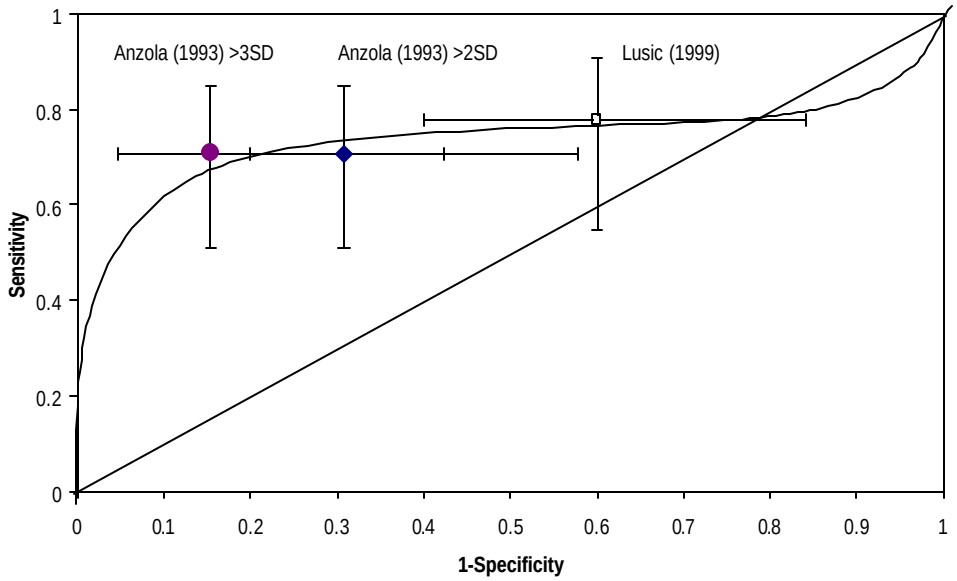


Figure 9. Blood prolactin: discrimination between epileptic and psychogenic seizures

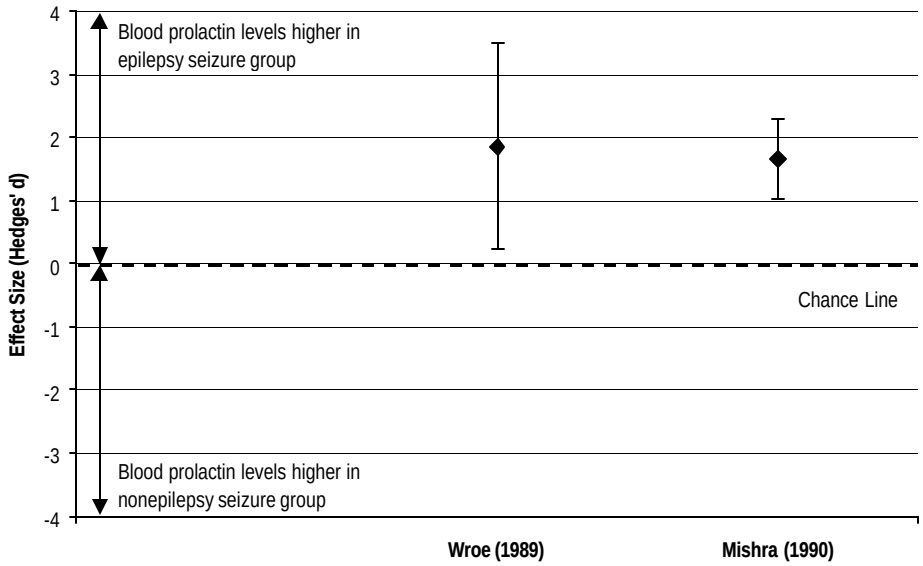
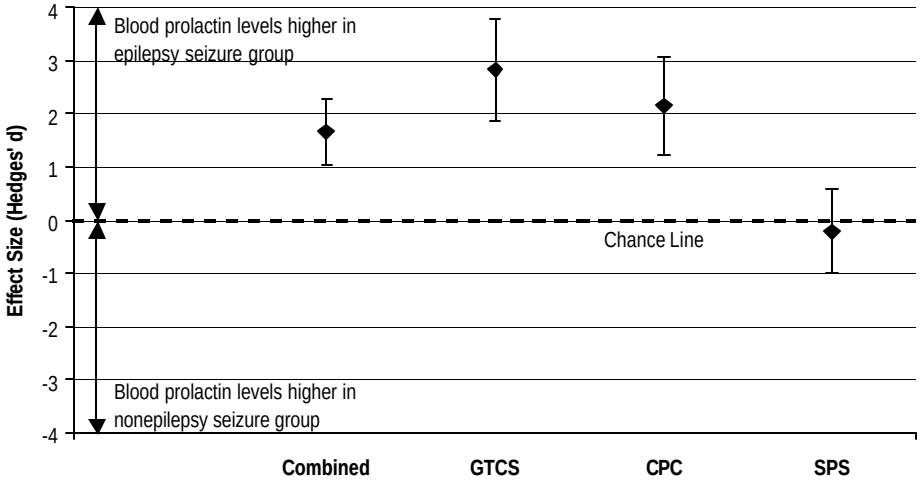


Figure 10. Blood prolactin: discrimination between different epileptic seizure types and psychogenic seizures

Data abstracted from Mishra (1990), GTCS: Generalized tonic-clonic seizures, CPS: Complex partial seizures, SPS: Simple partial seizures



Is seizure type in some patients with treatment-resistant epilepsy misdiagnosed in some patients?

There are two purposes to the present question. First, to establish whether there is evidence in the peer-reviewed literature to indicate that some patients believed to suffer from a specific seizure type actually suffer from a different seizure type (either alone or in combination with the originally diagnosed seizure type) and would, therefore, not be expected to respond satisfactorily to their current treatment regimen. The second purpose is to quantify, if relevant, the prevalence of these patients among the population of patients thought to suffer from a particular seizure type.

Question specific inclusion criteria

Articles were included for Question 2C if they met the general criteria for inclusion presented in the Methodology section, and if the article reported on a study that enrolled patients originally diagnosed as having a specific type of epileptic seizure (partial seizure, generalized seizure, absence seizure, etc).

Evidence base

No studies addressed Question 2C and met both the general and subquestion specific inclusion criteria listed above. Consequently, Question 2C could not be answered.

Which diagnostic modalities are useful in differentiating between different seizure types?

Because Question 2C cannot be answered in an evidence-based fashion, Question 2D could not be addressed in an evidence-based fashion.

Optimization of Antiepileptic Drugs

In this section of the Evidence Report, we addressed Key Question #3: *Is there evidence that patients with treatment-resistant epilepsy are not optimized at their current level of treatment?*

In the present question, we address whether patients described as having treatment-resistant epilepsy are receiving optimal dosages of the AED regimen prescribed for them. The available evidence for this question is derived from two types of studies. The first type is comprised of studies that assessed, using drug level monitoring, whether patients were truly treatment-resistant or at an otherwise optimized level of drug therapy. The second type is comprised of drug treatment studies that presented information on the pretrial or baseline status of the patients enrolled in a clinical trial. Patients in a clinical trial often receive optimized treatment as part of the trial and are therefore not representative of patients who are maintained on AEDs in clinical practice. Thus, the pretrial status of the patients is the best indication of whether drug optimization was part of routine clinical practice. We examined these groups of studies because they were the most likely type of studies to report the information necessary to address this question.

For the purposes of the present question, a drug regimen was defined as not optimized if a study reported enrolling any patients whose prior drug regimen: 1) had not been titrated, 2) was not in the therapeutic range, or 3) produced side effects. If a study reported that some patients had not received the maximal tolerated dosage, we considered this evidence of lack of titration, and therefore definitive evidence of lack of optimization. If the therapeutic range was not defined, we defined the range as either the therapeutic range of the maintenance dose or the blood concentration.[e] We considered patients receiving more than one AED to be in the upper end of the therapeutic range if at least one AED dosage or blood level was in this range. The Expert Panel and the Technical Experts formulated criteria 1 and 3. The second criterion, as originally suggested by the Technical Experts, specified only the upper end of the therapeutic range. This is a more stringent way to define optimization and may be inaccurate, as the maximum tolerable dosage for some patients may be below the upper end of the therapeutic range. We modified this criterion for this reason, and because several studies reported that not all patients were receiving drug doses in the therapeutic range. However, the possibility remains that certain patients outside the therapeutic range may have been optimized. Thus, the second and third criterion suggest the possibility of nonoptimization but do not provide definitive evidence of its existence. Therefore, we separately report patients who were not in the upper end of the therapeutic range

Question specific inclusion criteria

In addition to the general inclusion criteria (see Methodology section), we used the following criteria to determine whether a study was included:

1. The study must have reported information indicating that some patients in the study did not meet at least one of the criteria for optimization described above. Thus, we are seeking only evidence of nonoptimization, not a percentage of patients who are optimized.

[e] Ranges reported in Browne and Holmes.[390]

2. The study must have been published in 1975 or later. This ensured the use of evidence on standard AEDs as well as evidence on newer agents.
3. All drugs in the study must have been cleared for marketing in the United States by the Food and Drug Administration. If the study included some patients on non-FDA approved drugs, it was required to report data separately from patients on FDA-approved drugs, and we abstracted results only from the latter group of patients. This criterion was determined by the Expert Panel and the Technical Experts.

Excluded studies

We did not exclude any studies for reasons of quality.

Evidence base

We included 20 studies, all of which suggested that at least some patients may not have been optimized prior to study enrollment (Evidence Table 26). Six studies were conducted with the goal of assessing medical intractability or lack of optimized therapy through drug level monitoring; the remaining 14 studies were drug treatment studies that presented pretrial or baseline information concerning drug optimization.

Design and conduct of included studies

Internal validity

Since this question does not involve an analysis of results, but merely reporting of patient status before entering a study, there is only one relevant issue concerning the internal validity of these studies. An apparently nonoptimized drug regimen could result not only from an inadequate drug dosage, but also from a patient's lack of compliance with the prescribed regimen. If the nonoptimized patients in a study were actually noncompliant, this would alter the assumption that nonoptimization was primarily due to prescription of nonoptimized drug regimens. Four studies required that all patients in the study were compliant.[67-70] Two studies reported that some patients were suspected of noncompliance,[71,72] while in one study 8 of 35 patients admitted noncompliance (this study was not excluded because clearly other patients in the study were not optimized).[73] The remaining 13 studies reported no information concerning compliance.

External validity

Of the 20 studies mentioned above, seven were conducted in the United States and the remaining 13 were conducted in other countries (Evidence Table 26). Four of seven United States studies and six of 13 studies from other countries evaluated only adult patients. Two United States studies and six studies from other countries evaluated a study group of adult and pediatric patients. One United States study evaluated only pediatric patients, and one study from outside of the United States provided no information on the age range of its patient population.

Synthesis of study results

The summary of results is broken down according to the three criteria described in the introduction to this question. The relevant data for this question are presented in Evidence Table 26.

Did the study report any patients whose prior drug regimen had not been titrated?

Six studies (two United States and four outside the United States) reported information indicating that the prior drug regimen of some patients had not been titrated. The lack of reporting of titration information does not necessarily mean that few patients have their drug regimens titrated in clinical practice. Titration may be a common practice, but study investigators may not report it.

Did the study report any patients whose prior drug regimen was not in the therapeutic range?

Ten of 20 studies (three United States and seven from other countries) presented information indicating that the prior drug regimen was not in the therapeutic range for at least some patients in the studies. We further examined whether some studies enrolled patients whose prior drug regimen was not in the upper end of the therapeutic range.

Sixteen of 20 studies (five United States and 11 from other countries) presented information indicating that the prior drug regimen for some patients was not in the upper end of the therapeutic range.

Did the study report that there were any patients whose prior drug regimen produced side effects?

Four studies (two United States and two outside the United States) presented information indicating that drug side effects occurred in at least some patients on a prior drug regimen.

Drug Treatment Strategies

In this section of the Evidence Report, we addressed Key Question #4: *Which drug treatment strategy, 1) sequential monotherapy, 2) polytherapy, or 3) optimized current therapy leads to improved outcomes for patients with treatment-resistant epilepsy, and what are the relative improvements obtained with each strategy?*

In this question, we address three drug treatment strategies that could potentially benefit patients with treatment-resistant epilepsy. By definition, patients with treatment-resistant epilepsy have already received AEDs that were ineffective. Therefore, in the present question we are addressing whether any *changes* in patients' drug regimens can potentially reduce their seizures.

We define sequential monotherapy as switching patients to a single AED that none of the patients had yet received. According to the desires of the Partners, patients could have been receiving multiple prior drugs before initiation of monotherapy. Polytherapy is defined as the simultaneous administration of more than one AED. It typically involves the addition of a single novel AED to patients' drug regimens (referred to as "add-on" treatment). Finally, we define optimized current therapy as altering the dose of at least one drug in patients' drug regimens, or removing at least one drug from patients' drug regimens. In optimized current therapy, drug dose can be altered by changing the total daily dose, the number of doses in a given day, or the drug preparation (such as a slow-release preparation). According to the desires of the Partners, optimized current therapy can also consist of the removal of a drug from patients' regimens. We address each of these three strategies in separate subquestions.

This question addresses the safety and efficacy of drug strategies, not of particular drugs.[f] However, the literature on monotherapy is comprised primarily of trials that examine the effects of changing patients' treatment from a number of AEDs to a single specific drug (e.g., topiramate). Similarly, the literature on polytherapy is comprised primarily of trials that involve adding a specific drug (that patients had not previously received) to their existing regimen, and the literature on optimized current therapy is comprised primarily of trials that removed a specific drug from patients' drug regimens. Thus, the literature is comprised primarily of certain specific implementations of these strategies. Although the findings of the individual trials have limited generalizability, when considered in aggregate (as below) they provide the best available estimates of the effectiveness of the three strategies.

Question specific inclusion criteria

Although we divided this question into four subsections (one for each treatment strategy, and one for comparisons between strategies), we employed the same inclusion criteria for studies of each strategy. Thus, in addition to the general inclusion criteria described in the Methodology section, we included trials for this question if they met all of the following criteria:

1. The trial must have been published in 1975 or later. For school- and work-related outcomes, the trial must have been published in 1985 or later. Including trials since 1975, as well as more recent trials, facilitated incorporation of data from standard AEDs that may no longer be the focus of clinical research.

[f] We do not compare specific drugs as per the wishes of the Expert Panel and Technical Experts.

2. Before the trial, patients must have received unsuccessful treatment with at least one of the following drugs: carbamazepine, ethosuximide, phenobarbital, phenytoin, primidone, or valproate. The Technical Experts determined that these six drugs are standard AEDs.
3. All drugs received by patients during the trial must be cleared for marketing in the United States by the U.S. Food and Drug Administration (FDA). If the trial included some patients on non-FDA-approved drugs, it was required to report data separately from patients on FDA-approved drugs, and we abstracted results only from the latter group of patients. We included trials employing off-label usage of drugs for the treatment of epilepsy. Confining the question to only FDA-approved drugs was in accordance with the wishes of the Expert Panel and Technical Experts.
4. If a trial reported that some patients had been noncompliant, then results must have been reported separately for patients who were compliant. Noncompliant patients may not be treatment-resistant, and their seizure rates may drop during a trial. Whether the improvement in noncompliant patients was due to better compliance or to the beneficial effect of the trial drug cannot be determined. Therefore, the outcomes of noncompliant patients were not included.
5. If there were five or more placebo-controlled randomized trials on a specific drug treatment strategy, then other trials with other designs (e.g., trials that used a low AED dose as a control) were not considered. We adopted this criterion because results of placebo-controlled randomized trials are more easily interpreted compared to results of trials that employed other control groups.
6. Trial must be a Phase II or III efficacy trial. Earlier trials (Phase I) were not primarily intended to reduce seizures, and later trials (Phase IV) involved drugs whose effectiveness had already been documented by other trials.
7. Trials that used a crossover design must have reported results for the first period (i.e., before the crossover), or must have reported that seizure frequencies returned to baseline at the end of the washout period. In a crossover trial, the use of a drug at the start of a trial may have potentially influenced the effectiveness of a different drug used later in the trial. If seizure frequency returned to baseline at the end of the washout period, then the evidence suggests that the first drug is no longer active, and data for the second drug are interpretable. However, if a return to baseline was not reported, then we only abstracted data for the first period.

To include the maximum number of potentially relevant studies, we did not require studies to report patients' seizure frequencies at baseline. However, baseline seizure frequencies do provide a measure of the severity of patients' initial conditions, thus aiding in the interpretation of study findings. For example, suppose a treatment eliminated all seizures. Such an outcome would be more impressive if patients' baseline seizure frequencies were 20 per month than if frequencies were only five per month. Because the baseline frequency helps place the study results in proper context, ideally all studies would report this frequency.

Number of articles on each intervention

Applying the inclusion criteria yielded 55 studies describing the three drug strategies. There were 14 studies of sequential monotherapy, 30 studies of polytherapy, and 11 studies of optimized current therapy.

Sequential Monotherapy

Sequential monotherapy involves administering a single AED not yet received by any of the patients. Patients can receive any number of AEDs prior to the initiation of the new drug. However, all prior AEDs must be withdrawn from patients' drug regimens in order to investigate the effect of the novel monotherapy drug. In this section, we describe the evidence base for sequential monotherapy, assess the quality of these trials with respect to both internal and external validity, and analyze the trials' results for all relevant outcomes.

Excluded studies

Fourteen studies met the inclusion criteria. One of the 14 studies was excluded because the authors only reported a qualitative description of treatment efficacy.[74]

Evidence base

The evidence base for sequential monotherapy consists of 13 studies that enrolled 1,542 patients.

Design and conduct of included studies

Relevant design aspects of the 13 included studies appear in Evidence Tables 27 through 30. To assess the effect of sequential monotherapy, the ideal study would have randomly assigned patients to receive either a new drug as monotherapy, or to receive the same drug regimens used before the trial. None of the 13 studies employed this design. Twelve studies were randomized and controlled, but patients in the control groups did not receive their prestudy drug regimens. Instead, all prestudy drugs were withdrawn. In three studies, patients in the control groups received a placebo *alone*, and in the other nine studies, patients in the control groups received a low dose of a drug. Because these control groups do not address whether sequential monotherapy causes an improvement over patient's prestudy drug regimens, we did not abstract data from the control groups. Instead, we abstracted data from only the high-dose active-drug group in each of the 12 controlled studies. The 13th study did not have a control group, thus we abstracted data from the single group in that study. Among the 13 studies, eight drugs were given as monotherapy: felbamate (three studies), oxcarbazepine (three studies), gabapentin (two studies), lamotrigine (one study), primidone (one study), tiagabine (one study), topiramate (one study), and valproate (one study) (Table 7).

Internal validity

In evaluating internal validity, we determined whether the results were potentially biased by the threats to validity that are discussed in the Methodology section. Although other questions in this report consider the potential for attrition bias, we do not consider it here because attrition was a study outcome. As discussed earlier, the control groups of these studies are not relevant to the question. Consequently, for the purpose of this report, the studies can be viewed as case series and susceptible to several threats to internal validity (see Appendix B). All were potentially affected by both regression bias and extraneous event bias. Further, most studies were potentially affected by sample specification bias (12/13 studies) and measurement bias (10/11 studies that reported the method of seizure measurement).

External validity

In our appraisal of the external validity of studies of sequential monotherapy, we considered aspects of patient enrollment as well as the actual characteristics of patients in the studies. All patient characteristics appear in Evidence Tables 31 through 34. All 13 studies enrolled patients because of seizure type (partial seizures), thus the results of these studies are not applicable to the treatment of generalized seizures. Three studies enrolled adults only,[75-77] and the remaining 10 studies enrolled both children and adults. The mean age of patients in the studies ranged from 33.4 to 37 years. The proportion of patients who were female ranged from 0.43 to 0.63 and was greater than 0.50 in nine of the 11 studies that reported this characteristic. Median seizure frequency ranged from 5.5 to 13.4 seizures per month, and mean seizure frequency ranged from 6.3 to 70.7 seizures per month. The proportion of patients receiving two or more prior AEDs ranged from 0 to 1. The proportion was less than 0.5 in nine of the 11 studies that reported this patient characteristic. As a whole, then, the results of these studies apply primarily to adults with treatment-resistant epilepsy who experience partial seizures.

Synthesis of study results

In this section, we assess the results separately for each of the following outcomes: seizure frequency, adverse effects, quality of life, mood, cognitive function, functional status/ability, ability to return to work, ability to return to school, ability to hold a driver's license, and mortality. We included freedom from seizures as a seizure frequency outcome. The outcomes reported by each study are listed in Table 8. All outcomes from all of the studies appear in Evidence Tables 35 through 38. Seizure frequency and adverse effects were each reported by all 13 studies.

In cases where meta-analysis was feasible, we used random effects models. We employed these models because the included studies investigated different drugs. Therefore, these studies cannot be viewed as having been sampled from a population of studies with a fixed mean. Random effects models employ statistical methods that are most applicable when studies use different variations of a treatment.

Also, these studies reported data on an intent-to-treat basis. This was particularly important because they employed *a priori* exit criteria (such as doubling of monthly seizure frequency) to limit harms to patients. If any patient met an exit criterion, investigators removed the patient from the study and reinstituted the patient's prior AED regimen. Consequently, the analyses described below included all patients who were randomized to receive high-dose monotherapy.

Seizure frequency

Details of the seizure frequency results are presented in Evidence Table 35. The studies we included for this question reported 14 different measures of seizure frequency (Table 9). Only three seizure frequency outcomes were reported by five or more studies: the percentage of patients who were seizure-free during the study, the percentage of patients whose monthly seizure frequency doubled during the study (vs. baseline), and the percentage of patients whose highest two-day seizure frequency doubled during the study (vs. baseline). We emphasize that, because seizure frequency changes over time, a study's length of followup influences seizure frequency measurements. For example, a given patient is more likely to be seizure-free during a short-term study than a long-term study. Therefore, for each outcome, we considered the length of followup in the studies that reported the outcome.

Percentage of patients who were seizure-free during the study. This outcome was reported by six studies and ranged from 9 percent to 28 percent. Because of the lack of relevant control groups, we performed a threshold analysis (see the Methodology section for a discussion of this approach). In this analysis, we compared the results obtained in patients who received sequential monotherapy to those of a synthetic control group in which we varied the percentage of seizure-free patients. The percentage at which the difference between the monotherapy and "control" group became statistically nonsignificant is the threshold. The results of this analysis appear in Figure 11. Each summary estimate in the figure is based on Cohen's h. The summary estimate calculated at the 0 percent point on the graph (no patients in a synthetic control group were seizure-free) was 0.81 (CI: 0.64 to 0.98, p <0.000001) and corresponded to 16 percent (CI: 10 percent to 22 percent)[g] of patients experiencing no seizures during sequential monotherapy. The summary estimate became nonsignificant (no statistically significant difference between monotherapy and control patients in the number of patients becoming seizure-free) when the proportion of patients in the synthetic control group reached 10 percent.

We next performed a second threshold analysis of these data to test the effect of followup period on seizure-free status. Two of the six studies that reported seizure freedom employed short followup times (8 days in Bergey, Morris, Rosenfeld, et al.,[78] and 10 days in Schacter, Vasquez, Fisher, et al.).[79] In the other four studies, patients were followed for at least 16 weeks. The percentage of patients who were seizure-free was greater than 25 percent in both short-term studies, but was less than 14 percent in all of the long-term studies. Therefore, our second threshold analysis included only the four studies with longer followup times (Figure 12). The summary estimate calculated at the 0 percent point on the graph (no patients in a synthetic control group were seizure-free) was 0.67 (CI: 0.47 to 0.87, p <0.000001) and corresponded to 11 percent (CI: 5 percent to 18 percent) of patients experiencing no seizures during sequential monotherapy. The summary estimate became nonsignificant (no statistically significant difference between monotherapy and control patients in the number of patients becoming seizure-free) when the proportion of patients in the synthetic control group reached 6 percent.

In summary, approximately 11 percent of patients are seizure-free during long-term studies of sequential monotherapy. However, given the designs of these studies, whether the new drug actually *caused* any of the patients to become seizure-free during the study is not clear. Further, seizure frequencies change from month to month, and some patients with treatment-resistant epilepsy may experience periods without seizures, and some patients may have been misdiagnosed (Question 2). These latter patients may be more likely to become seizure-free. Even if some patients become seizure-free as a result of sequential monotherapy, the majority of patients (approximately 89 percent) continue to have seizures despite receiving a new drug as monotherapy. A firm conclusion about whether sequential monotherapy produces any new seizure-free patients would require the use of a relevant control group (i.e., continuation of prior drug regimens).

Percentage of patients whose monthly seizure frequency doubled. Freedom from seizures measures the percentage of patients who experienced maximum benefit. In contrast, seizure doubling indexes the percentage of patients who experienced significant harm. This outcome was reported by five studies and ranged from 9 percent to 29 percent. All five studies followed patients for at least 16 weeks, thus the concern about study duration does not apply to seizure

[g] The estimated percentages for each group were calculated by performing a random-effects meta-analysis in which the synthetic control group event rate was 0 for all trials. The summary Cohen's h from this meta-analysis was back-transformed into a percentage corresponding to the estimated percentage of patients who became seizure-free.

doubling. Due to the lack of relevant control groups, we performed a threshold analysis of this outcome (Figure 13). Each summary estimate in the figure is based on Cohen's h. The summary estimate calculated at the 0 percent point on the graph (no patients in a synthetic control group had a doubling of monthly seizure frequency) was 0.82 (CI: 0.64 to 0.99, p <0.000001) and corresponded to 16 percent (CI: 10 percent to 23 percent) of patients experiencing a doubling of monthly seizure frequency. The summary estimate became nonsignificant (no statistically significant difference between monotherapy and control patients in the number of patients experiencing a doubling of monthly seizure frequency) when the proportion of patients in the synthetic control group reached 10 percent.

Sequential monotherapy cannot be directly considered the cause of the doubling in seizure frequency because of the lack of a true control group. However, three factors do suggest a causal relation. First, at the beginning of the sequential monotherapy studies, all prestudy drugs were removed from patients' regimens and replaced with a new AED. Presumably, the original AEDs were already reducing seizure frequency. The removal of these drugs, therefore, may have caused seizures to increase. Second, a doubling of monthly seizure frequency was set by investigators as an *a priori* exit criterion. A doubling of seizure frequency resulted in immediate removal from the study, and all prestudy drugs were reinstituted. This suggests that investigators believed that the doubling of monthly seizure frequency was being caused by sequential monotherapy. Third, given the possibility that patients enter drug trials when they are relatively sick, they would not be expected to become even worse. Instead, based on regression-to-the-mean, reductions in seizure frequency would be expected. Each of these factors suggest that sequential monotherapy caused dramatic increases in seizures in some patients. A definitive conclusion about this possibility would require randomization of patients to either sequential monotherapy or a continuation of the prestudy drug regimen.

Percentage of patients whose highest two-day seizure frequency doubled. This outcome also measures the percentage of patients who experienced significant harm during studies of sequential monotherapy. It was reported by five studies, and ranged from 4 percent to 23 percent. All five studies followed patients for at least 16 weeks. Due to the lack of relevant control groups, we performed a threshold analysis of this outcome (Figure 14). Each summary estimate in the figure is based on Cohen's h. The summary estimate calculated at the 0 percent point on the graph (no patients in a synthetic control group had a doubling of two-day seizure frequency) was 0.76 (CI: 0.56 to 0.96, p <0.000001) and corresponded to 14 percent (CI: 8 percent to 21 percent) of patients experiencing a doubling of two-day seizure frequency. The summary estimate became nonsignificant (no statistically significant difference between monotherapy and control patients in the number of patients experiencing a doubling of two-day seizure frequency) when the proportion of patients in the synthetic control group reached 8 percent.

As stated previously, whether sequential monotherapy was the cause of doubling of two-day seizure frequency cannot be determined without a true control group. However, the same three factors discussed above apply to this outcome as well. Based on these factors, sequential monotherapy in some patients appears to have caused doubling in two-day seizure frequency. Combining the estimates of the two seizure increase outcomes that result in exiting a trial, 30 percent of patients in studies of sequential monotherapy experience a doubling in either monthly seizure frequency or two-day seizure frequency. Therefore, some patients may be experiencing large seizure increases as a direct result of sequential monotherapy. To provide a definite answer, randomizing patients to receive either sequential monotherapy or a continuation of the prestudy drug regimen would be necessary.

Adverse effects

In clinical practice, a physician prescribing an AED for a patient with treatment-resistant epilepsy must consider not only the possible reduction of seizure frequency, but also the possible adverse effects of the new drug. Before entering studies, patients with treatment-resistant epilepsy were already experiencing adverse effects from their prestudy antiepileptic drug regimens. None of the studies reported these patients' prestudy adverse effects, and none reported whether the adverse effects observed during the study were more or less severe compared to patients' prestudy adverse effects. This latter outcome would have been informative because patients (and physicians) seek to reduce adverse effects as well as seizure frequency.

All 13 included studies of sequential monotherapy reported adverse effects of the new drug treatment. The overall percentage of patients who experienced any side effects was reported by six studies and ranged from 53 percent to 95 percent (Table 10). Dizziness was the most common adverse effect in four studies, and headache was the most common adverse effect in two studies. All details of the adverse effects in the 13 studies appear in Evidence Table 36.

Percentage of patients who exited trials due to adverse effects. To summarize the available data on adverse effects, we focused on whether the adverse effects in a given patient were severe enough to warrant discontinuation of the new drug (i.e., trial exit). This outcome is a marker of treatment failure. All 13 included trials reported the percentage of patients who exited trials due to adverse effects, and it ranged from 0% to 29%. As with seizure frequency, due to the lack of relevant control groups we performed a threshold analysis (Figure 15). Each summary estimate in the figure is based on Cohen's h. The summary estimate calculated at the 0 percent point on the graph (no patients in a synthetic control group exited due to adverse effects) was 0.47 (CI: 0.24 to 0.71, $p < 0.000073$) and corresponded to 5 percent (CI: 1 percent to 12 percent) of patients exiting trials due to adverse effects. The summary estimate became nonsignificant (no statistically significant difference between monotherapy and control patients in the number of patients exiting trials due to adverse effects) when the proportion of patients in the synthetic control group reached 2 percent.

Quality of life

Only two studies of sequential monotherapy reported quality of life outcomes (Evidence Table 37).[80,81] Evidence Table 39 lists the scales and subscales used to measure quality of life in these studies. Due to the small number of studies, we did not perform meta-analyses of the results. There were no statistically significant changes in quality of life in either of the two studies. The lack of statistical significance may have been due to insufficient power. An estimate of the power of pre- vs. posttests would require knowledge of the correlation between baseline and outcome measurements. However, the authors did not report these correlations and therefore the power of this study to detect statistically significant quality of life changes could not be determined. Many of the subscales showed a nonsignificant improvement over baseline. However, these subscales are not independent (i.e., improvement on one subscale is likely to result in improvement in another subscale). Therefore, firm evidence-based conclusion about the influence of sequential monotherapy on quality of life cannot be based on these data.

Mood

Three studies of sequential monotherapy reported outcomes related to mood (Evidence Table 37).[80-82] Evidence Table 40 lists the scales and subscales used to measure mood in these studies. Each of the three studies investigated a different drug for sequential monotherapy. Two of the

three studies used the same set of scales (Dodrill, Arnett, Hayes, et al.[81] and Dodrill, Arnett, Shu, et al.[80]). As with quality of life, the small number of studies precluded any meta-analysis. None of the subscales in the study by Dodrill, Arnett, Hayes, et al.[81] showed statistically significant changes in mood. In the study by Dodrill, Arnett, Shu, et al.,[80] one of eight subscales (the Vigor-Activity subscale) exhibited a statistically significant decrement from baseline in mood. As discussed in the quality of life section, insufficient power may have prevented these studies from detecting changes in mood. However, incomplete reporting in the published literature prevents investigating this possibility.

Ketter, Malow, Flamini, et al.[82] reported that mood and psychiatric symptom scores changed after 2 weeks of sequential monotherapy in patients given felbamate. After the removal of all AEDs, patients' mood scores significantly worsened relative to baseline for seven of the 11 subscales. At both week 1 and week 2 of felbamate monotherapy, the decrements persisted. Thus, the initiation of felbamate monotherapy did not return patients' mood scores to baseline. With longer followup, patients' mood scores may potentially have returned to baseline or even improved over baseline. However, these findings suggest that the first phase of sequential monotherapy (i.e., the drug reduction phase) may cause significant worsening of mood and psychiatric symptom scores. However, because there was only one study reporting such changes, firm evidence-based conclusions cannot be drawn about the general effect of sequential monotherapy on mood.

Cognitive function

Only two studies of sequential monotherapy reported cognitive function (Evidence Table 38).[80,81] The two studies used the same subscales for measuring cognitive function (Evidence Table 41). Due to the small number of studies that reported the effect of sequential monotherapy on cognitive function, we did not perform meta-analyses of these data. In the study described by Dodrill, Arnett, Hayes, et al.[81], none of the 19 cognitive function subscales were significantly different from baseline. Of the 19 subscales in the study described by Dodrill, Arnett, Shu, et al.,[80] four showed a statistically significant improvement from baseline. These results may have been caused by a practice effect and not by tiagabine (see the discussion of instrumentation bias in the Methodology section). The power of these studies to detect changes in cognitive function could not be calculated because the authors did not report the correlations between baseline and outcome measurements.

Functional status/ability

No studies of sequential monotherapy reported this outcome.

Ability to return to work

No studies of sequential monotherapy reported this outcome.

Ability to return to school

No studies of sequential monotherapy reported this outcome.

Ability to hold a driver's license

No studies of sequential monotherapy reported this outcome.

Mortality

Five of the 13 studies of sequential monotherapy (38 percent) reported whether any patients died during the study. Three of the five studies reported that no patients died,[78,83,84] and the other two studies each reported one death.[68,85] The authors did not attribute either death to the treatment. The mortality rates in these five studies ranged from 0 percent to 2 percent (Evidence Table 42).

Table 7. Drugs and doses in studies of sequential monotherapy

Reference	Felbamate	Gabapentin		Lamotrigine	Oxcarbazepine	Primidone	Tiagabine	Topiramate	Valproate
	3600 mg/day	2400 mg/day	3600 mg/day	500 mg/day	2400 mg/day	750 mg/day	36 mg/day	1000 mg/day	150 µg/ml
Sachdeo (2001)[68]					✓				
Beydoun (2000)[86]					✓				
Kanner (2000)[87]						✓			
Schachter (1999)[79]					✓				
Gilliam (1998)[76]				✓					
Bergey (1997)[78]			✓						
Beydoun (1997)[85]									✓
Beydoun (1997)[83]		✓							
Sachdeo (1997)[84]								✓	
Devinsky (1995)[75]	✓								
Schachter (1995)[88]							✓		
Theodore (1995)[89]	✓								
Faught (1993)[77]	✓								
Totals	**3**	**1**	**1**	**1**	**3**	**1**	**1**	**1**	**1**

mg/day Maximum dose in milligrams per day
µg/ml Maximum dose in micrograms per milliliter

Table 8. Outcomes in studies of sequential monotherapy

Reference	Seizure Frequency	Adverse Effects	Quality of Life	Mood	Cognitive Function	Ability to Return to Work	Ability to Return to School	Ability to Hold a Driver's License	Mortality
Sachdeo (2001)[68]	✓	✓							✓
Beydoun (2000)[86]	✓	✓							
Kanner (2000)[87]	✓	✓							
Schachter (1999)[79]	✓	✓							
Gilliam (1998)[76]	✓	✓							
Bergey (1997)[78]	✓	✓							✓
Beydoun (1997)[85]	✓	✓							✓
Beydoun (1997)[83]	✓	✓	✓	✓	✓				✓
Sachdeo (1997)[84]	✓	✓							✓
Devinsky (1995)[75]	✓	✓							
Schachter (1995)[88]	✓	✓	✓	✓	✓				
Theodore (1995)[89]	✓	✓		✓					
Faught (1993)[77]	✓	✓							
Totals	**13**	**13**	**2**	**3**	**2**	**0**	**0**	**0**	**5**

Table 9. Seizure frequency outcomes in studies of sequential monotherapy

Reference	Mean Absolute Seizure Frequency	Median Absolute Seizure Frequency	Median Absolute Difference From Baseline	Median Percent Difference From Baseline	Number of Patients Seizure-free	Number of Patients With ≥ 75 Percent Reduction	Number of Patients With ≥ 50 Percent Reduction	Number of Patients With Any Reduction	Mean Rank of Seizure Frequency	Number of Patients With Doubling of Monthly Seizure Frequency	Number of Patients With Doubling of Highest Consecutive Two-Day Seizure Frequency	Median Time to Exit	Mean Time on Monotherapy	Risk Ratio of Time to Exit
Sachdeo (2001)[68]										✓	✓	✓		
Beydoun (2000)[86]					✓		✓			✓	✓	✓		
Kanner (2000)[87]					✓								✓	
Schachter (1999)[79]		✓			✓									✓
Gilliam (1998)[76]												✓		
Bergey (1997)[78]					✓								✓	
Beydoun (1997)[85]			✓	✓	✓	✓	✓	✓						
Beydoun (1997)[83]										✓	✓	✓	✓	
Sachdeo (1997)[84]					✓	✓	✓			✓	✓			✓
Devinsky (1995)[75]									✓					
Schachter (1995)[88]		✓					✓						✓	
Theodore (1995)[89]	✓													
Faught (1993)[77]										✓	✓			
Totals	**1**	**2**	**1**	**1**	**6**	**2**	**4**	**1**	**1**	**5**	**5**	**4**	**4**	**2**

Table 10. Overview of adverse effects of sequential monotherapy

Reference	Drug	Dose in mg/day	Percent of Patients Who Experienced Any Adverse Effect	Name of Most Commonly Experienced Adverse Effect	Percent of Patients Who Experienced This Adverse Effect
Sachdeo (2001)[68]	Oxcarbazepine	2400	NR	Headache	11% (5/45)
Beydoun (2000)[86]	Oxcarbazepine	2400	NR	Dizziness	46% (19/41)
Kanner (2000)[87]	Primidone	750	53% (16/30)	Irritability	37% (11/30)
Schachter (1999)[79]	Oxcarbazepine	2400	91% (46/51)	Nervous system	45% (23/51)
Gilliam (1998)[76]	Lamotrigine	500	75% (57/76)	Dizziness	20% (15/76)
Bergey (1997)[78]	Gabapentin	3600	73% (29/40)	Ataxia	20% (8/40)
Beydoun (1997)[85]	Valproate	150 µG/mL	NR	Tremor	64% (61/96)
Beydoun (1997)[83]	Gabapentin	2400	88% (80/91)	Dizziness	25% (23/91)
Sachdeo (1997)[84]	Topiramate	1000	NR	Paresthesia	58% (14/24)
Devinsky (1995)[75]	Felbamate	3600	NR	NR	NR
Schachter (1995)[88]	Tiagabine	36	95% (91/96)	Dizziness	35% (34/96)
Theodore (1995)[89]	Felbamate	3600	NR	NR	NR
Faught (1993)[77]	Felbamate	3600	NR	Headache	34% (19/56)

mg/day Milligrams per day
NR Not reported
µG/ml Micrograms per milliliter

Figure 11. Threshold analysis: sequential monotherapy and seizure freedom

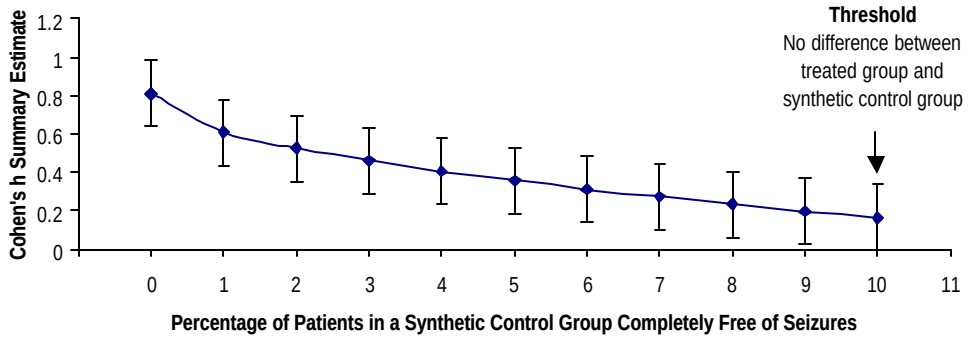


Figure 12. Threshold analysis: monotherapy and seizure freedom (long-term studies)

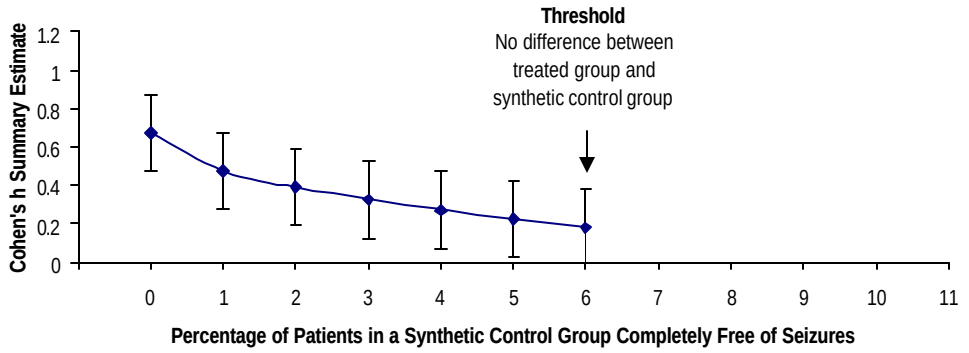


Figure 13. Threshold analysis: monotherapy and doubling of monthly seizure frequency

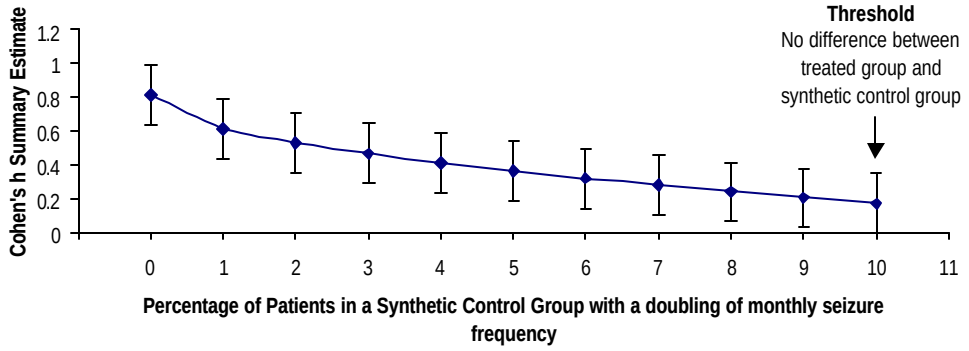


Figure 14. Threshold analysis: monotherapy and doubling of two-day seizure frequency

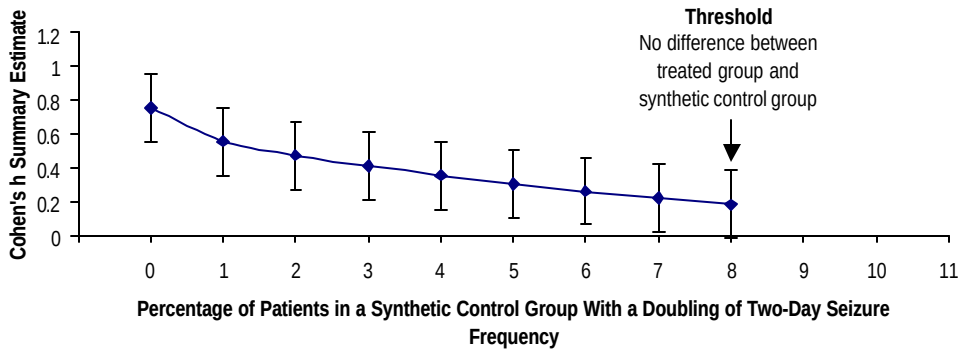
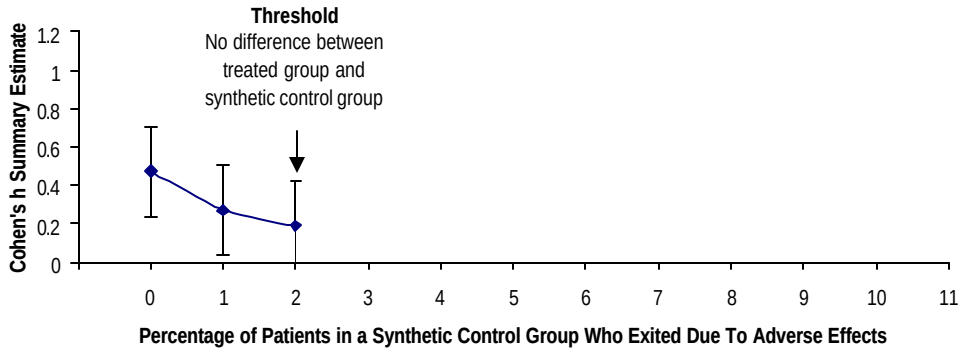


Figure 15. Threshold analysis: monotherapy and trial exits due to adverse effects



Polytherapy

Polytherapy is defined as the administration of a multiple-drug regimen in which at least one of the drugs is novel to each patient. As with sequential monotherapy, patients received any number of drugs prior to the initiation of a new drug. Most polytherapy interventions involve the addition of a single novel drug to patients' regimens (referred to as "add-on" treatment). In this section, we describe the evidence base for polytherapy, assess the quality of these trials with respect to both internal and external validity, and analyze the trials' results for all relevant outcomes.

Excluded studies

Thirty trials of polytherapy met the inclusion criteria. None were excluded for quality reasons.

Evidence base

The evidence base contained 30 trials that enrolled 4,834 patients.

Design and conduct of included studies

Aspects of the trial designs appear in Evidence Tables 43 through 46, and the patient characteristics appear in Evidence Tables 47 through 53. All 30 trials were randomized, placebo-controlled, add-on trials. In these trials, patients continued to take their pretrial drug regimens, and either a placebo or a new drug was added to those regimens. Nine add-on drugs were investigated in these trials: topiramate (9 trials), gabapentin (6 trials), lamotrigine (4 trials), levetiracetam (3 trials), tiagabine (2 trials), zonisamide (2 trials), felbamate (2 trials), oxcarbazepine (1 trial), and valproate (1 trial). No single dose of a given drug was used in all trials of that drug (Table 11). Of the 29 drug doses, 20 (69 percent) were employed by only one trial, and no drug dose was employed by more than four trials. Ten trials (33 percent) individualized the dose to each patient based on weight. These observations highlight the wide variation among trials' implementations of the polytherapy strategy.

Internal validity

For each trial of polytherapy, we determined whether the results were potentially biased by the factors noted in the Methodology section. Other questions in this report consider the potential for attrition bias, but for polytherapy, we did not consider it because attrition was a study outcome. All 30 trials of polytherapy were randomized and placebo-controlled. Thus, they were free from many potential threats to internal validity (see Appendix B). We meta-analytically tested selection bias with respect to several patient characteristics, and trials were free from potential selection bias in all but two cases (Evidence Table 54 through 58; also see discussion in Appendix B). All of the trials were free from five potential biases (sampling, regression, investigator, patient, and extraneous event). However, all of the trials had potential measurement bias. In addition, 90 percent of the trials had sample specification bias.

External validity

In our appraisal of the external validity of trials of polytherapy, we considered aspects of patient enrollment as well as the actual characteristics of patients in the trials (Evidence Tables 43 to 53). Twenty-seven trials (90 percent) enrolled patients because of seizure type: 25 for

partial seizures, and two for generalized seizures. Two trials enrolled only patients with Lennox-Gastaut syndrome, and one trial included patients with any seizure type or syndrome. Six trials enrolled children only, 23 trials enrolled adults only, and one trial enrolled both children and adults. In the six trials of children, the mean age ranged from 7.9 to 13.0, and in the 23 trials of adults, the mean age ranged from 29.4 to 38.0. The proportion of patients who were female ranged from 0.14 to 0.69, and was less than 0.50 in 21 of the 28 trials that reported this characteristic. Median seizure frequency ranged from 1 to 80 seizures per month, and mean seizure frequency ranged from 7.3 to 68.7 seizures per month. The proportion of patients who had received two or more prior AEDs ranged from 0 to 0.81. This proportion was greater than 0.5 in 15 of the 18 trials that reported this patient characteristic. As a whole, then, the characteristics of the patients in these studies are not particularly unusual.

Synthesis of study results

In this section, we assess the results separately for each of the relevant outcomes (Table 12). All reported outcomes appear in Evidence Tables 59 through 62. Seizure frequency and adverse effects were each reported by all 30 trials, whereas the other outcomes were not commonly reported.

In cases where meta-analyses were conducted, we used random-effects models because, as shown in Table 12, the trials employed a variety of drugs and doses. The trials are therefore not derived from a population of trials with a fixed mean. Our meta-analytic syntheses of trial results yield approximate estimates of the typical effect of adding a new AED to patients' prior AED regimens. However, these estimates have limited generalizability because the effect of a new AED may depend on the other AEDs in patients' regimens. Each trial employed a control group of patients who received an add-on placebo, but the prior regimens were different among different trials (and among patients in a single trial). Thus, the 30 trials did not administer the exact same "control" treatment. Because the treatments and controls differ across trials, the summary effect sizes from random-effects meta-analyses can only be used as approximate estimates of the effect of adding a new AED and may be best suited for use as starting points in future research. The actual effect on seizure frequency or adverse effects in any single patient is likely to depend on the specific drug to be added as well as characteristics of the AEDs already in use.

We performed all meta-analyses on an intent-to-treat basis. This means that we included all randomized patients in our analyses, not solely the patients who completed the trials. If a patient exited early from a trial and the authors did not report the relevant outcome for that patient, we assumed that seizure frequency did not decrease for that patient. This is a reasonable assumption because all patients who respond to a drug would likely be reported as responders.

Seizure frequency

The included trials reported 20 different measures of seizure frequency (Table 13). Seven measures were reported by five or more trials. One was a measure of absolute seizure frequency (median percentage reduction), and the remaining six were dichotomous measures. We did not analyze two of the dichotomous measures (75 percent or more reduction and 25 percent or more reduction) because they provided data that was effectively captured by other dichotomous measures (seizure-freedom, 50 percent or more reduction, and any reduction). The use of multiple seizure types and multiple study intervals necessitated that we adopt two selection rules

for abstracting data from an included study.[h] First, if a study reported the same seizure frequency measure for more than one seizure type, we selected the most general type for inclusion in any meta-analyses.[i] Second, if a study reported the same seizure frequency measure for different study intervals, we selected the longest interval for inclusion in any meta-analyses[j]. These selection rules permitted us to focus our analyses on the most general and widely reported seizure frequency measures.

In considering meta-analyses of the seizure frequency outcomes, nine of the 30 trials each contained three or more groups of patients. From each of these nine trials, therefore, multiple effect sizes can be computed (e.g., dose 1 vs. placebo, and dose 2 vs. placebo). Multiple effect sizes within a single trial are statistically dependent. Ideally, we would analyze these data using general linear models that account for this dependence. This, however, was precluded by the relative paucity of data. Therefore, to avoid this dependence, for each meta-analysis we selected only one drug dose from each trial. Thus, the effect size we computed for each trial was based on the difference between outcomes in one add-on drug group and the add-on placebo group. In some meta-analyses ("high-dose"), we selected the highest-dose group in each trial, whereas in other meta-analyses ("low-dose"), we selected the lowest-dose group in each trial. Trials with only one add-on drug group appear in both the high-dose and low-dose meta-analyses. Consequently, the high-dose meta-analysis was not independent of the low-dose meta-analysis.

A comparison between the results of a high-dose meta-analysis with those of the corresponding low-dose meta-analysis can be viewed as a form of sensitivity analysis. Larger effect sizes may be expected *a priori* in the high-dose meta-analysis (i.e., larger effects with higher doses). Performing both analyses permits us to estimate the robustness of the results. Although this approach allows us to estimate the effect of high- and low-dose polytherapy, it has the disadvantage that each meta-analysis uses only a subset of the available data. Consequently, some information is lost in our analysis.

Median percentage reduction. Twenty-four of the 30 included trials reported the median percentage reduction in seizures. However, none of these trials reported the dispersion about these medians (e.g., variances, standard deviations, interquartile ranges). Therefore, effect sizes could not be calculated and a meta-analysis was not conducted with these data.

Because the median percentage of seizure reduction was a commonly reported seizure frequency outcome, we plotted a summary of the published findings. This plot (Figure 16) depicts the 24 statistical comparisons to placebo that were reported.[k] Twenty-two of these 24 comparisons were statistically significant in favor of the add-on drug. The remaining two comparisons also favored the add-on drug, but the differences were not statistically significant. The range of medians was -18 percent to 13 percent for the groups that received add-on placebo, and 13 percent to 51 percent for the add-on drug groups (as a convention in the

[h] These selection rules were not necessary for sequential monotherapy because no such multiple-interval reporting occurred for the seizure frequency outcomes we analyzed.

[i] For example, if a study reported the number of patients free from all partial seizures as well as the number of patients free from secondarily generalized seizures (which is a special kind of partial seizure), we selected only the former outcome in meta-analyses.

[j] For example, if a study reported the number of patients seizure-free during the entire double-blind period as well as the number of patients seizure-free during the dose maintenance period alone (which is a subset of the double-blind phase), we selected only the former outcome in meta-analyses.

[k] Seventeen of the 24 trials reported a statistical comparison to placebo and there were 24 dose groups among these 17 trials. The plot includes all 24 dose groups. Because we did not perform a meta-analysis, both high-dose and low-dose groups are included in the plot.

epilepsy literature, negative numbers represent percentage increases from baseline, and positive numbers represent percentage decreases from baseline).

The data in Figure 16 provide evidence for a placebo effect: 14 of the 17 placebo groups (82 percent) had a median percentage reduction that was greater than zero (i.e., a beneficial effect indicated by the rightward shift on the x-axis). This percentage is significantly larger than 50 percent (two-tailed sign test, p = 0.013). The size of this placebo effect cannot be estimated because the trials did not report dispersion statistics for median percentage reduction. The observed medians, however, do indicate that patients with treatment-resistant epilepsy have fewer seizures when a placebo is added to their drug regimens. This placebo effect does not influence our investigations of polytherapy because all trials were placebo-controlled and therefore all effect sizes involved comparisons to placebo groups. However, the placebo effect does underscore the need for placebo controls in treatment trials involving patients with epilepsy, because if a treated group improves, part of that improvement may be due to the initiation of *any* medical intervention rather than to the intervention itself.

Seizure-freedom. Eighteen of the 30 included trials reported the percentage of patients who became seizure-free. Two of these trials, however, only reported seizure-freedom for a severe type of partial seizure (secondarily generalized seizures) that was experienced by only a subset of patients before the trials.[90,91] They did not report freedom from a seizure type that all patients had experienced before the trial. Thus, we analyzed the 16 trials of polytherapy that did report the latter kind of seizure-freedom results.

We first performed the high-dose meta-analysis. The percentage of patients who were seizure-free ranged from 0 percent to 9 percent in the high-dose groups and from 0 percent to 2 percent in the add-on placebo groups. The effect sizes are plotted in Figure 17, and the details of the random-effects meta-analysis appear in Evidence Table 63. The random-effects summary statistic (Cohen's h) was 0.29 (CI: 0.20 to 0.37). Patients who received a high-dose of add-on drug were statistically significantly more likely to become seizure-free compared to patients who received add-on placebo. The estimated summary percentages were 5 percent for the high-dose groups (CI: 3 percent to 7 percent) and 1 percent for the placebo groups (CI: 0 percent to 1.4 percent) as calculated from the back-transformed Cohen's h. Similar results were observed in the low-dose meta-analysis for seizure-freedom (Figure 18 and Evidence Table 64). The summary Cohen's h (0.28, CI: 0.20 to 0.36) was only slightly lower than the high-dose meta-analysis. The estimated summary percentage was 5 percent (CI: 3 percent to 7 percent) in the low-dose groups.

We performed four sensitivity analyses separately for the high-dose and low-dose meta-analyses. The sensitivity analyses involved recalculating the meta-analysis after separately removing the trial with the largest effect size, the smallest effect size, the largest sample size, and the smallest sample size. None of the four sensitivity analyses overturned our findings (Evidence Table 65 and Table 66).

In summary, the evidence suggests that adding a drug to patient's regimens increases the likelihood of becoming seizure-free. This finding occurred in both the high-dose and low-dose groups, and multiple sensitivity analyses did not overturn the results. However, seizure-freedom was analyzable for only 16 of the 30 trials of polytherapy.

50 percent reduction. Twenty-seven trials reported the percentage of patients who experienced 50 percent or more reduction in seizures. As with seizure-freedom, we performed both a high-dose meta-analysis and a low-dose meta-analysis. The range was 13 percent to 50 percent in the high-dose groups and 0 percent to 25 percent in the placebo groups. A plot of

the effect sizes appears in Figure 19, and the statistical details of the meta-analysis are in Evidence Table 67. The random effects summary statistic (Cohen's h) was 0.52 (CI: 0.43 to 0.62). Patients who received a high-dose of add-on drug were significantly more likely to experience 50 percent reduction compared to patients who received add-on placebo. The estimated summary percentages were 35 percent for the high-dose groups (CI: 31 percent to 38 percent) and 13 percent for the placebo groups (CI: 10 percent to 15 percent). We observed similar results with the low-dose meta-analysis (Figure 20 and Evidence Table 68). The random-effects summary Cohen's h was 0.45 (CI: 0.35 to 0.55), and the estimated summary percentage for the low-dose groups was 31 percent (CI: 27 percent to 36 percent).

As described previously, we performed four sensitivity analyses for both the high-dose and low-dose meta-analyses. None of these analyses overturned our findings (Evidence Tables 69 and 70).

These studies suggest that when a drug is added to patients' drug regimens, approximately one-third of patients will experience a 50 percent or more reduction in seizures. As mentioned above, however, the generalizability of this finding may be limited.

Any reduction. Five trials reported the percentage of patients who experienced any reduction in seizures. As with other measures, we performed both a high-dose meta-analysis and a low-dose meta-analysis. The range was 61 percent to 80 percent in the high-dose groups and 41 percent to 72 percent in the placebo groups. A plot of the effect sizes appears in Figure 21, and the statistical details of the meta-analysis are in Evidence Table 71. The random effects summary statistic (Cohen's h) was 0.37 (CI: 0.19 to 0.55). Patients who received a high-dose of add-on drug were significantly more likely to experience a reduction compared to patients who received add-on placebo. The estimated summary percentages were 70 percent for the high-dose groups (CI: 61 percent to 77 percent) and 52 percent for the placebo groups (CI: 44 percent to 61 percent). We obtained similar results with the low-dose meta-analysis (Figure 22 and Evidence Table 72). The random-effects summary Cohen's h was 0.31 (CI: 0.15 to 0.47), and the estimated summary percentage for the low-dose groups was 67 percent (CI: 59 percent to 74 percent).

We performed the four sensitivity analyses for both the high-dose and low-dose meta-analyses. None of these overturned our findings (Evidence Table 73 and 74).

These studies suggest that when certain AEDs are added to patients' drug regimens, approximately two-thirds of patients will experience some reduction in seizures. This analysis, like the previous one, may have limited generalizability.

Any increase. Six trials reported the percentage of patients who experienced any increase in seizures. One of these trials, however, reported this outcome for a specific seizure type that was experienced by only a subset of patients before the trial. Thus, we analyzed seizure increase data from the other five trials. As with other measures, we performed both a high-dose meta-analysis and a low-dose meta-analysis. The range was 16 percent to 38 percent in the high-dose groups and 28 percent to 44 percent in the placebo groups. A plot of the effect sizes appears in Figure 23, and the statistical details of the meta-analysis are in Evidence Table 75. The random effects summary statistic (Cohen's h) was 0.38 (CI: 0.23 to 0.53). Patients who received a high-dose of add-on drug were significantly less likely to experience an increase compared to patients who received add-on placebo. The estimated summary percentages were 21 percent for the high-dose groups (CI: 15 percent to 28 percent) and 39 percent for the placebo groups (CI: 32 percent to 46 percent). We observed similar results with the low-dose meta-analysis (Figure 24 and Evidence Table 76). The random-effects summary Cohen's h was 0.39 (CI: 0.22 to 0.57), and the

estimated summary percentage for the low-dose groups was 20 percent (CI: 15 percent to 27 percent).

We performed the four sensitivity analyses for both the high-dose and low-dose meta-analyses. None of these analyses overturned our findings (Evidence Table 77 and 78).

These data suggest that when certain AEDs are added to patients' drug regimens, approximately 20 percent of patients will experience an increase in seizures. This analysis, like the previous ones, may have limited generalizability.

Adverse effects

All 30 included studies of polytherapy reported adverse effects of the new drug treatment. The overall percentage of patients who experienced *any* side effects was reported by 16 studies, and ranged from 55 percent to 94 percent (Table 14). Somnolence was the most common adverse effect in nine studies, and dizziness was the most common adverse effect in four studies. All details of the adverse effects in the 30 studies appear in Evidence Table 60.

To summarize the available data on adverse effects, we focused on whether the adverse effects in a given patient were severe enough to warrant discontinuation of the new drug (i.e., trial exit). In trials of polytherapy, an add-on drug may be more likely or less likely to be discontinued due to adverse effects compared to add-on placebo.

Percentage of patients exiting trials due to adverse effects. All 30 trials of polytherapy reported this outcome. We meta-analyzed these data using the same methods that we used to analyze seizure frequency. The effect sizes are plotted in Figure 25, and the details of the high-dose meta-analysis appear in Evidence Table 79. The random-effects summary Cohen's h was significantly negative (-0.18, CI: -0.26 to -0.11). Thus, patients in the high-dose groups were significantly more likely to exit trials due to adverse effects compared to patients in placebo groups. The estimated summary percentages were 8 percent for the high-dose groups (CI: 6 percent to 10 percent) and 4 percent for the placebo groups (CI: 2 percent to 5 percent). Similar results were observed for the low-dose meta-analysis (Figure 26 and Evidence Table 80). The random-effects summary statistic was -0.16 (CI: -0.23 to -0.08), and the estimated summary percentage for the low-dose groups was 7 percent (CI: 5 percent to 9 percent).

We performed the same sensitivity analyses and they did not overturn any of our findings (Evidence Table 81 and 82).

Thus, adding a certain AED to a patient's drug regimen is more likely to cause adverse effects resulting in trial exit compared to adding a placebo. This finding persisted through multiple sensitivity analyses.

Tradeoff between seizure frequency and adverse effects. We next evaluated the tradeoff between seizure frequency and adverse effects in trials of polytherapy. In the section on seizure frequency, we concluded that adding a drug to a patient's regimen is more likely to reduce seizures compared to adding a placebo. However, in the section on adverse effects, we concluded that adding a drug is also more likely to cause adverse effects resulting in trial exit. To illustrate the tradeoff, we constructed a scatterplot in which the horizontal axis represented the effect size for 50 percent seizure reduction and the vertical axis represented the effect size for exiting the trial due to adverse effects (Figure 27). We inverted the vertical axis so that the ideal drug would fall in the upper right quadrant of the plot (corresponding to fewer seizures and fewer adverse effects). Forty groups of patients who received an add-on drug are included in the plot (corresponding to the 27 trials that reported both 50 percent seizure reduction and adverse effect attrition). Thirty-one of 40 patient groups (78 percent) were in the lower right quadrant (fewer seizures and more adverse effects), and seven groups (18 percent) were in the upper right

quadrant. This plot demonstrates the tradeoff between seizure frequency and adverse effects. However, reductions in both seizure frequencies and side effects also seem to occur.

Quality of life

Only two of the included trials of polytherapy reported quality of life (Evidence Table 61).[92,93] The two trials used different scales to measure quality of life (Evidence Table 83). Due to the small number of trials, we did not perform meta-analyses of the results. Instead, we created plots indicating the trials' results for all reported subscales. In the trial by Cramer, Arrigo, Van Hammee, et al.,[92] four of the nine subscales of quality of life showed a statistically significant advantage of levetiracetam over placebo. Each of the other five subscales showed a nonsignificant advantage of levetiracetam. A statistical power analysis of this trial was not possible due to the lack of reporting of measures of dispersion. The results of this trial suggest that polytherapy with levetiracetam improves some aspects of quality of life. In the trial by Dodrill, Arnett, Sommerville, et al.,[93] no statistically significant effect was found on any of the 10 subscales of quality of life. A statistical power analysis of this trial could not be conducted because these were pre-post comparisons and the authors did not report the correlations between baseline and outcome measurements. Because only two trials reported quality of life outcomes after polytherapy, evidence-based conclusions could not be made about the influence of polytherapy on quality of life.

Mood

One trial of polytherapy (add-on tiagabine) reported mood outcomes (Evidence Table 61).[93] The trial used eight subscales to measure mood (Evidence Table 84). None of the eight subscales showed a statistically significant improvement in mood after add-on tiagabine. A statistical power analysis of this trial was not possible because the authors did not report the correlations between baseline and outcome measurements. Because this is only one trial, drawing any evidence-based conclusions about whether polytherapy affects mood is not possible.

Cognitive function

Only one of the included trials of polytherapy reported cognitive function (Evidence Table 62).[93] The subscales for measuring cognitive function in this trial appear in Evidence Table 85. Because there was only one trial, we created a plot indicating its results for all reported subscales. Of the 19 subscales of cognitive function, only one (the Benton Visual Retention test, Form F) demonstrated a statistically significant effect. Patients in the placebo group improved from baseline more compared to patients who received tiagabine. A power analysis of this trial could not be conducted because the authors did not report the correlations between baseline and outcome measurements. Because only one trial addressed this issue, evidence-based conclusions cannot be made about the influence of polytherapy on cognitive function.

Functional status/ability

No trials of polytherapy reported this outcome.

Ability to return to work

No trials of polytherapy reported this outcome.

Ability to return to school

No trials of polytherapy reported this outcome.

Ability to hold a driver's license

No trials of polytherapy reported this outcome.

Mortality

Nine of the 30 trials of polytherapy (30 percent) reported whether any patients died during the trial. The mortality results of these trials are listed in Evidence Table 86. The mortality rates ranged from 0 percent to 2 percent. Five of the nine trials reported that no patients died, three trials each reported one death, and one trial reported two deaths. None of the authors attributed the deaths to the add-on drugs.

Table 11. Drugs and doses in trials of polytherapy

Reference	Drug	Trial Dose(s) [a]	Total Number of Trials That Used This Drug/Dose Combination
Faught (2001)[94]	Zonisamide	400	1
Ben-Menachem (2000)[95]	Levetiracetam	3000	2
Betts (2000)[96]	Levetiracetam	2000, 4000	1, 1
Cereghino (2000)[97]	Levetiracetam	1000, 3000	1, 2
Glauser (2000)[98]	Oxcarbazepine	Tailored to weight	1
Appleton (1999) [99]	Gabapentin	Tailored to weight	1
Biton (1999)[100]	Topiramate	Tailored to weight	3
Duchowny (1999)[30]	Lamotrigine	Tailored to weight	2
Elterman (1999) [101]	Topiramate	Tailored to weight	3
Korean Topiramate Study Group (1999)[102]	Topiramate	600	4
Sachdeo (1999) [103]	Topiramate	Tailored to weight	3
Uthman (1998)[104]	Tiagabine	16, 32, 56	1, 2, 1
Sachdeo (1997) [105]	Tiagabine	32[b], 32[c]	2, 1
Ben-Menachem (1996)[90]	Topiramate	800	2
Chadwick (1996)[106]	Gabapentin	1200	3
Faught (1996)[91]	Topiramate	200, 400, 600	1, 2, 4
Privitera (1996)[107]	Topiramate	600, 800, 1000	4, 2, 1
Sharief (1996)[108]	Topiramate	400	2
Tassinari (1996)[109]	Topiramate	600	4
Willmore (1996) [110]	Valproate	Tailored to weight	1
Anhut (1994)[111]	Gabapentin	900, 1200	2, 3
Messenheimer (1994)[112]	Lamotrigine	400	1
Bourgeois (1993)[113]	Felbamate	3600	1
Felbamate Study Group (1993)[114]	Felbamate	Tailored to weight	1
Matsuo (1993)[115]	Lamotrigine	300, 500	1, 1
McLean (1993)[116]	Topiramate	600, 1200, 1800	4, 1, 1
Schmidt (1993)[117]	Zonisamide	Tailored to weight	1
Sivenius (1991)[118]	Gabapentin	900	2
UK Gabapentin Study Group (1990)[119]	Gabapentin	1200	3
Jawad (1989)[70]	Lamotrigine	Tailored to weight	2

[a] Maximum dose in milligrams per day
[b] Based on 16 milligrams twice per day
[c] Based on 8 milligrams four times a day

Table 12. Outcomes in trials of polytherapy

Reference	Seizure Frequency	Adverse Effects	Quality of Life	Mood	Cognitive Function	Ability to Return to Work	Ability to Return to School	Ability to Hold a Driver's License	Mortality
Faught (2001)[94]	✓	✓							
Ben-Menachem (2000)[95]	✓	✓							
Betts (2000)[96]	✓	✓							
Cereghino (2000)[97]	✓	✓	✓						✓
Glauser (2000)[98]	✓	✓							✓
Appleton (1999)[99]	✓	✓							✓
Biton (1999)[100]	✓	✓							
Duchowny (1999)[30]	✓	✓							
Elterman (1999)[101]	✓	✓							✓
KTSG (1999)[102]	✓	✓							
Sachdeo (1999)[103]	✓	✓							
Uthman (1998)[104]	✓	✓	✓	✓	✓				
Sachdeo (1997)[105]	✓	✓							
Ben-Menachem (1996)[90]	✓	✓							✓
Chadwick (1996)[106]	✓	✓							✓
Faught (1996)[91]	✓	✓							✓
Privitera (1996)[107]	✓	✓							✓
Sharief (1996)[108]	✓	✓							
Tassinari (1996)[109]	✓	✓							
Willmore (1996)[110]	✓	✓							
Anhut (1994)[111]	✓	✓							✓
Messenheimer (1994)[112]	✓	✓							
Bourgeois (1993)[113]	✓	✓							
FSG (1993)[114]	✓	✓							
Matsuo (1993)[115]	✓	✓							
McLean (1993)[116]	✓	✓							
Schmidt (1993)[117]	✓	✓							
Sivenius (1991)[118]	✓	✓							

Table 12. Outcomes in trials of polytherapy (continued)

Reference	Seizure Frequency	Adverse Effects	Quality of Life	Mood	Cognitive Function	Ability to Return to Work	Ability to Return to School	Ability to Hold a Driver's License	Mortality
UKGSG (1990) [119]	✓	✓							
Jawad (1989)[70]	✓	✓							
Totals	**30**	**30**	**2**	**1**	**1**	**0**	**0**	**0**	**9**

Table 13. Seizure frequency outcomes in trials of polytherapy

Reference	Absolute Monthly Seizure Frequency		Absolute Difference From Baseline		Absolute Percent Difference From Baseline		Number of Patients With								Mean Response Ratio [a]	Mean Adjusted Response Ratio [a]
	Mean	Median	Mean	Median	Mean	Median	Seizure-free	³75 Percent Reduction	³50 Percent Reduction	³25 Percent Reduction	Any Reduction	Any Increase	³50 Percent Increase	³75 Percent Increase		
Faught (2001)[94]						✓	✓	✓	✓							
Ben-Menachem (2000)[95]		✓				✓	✓		✓							
Betts (2000)[96]		✓				✓	✓		✓							
Cereghino (2000)[97]						✓	✓	✓	✓							
Glauser (2000)[98]						✓	✓		✓							
Appleton (1999)[99]						✓	✓	✓	✓	✓		✓	✓	✓		✓
Biton (1999)[100]						✓	✓	✓	✓							
Duchowny (1999)[30]		✓				✓			✓	✓						
Elterman (1999)[101]						✓	✓	✓	✓							
KTSG (1999)[102]		✓			✓	✓	✓		✓							
Sachdeo (1999)[103]						✓	✓	✓	✓							
Uthman (1998)[104]		✓		✓		✓			✓			✓				
Sachdeo (1997)[105]		✓		✓					✓							
Ben-Menachem (1996)[90]						✓	✓	✓	✓						✓	✓

Table 13. Seizure frequency outcomes in trials of polytherapy (continued)

Reference	Absolute Monthly Seizure Frequency		Absolute Difference From Baseline		Absolute Percent Difference From Baseline		Number of Patients With								Mean Response Ratio [a]	Mean Adjusted Response Ratio [a]
	Mean	Median	Mean	Median	Mean	Median	Seizure-free	[3] 75 Percent Reduction	[3] 50 Percent Reduction	[3] 25 Percent Reduction	Any Reduction	Any Increase	[3] 50 Percent Increase	[3] 75 Percent Increase		
Chadwick (1996)[106]						✓			✓							
Faught (1996)[91]						✓	✓	✓	✓							
Privitera (1996)[107]						✓		✓	✓							
Sharief (1996)[108]						✓	✓	✓	✓							
Tassinari (1996)[109]		✓				✓	✓	✓	✓							
Willmore (1996)[110]	✓	✓	✓	✓			✓	✓	✓	✓						
Anhut (1994)[111]		✓				✓			✓		✓	✓	✓			
Messenheimer (1994)[112]						✓										
Bourgeois (1993)[113]																
FSG (1993)[114]					✓		✓	✓	✓	✓						
Matsuo (1993)[115]						✓	✓		✓	✓			✓		✓	✓
McLean (1993)[116]						✓		✓	✓	✓	✓	✓				
Schmidt (1993)[117]						✓	✓	✓	✓		✓					
Sivenius (1991)[118]		✓						✓	✓	✓	✓	✓				✓

Table 13. Seizure frequency outcomes in trials of polytherapy (continued)

Reference	Absolute Monthly Seizure Frequency		Absolute Difference From Baseline		Absolute Percent Difference From Baseline		Number of Patients With								Mean Response Ratio [a]	Mean Adjusted Response Ratio [a]
	Mean	Median	Mean	Median	Mean	Median	Seizure-free	³ 75 Percent Reduction	³ 50 Percent Reduction	³ 25 Percent Reduction	Any Reduction	Any Increase	³ 50 Percent Increase	³ 75 Percent Increase		
UKGSG (1990)[119]						✓		✓	✓	✓	✓	✓	✓	✓		
Jawad (1989)[70]	✓	✓														
Totals [b]	**2**	**11**	**1**	**3**	**2**	**24**	**18**	**17**	**27**	**8**	**5**	**6**	**4**	**2**	**2**	**4**

[a] The response ratio is the ratio (T-B)/(T+B) where T is the number of seizures a month during treatment and B is the number of seizures a month during baseline. Some authors adjusted the response ratio in order to account for differences between centers in multi-center trials (using ANOVA).

Table 14. Overview of adverse effects of polytherapy

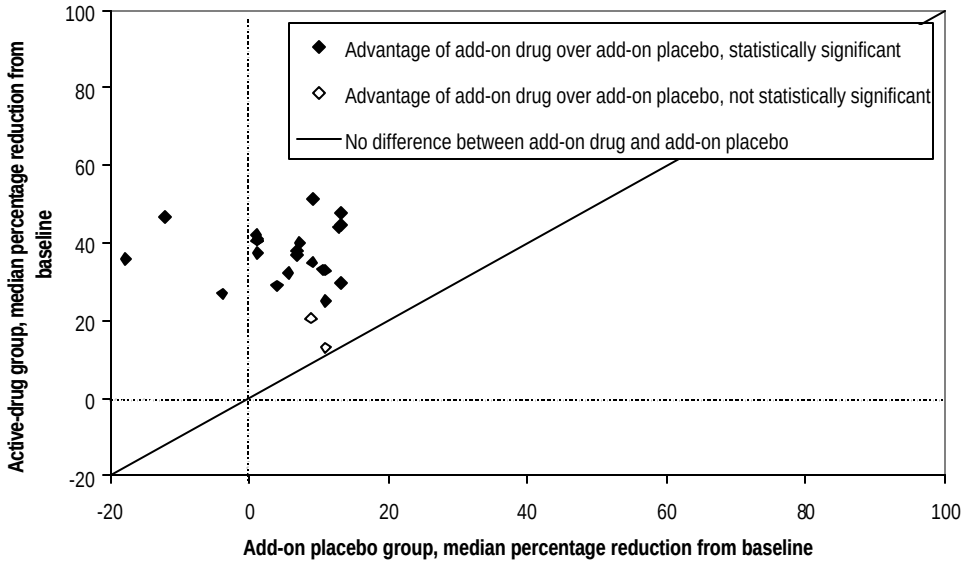
Reference	Drug and Dose (mg/day)	Percent of Patients Who Experienced Any Adverse Event	Name of Most Commonly Experienced Adverse Event	Percent of Patients Who Experienced This Adverse Event
Faught (2001)[94]	Zonisamide 400	NR	Somnolence	15% (18/118)
Ben-Menachem (2000)[95]	Levetiracetam 3000	55% (100/181)	Asthenia	13.8% (25/181)
Betts (2000)[96]	Levetiracetam 2000	83% (35/42)	Asthenia	31% (13/42)
Betts (2000)[96]	Levetiracetam 4000	84% (32/38)	Somnolence	45% (17/38)
Cereghino (2000)[97]	Levetiracetam 1000	89% (87/98)	Infection	28% (27/98)
Cereghino (2000)[97]	Levetiracetam 3000	89% (90/101)	Infection	27% (27/101)
Glauser (2000)[98]	Oxcarbazepine 1800	91% (125/138)	Vomiting	36% (50/138)
Appleton (1999)[99]	Gabapentin 1800	NR	Viral infection	11% (13/119)
Biton (1999)[100]	Topiramate 400	NR	Upper respiratory tract infection	41% (16/39)
Duchowny (1999)[30]	Lamotrigine 750	94% (92/98)	Somnolence	24% (24/98)
Elterman (1999)[101]	Topiramate 400	NR	Upper respiratory tract infection	41% (17/41)
KTSG (1999)[102]	Topiramate 600	81% (74/91)	Anorexia	21% (19/91)
Sachdeo (1999)[103]	Topiramate 600	NR	Somnolence	42% (20/48)
Uthman (1998)[104]	Tiagabine 16	NR	Nervous system	69% (42/61)
Uthman (1998)[104]	Tiagabine 32	NR	Nervous system	70% (62/88)
Uthman (1998)[104]	Tiagabine 56	NR	Nervous system	77% (44/57)
Sachdeo (1997)[105]	Tiagabine 32	NR	Nervousness	10.5% (11/105)
Ben-Menachem (1996)[90]	Topiramate 800	NR	Fatigue	79% (22/28)
Chadwick (1996)[106]	Gabapentin 1200	67% (39/58)	Somnolence	12% (7/58)
Faught (1996)[91]	Topiramate 200	NR	Dizziness	36% (16/45)
Faught (1996)[91]	Topiramate 400	NR	Dizziness	33% (15/45)
Faught (1996)[91]	Topiramate 600	NR	Dizziness	35% (16/46)
Privitera (1996)[107]	Topiramate 600	NR	Fatigue	38% (18/48)
Privitera (1996)[107]	Topiramate 800	NR	Abnormal thinking	44% (21/48)
Privitera (1996)[107]	Topiramate 1000	NR	Dizziness	38% (18/47)
Sharief (1996)[108]	Topiramate 400	NR	Somnolence	35% (8/23)
Tassinari (1996)[109]	Topiramate 600	NR	Headache	27% (8/30)
Willmore (1996)[110]	Valproate 90 mg/kg	NR	Nausea	48% (37/77)

Table 14. Overview of adverse effects of polytherapy (continued)

Reference	Drug and Dose (mg/day)	Percent of Patients Who Experienced Any Adverse Event	Name of Most Commonly Experienced Adverse Event	Percent of Patients Who Experienced This Adverse Event
Anhut (1994)[111]	Gabapentin 900	63% (33/52)	Somnolence	22% (24/111)
Anhut (1994)[111]	Gabapentin 1200	68% (76/111)	Somnolence	13% (7/52)
Messenheimer (1994)[112]	Lamotrigine 400	NR	Rash	7% (3/44)
Bourgeois (1993)[113]	Felbamate 3600	NR	Headache	40% (12/30)
FSG (1993)[114]	Felbamate 3600	NR	Anorexia	49% (18/37)
Matsuo (1993)[115]	Lamotrigine 300	NR	Headache	32% (23/71)
Matsuo (1993)[115]	Lamotrigine 500	NR	Dizziness	54% (39/72)
McLean (1993)[116]	Gabapentin 600	88% (89/101)	Dizziness	25% (13/53)
McLean (1993)[116]	Gabapentin 1200	91% (49/54)	Somnolence	36% (36/101)
McLean (1993)[116]	Gabapentin 1800	87% (46/53)	Somnolence	20% (11/54)
Schmidt (1993)[117]	Zonisamide 20 mg/kg	59% (42/71)	Fatigue	23% (16/71)
Sivenius (1991)[118]	Gabapentin 900	NR	Drowsiness	25% (4/16)
UKGSG (1990)[119]	Gabapentin 1200	62% (38/61)	Somnolence	14.8% (9/61)
Jawad (1989)[70]	Lamotrigine 400	NR	NR	NR

mg/day Milligrams per day
NR Not reported
mg/kg Milligrams per kilogram

Figure 16. Median percentage reduction in seizures after polytherapy



Note: In this plot, positive numbers represent *reductions* in seizures, whereas negative numbers represent *increases* in seizures.

Figure 17. Forest plot: polytherapy and seizure-freedom (high-dose)

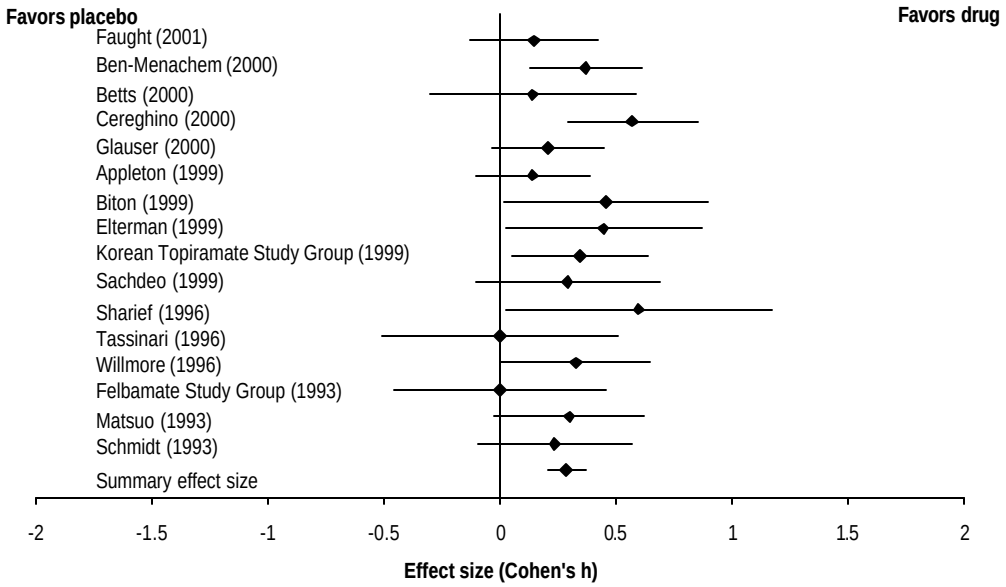


Figure 18. Forest plot: polytherapy and seizure-freedom (low-dose)

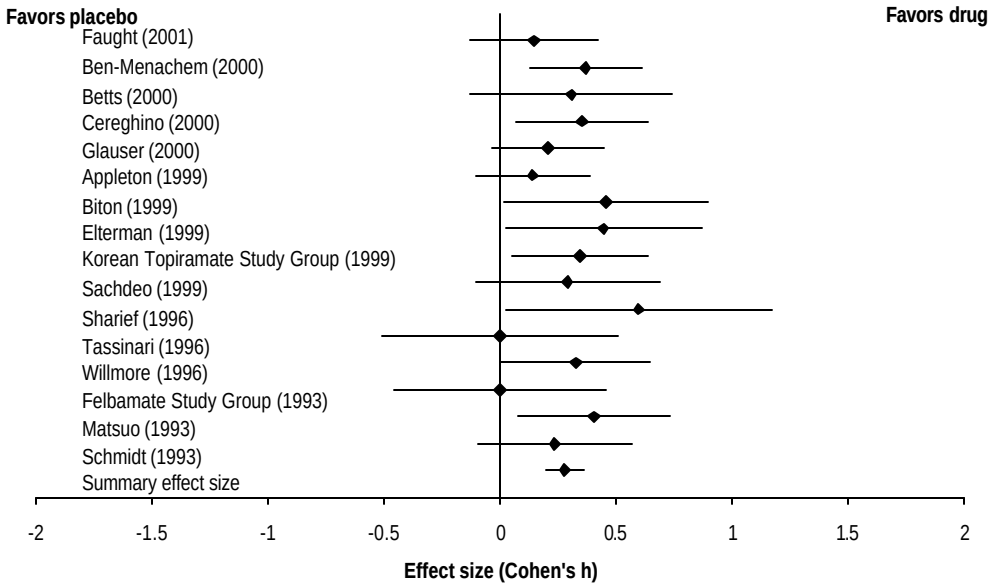


Figure 19. Forest plot: polytherapy and 50 percent seizure reduction (high-dose)

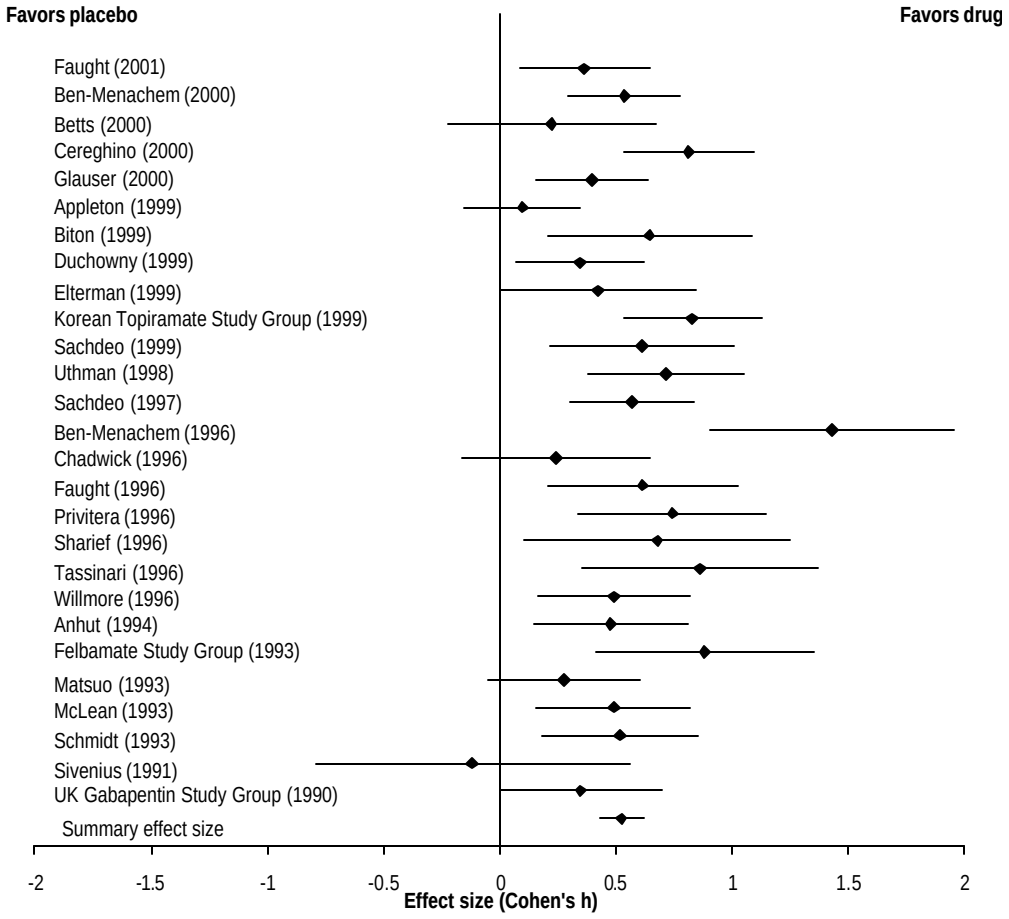


Figure 20. Forest plot: polytherapy and 50 percent seizure reduction (low-dose)

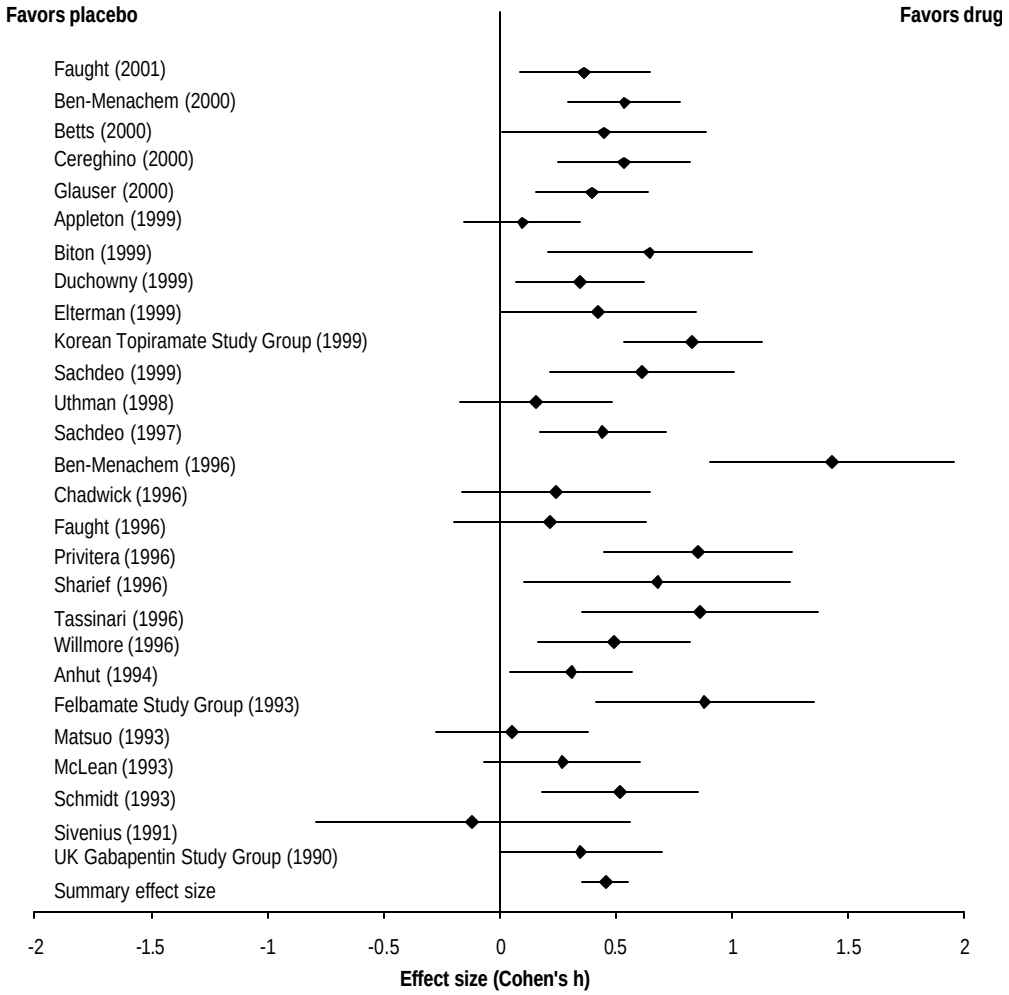


Figure 21. Forest plot: polytherapy and any seizure reduction (high-dose)

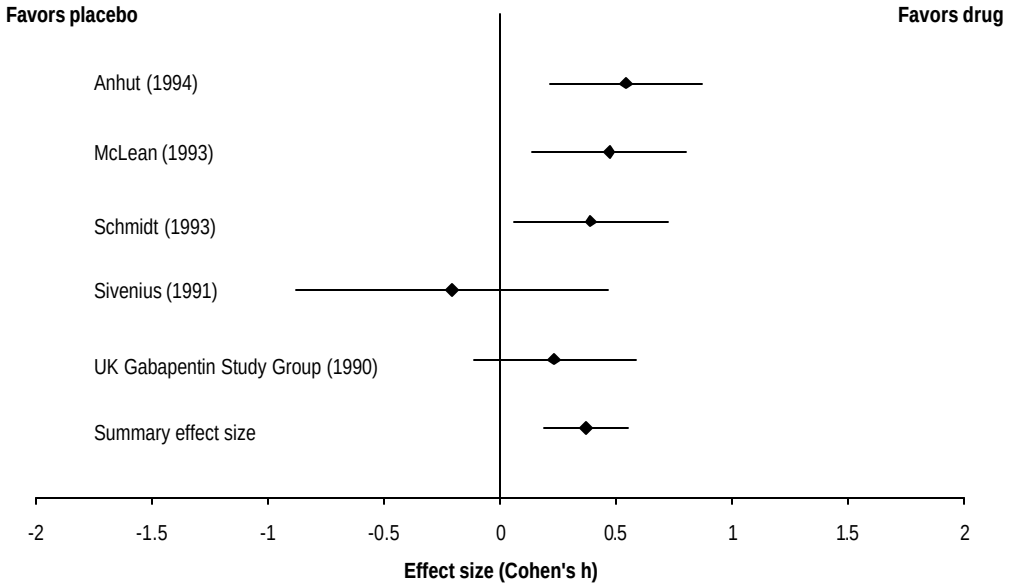


Figure 22. Forest plot: polytherapy and any seizure reduction (low-dose)

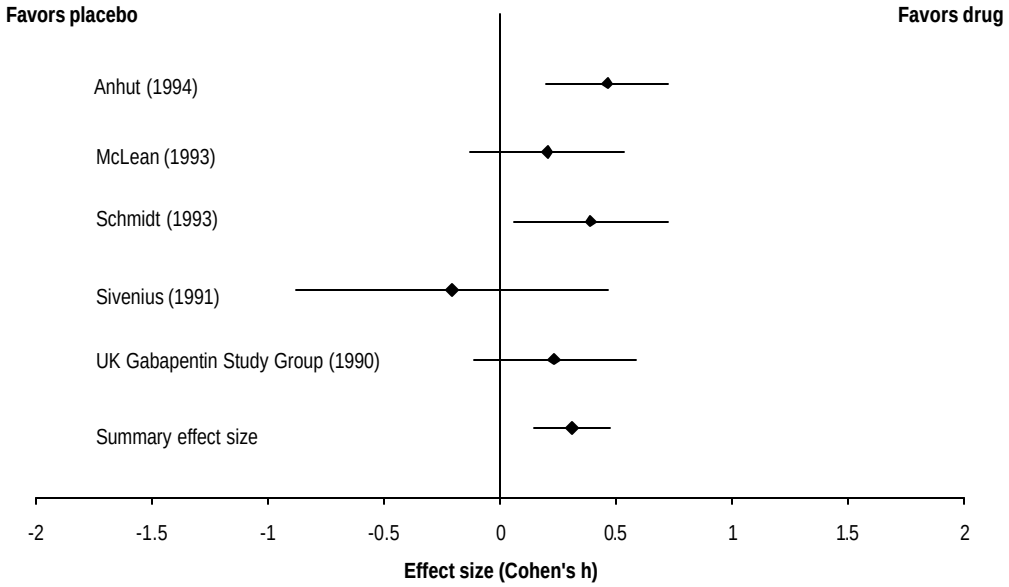


Figure 23. Forest plot: polytherapy and any seizure increase (high-dose)

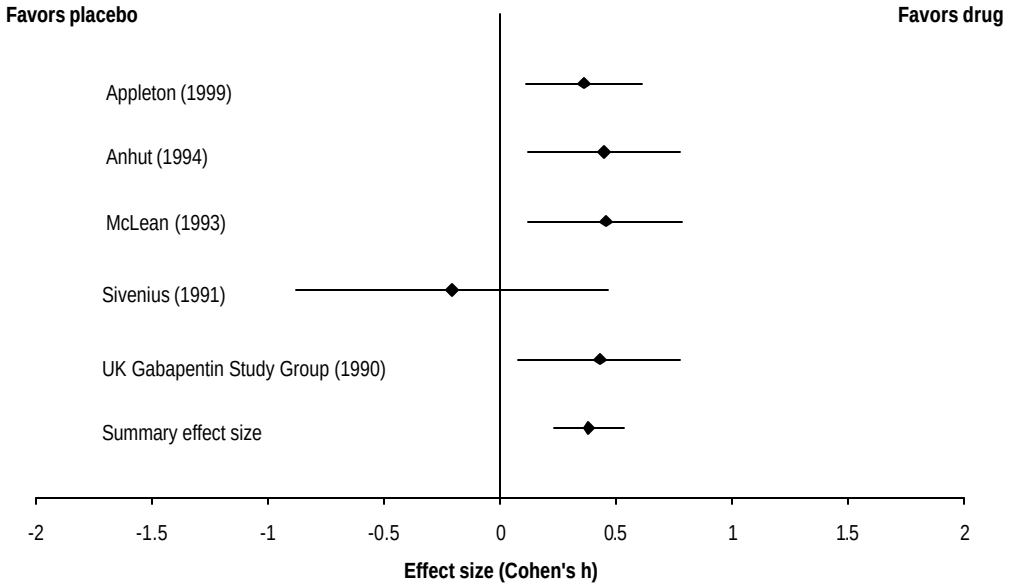


Figure 24. Forest plot: polytherapy and any s eizure increase (low-dose)

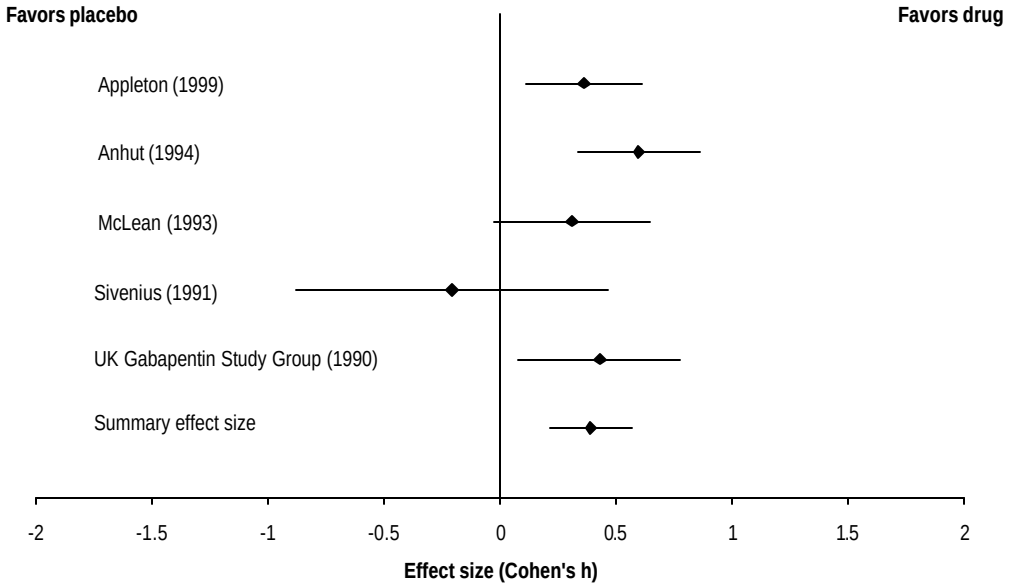


Figure 25. Forest plot: polytherapy and trial exits due to adverse effects (high-dose)

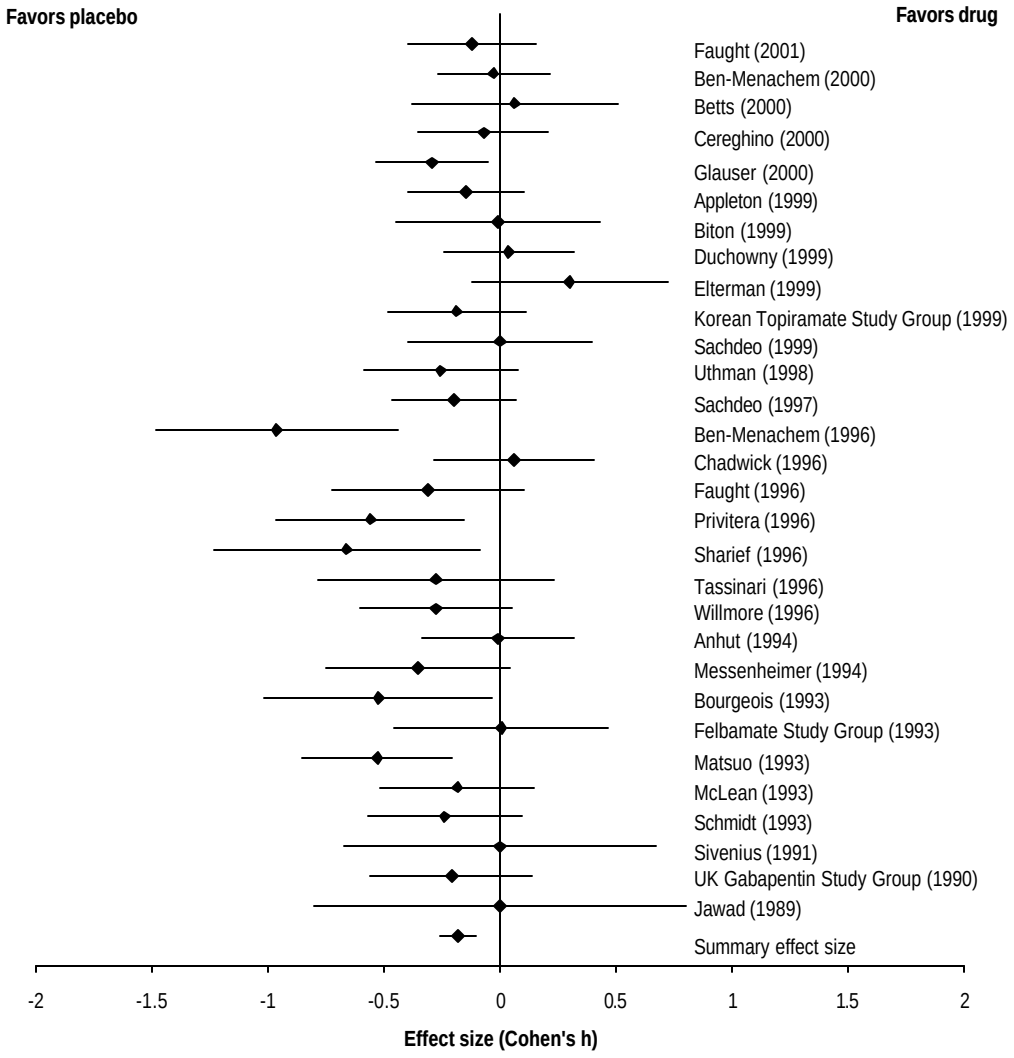


Figure 26. Forest plot: polytherapy and trial exits due to adverse effects (low-dose)

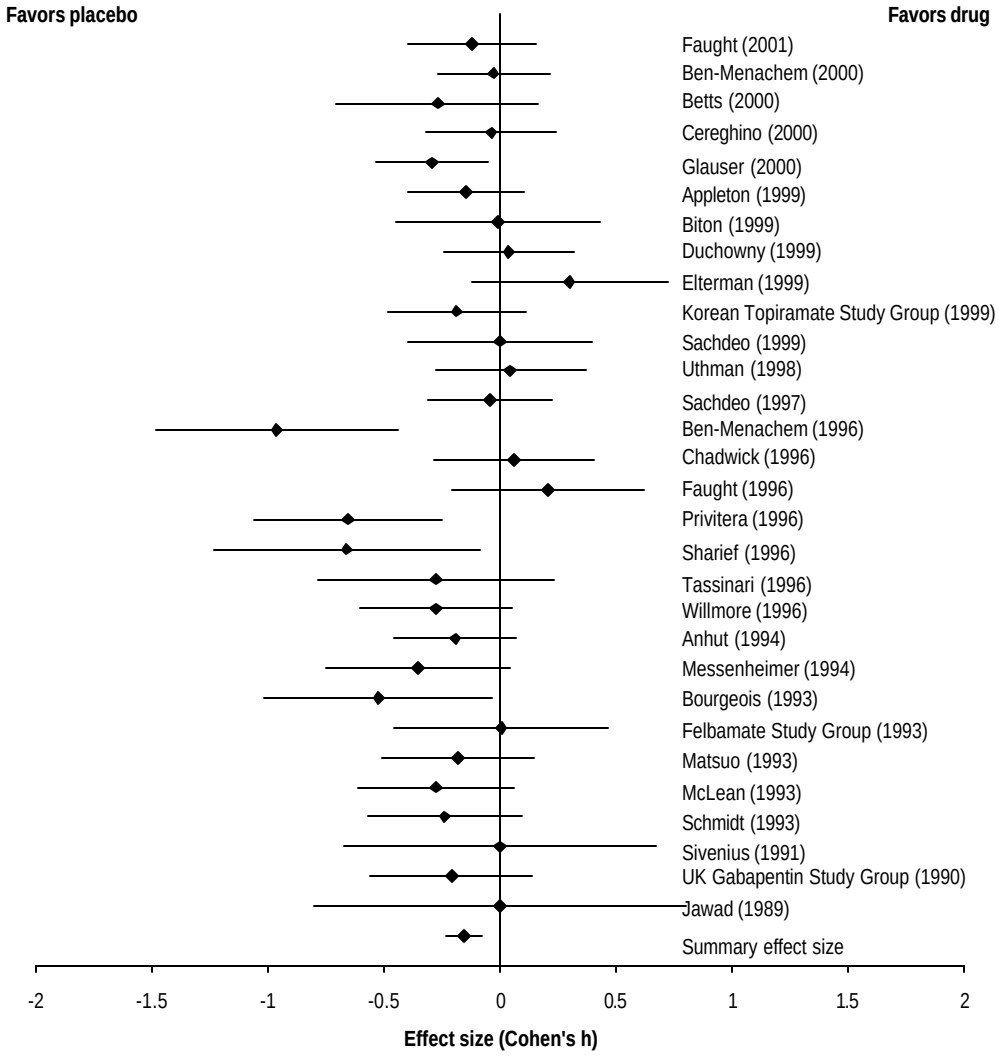
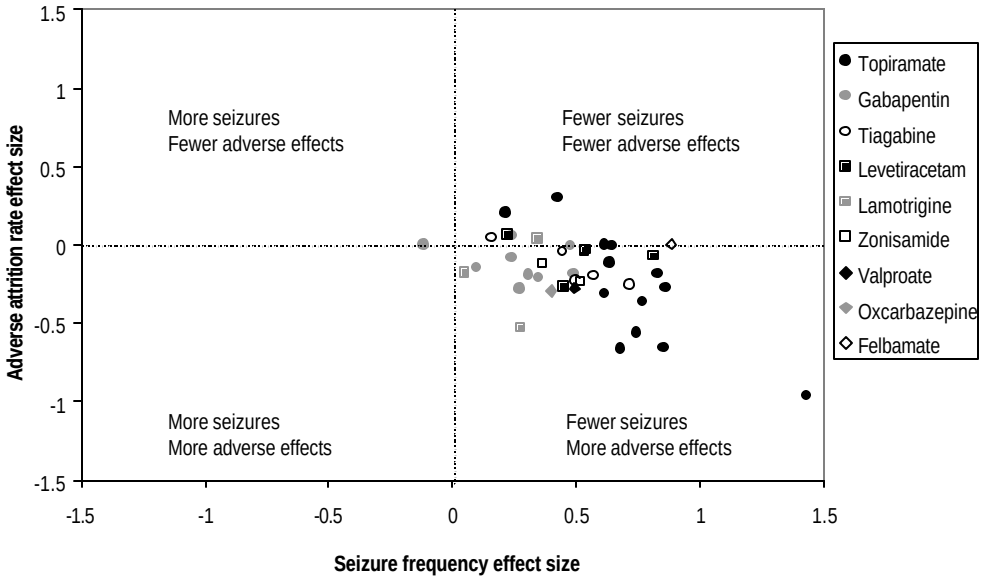


Figure 27. Tradeoff between seizure frequency and adverse effects



Optimization of Current Drug Therapy

The previous two parts of the present question addressed strategies related to the use of new AEDs or new combinations of AEDs in patients with treatment-resistant epilepsy. In this section, we assess strategies designed to optimize the effectiveness of a patient's current drug regimen. Strategies designed to optimize current drug therapy seek to improve patient outcomes by either; (1) reducing seizure frequency without increasing the incidence (or intensity) of the side effects associated with AED treatment or, (2) by reducing the side effects of AED without increasing seizure frequency, seizure severity, or the onset of a new seizure type. Ideally, a drug regimen would both decrease seizure frequency and reduce side effects. However, as shown above, this rarely occurs in patients with treatment-resistant epilepsy, and a trade-off exists between the intensity of drug treatment and the incidence and severity of associated side effects. Thus, in order for an optimization strategy to be of value, it must either lead to reductions in seizure frequency or reductions in side effects (and/or improvements in quality of life, cognitive function, and mood) while not leading to increases in the other.

Published literature describes three different methods for optimizing drug therapy in patients with treatment-resistant epilepsy: (1) increasing the dose of the current drug (or drugs) to maximumtolerable levels, (2) modifying the frequency of dosing, and (3) reducing the total number of drugs. In this subquestion, we evaluate the literature pertaining to all three of these strategies.

Number of studies addressing each drug optimization strategy

Eleven included articles addressed one of the three drug optimization strategies presented above (Evidence Table 87). Eight of the eleven articles described studies that assessed the drug reduction strategy, two articles described studies that assessed themaximum tolerable dose strategy, and one article described a study that assessed the dosing frequency strategy.

As discussed in the Methodology section of this report, only treatment strategies that were addressed by at least five acceptable studies were evaluated. One of the three drug -optimization strategies, the drug reduction strategy, was addressed by enough studies to meet this criterion. The prerequisite number of studies did not address the remaining two strategies, even when the inclusion criteria were relaxed to allow for the inclusion of retrospective studies. Consequently, we do not include further information concerning implementation of either the maximal tolerable dose or the optimized dosing frequency strategies.

Drug Reduction Strategy

The goal of drug reduction strategy is to reduce the number of AEDs without increasing seizure frequency above some unacceptable level. As implied above, this strategy is based on the (reasonable) assumption that reducing the number of AEDs taken by a patient should result in reduced side effects, which will lead to increased quality of life, improved cognitive function, improvements in mood, and reduced costs.[1]

[1] An evaluation of costs associated with the treatments assessed in this report is beyond the scope of the current report.

Excluded articles

We excluded one of the eight articles that both met the general and question-specific inclusion criteria. This article and the reason for its exclusion are presented in Evidence Table 88.

Evidence base

After the exclusion of one study, seven articles remained.[120-126] These studies included data collected from 311 patients. Details of the studies described by these articles are presented in Evidence Tables 89 through 98.

All of the studies included in the present evidence base were prospective, three were controlled,[121,122,126] and the remaining four utilized a case series design.[120,123-125] Two of the three controlled trials were single-blinded,[m] and not randomized.[121,126] The remaining controlled trial was randomized and double blinded.[122] However, this study randomized patients within the drug reduction arm to drug reduction at either a slow rate or a fast rate, and patients were not randomly allocated to the two principal arms of the study, the drug reduction and the control arms. Since the primary objective of the present subquestion is to determine whether implementation of the drug reduction strategy leads to improved patient outcomes, this study, for the purposes of this section of the report, must be considered a nonrandomized controlled trial.

Design and conduct of included studies

This section presents the findings of our systematic assessment of the quality of the seven studies that assessed the effectiveness of the drug reduction strategy. This systematic assessment consisted of an appraisal of both the internal and external validity of each included study.

Internal validity

Measurement bias, regression to the mean, extraneous event bias, and sample specification bias were potentially present in all seven studies. Patient reporting bias and investigator reporting bias may have been present in six studies. Selection bias potential affected the three controlled trials. Sampling bias may have been present in the six studies that did not report how patients were enrolled in the study. Attrition bias was a potential factor in one study with more than a 10 percent attrition rate. These potential biases with respect to this question are discussed in detail in Appendix B.

External validity

Details of the patient characteristics that were reported by each of the articles in the present evidence base are presented in Evidence Tables 93 through 98.

The range of ages covered by each of the studies in the present evidence base tended to be broad, and, although no study exclusively enrolled adults, six of the seven studies enrolled mainly adults.[120-122,124-126] The remaining study enrolled solely children.[123] We were unable to determine the upper age of the patients in the study described by Callaghan, O'Dwyer, and Keating[124] because of inconsistent reporting (the reported mean patient age was 26 years but the range was reported as 6 to 24 years). The duration of epilepsy suffered by the patients in the included studies varied considerably with durations ranging from less than 1 year to well over 60 years.

---

[m] Investigators, but not patients are blinded to treatment regimen.

The proportion of females in each of the studies included in the present evidence base varied considerably between studies (from under 25 percent to over 80 percent). One study did not report the sex ratio (Schmidt[125]).

Two of the seven studies included for this question did not restrict their patient sample by age or seizure type.[124,126] The remaining five studies enrolled patients because they were considered representative of a specific subpopulation of patients with treatment-resistant epilepsy. Three of the studies recruited institutionalized patients with severe epilepsy and multiple cognitive and/or behavioral deficits.[120-122] Two of the studies recruited patients because they suffered from a particular seizure type.[120,125]

Although all studies included in the present evidence base investigated a common optimization strategy (the drug reduction strategy), each study did so in a different way. For example, the aim of Specht, Boenigk, Wolf, et al.[123] was to evaluate the effects of the removal of all patients in their study from a single drug (clonazepam), whereas the aim of the study by Callaghan, O'Dwyer, and Keating[124] was to reduce all patients in their study from polytherapy to monotherapy or, if this was not possible, to two AEDs. Because the evidence base pertaining to drug reduction strategy was small, quantitative analyses could not be performed that would indicate whether the findings of the individual studies were similar. Without evidence to demonstrate such similarity, conclusions about the effectiveness of the drug reduction strategy as a whole are not possible. Instead, each variation of the drug reduction strategy must be considered separately, and the findings of each individual study may only be generalized to patients with characteristics similar to those included in that study.

Synthesis of study results

The assessment of study quality presented above indicates that, given the present evidence base, definitive conclusions cannot be drawn about whether implementation of the drug reduction strategy is effective in improving outcome in patients with treatment-resistant epilepsy. Acknowledging this, we have instead evaluated the available data with the aim of determining whether the implementation of this strategy may plausibly be effective in improving outcome among patients with treatment-resistant epilepsy.

Not all of the outcomes listed by the Technical Expert Panel (see question-specific inclusion criteria above) were reported on in all of the articles in the present evidence base. The reported outcome measures and the articles that contained data pertaining to these outcome measures are presented in Table 15.

Seizure frequency outcomes

As stated previously, the goal of the drug reduction strategy is to remove a drug (or drugs), thereby reducing the occurrence of (or the risk for) adverse effects associated with the use of AEDs. This goal must be accomplished without increasing seizure frequency to unacceptably high levels. Although reductions in seizure frequency are desirable and may indeed occur, they are not, in this instance, to be expected. Consequently, studies needed only to demonstrate that implementation of the drug reduction strategy resulted in other benefits such as reductions in adverse events, increases in cognitive function, increases in quality of life, reduced cost, etc. This means that trials that evaluate changes in seizure frequency that result from drug reduction strategy must also demonstrate, through hypothesis testing, that clinically meaningful increases in seizure frequency did not occur.

In such trials, which are akin to studies of therapeutic equivalence,[n] classical hypothesis testing (with the usual null hypothesis that there is no difference between the interventions) is inappropriate.[127-131] This is because the desired result of a bioequivalence study would be to prove the null hypothesis by showing that no increases in seizure frequency occurred in the treatment group when compared to the comparison group.[o] An alternative hypothesis allows meaningful statistical analyses to be performed. In this instance, the alternative hypothesis is that seizure frequency increases in the treatment group will be less than a prespecified level, δ, above the seizure frequencies seen in the comparison group (H_A: $X_{DRS}-X_c<\delta$; where H_A = alternative hypothesis). Thus, to demonstrate that implementation of a drug reduction strategy does not lead to increases in seizure frequency, any difference in seizure frequency between the treatment group and the comparison group (along with its 95 percent confidence intervals) must fall entirely below δ. Confidence intervals that extend above δ indicate that the alternative hypothesis has not been refuted and implementation of the strategy may lead to increases in seizure frequency.

As is the case with conventional hypothesis testing, a study should be designed with adequate power to avoid the possibility of making Type II statistical errors. As shown in Table 16, when performing hypothesis testing using the alternative hypotheses, the "standard" rules of a Type I error and a Type II error become reversed. Thus, a Type I error is made if the difference $X_{DRS}-X_c$ is less than δ when, in fact, the difference is greater than or equal to δ, and a Type II error is made when the difference is greater than or equal to δ when it is actually less than δ.

Given the information above, the seizure outcomes of importance in this evaluation are those that assess increases in seizure frequency. Outcomes that asses s improvements in seizure frequency (proportion of patients seizure-free, proportion of patients achieving a greater than 50 percent decrease in seizure frequency), though interesting, are of secondary importance. As a result, we have focused this sectionof the report on three seizure frequency outcomes (absolute seizure frequency, percentage change in seizure frequency, and proportion of patients with an increase in seizure frequency). Data pertaining to the remaining seizure frequency outcomes are summarized in Evidence Table 99 but are not discussed further.

Absolute seizure frequency. Two of the three controlled trials included for this subquestion presented data on (mean or median) absolute seizure frequency. Two studies are too few to allow a quantitative analysis to be performed. As a result, we present the findings of our semi - quantitative analysis of the available data. These data are presented in Evidence Table 99.

As discussed above, to demonstrate that seizure frequency does not increase in patients using a drug reduction strategy, the strategy must be shown not to cause clinically important increases in seizure frequency ($X_{DRS}- X_{c} \leq \delta$). This requires the authors to explicitly state what they consider a meaningful increase in seizure frequency (δ). Based on this seizure frequency, they should then state the size of the study (power) necessary to overturn the null hypothesis that seizure frequency in patients who received drug reduction will increase above this predefined seizure frequency.[p] Neither of the two controlled trials that reported this outcome stated what

[n] These studies are also known as studies on noninferiority or studies of bioequivalence

[o] In other words, trying to prove that $X_{DRS}-X_c = 0$, where X_c = mean seizure frequency in control group and, X_{DRS} = mean seizure frequency in the drug reduction group

[p] Power calculations for the testing the null hypothesis of a study of bioequivalence have been developed[127,391] and are specified in terms of a one-sided confidence interval for the difference $X_{DRS}-X_c$, with a specified probability $1-\beta$ that the interval will not include the predefined seizure increase.

they considered to be a clinically important difference in seizure frequency, nor did they perform a power analysis.[q]

All of the statistical analyses presented in these two articles tested the traditional null hypothesis that no between-groups differences in seizure frequency exist. Thus, their analyses essentially attempted to prove the null hypothesis that there was no change in seizure frequency. As discussed above, this is inappropriate.

Because the investigators did not determine the power of their study and because their statistical analyses were not appropriate for the clinical question of interest, the seizure frequency analyses in the articles are of limited value. However, summary data from these studies may still be used to provide some useful information. This can be accomplished by calculating the mean difference in seizure frequency (and its CI) between the drug-reduction group and the control group for each study. The upper CI of this difference can then be used to determine the maximum magnitude of increase in seizure frequency that will not lead to the alternative hypothesis (H_A: $X_{DRS} - X_c \leq \delta$) being accepted over the null hypothesis (H_0: $X_{DRS} - X_c > \delta$). Such an approach, however, requires that the study report seizure frequency data in such a way that a difference can be calculated.

Duncan, Shorvon, and Trimble[122] summarized their seizure frequency data in terms of mean seizure frequencies along with its standard deviation. The range cannot be used to calculate a valid standardized between-groups difference. No other measures of dispersion were reported. As a result, the seizure frequency data presented by Duncan, Shorvon, and Trimble[122] cannot be used to determine whether implementation of the drug reduction strategy leads to clinically important increases in seizure frequency.

Unlike Duncan, Shorvon, and Trimble,[122] Thompson and Trimble[126] presented mean seizure frequency along with its dispersion (expressed in terms of standard deviations). However, the analysis described above still cannot be performed, because the technique is sensitive to pretreatment differences in seizure frequency. Although no statistically significant between-groups differences in seizure frequency data were detected at pretreatment, a between-groups differences in seizure frequency at baseline did exist, and these differences were large enough to lead to biased posttreatment effect size estimates. For example, the mean pretreatment frequency for partial seizures in the drug reduction group was 21.1 (SD: 34.6) seizures per week compared to 6.8 (SD: 9.7) per week in the control group. Thus, patients in the drug reduction arm were experiencing more than three times the number of seizures per week compared to the patients in the control arm at study onset. Consequently, the study is biased against finding that the implementation of the drug reduction strategy will lead to increases in seizure frequency.

To summarize, the data from the currently available controlled trials could not be used to draw evidence-based conclusions about whether or not implementation of the drug reduction strategy leads to increases in seizure frequency.

Although none of the four included case series reported on this outcome, two studies did present individual patient data that allowed us to summarize the seizure frequency data both pre and post implementation of the drug reduction strategy. These data are presented in Evidence Table 99. They do not suggest that seizure frequencies increase following implementation of the drug reduction strategy. However, because these data originate from two uncontrolled studies and, because seizure frequency in patients with treatment-resistant epilepsy commonly demonstrates regression to the mean (see Methodology section),[23,132] this observation does not

[q] Because there is no consensus in the literature about what defines a clinically important increase in seizure frequency, we are precluded from performing our own power analyses of these data.

provide convincing evidence to support the contention that implementation of the drug reduction strategy does not lead to increases in seizure frequency.

As will be seen in the following sections, other seizure frequency-based outcomes suggest that drug reduction strategy may lead to large increases in seizure frequency in some patients.

Mean or median percentage change in seizure frequency. None of the three included controlled trials presented data on the percentage change in seizure frequency following implementation of the drug reduction strategy. Thus, conclusions about this outcome can only be based on case series data.

Two of the four case-series studies presented individual patient data that allowed us to assess this outcome. These data show that the median percentage change in seizure frequency from baseline was –0.09 percent[r] (Range: -100 percent to 412 percent) in the study by Specht, Boenigk, Wolf, et al.[123] and –0.12 percent (Range: -100 percent to 2,678 percent) in the study by Schmidt.[125] In both studies, more than 50 percent of the patients experienced a reduction in seizure frequency following implementation of the drug reduction strategy, and just under 50 percent of the patients experienced an increase in seizure frequency from baseline (43 percent of patients in the study by Specht, Boenigk, Wolf, et al.[123] and 47 percent of patients in the study of Schmidt[125]). The proportion of patients who experience an increase in seizure frequency from baseline is addressed in more detail in the following section of the report.

Thus, these data suggest that a high proportion of patients (close to 50 percent) may experience increases in seizure frequency following the implementation of the drug reduction strategy. The data also suggest that some patients may experience decreases in seizure frequency. Given that regression to the mean is known to influence seizure frequency data,[23,132] some of these observed reductions in seizure frequency were probably a manifestation of this bias. The only other possible explanation is that the withdrawn drug was somehow causing seizures.

Proportion of patients with an increase in seizure frequency. None of the three included controlled trials presented data on the proportion of patients with an increase in seizure frequency following implementation of the drug reduction strategy. Thus, conclusions on this outcome can only be based on data from case series.

One of the four case series presented data on increases in seizure frequency. Callaghan, O'Dwyer, and Keating[124] reported that three of the 35 patients (9 percent) included in their study demonstrated an increase in seizure frequency. This information, however, is of limited value because the authors did not define what they meant by "worse." Consequently, the magnitude of the reported increase in seizure frequency in these three patients cannot be determined, and no conclusions can be drawn as to whether these increases were clinically important.

Two other articles presented individual patient data that allowed us to calculate the proportion of patients with an increase in seizure frequency (Schmidt[125] and Specht, Boenigk, Wolf, et al.[123]). Because the magnitude of a clinically important increase in seizure frequency remains ambiguous, we believed that arbitrarily reporting the proportion of patients above any single frequency was inappropriate. Instead, we calculated the proportion of patients that demonstrated increases in seizure frequency above a series of percentage increases from baseline (thresholds). These data, which are presented in Evidence Table 99, are summarized in Figure 28.

---

[r] By convention a negative sign is used to indicate that seizure frequency has increased. This is because the primary outcome of interest in a treatment trial is usually the percentage *reduction* in seizure frequency. However, in this case we are interested in *increases* in seizure frequency. Consequently, we use a minus sign to indicate a reduction in seizure frequency. Thus a percentage change in seizure frequency of -100 percent indicates that a patient is seizure-free.

This figure shows that a statistically significant proportion of patients in both case series exhibited large (>100 percent) increases in seizure frequency when compared to baseline (27.8 percent of patients in the study by Schmidt[125] and 8.6 percent in the study by Specht, Boenigk, Wolf, et al.[123]). In neither study did these patients have unusually low seizure frequency rates at the onset of the study, suggesting that implementation of the drug reduction strategy will result in increased seizure frequency in a significant proportion of patients.

Mood

Two of the three controlled trials presented data on changes in mood following the implementation of a drug reduction strategy. Two studies are too few to allow a quantitative analysis of the available data to be performed. As a result, we present the findings of our semi-quantitative analysis of the available data. These data are presented in Evidence Table 99.

Both Thompson and Trimble[126] and Duncan, Shorvon, and Trimble[122] presented mood data collected using two validated self-administered psychometric instruments. These instruments were the Middlesex Hospital Questionnaire (MHQ) and the Mood Adjective Checklist (MACL).

The MHQ is a self-administered questionnaire that measures six domains and provides a composite score. This instrument is commonly used as an aid in the diagnosis of clinical depression. The six domains that are assessed include: Free-floating anxiety (F-FA), phobic anxiety (PHO), obsessive-compulsive (OBS), somatic anxiety (SOM), depressive traits (DEP), and hysteric (HYS) traits. Although Duncan, Shorvon, and Trimble[122] presented data for all six domains, Thompson and Trimble[126] only reported on two (F-FA and DEP). Neither Duncan, Shorvon, and Trimble[122] nor Thompson and Trimble[126] found a statistically significant between-groups difference in any of the domains measured using the MHQ following completion of drug reduction. Nor were any trends in the data detected that would indicate that mood either improved or deteriorated following drug reduction.

The MACL is a standardized scale commonly used to detect alterations in mood across five domains. These domains provide measures of anxiety, fatigue, hostility, vigor, and depression, along with a composite score. Although both studies measured mood alterations using this instrument, only Thompson and Trimble[126] presented relevant data in their article. Again, as was the case with reporting of the data obtained using the MHQ, Thompson and Trimble[126] did not report data for all of the measured domains (in this case data for the domain "hostility" was not reported) and no explanation was provided as to why this was the case. Analysis of data abstracted from Thompson and Trimble[126] did not find statistically significant between-group differences in any of the domains measured using the MACL. Nor were any trends in the data detected that would indicate that mood either improved or deteriorated following drug reduction. This finding was corroborated by Duncan, Shorvon, and Trimble[122] who reported that, "There were no statistically significant differences between the four groups[s] on the anxiety, depression, fatigue, vigor, or hostility subscales of the Mood Adjective Checklist."[122]

Cognitive function

All three of the controlled trials included for the present subquestion presented data on changes in cognitive function following implementation of a drug reduction strategy when compared to a control group comprised of patients who were maintained on their current

s Duncan, Shorvon, and Trimble[122] are referring to the four arms of their study (control group, phenytoin-removed group, carbamazipine-removed group, and sodium valproate-removed arm.

polytherapy drug regimen. Three studies are too few to allow a quantitative analysis to be performed. As a result, we present the findings of our semi-quantitative analysis of the available data. These data are presented in Evidence Table 99.

All three studies measured cognitive function using a series of standardized clinical tests. These tests included tests of concentration and attention, memory, and tests of psychomotor performance.

Tests of concentration/attention. All three controlled trials measured concentration/attention before and after the implementation of the drug reduction strategy. These data are summarized in Figure 29. May, Bulmahn, Wohlhuter et al.[121] used the d2 test and the modified version of the Frankfurt Concentration Test for Children (FCTC). Duncan, Shorvon, and Trimble[122] used the Letter Cancellation Task (LCT), and Thompson and Trimble[126] used the Stroop test (ST) and a test of visual scanning speed (VSS).

Data from only one of the three studies, Duncan, Shorvon, and Trimble,[122] suggested that concentration improvement was statistically significant among patients who had undergone drug-reduction when compared to patients in the control group. The only statistically significant posttreatment benefit was seen in patients who were removed from sodium valproate ($t = 4.245$; $p = 0.000108$[t]) followed by patients who were removed from phenytoin ($t = 1.965$; $p = 0.056$). Assessment of the pretreatment LCT data, however, suggested the presence of selection bias, with patients who were removed from sodium valproate having statistically significantly higher baseline LCT scores compared to those in the control group ($t = 3.404$; $p = 0.00140$). Thus, the posttreatment between-groups difference was essentially the same as the pretreatment difference.

No such bias was found to have affected the LCT scores on removal of phenytoin and these data suggest that removal of phenytoin may lead to an improvement in concentration/attention in some patients. However, interpretation of the importance of a mean improvement of 18 points is difficult because the authors did not indicate if such a between-groups difference was clinically important.

May, Bulmahn, Wohlhuter et al.[121] argued that their FCTC data showed a statistically significant between-groups difference in patients in the drug reduction arm (all of whom had phenytoin removed). Figure 30 shows graphically their reported pre-and posttreatment FCTC data. The data, as presented in the article, can lead to different conclusions. Changes in FCTC score seen from baseline between the two arms of their study were compared instead of the posttreatment data alone. Because FCTC scores improved in the reduction group and declined in the control group, the comparison found a significant between-groups difference. As shown in Figure 30, the changes in FCTC could reasonably be argued to be due to regression to the mean rather than an effect of treatment.

Memory. All three controlled trials measured memory before and after the implementation of the drug reduction strategy. May, Bulmahn, Wohlhuter et al.[121] measured memory using a digit span and an immediate recall of pictures, and a delayed-recall task at the end of the test session that were taken from the Lern- und Gedachtnis-test (LGT-3). Duncan, Shorvon, and Trimble[122] measured memory using a digit span task derived from the Wechsler Adult Intelligence Scale. Thompson and Trimble[126] used an immediate-recall and delayed-recall of pictures task that they developed and validated themselves.[133-135] Data on the effects of drug reduction on memory collected in these studies are summarized in Figure 31.

When considered as a whole, these data do not provide evidence that drug reduction leads to improved memory. Although the data from Duncan, Shorvon, and Trimble[122] suggest that drug

[t] Statistical analysis performed by ECRI using available data

reduction may lead to statistically significant improvements in short-term memory (as measured by digital scanning backwards) in some patients who were removed from sodium valproate, these results may be biased. This is illustrated by Figure 32, which shows that a pretreatment difference in short-term memory existed between patients removed from sodium valproate and patients in the control group. Although this difference was not statistically significant ($t = 1.842$; $p = 0.072$; Hedges' $d = 0.53$; CI: –0.05 to 1.11), it is large enough to have biased the posttreatment between-groups effect size data. Indeed, Duncan, Shorvon, and Trimble[122] reported that their statistical analyses showed that removal of phenytoin, carbamazepine, or sodium valproate did not lead to improvements in short-term memory.

Psychomotor function. All three controlled trials measured psychomotor function before and after drug reduction. May, Bulmahn, Wohlhuter et al.[121] used the pegboard, a pursuit rotor, and tapping. Duncan, Shorvon, and Trimble[122] used tapping alone, as did Thompson and Trimble.[126] Data on the effects of drug reduction on psychomotor function in these studies are summarized in Figure 33. These posttreatment, between-groups effect sizes do not provide evidence that implementation of the drug reduction strategy results in improved psychomotor function.

Again, these findings contradict the conclusions drawn by the authors. May, Bulmahn, Wohlhuter et al.[121] reported that their data demonstrated a statistically significant improvement in psychomotor function when measured using finger tapping (with the dominant hand) and pursuit rotor failure (again using the dominant hand). Figure 34 shows graphically the pre-and posttreatment data reported in these studies. These data can lead to different conclusions. As stated above, the discrepancy is due to comparing the changes in psychomotor function from baseline between the two arms of the study instead of comparing the posttreatment data alone. As shown by such a comparison in Figure 34, changes in psychomotor function in the drug reduction arm of the study could reasonably be argued to be due to regression to the mean rather than treatment.

Adverse events

Identification of treatment-related morbidities can only be achieved by comparing reported adverse event rates in patients who underwent drug reduction against a control group comprised of patients who were maintained on their current treatment regimen. Although case series identify possible adverse events that may be associated with a treatment, their data cannot be used to draw evidence-based conclusions about whether these adverse events are a consequence of the drug reduction strategy. As a result, we only considered adverse events data abstracted from controlled trials in this section of the report. However, adverse events data abstracted from the four case series included in the present evidence base are tabled in Evidence Table 100.

One of the three controlled trials reported relevant data on adverse events. Patients included in the drug reduction arm of the study by Duncan, Shorvon, and Trimble[122] suffered no additional adverse events compared to patients included in the control arm. Thus, although the patients undergoing drug reduction did experience some adverse events, these adverse events cannot be attributed to drug reduction strategy used in this study.

Mortality

No patients were reported to have died during any of the seven included studies. Thus, no evidence exists to suggest that implementation of the drug reduction strategy leads to increased mortality.

Table 15. Outcomes in studies of drug reduction

Reference	Reported Seizure Outcomes					Reported Nonseizure Outcomes			
	Difference in Absolute Seizure Frequency	Percent Change in Seizure Frequency	Percent Patients Seizure-free	Percent of Patients With >50 Percent Reduction in Seizure Rate	Percent of Patients With an Increase in Seizure Frequency	Mood	Cognitive Function	Adverse Events	Mortality
Controlled trials performed outside of the United States									
May (1992)[121]	✓[a]						✓		✓
Duncan (1990)[122]	✓					✓[b]	✓[c]	✓[d]	✓
Thompson (1982)[126]	✓					✓	✓		✓
Case series performed in the United States									
Mirza (1993)[120]			✓	✓				✓	✓
Case series performed outside of the United States									
Specht (1989)[123]	✓[e]	✓[e]	✓[e]	✓[e]	✓[e]			✓	✓
Callaghan (1984)[124]			✓	?	✓			✓	✓
Schmidt (1983)[125]	✓[e]	✓[e]	✓[e]	✓[e]	✓[e]			✓	✓
Number of articles addressing outcome	**4[f]**	**2**	**4**	**4**	**3**	**2**	**3**	**5**	**7**

[a] May, Bulmahn, Wolhlhuter et al.[121] reported that no statistically significant between groups differences in seizure frequency were seen but did not present any data
[b] Mood data abstracted from Kendrick, Duncan, and Trimble[136]
[c] Cognitive function data abstracted from Duncan, Shorvon, and Trimble[137]
[d] Adverse events data abstracted from Duncan, Shorvon, and Trimble[138]
[e] Data calculated by ECRI from individual patient data
[f] Does not include May, Bulmahn, Wolhlhuter et al.[121] (see footnote a above).

Table 16. Possible decisions based on hypothesis test

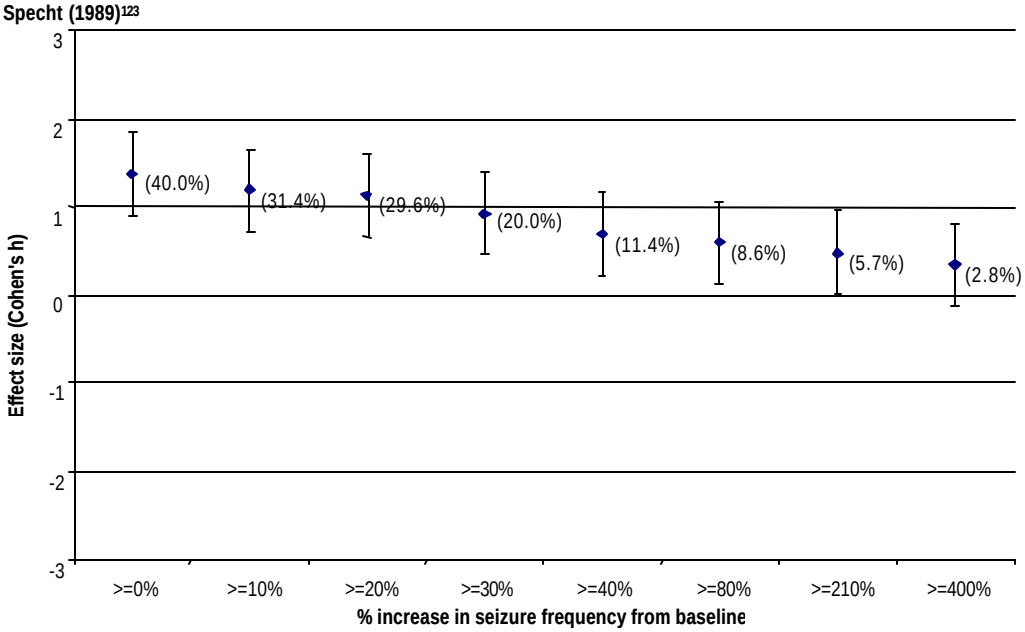
True Difference $X_{DRS} - X_c$	Testing H_0: $X_c = X_{DRS}$		Testing H_0': $X_{DRS} \geq X_c + d$	
	Fail to Reject (good for DRS)	Reject (bad for DRS)	Reject (good for DRS)	Fail to Reject (bad for DRS)
$X_{DRS} - X_c = 0$ (good for DRS)	Correct Decision	Type I error	Correct Decision	Type II error
$X_{DRS} - X_c = \delta$ (bad for DRS)	Type II error	Correct Decision	Type I error	Correct Decision

Adapted from Blackwelder[128]
DRS Drug reduction study
H_0 Null hypothesis (standard)
H_0' Null hypothesis (therapeutic equivalence studies)
X_c Mean seizure frequency in control group
X_{DRS} Mean seizure frequency in drug reduction therapy group
δ Predefined difference in mean seizure frequency above which use of drug reduction study is unacceptable

Figure 28. Increase in seizure frequency and drug reduction strategies

Percentage presented in parentheses is the actual proportion of patients with seizure frequencies greater than the percent increase in seizure frequency shown on the X-axis. The diamond and error bars represent the effect size and 95% CI.

Specht (1989)[123]



Schmidt (1983)[125]

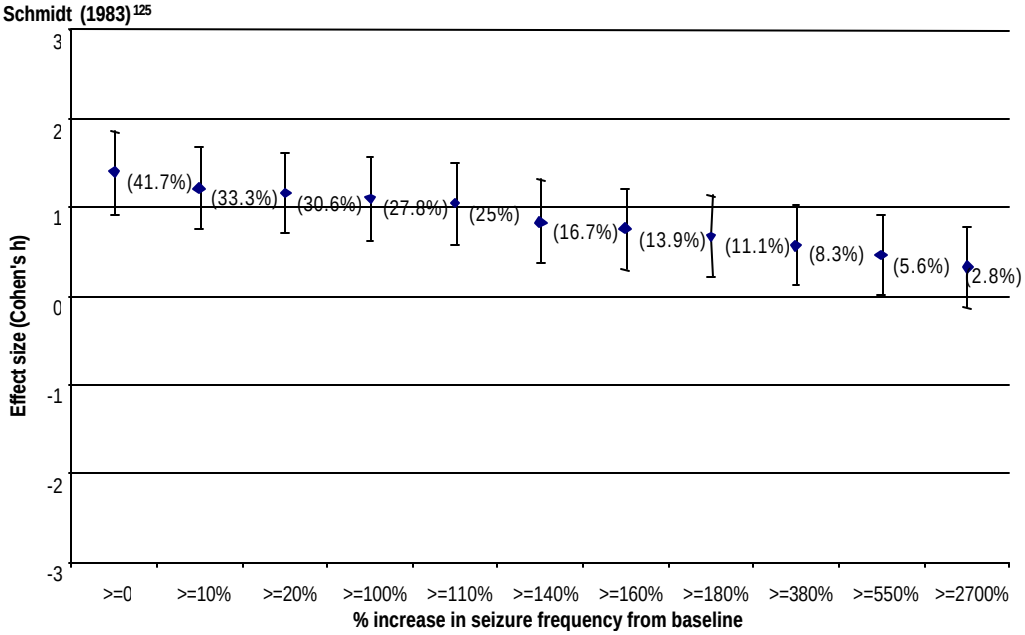
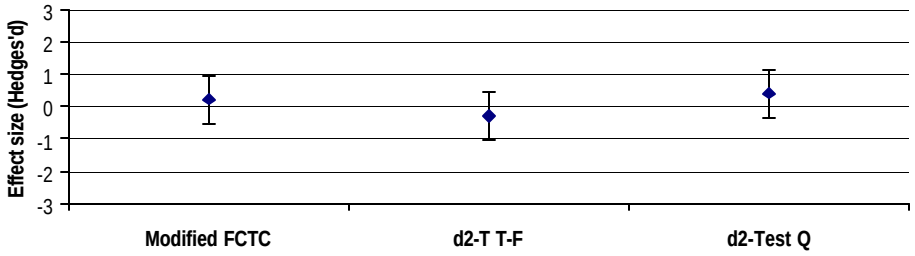
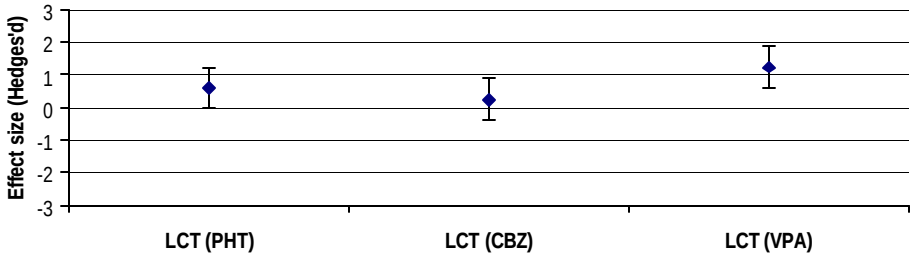


Figure 29. Drug reduction strategies and tests of concentration/attention

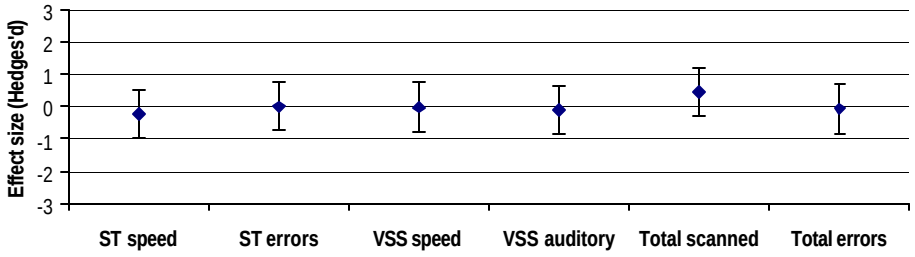
May (1992) [121]



Duncan (1990)[122]



Thompson (1982)[126]



FCCT	Frankfurt Concentration Test for Children
LCT	Letter cancellation task
PHT	Phenytoin vs. Control
CBZ	Carbamezapine vs. Control
VPA	Valprioc acid vs. Control
d2 T T-F	d-2 test total number minus failures
d2 Test Q	d-2 test failure quotient
ST	Stroop Test
VSS	Visual scanning speed

Figure 30. Drug reduction strategies and the Frankfurt Concentration Test for Children

Pre- and posttreatment Frankfurt Concentration Test for Children data from May (1992)[121]

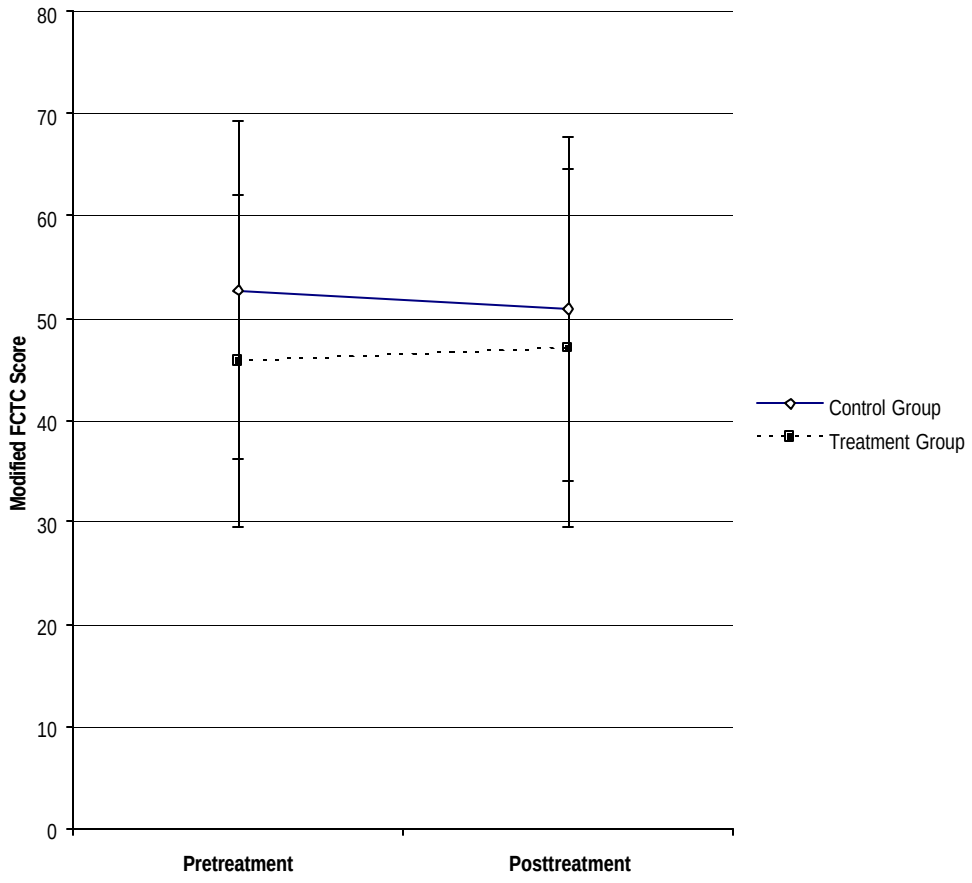
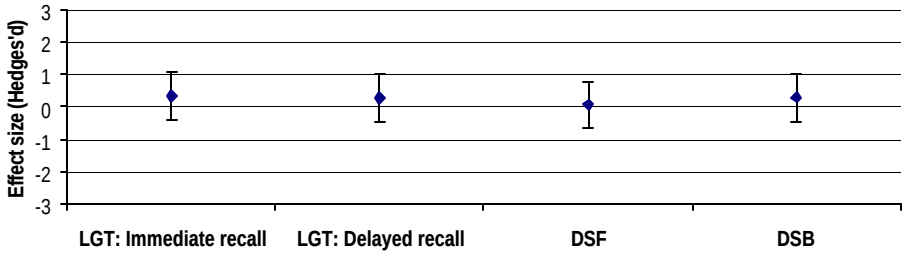
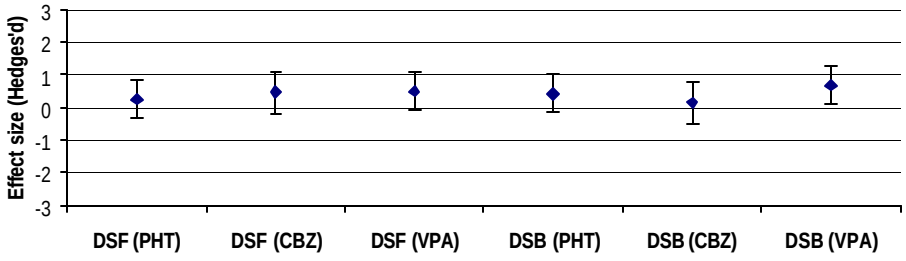


Figure 31. Drug reduction strategies and tests of memory

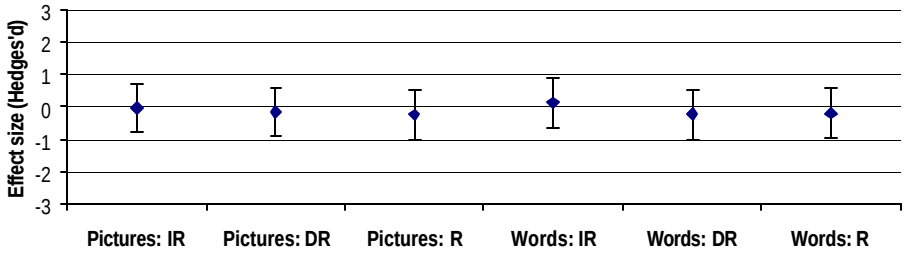
May (1992) [121]



Duncan (1990)[122]



Thompson (1982)[126]



Abbreviations:

Deleted: <sp>

LGT	Lern- und Gedachtnis Test
DSF	Digit scan forwards
DSB	Digit scan backwards
PHT	Phenytoin vs. Control
CBZ	Carbamezapine vs. Control
VPA	Valprioc acid vs. Control
IR	Immediate recall
DR	Delayed recall
R	Recognition

Figure 32. Drug reduction strategies and digital scanning score

Data from Duncan (1990)[122] showing effects of valprioc acid removal on digital scanning score

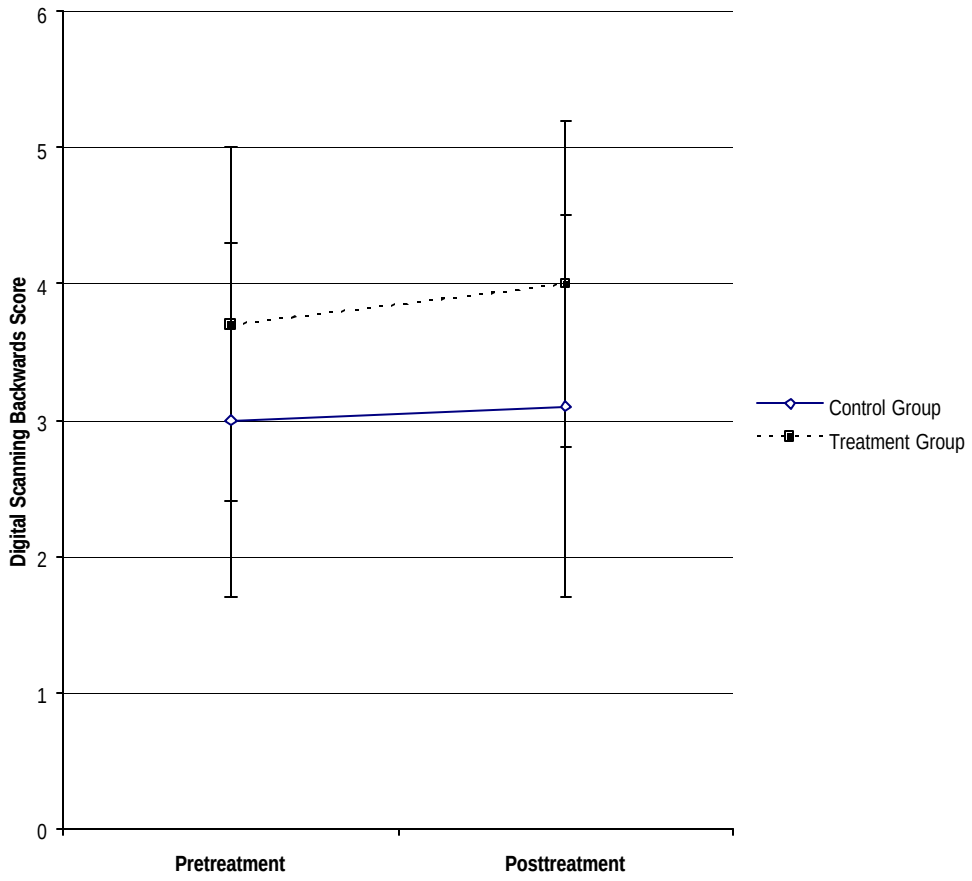
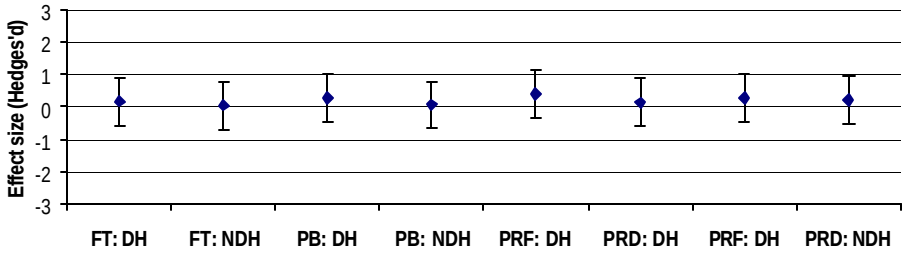
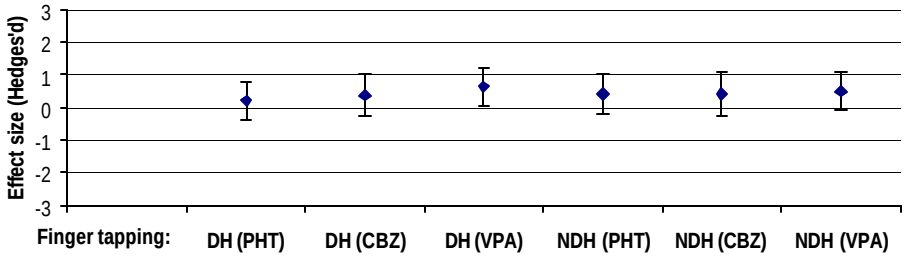


Figure 33. Drug reduction strategies and tests of psychomotor function

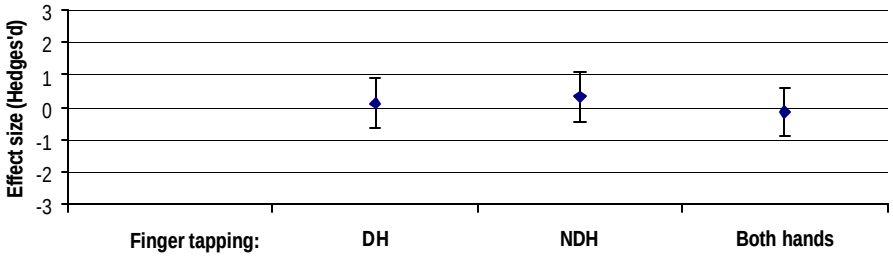
May (1992) [121]



Duncan (1990) [122]



Thompson (1982) [126]

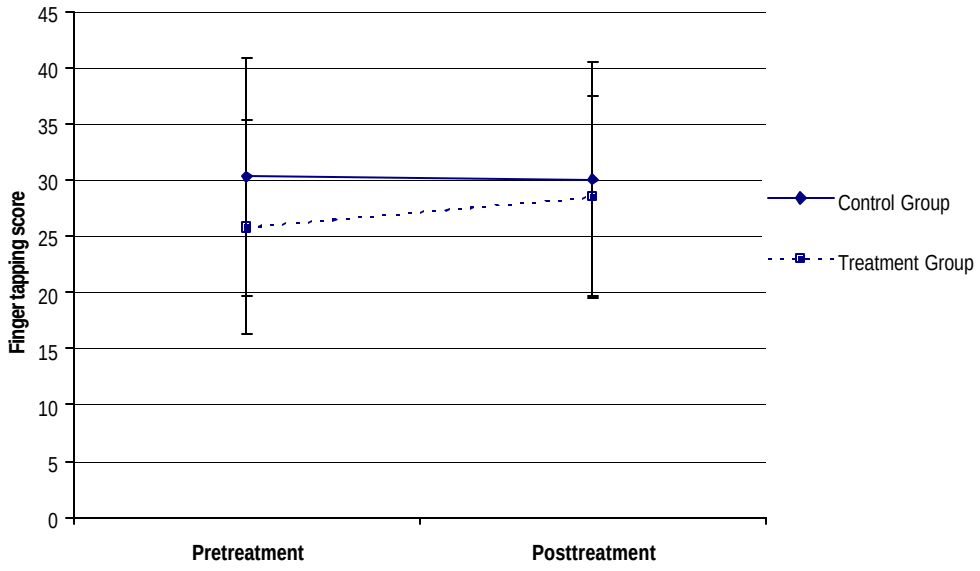


FT	Finger tapping
DH	Dominant hand
NDH	Non-dominant hand
PB	Pegboard
PRF	Pursuit Rotor Failure
PFD	Pursuit Failure Duration
PHT	Phenytoin vs. Control
CBZ	Carbamezapine vs. Control
VPA	Valprioc acid vs. Control

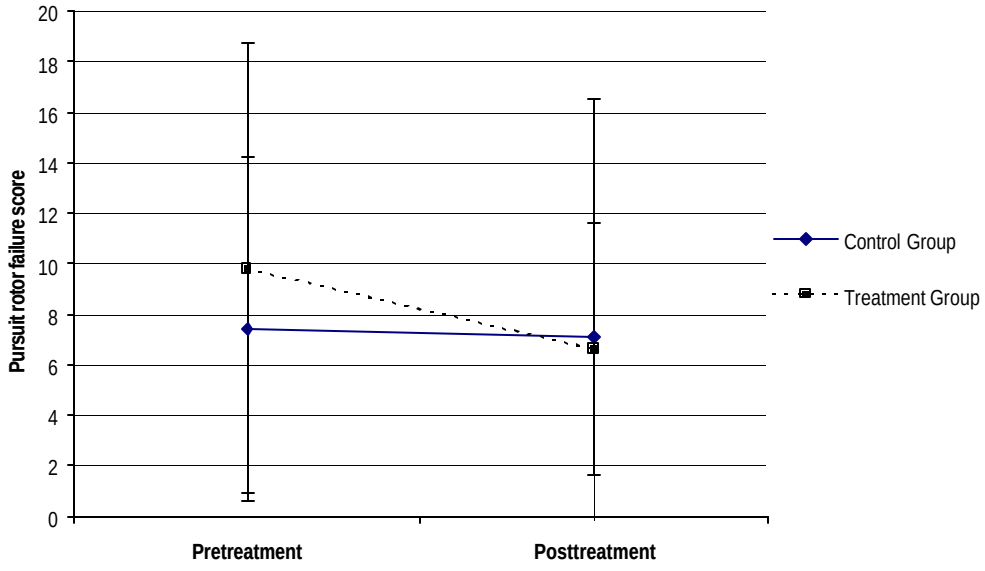
Figure 34. Drug reduction strategies and psychomotor function

Pre-and posttreatment psychomotor function data presented by May (1992) [121]

Finger tapping with dominant hand



Pursuit rotor failure of dominant hand



Comparisons of Drug Strategies

None of the trials that met the inclusion criteria directly compared the drug strategies. Indirectly comparing two of the drug strategies is possible if the patients enrolled in all of the trials of all strategies were similar. The drug reduction strategy cannot be compared with the other two strategies, because of the differing intentions of investigators in these latter trials. The intent of the trials of polytherapy and sequential monotherapy was to reduce seizures without causing adverse effects, whereas the intent of the trials of drug reduction therapy was to reduce the number of drugs without increasing seizures.

To determine whether trials of polytherapy and sequential monotherapy enrolled similar patients, we compared the number of drugs given to patients receiving these strategies. Differences in the number of drugs likely mean that the severity of epilepsy in patients who received polytherapy was different from that in patients who received sequential monotherapy. Among 11 trials of sequential monotherapy that reported this percentage, two (18 percent) reported that more than half of the patients were receiving two or more prior AEDs. By contrast, among the 18 trials of polytherapy that reported this percentage, 16 (89 percent) reported that more than half of the patients were receiving two or more prior AEDs. These percentages are significantly different ($\chi^2(1) = 14.5$, $p = 0.00014$). Thus, patients who received polytherapy had more severe epilepsy compared to patients who received sequential monotherapy. This difference precludes comparison of the *quantitative* results of the two strategies.

A qualitative comparison, however, suggests that polytherapy is clinically preferable to sequential monotherapy. In the section on sequential monotherapy, the evidence indicated that some patients had harmful increases in seizures as a direct result of the treatment, and whether sequential monotherapy caused any patients to become seizure-free could not be determined. In short, sequential monotherapy appeared more likely to be harmful than beneficial. By contrast, the reverse was true for polytherapy. Adding a drug reduced seizures by 50 percent in many patients, whereas adverse effects causing trial exits were rare. By inference, this suggests that polytherapy is preferable to sequential monotherapy.

Further, as discussed above, patients who received polytherapy had been receiving more drugs before the trial, thus they likely had more severe epilepsy. Patients with more severe epilepsy are, by definition, more difficult to treat. Thus, even in a more difficult-to-treat population, polytherapy helped many patients. This finding underscores the qualitative conclusion that polytherapy is preferable to sequential monotherapy for patients with treatment-resistant epilepsy.

Nondrug Treatments

In this section of the Evidence Report, we addressed Key Question #5*: Which methods of nondrug treatment for epilepsy after initial treatment failure lead to improved outcomes for patients with treatment-resistant epilepsy?* This question is divided into two parts. The first part addresses surgical interventions and the second addresses nondrug, nonsurgical interventions.

Surgical Interventions

In this section, we address the efficacy of surgical intervention when treatment with AEDs has failed to produce adequate seizure control. Patients who receive surgery have been determined to be treatment-resistant as part of their presurgical evaluation. For most patients, only a single surgical option will be available due to the nature and location of the lesion or condition responsible for generating their seizures. In patients undergoing temporal lobectomy, hemispherectomy, or corpus callosotomy, some variations in procedures are available.

A list of the specific surgical interventions and outcome measures addressed under surgical interventions is presented in the following section.

Question specific inclusion criteria

We included articles if they met the general inclusion criteria detailed in the Methodology section, and if they met the following question-specific criteria:

1. All seizure frequency outcomes were reported before and 2 or more years after surgery, except for studies of multiple subpial transection (MST). This followup period was recommended by the Expert Panel and Technical Experts, who noted that because surgery is irreversible, relatively long-term data are of primary interest. However, because MST is a relatively new surgical procedure with a limited reference base, a minimum of 6 months followup was used to increase the size of the evidence base for this procedure.
2. The study was published in 1985 or later. This cutoff was used because the Expert Panel and Technical Experts noted that surgical treatments for epilepsy have substantially changed since this date.
3. One of the following specific interventions, as recommended by the Expert Panel and Technical Experts, was examined:
   a. Anterior temporal lobe resection
   b. Frontal lobe resection
   c. Parietal lobe resection
   d. Occipital lobe resection
   e. Cerebral hemispherectomy
   f. Corpus callosotomy
   g. MST separate from or in combination with other resections

Number of studies addressing surgical intervention

The order of the material presented in this section differs from that presented in the discussion of other interventions. This change in organization was necessary because we required a minimum 2-year followup period for most outcome measures used to evaluate

surgery. The only exceptions were outcomes for mood (depression and psychosis), cognitive function (IQ and memory), and complications and mortality related to surgery. We shortened the minimum required followup time for these outcomes because they may manifest themselves relatively early after surgery.

Different studies make up the evidence base for each outcome. We will separately discuss each outcome and its specific evidence base under each surgical intervention rather than examining all of the studies in the evidence base for a single intervention. This section, number of studies addressing surgical intervention, presents an overview of all of the studies meeting our inclusion criteria for each surgical intervention examined under Key Question #5. The actual evidence base for each intervention and outcome will be discussed separately later in this report.

One hundred and seventy-nine studies met our inclusion criteria for surgical intervention. We provide a listing of each study meeting the inclusion criteria for each surgical intervention in Evidence Table 102 and a summary in Table 17. Only two studies each were found to meet our inclusion criteria for parietal lobe and occipital lobe surgery for the treatment of epilepsy. Consequently, we did not assess these interventions.

Evidence Tables 103 to 108 provide general information on each of the studies examined in this report organized into tables according to surgical intervention or reporting of control patients. The information in these tables includes the years during which the studies were conducted, the country in which the study was conducted, the primary center where the surgery was performed, if the study was conducted in multiple centers, whether patients were selected retrospectively or prospectively, and the study design.

Temporal Lobe Surgery

Temporal lobe surgery is intended to eliminate complex partial seizures by removing the lesion or epileptogenic area responsible for the development of these seizures. Complex partial seizures with or without secondary generalization are the most common seizure type associated with temporal lobe epilepsy.[139] The second most common seizure type is a simple partial seizure, which is commonly experienced as the patient's typical aura.

Temporal lobe surgery candidates constitute the largest group of epilepsy surgery patients.[140] Preoperative evaluation determines the type of lesion (tumor, vascular malformation, mesial temporal sclerosis, or other known or unknown etiology). The actual procedure depends on the location of the lesion (deep or superficial) and the extent to which tissue is to be removed.[141-143] An en bloc anterior temporal lobectomy is a standardized operative procedure in which 4.5 to 5.0 cm of the anterior lateral temporal lobe neocortex is removed along with the amygdala, the anterior aspect of the parahippocampal gyrus, and the hippocampus in the medial portion of the temporal lobe. Neocortical lesionectomy is used when the lesion, usually a tumor or vascular malformation, is contained entirely in the neocortex of the temporal lobe. Selective amygdalohippocampectomy (AH) involves the removal of the amygdala and hippocampus only. Intraoperative EEG readings may be used to "tailor" the extent of tissue resection by defining a zone of frequent interictal spiking. The use of this technique may result in more or less tissue being removed compared to the "standard" approach. Another modification to the standard approach is to remove less than 4.5 cm of the anterior temporal lobe and is referred to as "partial" resection. The Evidence Tables pertaining to temporal lobe surgery will refer to these procedures as standard, tailored, partial, amygdalohippocampectomy, and neocortex.

Seizure-free

Several outcome measurements examined in other questions of this report, such as changes in the proportion of patients experiencing at least a 50 percent reduction in seizure frequency, are not included in our examination of surgical intervention because they are rarely (if ever) reported in studies of epilepsy surgery.

Excluded studies

We excluded one study of temporal lobe surgery reporting seizure-free outcome measures from the evidence base because of poor quality. This study and the reason for its exclusion are listed in Evidence Table 109.

Evidence base

Among the 105 studies of temporal lobe surgery meeting our inclusion criteria, 73 reported some sort of seizure-free outcome measurement. Studies of temporal lobe surgery used four different outcome measurements when reporting a patient as "seizure-free." Each outcome measurement results in a different set of patients being considered "seizure-free" and therefore the data collected under each outcome measurement must be evaluated separately.

The most often used outcome measurement among the 73 studies in our evidence base was Engel class I, which was reported in 33 studies (Table 18). Engel class I is part of a four-part system for evaluating the success of surgery in patients with epilepsy.[144] In this class, patients are considered "seizure-free" if they fit into one of four categories. The four categories are

completely seizure-free since surgery (free of both complex and simple partial seizures); aura only since surgery (the patient is free of complex partial seizures but still has simple partial seizures); some seizures after surgery, but seizure-free for at least 2 years; and atypical generalized convulsions with AED withdrawal only.

The other three outcome measurements for "seizure-free" all assume that patients are free of complex partial seizures at the time of examination, but differ on whether they consider a patient "seizure-free" if they still have simple partial seizures (auras). Twenty studies specifically considered patients as "seizure-free" if they were free of both complex partial seizures and simple partial seizures (Table 18). In this report, we will refer to this group of patients as seizure-free with no auras. Twenty-six studies specifically considered patients as "seizure-free" if they were free of complex partial seizures, but patients could still have simple partial seizures and be considered "seizure-free" (Table 18). Therefore, this outcome measurement combines patients who are free of both complex and simple partial seizures with patients who are free of complex partial seizures but still have auras. In this report, we will refer to this group of patients as seizure-free with auras. Studies using the fourth outcome measurement, "seizure-free" did not state whether such patients experienced auras. Sixteen studies used this outcome measurement (Table 18). Since these studies do not report if their "seizure-free" patients do or do not have auras, these studies are probably a combination of studies reporting seizure-free with no auras and studies reporting seizure-free with auras. In this report, we will refer to this group of patients as seizure-free undefined.

Studies using Engel class I, which has the least restrictive means of determining if a patient is seizure-free, may be expected to report the largest percentage of seizure-free patients. Seizure-free with auras is similar to Engel class I, but somewhat more restrictive. Studies using this outcome measurement may be expected to report slightly fewer patients as seizure-free compared to Engel class I. Seizure-free with no auras is the most restrictive, and studies using this outcome measurement may be expected to report the smallest percentage of seizure-free patients. Because studies using seizure-free undefined may be a combination of studies using seizure-free with no auras and seizure-free with auras, these studies may be expected to report a percentage of seizure-free patients somewhere between studies reporting seizure-free no auras and studies reporting seizure-free with auras.

The 73 studies of temporal lobe surgery examined 3,978 patients. Twenty studies with 734 patients reported seizure-free with no auras, 26 studies with 1,396 patients reported seizure-free with auras, 16 studies with 977 patients reported seizure-free undefined and 33 studies with 1,549 patients reported Engel class I. If a study reported separate outcome and patient information according to a specific age group, type of surgery, or pathology, these data are presented separately in Evidence Table 110 and are considered a separate study in any analysis. Sixteen studies reported more than one of the four seizure-free categories, but no studies reporting seizure-free as undefined with respect to auras also reported one of the other categories. Of the studies that reported more than one outcome, five studies reported seizure-free with no auras, seizure-free with auras, and Engel class I, nine studies reported seizure-free with no auras and seizure-free with auras, and twelve studies reported Engel class I with either seizure-free with no auras or seizure-free with auras.

In addition to the studies of temporal lobe surgery, our evidence base also includes 12 studies that report seizure frequency outcome measurements for a total of 749 surgery "control" patients. Table 19 presents a listing of the seizure-free categories used by each of these studies. Seven reported seizure-free without reference to auras, three studies reported both seizure-free with

no auras and seizure-free with auras, one study reported only seizure-free with no auras, and a single study reported Engel class I along with seizure-free with no auras and seizure-free with auras.

Design and conduct of included studies

Once a patient has been identified as a suitable candidate for surgery, withholding surgery may be considered unethical. Consequently, the literature on surgical interventions consists mainly of uncontrolled trials in which all patients receive a single treatment and patients are not randomized to a nonsurgery group or a group receiving an alternative treatment approach. These studies generally do not provide a control group against which to evaluate the efficacy of surgery. The remainder of this section presents an assessment of the quality of the evidence base used to draw conclusions about the effectiveness of temporal lobe surgery in patients with treatment-resistant epilepsy. Our assessment consists of an appraisal of each study's internal and external validity.

Internal validity

Internal validity refers to the strength of the presumed causal relationship between the intervention and the outcome of interest.[21] For studies of surgical intervention in treatment-resistant epilepsy, one presumed relationship is between the surgical removal of tissue and changes in posttreatment seizure frequency. Table 20 lists the study designs in the evidence base for seizure-free outcome measurements after temporal lobe surgery. These studies are exclusively case series. Case series have a number of biases that can weaken the internal validity of a study. These biases can be ruled out if they are considered implausible in the particular context of a given study or they are plausible but did not actually occur.[21] Specific aspects of internal validity are discussed in the Methodology section of this document.

All of the studies discussed in this section on seizure frequency outcomes potentially have the following biases: extraneous event bias, investigator reporting bias, and patient reporting bias. Attrition bias and maturation bias are of specific importance to studies of surgery.

Attrition bias refers to the loss of patients, for any reason, before the minimum 2-year followup period. All studies with retrospective patient enrollment have this bias because they only record outcomes for patients with the minimum 2-year followup period. Only 10 of the 73 studies of seizure frequency outcomes had prospective as opposed to retrospective patient enrollment. The effect of attrition bias in the surgical studies considered in this report was limited by the requirement that studies report consecutive patients.

Maturation bias refers to individuals who received surgery but would have eventually "outgrown" the disease without surgical intervention. This seems implausible since surgery candidates often wait for more than a year before undergoing surgery and individuals may wait on average for 20 years from the onset of seizures before considering a surgical option.[145] A randomized controlled trial of temporal lobe surgery reported that 8 percent of control patients became free of complex partial seizures during a 1-year waiting period prior to surgery.[145] This finding suggests that maturation may occur, but that it affects only a small proportion of surgical patients.

External validity

As previously discussed, candidates for epilepsy surgery must complete an extensive presurgical evaluation to determine their suitability for surgery. Patients with temporal lobe epilepsy usually have a specific focal lesion and experience complex partial seizures with or

without secondary generalized seizures.[139] Therefore, the patients in published studies of surgery for temporal lobe epilepsy should be representative of all patients considering this surgery. However, many publications of epilepsy surgery select a specific patient population based on age or pathology, or use only one variation of a surgical technique. The results of these studies may or may not be generalizable to all temporal lobe surgery patients.

The specific patient characteristics of temporal lobe surgery patients reported in each publication in the evidence base for seizure-free outcomes are presented in Evidence Table 110. Age at surgery, age at seizure onset, and duration of epilepsy prior to surgery are commonly reported patient characteristics.

Among the 21 studies reporting seizure-free with no auras, 20 reported a mean age at surgery. The mean age at surgery in these studies varied from 9.4 years to 35 years, with only two studies having a mean less than 20 years of age. The range for age at surgery varied from 3 years to 62 years of age. Two studies examined only patients who were less than 20 years of age. Age at seizure onset was reported in 10 studies. The mean age at onset in these studies varied from 4 to 21 years of age with a range of less than a year to 44 years of age. Duration of epilepsy prior to surgery was reported in 11 studies. The mean duration varied from 6 to 19 years and the range varied from 1 to 45 years.

The patient characteristics for studies reporting seizure-free with auras were similar to the studies reporting seizure-free with no auras. Among the 26 studies reporting seizure-free with auras, 23 reported a mean age at surgery. The mean age at surgery varied from 8.3 years to 37 years with five studies having a mean less than 20 years of age. The range for age at surgery varied from 1 year to 86 years of age. Four studies examined only patients who were less than 20 years of age. Age at seizure onset was reported in 15 studies. The mean age at onset varied from 1 to 25 years of age with a range of less than a year to 62 years of age. Duration of epilepsy prior to surgery was reported in 11 studies. The mean duration varied from 5 to 26 years and the range varied from less than a year to 81 years.

Based on the distribution of patient characteristics, this evidence base seems to be generalizable to temporal lobe surgery patients in clinical practice.

Synthesis of study results

We will separately discuss each of the four “seizure-free” outcome measurements because, as mentioned above, each outcome measurement refers to a different group of “seizure-free” patients. We begin our analysis with studies reporting patients as seizure-free with no auras. This is the most restrictive group, but the ultimate goal of surgery is to be completely seizure-free. Next, we analyze studies reporting patients as seizure-free with auras. This patient population is free of complex partial seizures. Our analysis of the studies reporting Engel class I follows our analysis of the more restrictive “seizure-free” outcome measurements. Studies that did not report if auras were considered in their calculation of the number of patients who were seizure-free after surgery are analyzed last.

Meta-analytic threshold analysis of studies reporting seizure-free with no auras

Evidence Table 111 presents the actual patient counts, percentages, and calculated effect sizes for each study used in this analysis. The individual study effect sizes (Cohen’s h) presented in this Evidence Table were based on no patients in a synthetic control group becoming seizure-free with no auras. Figure 35 presents a forest plot of these effect sizes to show the extent of variation between studies, but no scale is provided because these effect sizes were not calculated using actual control groups.

The results of our threshold analysis of studies reporting seizure-free with no auras appear in Figure 36. Each summary estimate in the threshold analysis is based on Cohen's h. The summary estimate calculated at the 0 percent point (no patients in a synthetic control group became seizure-free with no auras) was 1.67 (CI: 1.57 to 1.77, p <0.000001) and corresponded to 55 percent (CI: 50 percent to 60 percent)[u] of patients becoming completely seizure-free after surgery. The summary estimate became nonsignificant (no statistically significant difference between surgery and control patients in the number of patients becoming seizure-free) when the proportion of patients in the synthetic control group reached 50 percent. There was no statistically significant heterogeneity among the studies in the threshold analysis (Q = 11.9, p = 0.92).

This analysis suggests that, after temporal lobe surgery, approximately 55 percent of patients will be completely seizure-free. However, this calculation was based on no patients in similar studies becoming seizure-free without surgery, so it does not estimate the percentage of patients who become seizure-free *because* of surgery. Some patients may become seizure-free without surgery. Readers are asked to consider the plausibility of 50 percent of temporal lobe epilepsy patients becoming completely seizure-free without benefit of surgery. They should also consider the above-noted difficulties with the internal validity of these studies, difficulties that could cause the threshold to decrease.

Meta-analytic threshold analysis of studies reporting seizure-free with auras

Evidence Table 112 presents the actual patient counts, percentages, and calculated effect sizes for each study used in this analysis. The individual study effect sizes (Cohen's h) presented in the Evidence Table were based on no patients in a synthetic control group becoming seizure-free with auras. Figure 37 presents a forest plot of these effect sizes to show the extent of variation between studies, but no scale is provided because these effect sizes were not calculated using actual control groups.

The results of our threshold analysis of studies reporting seizure-free with auras appear in Figure 38. The summary estimate calculated at the 0 percent point was 1.95 (CI: 1.87 to 2.02, p <0.000001) and corresponded to 68 percent (CI: 65 percent to 72 percent) of patients becoming free of complex partial seizures after surgery. The summary estimate became nonsignificant when the proportion of patients in the synthetic control group reached 65 percent. There was no statistically significant heterogeneity among the studies in the threshold analysis (Q = 24.2, p = 0.57).

This analysis suggests that, after temporal lobe surgery, approximately 68 percent of patients will be free of complex partial seizures (some patients may still have auras). However, this calculation was based on no patients becoming seizure-free without surgery, so it does not estimate net health benefit of surgery. Some patients may become seizure-free without surgery. The threshold analysis suggests that approximately 65 percent of patients in similarly designed studies would have to become seizure-free without surgery before surgery could be considered ineffective. Readers are asked to consider the plausibility of temporal lobe epilepsy patients achieving this threshold level without benefit of surgery. They should also consider the above-noted difficulties with the internal validity of these studies, difficulties that could cause the threshold to decrease.

To evaluate the plausibility of these threshold levels occurring among surgical candidates who do not receive surgery, we examined seizure rates in the available literature on such

[u] Computed from a back-transformation of Cohen's h.

patients. Of the twelve studies reporting seizure-free outcome measurements for surgery control patients, only three reported both seizure-free with no auras and seizure-free with auras (Evidence Table 113). An additional study reported just seizure-free with no auras. Estimates of the percentage of control patients likely to become seizure-free with no auras varied from 0 percent to 20 percent. The estimates for seizure-free with auras varied from 7.5 percent to 27 percent. These differences in seizure rates are most likely due to differences in the patients considered in each study. Patients may have refused surgery or were considered unsuitable for surgery and then were reported as "control" patients. Several studies did not report the reasons why patients did not receive surgery (Evidence Table 114). Therefore, although these data suggest that temporal lobe surgery is effective, the patients in these studies may not be comparable to the surgical patients from the studies used in our meta-analysis.

Comparison of meta-analytic threshold results to findings of a randomized controlled trial of temporal lobe surgery

An RCT conducted by Wiebe, Blume, Girvin, et al.[145] at the London Health Sciences Center at the University of Western Ontario examined seizure-free outcomes in patients who were randomized to temporal lobe surgery or required to wait 1 year before receiving surgery. A 1-year wait before undergoing preoperative investigations is routine practice at this institution. Therefore, randomizing patients to a wait list of 1 year was considered ethical. Patients were older than 16 years of age and continued to have at least monthly seizures despite the use of one or more AEDs. Patients randomized to surgery underwent a standard anterior temporal lobectomy. All patients were evaluated every 3 months for 1 year, and two epileptologists who were blinded to the identity of the patients and their treatment groups judged the adequacy of treatment through written clinical information.

Both seizure-free with no auras and seizure-free with auras were used to define the seizure-free status of the patients in this study. This study was not included in our analyzes of seizure-free data because the followup period was only 1 year. In a group of 40 control patients, one patient became seizure-free with no auras and two additional patients became seizure-free with auras for a total of three seizure-free patients (7.5 percent). Based on this study's findings, the synthetic control group levels of 50 percent and 65 percent needed to overturn the results of our threshold analysis seem unlikely to be achieved in a clinical setting.

Wiebe, Blume, Girvin, et al.[145] reported that among the 40 surgery patients 38 percent were completely free of seizures and 58 percent were free of seizures impairing awareness (seizure-free with or without auras). These results are somewhat lower than our meta-analytic estimates of 55 percent (CI: 50 percent to 60 percent) and 68 percent (CI: 65 percent to 72 percent), respectively, based on studies with a minimum 2 year followup. These results do fit within the range of results reported for studies that were included in the analysis (Evidence Table 111 and 112).

Factors that may influence seizure-free outcomes

The lack of statistically significant heterogeneity among the effect sizes in the studies reporting seizure-free no auras and seizure-free with auras indicates that several covariates, such as the surgical procedures, country where the study was performed, and specific pathology reported by each study, did not have large in fluences on the success of surgery.[v] For example, if tailored temporal lobectomy had produced many more completely seizure-free patients

[v] These results are limited by the statistical power of our meta-analysis, so small or moderate differences might still exist.

compared to standard temporal lobectomy, then our meta-analysis of studies reporting seizure-free no auras would have shown significant heterogeneity. This was not the case. The same can be said for studies examining only specific pathologies. We did not find that studies examining only patients with mesial temporal sclerosis, tumors, or vascular malformations had differing effect sizes. However, during the original organization of this project, the Expert Panel expressed an interest in knowing if certain study level factors influenced surgical outcomes. Therefore, we regrouped studies according to specific covariates (United States versus other countries, studies of mesial temporal sclerosis only versus studies examining various pathologies, and studies of standard temporal lobectomy versus studies of tailored temporal lobectomy versus studies of other surgical procedures). Evidence Table 115 and 116 show the summary effect size estimates based on seizure-free with no auras and seizure-free with aura outcome measurements, respectively. The recalculated summary estimates showed no statistically significant effect of any of these covariates.

Meta-analytic threshold analysis of studies reporting Engel class I

Evidence Table 117 presents the actual patient counts, percentages, and calculated effect sizes for each study in this analysis. The individual study effect sizes (Cohen's h) presented in the Evidence Table were based on no patients in a synthetic control group achieving Engel class I. Figure 39 presents a forest plot of these effect sizes to show the extent of variation between studies, but no scale is provided because these effect sizes were not calculated using actual control groups.

Our threshold analysis of studies reporting Engel class I found statistically significant heterogeneity among the effect sizes indicating a large amount of variation among study results ($Q = 77.7$, $p = 0.00002$). Therefore, the summary estimates in any threshold analysis of these data were not calculated. Rather, we sought to "explain" the source(s) of heterogeneity using meta-regression.

Despite the heterogeneity, all of the effect sizes (based on a Cohen's h with no control patients achieving Engel class I) in these studies were statistically significant. Therefore, these studies indicate that temporal lobe surgery is effective in producing seizure-free patients. The heterogeneity prevents an accurate estimation of the overall percentage of patients likely to achieve Engel class I status after surgery.

Meta-regression. In our meta-regression of the 33 studies reporting Engel class I, we again computed Cohen's h assuming a synthetic control group that did not experience any changes in the outcome of interest. Our prior analysis of studies reporting seizure-free no auras and seizure-free with auras suggested that the type of surgical procedures used in each study and the pathology examined in each study does not influence the estimate of the number of patients likely to become seizure-free. Therefore, we did not enter surgical procedures or pathology into this meta-regression. We instead looked for sources of heterogeneity due to differences in usage of the Engel classification system between countries and possible shifts in usage over time. Usage refers to differences in the interpretation of which patients belong in Engel class I. We entered into the meta-regression whether the study was performed in the United States, the year the study started, and the year the study ended. The data used in the meta-regression is presented in Evidence Table 118 and the results of the meta-regression appear in Evidence Table 119.

None of the three variables in our meta-regression explained the heterogeneity when used in one-, two-, or three-predictor models. Figure 40 graphically presents the results of the meta-regression. The dotted line on the graph represents the level of reduction in heterogeneity needed to obtain a statistically insignificant Q_E in any of the models. The meta-regressions failed to

reach or pass this line. Therefore, the heterogeneity among studies using Engel class I is not explained by differences in usage between the United States and other countries or due to shifts in usage over time. Consequently, a summary estimate that is adjusted for the sources of heterogeneity among study results could not be derived, and there is no ready explanation for why the results of these studies differ.

Meta-analytic threshold analysis of studies reporting seizure-free undefined

Evidence Table 120 presents the actual patient counts, percentages, and calculated effect sizes for each study in this analysis. The individual study effect sizes (Cohen's h) presented in the Evidence Table were based on no patients in a synthetic control group becoming seizure-free undefined. Figure 41 presents a forest plot of these effect sizes to show the extent of variation between studies, but no scale is provided because these effect sizes were not calculated using actual control groups.

Our threshold analysis of studies reporting seizure-free undefined found statistically significant heterogeneity among the effect sizes ($Q = 43.4$, $p = 0.00002$). Therefore, the summary estimates in any threshold analysis of these data were not calculated. Rather, we sought to "explain" the source(s) of heterogeneity using meta-regression.

Despite the heterogeneity, all of the effect sizes calculated from studies reporting seizure-free undefined were statistically significant. Therefore, these studies indicate that temporal lobe surgery is effective in producing seizure-free patients. The heterogeneity prevents an accurate estimation of the overall percentage of patients likely to achieve seizure-free status after surgery.

Meta-regression. In our meta-regression of the 16 studies reporting seizure-free undefined, we computed Cohen's h again assuming a synthetic control group that did not experience any changes in the outcome of interest. Our prior analysis of studies reporting seizure-free no auras and seizure-free with auras indicates that the type of surgical procedures used in each study and the pathology examined in each study did not influence the estimate of the number of patients who were likely to become seizure-free. Therefore, we did not enter surgical procedures or pathology into this meta-regression. Since this outcome is probably a combination of patients who are seizure-free no auras and seizure-free with auras, the heterogeneity is most likely due to differences in usage between studies. We therefore looked for sources of heterogeneity due to differences in usage between countries and possible shifts in usage over time. We entered into the meta-regression whether the study was performed in the United States, the year the study started, and the year the study ended. The data used in the meta-regression is presented in Evidence Table 121 and the results of the meta-regression appear in Evidence Table 122.

None of the three variables in our meta-regression explained the heterogeneity when used in one-, two-, or three-predictor models. Figure 42 graphically presents the results of the meta-regression. The dotted line on the graph represents the level of reduction in heterogeneity needed to obtain a statistically insignificant Q_E in any of the models. The meta-regressions failed to reach or pass this line. Therefore, the heterogeneity among studies using seizure-free undefined is not explained by differences in usage between the United States and other countries or due to shifts in usage over time. Consequently, a summary estimate that is adjusted for the sources of heterogeneity among study results could not be derived, and there is no ready explanation for why the results of these studies differ.

Analysis of nested case-control studies

Within any single study, seizure-free outcome measures may have been analyzed by the authors for variables that influenced the success of surgery. We term studies that reported these

findings as nested case-control studies. Unlike actual controlled studies, no patients in these studies are untreated. Rather, following treatment, patients are divided into those with successful outcomes and those without, and then various patient characteristics or other variables are compared for differences between the successful patients and nonsuccessful patients. Variables commonly examined for their influence on surgical success are age at surgery, age at first seizure, duration of epilepsy prior to surgery, gender, location of surgery (left vs. right temporal lobe), and type of pathology. Evidence Tables 123 and 124 present the findings, both statistically significant and nonsignificant, reported by each of the nine nested case-control studies in our evidence base for seizure-free measurements. Nested case-control studies using multiple or logistic regression to control for covariates in their analysis provide a more reliable estimate of the correlation between surgical success and patient characteristics compared to studies using univariate approaches. For this reason, our evidence table listed whether a study used multiple regression or a univariate test (t-test or chi-square test) in their analysis. Among the nine studies, two reported using multiple regression (Blume, Desai, Girvin, et al.,[146] and Cutfield and Wrightson[147]). Blume, Desai, Girvin, et al.,[146] in a study of 125 patients, found that younger age at surgery favored outcomes that are more successful. Cutfield and Wrightson,[147] in a study of 26 patients, did not find any patient characteristics that favored successful surgery. Only one of the seven studies using univariate procedures, Hennessy, Elwes, Honavar, et al.,[148] also found that younger age significantly favored successful surgery.

Meta-analysis of patient characteristics

Dodrill, Van Belle, and Wilkus[149] have pointed out that small sample sizes have lead to inconsistency in the conclusions reached about the significance of most variables believed to influence surgical outcomes. Therefore, many individual nested case-control studies may not be able to detect clinically meaningful effects.

To address this difficulty, we performed several separate meta-analyses. Table 21 presents a list of the 24 studies of temporal lobe surgery that provided data for these analyses. All of these studies were included in the previous meta-analyses examining the efficacy of surgery based on one of the four outcome measurements for reporting patients as seizure-free. At least five studies reported one or more of the following continuous variables: individual patient data for age at surgery, age at seizure onset, or duration of epilepsy prior to surgery. At least five studies separately reported one of the following dichotomous variables for patients who received successful and nonsuccessful surgery: the number of males versus female, left side surgeries versus right side surgeries, patients with simple partial seizures versus patients without simple partial seizures, or patients with secondarily generalized seizures versus patients without secondarily generalized seizures. Success was based on any of the four “seizure-free” outcome measurements.

We calculated a point-biserial correlation (r_{pb}) from the individual patient data in each study reporting the age at surgery, age at seizure onset, and duration of epilepsy prior to surgery, and then combined these in a separate meta-analysis for each variable. The coefficient was calculated so that a positive correlation indicated that an older age or longer duration favored a successful outcome and a negative result indicated that a younger age or shorter duration favored a successful outcome. For the other patient characteristics, we calculated Cohen’s h so that a positive effect size indicated that males, the left side, patients with simple partial seizures, or patients with secondarily generalized seizures had more successful surgery compared to females, the right side, patients without simple partial seizures, or patients without secondarily generalized seizures.

Our summary estimates are not adjusted for the influence of the other potentially important covariates in a study. An analysis using hierarchical modeling would be useful to search for factors that influence surgical outcomes by combining the patient-level data across studies, but such an analysis is beyond the scope of this report.

Age at surgery. In our first meta-analysis of "predictors" of surgical success, we sought to determine whether different outcomes were obtained in patients of different ages at the time they receive surgery. Individual ages at surgery for patients with successful and nonsuccessful surgery were reported in 18 studies with 297 patients. Evidence Table 125 presents the definition used for successful surgery and the point-biserial correlation calculated in each of the 18 studies. Figure 43 presents a forest plot of the correlations. The meta-analysis produced a summary estimate that was not statistically significant (r_{pb} = 0.02, CI: –0.11 to 0.14, p = 0.81) suggesting that age at surgery had no influence on the success of surgery in these studies. The effect sizes in this meta-analysis were not heterogeneous (Q = 10.7, p = 0.91).

We performed a sensitivity analysis to show that a single study did not have excessive influence over the results of the analysis. This ensures that our conclusion (no effect of age on success of surgery) cannot be overturned by the removal of just one study. The summary estimate and other statistics did not change because of the sensitivity analysis. The correlation between surgical success and age at surgery changed by no more than 0.02 due to removal of studies during the sensitivity analysis. The summary estimate remained statistically nonsignificant. The results of the sensitivity analysis and the original meta-analysis are presented in Evidence Table 126.

Age at seizure onset. In our second meta-analysis of "predictors" of surgical success, we sought to determine whether different outcomes were obtained in patients of different ages at seizure onset. Individual ages at seizure onset for patients with successful and nonsuccessful surgery were reported in 13 studies with 207 patients. Evidence Table 127 presents the definition used for successful surgery and the point-biserial correlation in each of the 13 studies. Figure 44 presents a forest plot of the correlations. The meta-analysis produced a summary estimate that was not statistically significant (r_{pb} = -0.11, CI: -0.26 to 0.04, p = 0.16) suggesting that age at seizure onset had no influence on the success of surgery in these studies. The effect sizes in this meta-analysis were not heterogeneous (Q = 7.2, p = 0.89).

We performed a sensitivity analysis to ensure that a single study did not have excessive influence over the results of the analysis. The summary estimate and other statistics did not change because of the sensitivity analysis. The correlation coefficient changed by no more than 0.03 due to removal of studies during the sensitivity analysis. The summary estimate remained statistically nonsignificant. The results of the sensitivity analysis and the original meta-analysis are presented in Evidence Table 128.

Duration of epilepsy prior to surgery. In this meta-analysis, we sought to determine whether different outcomes were obtained in patients with different durations of epilepsy prior to the time they receive surgery. Individual durations of epilepsy prior to surgery for patients with successful and nonsuccessful surgery were reported in 12 studies with 192 patients. Evidence Table 129 presents the definition used for successful surgery and the point-biserial correlation in each of the 12 studies. Figure 45 presents a forest plot of the effect sizes. The meta-analysis produced a summary estimate that was not statistically significant (r_{pb} = 0.15, CI: -0.01 to 0.30, p = 0.06) suggesting that duration of epilepsy prior to surgery did not influence the success of surgery in these studies. The effect sizes in this meta-analysis were not heterogeneous (Q = 15.9, p = 0.20).

We then performed sensitivity analyses on these results. When the study with the largest negative effect size (favors shorter duration of epilepsy) was removed, the summary estimate became statistically significant (r_{pb} = 0.20, CI: 0.04 to 0.35, p = 0.02). The effect sizes remained homogenous when this study was removed (Q = 9.4, p = 0.58). Thus, without this study in the analysis, patients with a longer duration of epilepsy prior to surgery appear to have a slightly better chance of having successful surgery compared to patients with a shorter duration of epilepsy prior to surgery. The removal of other studies during the sensitivity analysis changed the correlation by no more than 0.03. Therefore, patients with a longer duration of epilepsy prior to surgery appear to have a tendency towards better outcomes after surgery, but this tendency is not robust. The results of the sensitivity analysis as well as the original meta-analysis are presented in Evidence Table 130.

Gender. We next investigated whether a greater percentage of males compared to females had successful surgery. The number of male and female patients among patients with successful and nonsuccessful surgery was reported in 15 studies with 306 patients. Evidence Table 131 presents the individual number of male and female patients and the number of successful surgeries in each, the definition used for successful surgery, and the Cohen's h in each of the 15 studies. Figure 46 presents a forest plot of these effects. The meta-analysis produced a statistically significant Q statistic (27.9, p = 0.015), so the summary effect size is not meaningful. Two of the 15 studies showed a statistically significant increase in the number of female patients with successful outcomes compared to male patients. Of the remaining 13 studies, eight favored male patients and five favored female patients, although none of these studies showed a statistically significant difference.

To "explain" this heterogeneity, we performed 36 meta-regressions (see the Methodology section for a description of our approach to meta-regression). Of these, no one-predictor model explained the heterogeneity, and five two-predictor models did. No clear "best" model was obvious among these five models. Consequently, no obvious explanation for the variation among these studies is apparent, and why surgery is more or less successful in males compared to females in these studies is unclear. All of the study and patient characteristics used in our meta-regression are presented in Evidence Table 132. The meta-regressions are presented in Evidence Table 133 and Figure 47.

Location of surgery. In this meta-analysis, we sought to determine whether surgery was more successful in patients who had surgery in the left temporal lobe or the right temporal lobe. The percentage of left-sided and right-sided operations among patients with successful and nonsuccessful surgery was reported in 19 studies with 404 patients. Evidence Table 134 presents the number of left-sided and right-sided operations and the number of successful patients in each, the definition used for successful surgery, and the Cohen's h calculated in each of the 19 studies. Figure 48 presents a forest plot of these effects. The meta-analysis produced a summary estimate that was not statistically significant (-0.07, CI: -0.27 to 0.13, p = 0.49), suggesting that location of surgery had little or no influence on the success of surgery. The effect sizes in this meta-analysis were not heterogeneous (Q = 17.9, p = 0.46).

The summary estimate and other statistics did not change because of the sensitivity analysis. The back-transformed estimate for the difference between the percentage of left side surgery patients who achieved successful surgery and the percentage of right side surgery patients who achieved successful surgery was 0 regardless of the studies that were removed during the sensitivity analysis. The sensitivity analysis and the original meta-analysis are presented in Evidence Table 135.

Simple partial seizures. We next compared surgical success rates in patients with simple partial seizures to success rates in patients without simple partial seizures. The number of patients with simple partial seizures among patients with successful and nonsuccessful surgery was reported in five studies with 131 patients. Evidence Table 136 presents the number of patients with and without simple partial, the number of successful surgeries in each, the definition used for successful surgery, and the Cohen's h calculated in each of the five studies. Figure 49 presents a forest plot of these effects. The meta-analysis produced a summary estimate that was not statistically significant (0.10, CI: -0.30 to 0.51, p = 0.62), suggesting that the presence of simple partial seizures had no influence on the success of surgery. The effect sizes in this meta-analysis were not heterogeneous (Q = 7.1, p = 0.13).

The summary estimate and other statistics showed only small changes because of the sensitivity analysis. The back-transformed estimates for the difference between the percentage of patients with simple partial seizures who achieved successful surgery and the percentage of patients without simple partial seizures who achieved successful surgery varied between –9 and 1 as studies were removed during the sensitivity analysis. The summary estimate did not become statistically significant when studies were removed. The sensitivity analysis and the original meta-analysis are presented in Evidence Table 137.

Secondarily generalized seizures. In our final meta-analysis on characteristics that may "predict" successful temporal lobe surgery, we examined whether patients with secondarily generalized seizures had different outcomes compared to patients without secondarily generalized seizures. The number of patients with or without secondarily generalized seizures among patients with successful and nonsuccessful surgery was reported in seven studies with 256 patients. Evidence Table 138 presents the individual number of patients with and without secondarily generalized seizures and the number of successful surgeries in each, the definition used for successful surgery, and the Cohen's h calculated in each of the seven studies. Figure 50 presents a forest plot of these effects. The meta-analysis produced a statistically significant Q statistic (31.8, p = 0.00002) so the summary effect size was not meaningful.

Two studies reported that patients without secondarily generalized seizures have better outcomes, one study reported that patients with secondarily generalized seizures have better outcomes, and four studies reported no differences in outcomes between patients with or without secondarily generalized seizures.

To "explain" this heterogeneity, we performed 51 meta-regressions. Of these, no models explained the heterogeneity. Consequently, no obvious reason is apparent for why some studies had different results compared to other studies, and whether surgery is more or less successful in patients with secondarily generalized seizures is unclear. All of the study and patient characteristics used in our meta-regression are presented in Evidence Table 139. The results of the meta-regression are presented in Evidence Table 140 and Figure 51.

Quality of Life Outcome Measurements

Evidence base

The Epilepsy Surgery Inventory global score was reported in one study with 47 patients[150] and the Quality of Life in Epilepsy global score was reported in one study with 90 patients.[151]

Design and conduct of included studies

All of the previously discussed biases about the internal validity of studies that reported seizure-free outcome measures potentially apply to the studies that reported quality of life outcome measurements.

Synthesis of study results

Evidence Table 141 presents a summary of the findings in each of the studies reporting quality of life measurements. No statistically significant change was found between the baseline Epilepsy Surgery Inventory overall score and the overall score 2 years after surgery.[150] However, the authors did report that patients with low baseline scores showed the greatest improvement after surgery. This suggests the presence of regression to the mean. Patients who received surgery did show a statistically significant improvement in the Quality of Life in Epilepsy global score 2 years after surgery both compared to baseline and a control group of patients.[151] The entire improvement in global score was contributed by patients who became completely seizure-free. Once again, though, regression to the mean cannot be ruled out as an explanation for these results.

Employment Outcome Measurements

Evidence base

Among the 105 studies of temporal lobe surgery in our evidence base, five reported some form of employment data that met our inclusion criteria. Studies must have reported the number of patients not able to obtain work prior to surgery and the number of patients able to obtain work after surgery, or must have reported the number of patients working prior to surgery and the number of patients not able to remain at work after surgery. The five studies had 318 patients. Four of the studies were conducted in the United States and the fifth study was from Canada. Table 22 presents a listing of the five studies reporting employment data.

Design and conduct of included studies

All of the previously discussed biases about the internal validity of studies that reported seizure-free outcome measures potentially apply to the studies that reported employment outcome measurements. In particular, the lack of a precise definition of who is employed may lead to inconsistencies in the reporting of this outcome.

Synthesis of study results

Although each of the five included studies evaluated more than 10 patients, in three studies fewer than 10 patients were reported to be in the “not able to obtain work prior to surgery” category or in the “working prior to surgery” category (Evidence Table 142). The other patients in the study were not actively seeking employment, were of preschool age or in school. Therefore, we did not perform a meta-analysis of these data. The studies do show that some patients who were unable to obtain employment prior to surgery do find employment after surgery. In the two studies with more than 20 patients unable to obtain work prior to surgery, 7 out of 20 patients and 15 out of 28 patients were able to obtain employment after surgery. In two studies with more than 30 patients, 57 out of 67 patients and 30 out of 33 patients working prior to surgery were able to maintain employment after surgery. A third study with 13 patients showed that nine patients remained working after surgery. While 85 percent (96 out of 113) of

the patients in these latter three studies were able to remain employed after surgery, 15 percent of the patients were not able to maintain their employment.

Education Outcome Measurements

Evidence base

Return to (or ability to remain in) school was reported in two studies with 37 patients, however only one study had more than 10 patients of school age.[152,153]

Design and conduct of included studies

All of the previously discussed cautions about the internal validity of studies that reported seizure-free outcome measures apply to the two studies that reported education outcome measurements.

Synthesis of study results

Evidence Table 143 presents a summary of the findings in the two studies reporting education outcome measurements. These studies reported that all patients attending school prior to surgery remained in school after surgery.[152,153]

Ability to Obtain a Driver's License Outcome Measurements

Evidence base

Only Reeves, So, Evans et al.,[154] who studied 134 patients, reported on the ability of patients to obtain a drivers license after surgery.

Design and conduct of included studies

All of the previously discussed cautions about the internal validity of studies that reported seizure-free outcome measures apply to this study.

Synthesis of study results

Evidence Table 144 presents a summary of the findings in the study reporting ability to obtain a driver's license. Surgery was reported to have produced a statistically significant increase in the number of patients able to drive.[154]

Mood Outcome Measurements - Depression

Epilepsy has been associated with an increased incidence and prevalence of behavioral disorders and in particular with anxiety and depression.[4,155] New cases of depression have been associated with temporal lobe surgery[156] and the National Institutes of Health Consensus Development Conference Statement: Surgery for Epilepsy has recommended that symptoms of anxiety and depression be assessed following surgery.[140] The following section evaluates studies that reported new cases of depression after temporal lobe surgery.

Evidence base

Among the 105 studies of temporal lobe surgery meeting our inclusion criteria, 10 reported whether their patients experienced new cases of depression after surgery. These patients had not been diagnosed with clinical depression prior to surgery. The 10 studies examined 597 patients.

Table 23 lists the studies. Evidence Table 145 provides study information including the methods of diagnosis for depression reported in each study. Five studies used the Diagnostic and Statistical Manual of Mental Disorders, 3rd or 4th edition (DSM-III, IV) criteria, one study used the International Classification of Disease 10th revision (ICD-10), two studies used the Center for Epidemiological Studies-Depression Scale (CES-D), and two studies reported diagnose by a psychiatrist. Only the RCT by Wiebe, Blume, Girvin, et al.[145] provided data on a control group comparable to the patients receiving surgery.

Although these 10 studies reported new cases of depression after surgery, they did not report the actual number of patients who were either clinically depressed or free of depression prior to surgery. Patients were not excluded from surgery for clinical depression in these studies. Therefore, our analysis uses the total number of patients receiving surgery rather than the actual number of patients free of depression prior to surgery.

Design and conduct of included studies

Internal validity

All but one of the 10 studies in the evidence base for new cases of depression are uncontrolled studies of case series design. Therefore, these uncontrolled studies have the same concerns with regard to internal validity as previously discussed with regard to seizure-free outcomes. Attrition bias may not be a major concern in these studies because all patients were examined during the relatively short followup periods (no more than 1 year).

Depression occurs in patients with epilepsy, both before and after surgery. Therefore, the lack of control patients in most of these studies prevents any determination of whether the effect of surgery is to increase or decrease the incidence of depression. The analysis of the studies can only provide an estimate of the number of patients likely to experience depression after surgery.

External validity

The specific patient characteristics of temporal lobe surgery patients reported in each study are presented in Evidence Table 146. The patients in these studies were between 20 and 50 years old at the time of surgery, the mean age of seizure onset was between 9 and 16 years of age, and the mean duration of epilepsy prior to surgery was approximately 18 years.

Based on the distribution of patient characteristics, this evidence base seems to be generalizable to temporal lobe surgery patients in clinical practice.

Synthesis of study results

Meta-analytic threshold analysis of depression outcome measurements

Evidence Table 147 presents the actual patient counts, percentages, and calculated effect sizes for each study in this analysis. The individual study effect sizes (Cohen's h) presented in the Evidence Table were based on no patients in a synthetic control group becoming clinically depressed after surgery. Figure 52 presents a forest plot of these effect sizes to show the extent of variation between studies, but no scale is provided because these effect sizes were not calculated using actual control groups.

All of the studies reported a statistically significant occurrence of new cases with a range of 4 percent to 24 percent. Our threshold analysis of studies reporting new cases of depression found statistically significant heterogeneity among the study results ($Q = 18.0$, $p = 0.035$). Therefore, we did not compute the summary estimates in any threshold analysis of these data. Rather, we sought to "explain" the source(s) of heterogeneity using meta-regression.

Meta-regression. To "explain" this heterogeneity, we performed 13 meta-regressions. Of these, no models explained the heterogeneity. Consequently, no obvious reason is apparent to explain why some studies reported more new cases of depression than other studies, and whether surgery is more or less responsible for new cases of depression is unclear. All of the study and patient characteristics used in our meta-regression are presented in Evidence Table 148. The results of the meta-regression are presented in Evidence Table 149 and Figure 53.

Because all but one study lacked a control group, these studies do not provide evidence that surgery was directly responsible for the new cases of depression or that surgery reduced the incidence of depression. This is highlighted by the results of the one RCT among these studies. Wiebe, Blume, Girvin, et al.,[145] using the Center for Epidemiological Studies-Depression Scale, reported that 8 out of 40 control patients (20 percent) developed depression during the year preceding their surgical.

Mood Outcome Measurements - Psychosis

Besides depression, treatment-resistant epilepsy has been associated with a variety of psychiatric disorders.[155] Surgery for treatment-resistant epilepsy may also have psychiatric consequences. The following section evaluates studies that reported new cases of psychotic disorders (primarily schizophrenia and bipolar disorder) after temporal lobe surgery.

Evidence base

Among the 105 studies of temporal lobe surgery meeting our inclusion criteria, six reported whether their patients experienced new cases of psychosis after surgery. The six studies examined 385 patients. Four of the six studies are also part of the evidence base for depression discussed earlier. Table 24 lists the studies and Evidence Table 150 provides study information including the methods of diagnosis in each. Four of the studies reported using specific criterion, while two studies reported using evaluations by a psychiatrist only. Only the RCT by Wiebe, Blume, Girvin, et al.[145] provided data from a control group.

Although these six studies reported new cases of psychosis after surgery, they did not report the actual number of patients who had a psychotic disorder or were free of psychotic disorders prior to surgery. Two of the studies excluded patients who had chronic psychosis and the remaining four studies did not exclude patients with psychiatric disorders. Therefore, our analysis uses the total number of patients receiving surgery rather the actual number of patients free of psychosis prior to surgery.

Design and conduct of included studies

Internal validity

All but one of the six studies in the evidence base for assessing new cases of psychosis are uncontrolled studies of case series design. Therefore, these studies have the same concerns with regard to internal validity as previously discussed with regard to seizure-free outcomes. In particular, variations in the use of any of the specific criteria, or variations in individual psychiatrists could lead to inconsistencies in the reporting of this outcome. Attrition bias is not a concern because all patients were examined after surgery.

Psychosis can occur in patients with epilepsy, both before and after surgery. Therefore, the lack of control patients in most of these studies prevents any determination of whether the effect of surgery is to increase or decrease the incidence of psychosis. Our analysis of these studies can only provide an estimate of the number of patients likely to experience psychosis after surgery.

External validity

The specific patient characteristics of temporal lobe surgery patients reported in each study are presented in Evidence Table 151. The patients in these studies were between approximately 20 to 40 years old at the time of surgery, the mean age of seizure onset was between 10 and 15 years of age, and the mean duration of epilepsy prior to surgery was approximately 18 years.

Based on the distribution of patient characteristics, this evidence base seems to be generalizable to temporal lobe surgery patients in clinical practice.

Synthesis of study results

Meta-analytic threshold analysis of psychosis outcome measurements

Evidence Table 152 presents the actual patient counts, percentages, and calculated effect sizes for each study in this analysis. The individual study effect sizes (Cohen's h) presented in the Evidence Table were based on a control group in which no patients develop psychosis. Figure 54 presents a forest plot of these effect sizes to show the extent of variation between studies, but no scale is provided because these effect sizes were not calculated using actual control groups.

The results of our threshold analysis of studies reporting new cases of psychosis appear in Figure 55. Each summary estimate in the graph is Cohen's h. The summary estimate calculated at the 0 percent point (no patients in a synthetic control group developed psychosis after surgery) was 0.37 (CI: 0.23 to 0.51, $p < 0.000001$) and corresponded to 3 percent (CI: 1 percent to 6 percent) of patients developing psychosis after surgery. The summary estimate became nonsignificant (no statistically significant difference between surgery and control) when the proportion of patients in the synthetic control group reached 2 percent. There was no statistically significant heterogeneity in the threshold analysis ($Q = 6.5$, $p = 0.26$).

Because all but one study lacked a control group, these studies do not provide evidence that surgery was directly responsible for the new cases of psychosis or that surgery *caused* an increase or decrease in the incidence of psychosis. This can be seen in the one RCT among these studies. Wiebe, Blume, Girvin, et al.[145] reported that 1 out of 40 control patients (2.5 percent) developed psychosis during the year preceding their surgical treatment compared to 1 out of 36 surgery patients (2.8 percent). This percentage of new cases of psychosis among control patients suggests that surgery may not be responsible for all new cases of psychosis after surgery. Nevertheless, our analysis provides an estimate of the number of new cases that may be expected after temporal lobe surgery, regardless of cause.

Cognitive Function Outcome Measurements - IQ

Treatment-resistant epilepsy may be associated with a slow progressive cognitive deterioration. A study of 209 patients with temporal lobe epilepsy reported that patients with a duration of greater than 30 years performed worse on full scale IQ tests compared to patients with less than 30 years duration.[157] Due to the nature of the procedure, patients contemplating temporal lobe surgery may also be concerned with the potential for loss of intellectual functioning after surgery. The following section evaluates studies that reported both the number of patients to have a significant change in IQ (increase or decrease) and the pre- and postsurgery mean IQs. The authors of these studies defined a clinically significant increase or decrease in IQ as a change of at least one to two standard errors, and our analysis, therefore, incorporated this definition.

Evidence base

Among the 105 studies of temporal lobe surgery meeting our inclusion criteria, six reported if their patients experienced a significant decrease or increase in IQ after surgery as well as reported the mean pretest and posttest IQ. The six studies examined 449 patients. Table 25 lists the studies, and Evidence Table 153 provides study information including the methods used in each. We abstracted and analyzed the verbal IQ scores from each study because these data were reported in all six studies. Only the study by Chelune, Nagle, Lueders, et al.[158] provided data on a control group.

Design and conduct of included studies

Internal validity

All but one of the six studies in the evidence base for assessing changes in IQ are uncontrolled case series. Therefore, these studies have the same concerns with regard to internal validity as previously discussed with regard to seizure-free outcome reporting. Investigator bias and patient reporting bias may be reduced (but not eliminated) due to the use of a standardized intelligence test (Wechsler Intelligence Scale) and a predefined cutoff determining when a patient's IQ has undergone a significant change. Attrition bias is not a concern because all patients were examined after surgery.

Decreases in IQ scores can occur in patients with epilepsy, both before and after surgery. Therefore, the lack of control patients in most of these studies limits our ability to determine whether surgery decreased IQ scores. However, since increases in IQ scores are unlikely to occur spontaneously, any increase in IQ scores after surgery are likely to be a consequence of surgery. Our analysis provides an estimate of the number of patients likely to experience either an increase or decrease in IQ after surgery.

External validity

The specific patient characteristics of temporal lobe surgery patients reported in each study are presented in Evidence Table 154. Three of the studies examined only children and adolescents while the other three studies examined only adults. The children were approximately 5 to 15 years old at the time of surgery, while the adults were between 20 and 40 years old at the time of surgery. The mean age of seizure onset was approximately 5 years of age for the children and approximately 10 to 15 years for the adults. The mean duration of epilepsy prior to surgery was approximately 10 years in all six studies with a broad range of between 1 to 17 years.

Based on the distribution of patient characteristics, this evidence base seems to be generalizable to temporal lobe surgery patients in clinical practice.

Synthesis of study results

Meta-analytic threshold analysis of decreases in IQ after surgery

Evidence Table 155 presents the actual patient counts, percentages, and calculated effect sizes for each study in this analysis. The individual study effect sizes (Cohen's h) presented in the Evidence Table were based on a control group in which no patients experience a clinically significant decrease in IQ. Figure 56 presents a forest plot of these effect sizes to show the extent of variation between studies, but no scale is provided because these effect sizes were not calculated using actual control groups.

The results of our threshold analysis of studies reporting patients with clinically significant decreases in IQ after surgery appear in Figure 57. Each summary estimate in the graph is

Cohen's h. The summary estimate calculated at the 0 percent point (no patients in a synthetic control group showed a significant decrease in IQ) was 0.65 (CI: 0.52 to 0.78, p <0.000001) and corresponded to 10 percent (CI: 7 percent to 14 percent) of patients experiencing a clinically significant decrease (equal to 1 to 2 standard deviation units) in IQ after surgery. The summary estimate became nonsignificant (no statistically significant difference between surgery and control) when the proportion of patients in the synthetic control group reached 7 percent. There was no statistically significant heterogeneity in the threshold analysis (Q = 2.3, p = 0.81).

Meta-analytic threshold analysis of increases in IQ after surgery

Evidence Table 155 presents the actual patient counts, percentages, and calculated effect sizes for each study in this analysis. The individual study effect sizes (Cohen's h) presented in the Evidence Table were based on a control group in which no patients experience a clinically significant increase in IQ. Figure 58 presents a forest plot of these effect sizes to show the extent of variation between studies, but no scale is provided because these effect sizes were not calculated using actual control groups.

The results of our threshold analysis of studies reporting patients with clinically significant increases in IQ after surgery appear in Figure 59. Each summary estimate in the graph is Cohen's h. The summary estimate calculated at the 0 percent point (no patients in a synthetic control group showed a significant increase in IQ) was 0.74 (CI: 0.61 to 0.88, p <0.000001) and corresponded to 13 percent (CI: 9 percent to 18 percent) of patients experiencing a clinically significant increase in IQ after surgery. The summary estimate became nonsignificant (no statistically significant difference between surgery and control) when the proportion of patients in the synthetic control group reached 10 percent. There was no statistically significant heterogeneity among the effect sizes in the threshold analysis (Q = 4.3, p = 0.51).

Because all but one study lacked a control group, these studies do not provide evidence that surgery was directly responsible for the decreases in individual patient IQ scores, although a case can be made that surgery is responsible for any increases in IQ. The one study with a control group, Chelune, Nagle, Lueders, et al.,[158] reported two patients with clinically significant increases and two patients with clinically significant decreases in verbal IQ out of 40 control patients (5 percent each) compared to eight patients with clinically significant increases and eight patients with clinically significant decreases in verbal IQ out of 96 surgery patients (8.3 percent each). These percentages for increases and decreases among the control patients are lower than the percentages needed to overturn the conclusions of our threshold analysis suggesting that surgery may plausibly be responsible for changes in IQ.

Meta-analysis of changes in mean IQ after surgery

We also performed a meta-analysis of the data on mean pretest and posttest verbal IQ scores from these same studies (Evidence Table 156). We excluded from the analysis one study that did not report a measure of dispersion for the means. This analysis used Hedges' d as an effect size. A forest plot of the results of this meta-analysis is presented in Figure 60. The meta-analysis produced a summary estimate that was not statistically significant (-0.05, CI: -0.21 to 0.11, p = 0.53), suggesting no dramatic changes in mean IQ after surgery. The effect sizes in this meta-analysis were not heterogeneous (Q = 1.5, p = 0.82).

The summary estimate showed only small changes during the sensitivity analysis and remained statistically nonsignificant. The results of the sensitivity analysis as well as the original meta-analysis are presented in Evidence Table 157.

Cognitive Function Outcome Measurements - Memory

Temporal lobe surgery usually requires the removal of the hippocampus, a part of the brain important to memory capacity. Therefore, memory function is at risk whenever this procedure is performed.[159] The following section evaluates studies that reported both the number of patients with a significant change in memory function (increase or decrease) and the pre- and postsurgery mean memory scores.

Evidence base

Among the 105 studies of temporal lobe surgery meeting our inclusion criteria, five studies reported individual changes in one of the measurements in the Wechsler Memory Scale as well as reported the mean score before and after surgery. The five studies had 342 patients. Only two of the five studies reported the same portion of the Wechsler Memory Scale. Therefore, we did not perform a meta-analysis of these data. One study, Chelune, Nagle, Lueders, et al.,[158] provided data on a control group. Table 26 presents a listing of the five studies reporting memory changes. Study information and the portion of the Wechsler Memory Scale used in each study are presented in Evidence Table 158.

Synthesis of study results

Evidence Table 159 presents the finding for the five studies in the evidence base for this section. Patients experienced both increases and decreases in memory function, but the individual percentages in each study varied widely (Figure 61). The range of patients who showed an increase was 1 percent to 34 percent and the range of patients who showed a decrease was 9 percent to 62 percent.

To further explore these data, we calculated individual study results (using Cohen's h) by assuming that a control group would have no patients experiencing an increase or decrease in memory score. Statistically significant effects indicate that the percentage of patients experiencing an increase or decrease in memory score was significantly different from zero. All five studies showed statistically significant percentages of patients with memory decreases, and four studies showed statistically significant percentages of patients with memory increases (Figure 62).

Complications Due to Surgery

Serious permanent complications and transient complications are an inherent part of surgery. Temporal lobe surgery can result in various forms of paralysis due to obstruction of blood vessels or other damage to brain tissue. The following section evaluates studies that reported cases of serious permanent complications. We considered moderate to severe permanent neurological deficits, especially hemiplegia, to be serious complications. We considered all other reported surgical complications to be mild or transient. Development of postsurgical depression or psychosis, and declines in IQ or memory are not considered in this section because we examined them separately (see above).

Evidence base

Among the 105 studies of temporal lobe surgery meeting our inclusion criteria, 40 studies reported on complications due to surgery. The 40 studies examined 2091 patients (Table 27). We abstracted data on serious permanent complications only if the publication specifically

reported such a complication or specifically reported that no such complications occurred. We abstracted data on mild or transient complications only from studies reporting data on serious permanent complications. Six of the 40 studies did not report on the occurrence of mild or transient complications.

Design and conduct of included studies

Internal validity

The complications reported by these studies could only have occurred because of surgery, so the internal validity with regard to the cause and effect is not in question. However, some potential biases are still present. Investigator reporting bias may have affected the reporting of mild or transient complications because they may not be regarded as important by some investigators. Attrition bias is not a concern because all patients were examined after surgery. Maturation bias is also not a concern when reporting complications.

External validity

The specific patient characteristics of temporal lobe surgery patients reported in each study are presented in Evidence Table 160. The 40 studies in the evidence base cover a wide range of patient ages at surgery, onset of seizures, and duration of epilepsy. Eleven studies enrolled patients with a mean age at surgery of less than 20 years with no patient exceeding 22 years of age. Twenty-eight studies enrolled patients with a mean age at surgery of greater than 20 years with youngest and oldest ages that varied between 1 year and 86 years. Twelve studies had a mean age of seizure onset of less than 10 years of age and 15 studies had a mean age of seizures onset after 10 years of age. The range of seizure onset varied from less than 1 year of age to 62 years of age. Ten studies reported a mean duration of epilepsy of less than 10 years and 18 studies reported a mean duration of epilepsy of greater than 10 years. The range for duration of epilepsy varied between less than 1 year and 81 years.

Of the 40 studies, three included patients who received surgery starting in the 1940s and 1950s[160-162] and three included patients who received surgery starting in the 1960s (Table 27).[147,163,164]

Based on the distribution of patient characteristics, this evidence base seems to be generalizable to temporal lobe surgery patients in clinical practice.

Synthesis of study results

Evidence Table 161 presents a study-by-study list of the complications reported in each of the forty studies in the evidence base. Among the 2,091 temporal lobe surgery patients, 42 serious permanent complications were reported. This corresponds to 2 percent of the patients or 20 serious complications per 1,000 surgery patients. If the six studies which included patients from the 1940s, 1950s, and 1960s are removed, the number of serious complications was 32 out of 1,534 patients or 2.1 percent of patients.

Seventy-nine mild or transient complications were reported among 1,339 patients, which correspond to 6 percent or 59 complications per 1,000 surgery patients. The number of mild or transient complications may be underestimated by these data because of differences in reporting these complications across studies. Clinician judgment as to the importance of reporting various mild or transient complications will likely vary across studies, whereas, the occurrence of permanent paralysis will usually warrant reporting.

Surgery-related Mortality

Any surgical procedure may result in such serious complications that death results. The following section evaluates studies that reported deaths due to temporal lobe surgery.

Evidence base

Among the 105 studies of temporal lobe surgery meeting our inclusion criteria, 38 studies reported a death due to surgery or specifically reported that no deaths occurred due to surgery. The 38 studies examined 2,065 patients (Table 28). Only four of these studies were not included in the evidence base for our analysis on complications (see above).

We abstracted only deaths specifically reported to be caused by surgery. Deaths as a result of invasive presurgical diagnostic procedures were not included.

Design and conduct of included studies

Internal validity

The deaths reported here could only have occurred through surgery. Investigator reporting bias, attrition bias, and maturation bias are not a concern when reporting surgery-related mortality.

External validity

The specific patient characteristics of temporal lobe surgery patients reported in each study are presented in Evidence Table 162. The 38 studies in the evidence base cover a wide range of ages at surgery, onset of seizures, and duration of epilepsy. Three studies enrolled patients with a mean age at surgery of less than 10 years and seven studies had a mean age at surgery between 10 and 15 years. The oldest patients in these ten studies did not exceed 22 years of age. Twenty-six studies enrolled patients with a mean age at surgery of greater than 21 years with ranges that varied between the youngest patients being 1 year of age to the oldest patient being 74 years of age. Mean age of seizure onset, reported in 22 studies, was between 2 and 20 years of age. The range of seizure onset varied from less than 1 year of age to 49 years of age. Mean duration of epilepsy, reported in 24 studies, was between 2 and 20 years. The range for duration of epilepsy varied from less than 1 year to 53 years.

Of the 38 studies, three included patients who received surgery starting in the 1940s and 1950s[160-162] and two included patients who received surgery starting in the 1960s (Table 28).[147,163]

Based on the distribution of patient characteristics, this evidence base seems to be generalizable to temporal lobe surgery patients in clinical practice.

Synthesis of study results

Among the 2,065 temporal lobe surgery patients, five deaths were reported (0.24 percent or 2.4 deaths per 1,000 patients). The five deaths were reported in four studies, all of which had more than 70 patients (Table 30). The study reporting two deaths enrolled patients from 1957 to 1988.[162] Six studies with 70 or more patients reported no deaths. Twenty-eight studies had a sample size of less than 70 patients. With a potential incidence rate of 1 or 2 deaths per 1,000 patients, studies with small sample sizes are not likely to report a death due to surgery. If the studies with patients from the 1940s, 1950s, and 1960s are removed, three surgery-related deaths occurred among 1,608 patients (0.19 percent).

Table 17. Epilepsy surgery studies

Intervention	Total Number of Studies [a]
Temporal Lobe Surgery	105
Corpus Callosotomy	26
Frontal Lobe Surgery	18
Hemispherectomy	11
Multiple Subpial Transection	10
Parietal Lobe Surgery	2
Occipital Lobe Surgery	2
Surgical Controls	12

[a] Seven studies reported on more than one surgery category and are therefore double counted in this table

Table 18. Temporal lobe surgery: seizure-free outcome reporting

Reference	Undefined	No Auras	With Auras	Engel Class I	Reference	Undefined	No Auras	With Auras	Engel Class I
Bouilleret (2002)[165]		✓		✓	Son (1999)[166]				✓
Alsaadi (2001)[167]				✓	Maher (1998)[168]		✓		✓
Boling (2001)[169]				✓	Radhakrishnan (1998) [170]			✓	
Hennessy (2001) [171]			✓		Szabo (1998)[172]		✓		✓
Hennessy (2001) [148]			✓		Bizzi (1997) [173]				✓
Jan (2001)[174]			✓		Cappabianca (1997) [175]	✓			
Kanemoto (2001)[176]				✓	Casazza (1997)[177]		✓		✓
Schramm (2001)[178]				✓	Ho (1997) [179]			✓	
Sotero de Menezes (2001)[180]			✓	✓	Keene (1997)[181]				✓
Verma (2001) [182]			✓		Kilpatrick (1997) [183]			✓	✓
Wilson (2001) [184]		✓	✓		McLachlan (1997) [185]		✓		
Dupont (2000)[186]		✓		✓	Schwartz (1997) [187]		✓	✓	
Eberhardt (2000)[188]		✓	✓	✓	Silander (1997) [189]	✓			
Foldvary (2000) [190]				✓	Sisodiya (1997)[191]		✓	✓	✓
Holmes (2000) [192]		✓			Adam (1996) [193]		✓	✓	✓
Iannelli (2000)[194]				✓	Goldstein (1996) [195]	✓			
Markand (2000) [151]		✓	✓	✓	Holmes (1996)[196]		✓		
Rao (2000)[197]		✓			Sirven (1996)[198]	✓			
Robinson (2000)[199]				✓	Acciarri (1995)[200]	✓			
Assaf (1999) [201]			✓		Berkovic (1995)[202]				✓
Eriksson (1999) [203]			✓		Davies (1995)[163]				✓
Henry (1999)[204]	✓				Jooma (1995) [205]				✓
Holmes (1999) [206]		✓			Jooma (1995) [207]				✓
Mathern (1999)[208]	✓				Liu (1995)[209]			✓	
Mitchell (1999)[210]			✓		Renowden (1995) [211]		✓		
Rossi (1999) [212]		✓			Thadani (1995) [213]				✓
Salanova (1999)[214]				✓	Vossler (1995)[215]			✓	

Table 18. Temporal lobe surgery: seizure-free outcome reporting (continued)

Reference	Undefined	No Auras	With Auras	Engel Class I
Blume (1994)[146]			✓	✓
Guldvog (1994)[160]	✓			
Guldvog (1994)[161]	✓			
Berkovic (1991)[216]				✓
Hopkins (1991)[217]	✓			
Rasmussen (1991)[218]	✓			
Wieser (1991)[219]				✓
Bidzinski (1990)[162]	✓			
Mizrahi (1990)[152]	✓			
Walczak (1990)[164]			✓	
Yeh (1990)[220]			✓	
Sperling (1989)[221]				✓
Estes (1988)[222]			✓	
Bladin (1987)[223]	✓			
Cutfield (1987)[147]	✓			
Drake (1987)[224]	✓			
Lieb (1986)[225]				✓
Meyer (1986)[226]		✓	✓	
Delgado-Escueta (1985)[153]		✓	✓	✓

Table 19. Temporal lobe surgery: seizure-free outcome reporting in "control" patients

Reference	Seizure-Free Outcome Measurements			
	Undefined	No Auras	With Auras	Engel Class I
Bauer (2001)[227]	✓			
Kumlien (2001)[228]	✓			
Wiebe (2001)[145]		✓	✓	
Markand (2000) [151]		✓		
Holmes (1998) [229]	✓			
Wolf (1998) [230]		✓	✓	✓
McLachlan (1997) [185]	✓			
Hermanns (1996)[231]	✓			
Vickrey (1995)[232]		✓	✓	
Guldvog (1991)[233]	Reported changes in seizure frequency only			
Huttenlocher (1990)[234]	✓			
Harbord (1987)[235]	✓			

Table 20. Temporal lobe surgery: study designs

Study designs for studies of temporal lobe surgery reporting seizure-free outcomes

Prospective Nested Case-controlled Studies[a]	Retrospective Nested Case-controlled Studies	Prospective Case Series Studies	Retrospective Case Series Studies
5	30	5	33

[a] Nested case-controlled studies are defined as any study reporting patient characteristics (age at treatment, age at seizure onset, duration of epilepsy prior to treatment, etc.) separately for patients with good outcomes (seizure-free, Engel Class I, etc.) and patients with poor outcomes. Nested case-controlled studies are also considered case series studies because all patients received the same treatment.

Table 21. Temporal lobe surgery: individual patient data

Studies of temporal lobe surgery reporting individual patient data for patients with successful and nonsuccessful surgery

Reference	Age at Treatment	Age at Seizure Onset	Duration of Epilepsy Prior to Treatment	Gender	Side of Surgery	Simple Partial Seizures	Secondarily Generalized Seizures
Bouilleret (2002)[165]	✓	✓	✓	✓	✓	✓	
Hennessy (2001)[148]				✓	✓	✓	✓
Hennessy (2001)[171]							✓
Sotero de Menezes (2001)[180]	✓				✓		
Verma (2001)[182]	✓	✓	✓	✓	✓		
Eberhardt (2000)[188]	✓			✓	✓		
Holmes (1999)[206]	✓			✓	✓		
Szabo (1998)[172]	✓	✓	✓	✓	✓		
Kilpatrick (1997)[183]	✓			✓	✓		
Schwartz (1997)[187]	✓			✓	✓		
Sisodiya (1997)[191]	✓			✓	✓		
Adam (1996)[193]		✓			✓		
Goldstein (1996)[195]					✓		
Jooma (1995)[205]	✓	✓	✓				
Liu (1995)[209]	✓	✓	✓	✓	✓		✓
Vossler (1995)[215]	✓	✓	✓	✓	✓		
Blume (1994)[146]	✓						
Berkovic (1991)[216]	✓	✓	✓	✓	✓	✓	✓
Hopkins (1991)[217]	✓	✓	✓				
Mizrahi (1990)[152]	✓	✓	✓		✓		
Yeh (1990)[220]	✓	✓	✓	✓	✓	✓	✓
Estes (1988)[222]					✓		
Drake (1987)[224]	✓	✓	✓	✓		✓	✓
Delgado-Escueta (1985)[153]	✓	✓	✓	✓	✓	✓	✓

Table 22. Temporal lobe surgery: employment studies

Reference	Number of Patients	Country	Years Study Conducted	Mean Age at Surgery	Youngest Patient	Oldest Patient
Boling (2001)[169]	18	Canada	1981-1999	54	50	64
Reeves (1997) [154]	134	United States	1988-1991	31		
Sperling (1995)[236]	73	United States	1986-1990	33.2		
Mizrahi (1990)[152]	22	United States	1980-1986	21	7	36
Delgado-Escueta (1985)[153]	15	United States	1972-1983	26.5	12	39

Table 23. Temporal lobe surgery: new cases of depression after surgery

Reference	Number of Patients	Country	Years Study Conducted	Number of Cases
Kanemoto (2001)[176]	52	Japan	1987-1999	2
Kohler (2001)[237]	58	United States	1986-1999	6
Nees (2001)[238]	50	England	1992-1994	14
Wiebe (2001)[145]	36	Canada	1996-2000	7
Anhoury (2000)[239]	109	England	1988-1997	26
Derry (2000) [240]	39	Canada	1996-1998	4
Altshuler (1999) [156]	49	United States	1974-1990	5
Ring (1998)[241]	60	England	1995-1996	7
Naylor (1994)[242]	37	Denmark	1987-1991	2
Bladin (1992)[243]	107	Australia	1975-1991	5

Table 24. Temporal lobe surgery: new cases of psychosis after surgery

Reference	Number of Patients	Country	Years Study Conducted	Number of Cases
Kanemoto (2001)[176]	52	Japan	1987-1999	7
Wiebe (2001)[145]	36	Canada	1996-2000	1
Anhoury (2000) [239]	109	England	1988-1997	3
Blumer (1998) [244]	44	United States	1994-1995	2
Naylor (1994)[242]	37	Denmark	1987-1991	0
Bladin (1992) [243]	107	Australia	1975-1991	3

Table 25. Temporal lobe surgery: changes in IQ

Studies of temporal lobe surgery reporting both the number of patients with IQ changes after surgery and the pretreatment and posttreatment mean IQ.

Reference	Number of Patients	Country	Years Study Conducted	Number of Decreases	Number of Increases	Mean Age at Surgery	Youngest Patient	Oldest Patient
Miranda (2001) [245]	50	Canada	1976-1998	7	7	13.3	6.4	18.3
Robinson (2000)[199]	21	United States	1993-1998	1	4	15.4	9.4	21.7
Westerveld (2000)[246]	82	United States		8	7	14.4	6	17
Chelune (1993)[158]	96	United States	1990-1991	8	8	29.4		
Ivnik (1988)[247]	141	United States	1972-1987	13	27	28		
Powell (1985)[248]	59	England	1973-1984	10	8	25.5	15	

Table 26. Temporal lobe surgery: changes in memory

Studies of temporal lobe surgery reporting individual changes in patient memory after surgery

Reference	Number of Patients	Country	Years Study Conducted	Number of Decreases	Number of Increases	Mean Age at Surgery	Youngest Patient	Oldest Patient
Canizares (2000)[249]	33	Spain	1998-1999	3	10	30.9		
Chelune (1993)[158]	96	United States	1990-1991	28	1	29.4		
Ivnik (1988)[247]	141	United States	1972-1987	48	48	28		
Ojemann (1985)[250]	13	United States	1983-1983	8	3	28.9	17	49
Powell (1985)[248]	59	England	1973-1984	8	13	25.5	15	

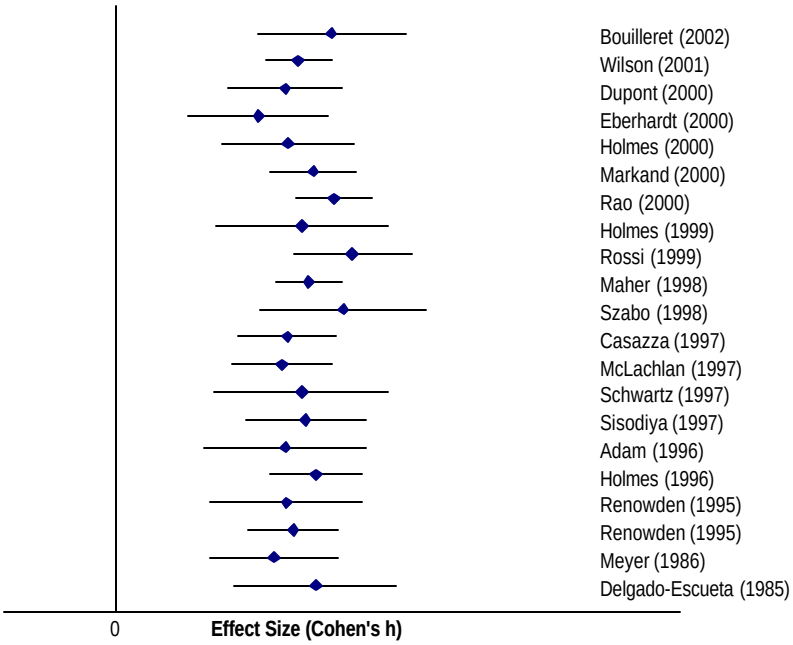
Table 27. Temporal lobe surgery: complications due to surgery

Reference	N	Years Study Conducted	Country	Permenant Complications	Reference	N	Years Study Conducted	Country	Permenant Complications
Boling (2001)[169]	18	1981-1999	Canada	0	Acciarri (1995)[200]	10	1975-1992	Italy	0
Schramm (2001)[178]	61	1993-1999	Germany	0	Davies (1995)[163]	12	1969-1988	England	1
Sotero de Menezes(2001)[180]	15	1978-1993	United States	1	Jooma (1995)[205]	30	1985-1992	United States	2
Wiebe (2001)[145]	36	1996-2000	Canada	2	Liu (1995)[209]	22	1983-1990	United States	2
Iannelli (2000)[194]	37	1981-1997	Italy	1	Wyler (1995)[142]	70	1990-1992	United States	3
Rao (2000)[197]	164	1995-1998	India	1	Blume (1994)[146]	125	1974-1989	Canada	1
Robinson (2000)[199]	21	1993-1998	United States	0	Guldvog (1994)[160]	64	1952-1988	Norway	0
Wurm (2000)[251]	16	1997-1998	Austria	1	Guldvog (1994)[161]	35	1949-1988	Norway	6
Altshuler (1999)[156]	49	1974-1990	United States	5	Hopkins (1991)[217]	11	1978-1988	Australia	0
Leung (1999)[252]	11	1994-1998	Hong Kong	1	Bidzinski (1990)[162]	320	1957-1988	Poland	2
Parrent (1999)[253]	19	1994-1997	Canada	0	Mackenzie (1990)[254]	30	1983-1989	Australia	0
Salanova (1999)[214]	145	1984-1995	United States	2	Mizrahi (1990)[152]	22	1980-1986	United States	0
Son (1999)[166]	71	1994-1999	South Korea	2	Walczak (1990)[164]	100	1964-1985	United States	1
Visudhiphan (1999)[255]	14	1993-1998	Thailand	0	So (1989)[256]	48	1973-1987	Canada	0
Radhakrishnan (1998)[170]	175	1988-1991	United States	2	Cutfield (1987)[147]	26	1961-1980	New Zealand	0
Wyllie (1998)[257]	72	1990-1996	United States	0	Drake (1987)[224]	16	1974-1986	Canada	0
Bizzi (1997)[173]	14	1990-1994	United States	1	King (1986)[258]	23	1981-1983	United States	0
Blume (1997)[259]	14	1977-1994	Canada	0	Meyer (1986)[226]	50	1970-1983	United States	0
Kilpatrick (1997)[183]	36	1993-1995	Australia	0	Carey (1985)[260]	24	1975-1984	Ireland	0
Adam (1996)[193]	30	1991-1994	France	2	Delgado-Escueta(1985)[153]	15	1972-1983	United States	1

Table 28. Temporal lobe surgery: surgery-related mortality

Reference	N	Years Study Conducted	Country	Deaths
Boling (2001)[169]	18	1981-1999	Canada	0
Schramm (2001)[178]	61	1993-1999	Germany	0
Wiebe (2001)[145]	36	1996-2000	Canada	0
Iannelli (2000)[194]	37	1981-1997	Italy	0
Rao (2000)[197]	164	1995-1998	India	1
Robinson (2000)[199]	22	1993-1998	United States	0
Wurm (2000)[251]	16	1997-1998	Austria	0
Altshuler (1999)[156]	49	1974-1990	United States	0
Leung (1999)[252]	11	1994-1998	Hong Kong	0
Parrent (1999)[253]	19	1994-1997	Canada	0
Salanova (1999)[214]	145	1984-1995	United States	0
Son (1999)[166]	71	1994-1999	South Korea	0
Visudhiphan (1999)[255]	14	1993-1998	Thailand	0
Wyllie (1998)[257]	72	1990-1996	United States	1
Bizzi (1997)[173]	14	1990-1994	United States	0
Blume (1997)[259]	14	1977-1994	Canada	0
Kilpatrick (1997)[183]	36	1993-1995	Australia	0
Adam (1996)[193]	30	1991-1994	France	0
Acciarri (1995)[200]	10	1975-1992	Italy	0
Berkovic (1995)[202]	135	1986-1991	Australia	0
Davies (1995)[163]	12	1969-1988	England	0
Liu (1995)[209]	22	1983-1990	United States	0
Wyler (1995)[142]	70	1990-1992	United States	0
Blume (1994)[146]	125	1974-1989	Canada	0
Guldvog (1994)[160]	64	1952-1988	Norway	0
Guldvog (1994)[161]	35	1949-1988	Norway	0
Bladin (1992)[243]	107	1975-1991	Australia	0
Elwes (1991)[261]	108	1976-1987	England	1
Hopkins (1991)[217]	11	1978-1988	Australia	0
Bidzinski (1990)[162]	320	1957-1988	Poland	2
Mizrahi (1990)[152]	22	1980-1986	United States	0
Yeh (1990)[220]	12	1982-1986	Japan	0
So (1989)[256]	48	1973-1987	Canada	0
Cutfield (1987)[147]	26	1961-1980	New Zealand	0
Drake (1987)[224]	16	1974-1986	Canada	0
Meyer (1986)[226]	50	1970-1983	United States	0
Carey (1985)[260]	24	1975-1984	Ireland	0

Figure 35. Forest plot: temporal lobe surgery and seizure-free with no auras



A scale is not shown because the effect sizes were not calculated with actual control groups

Figure 36. Threshold analysis: temporal lobe surgery and seizure-free with no auras

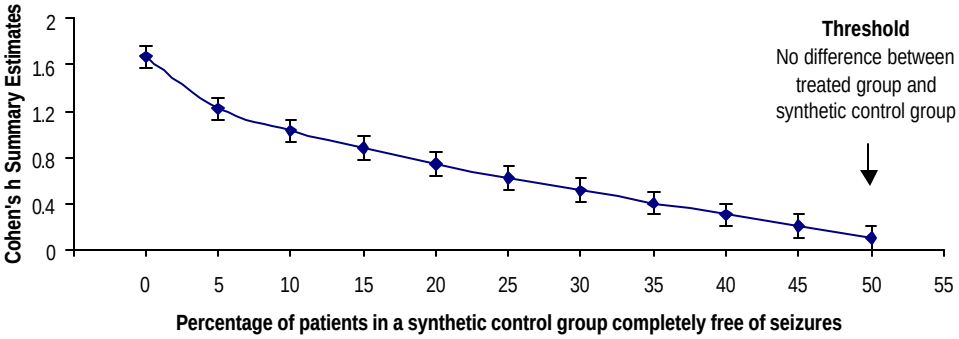
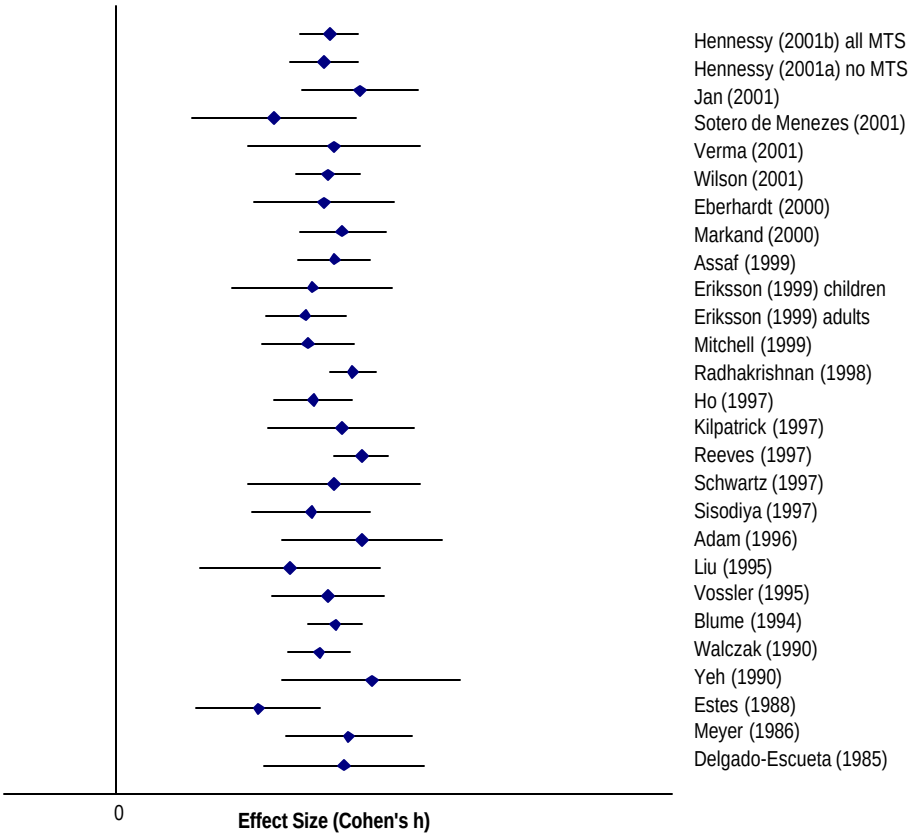


Figure 37. Forest plot: temporal lobe surgery and seizure-free with auras



A scale is not shown because the effect sizes were not calculated with actual control groups
MTS = Patients with mesial temporal sclerosis

Figure 38. Threshold analysis: temporal lobe surgery and seizure-free with auras

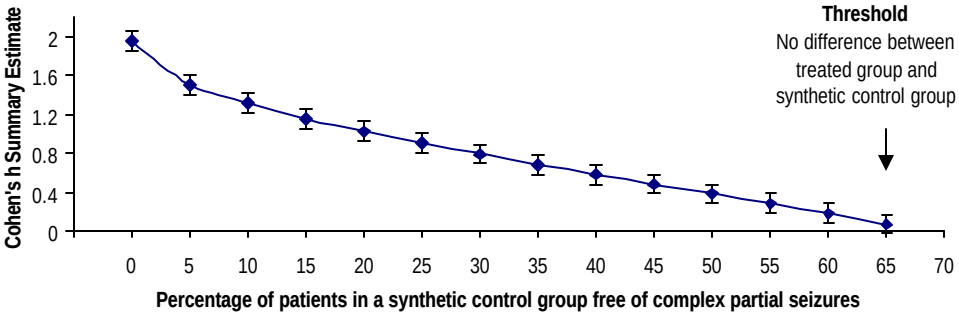
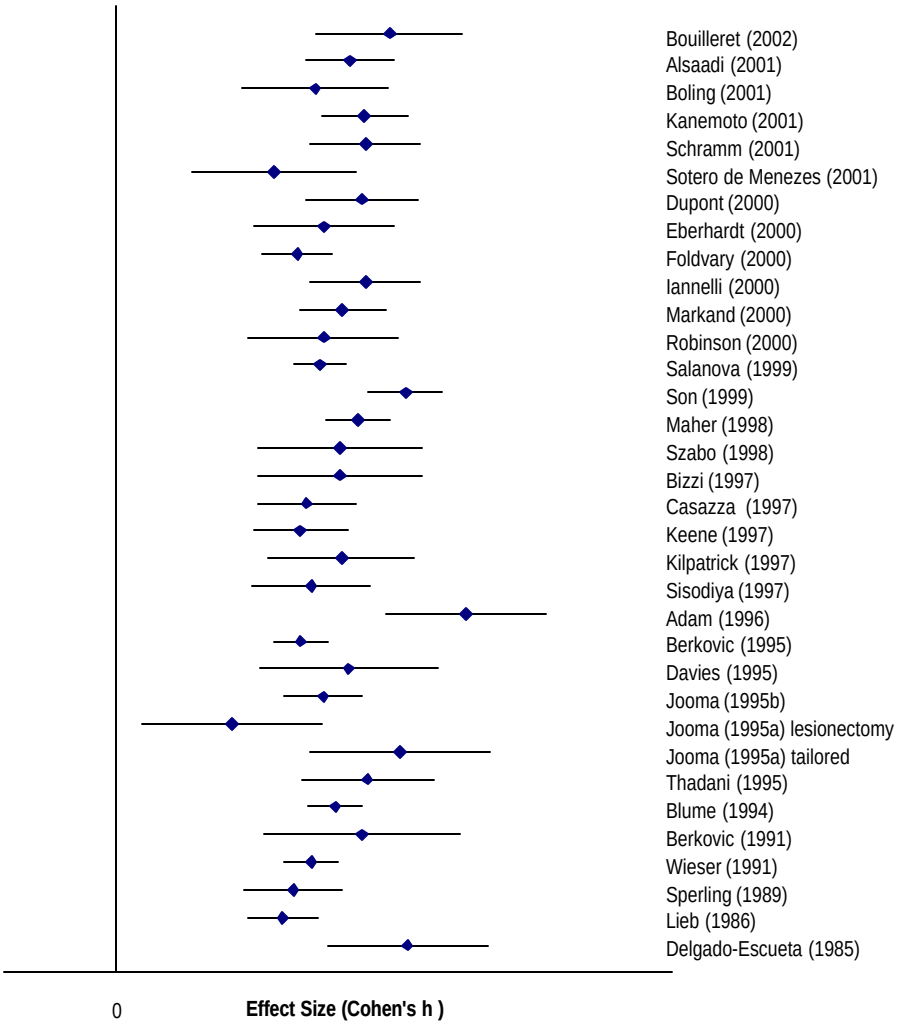


Figure 39. Forest plot: temporal lobe surgery and Engel Class I



A scale is not shown because the effect sizes were not calculated with actual control groups

Figure 40. Meta-regression: temporal lobe surgery and Engel class I

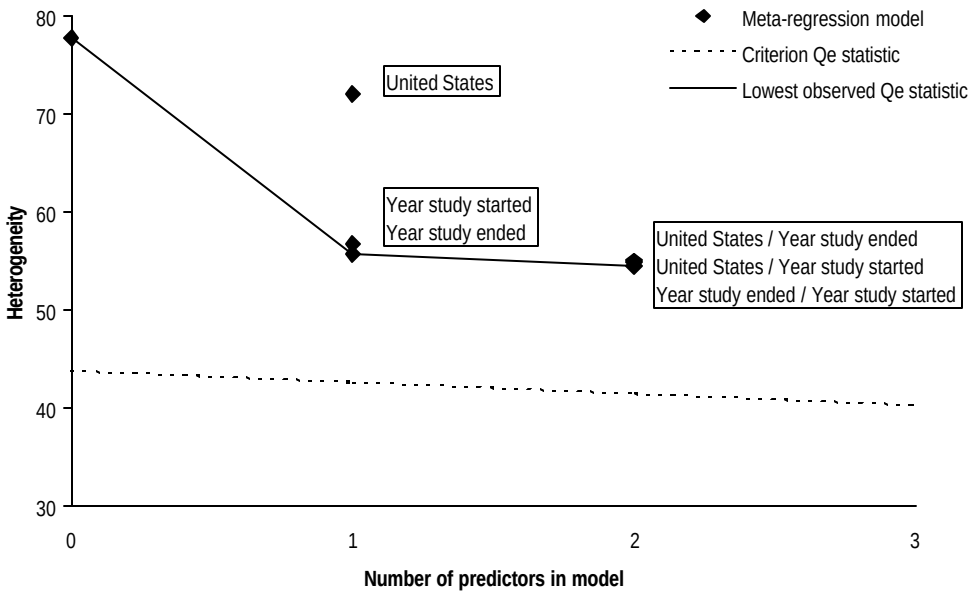
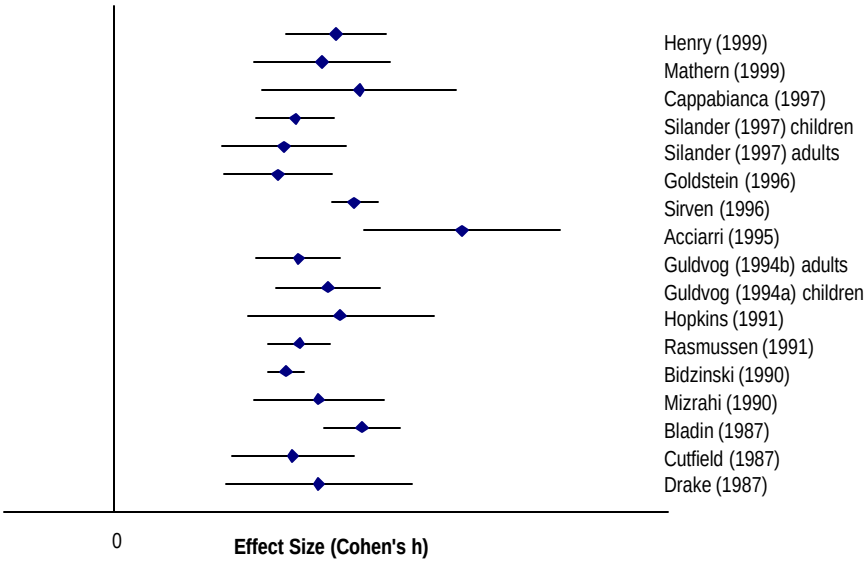


Figure 41. Forest plot: temporal lobe surgery and seizure-free undefined



A scale is not shown because the effect sizes were not calculated with actual control groups

Figure 42. Meta-regression: temporal lobe surgery and seizure-free undefined

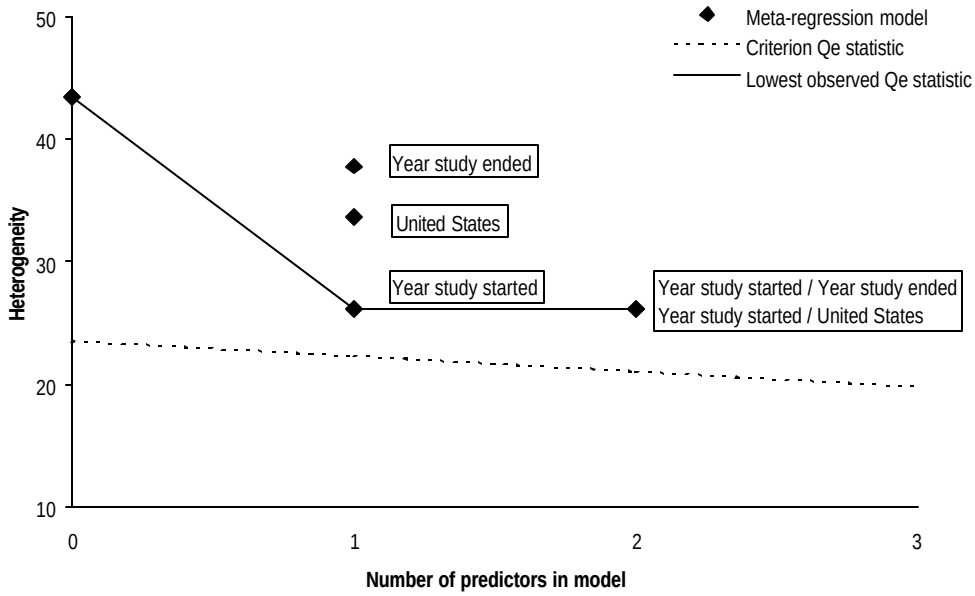


Figure 43. Forest plot: temporal lobe surgery and patient age at surgery

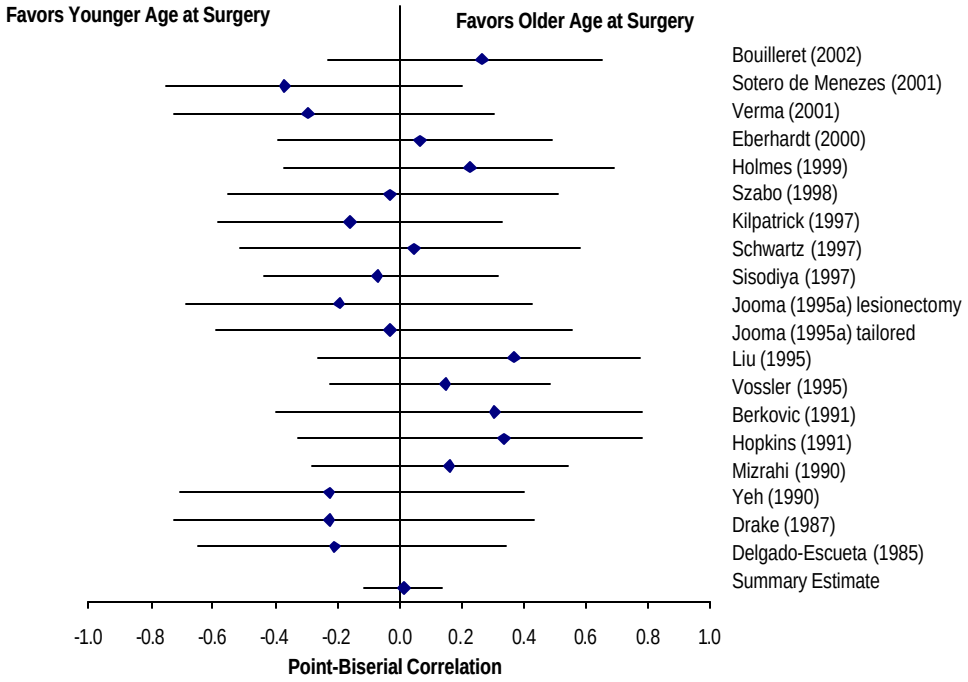


Figure 44. Forest plot: temporal lobe surgery and patient age at onset of seizures

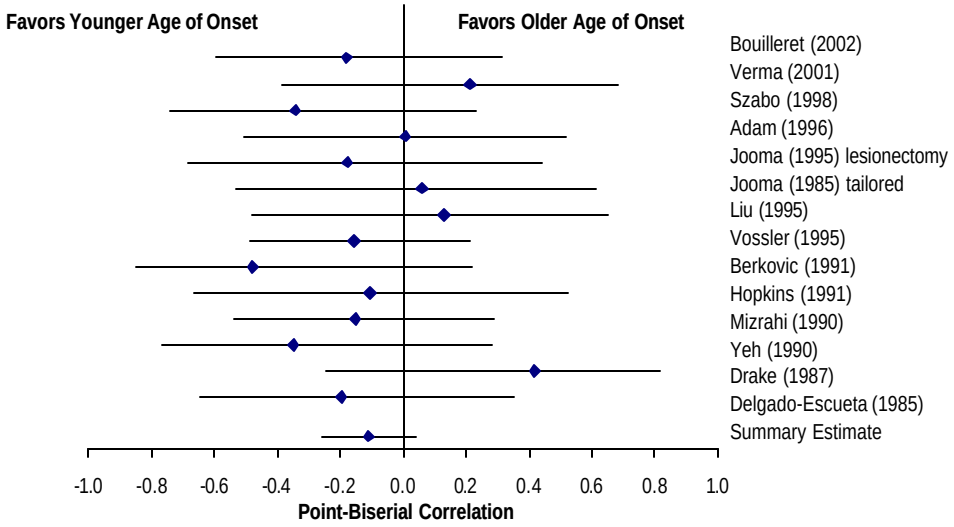


Figure 45. Forest plot: temporal lobe surgery and duration of epilepsy prior to surgery

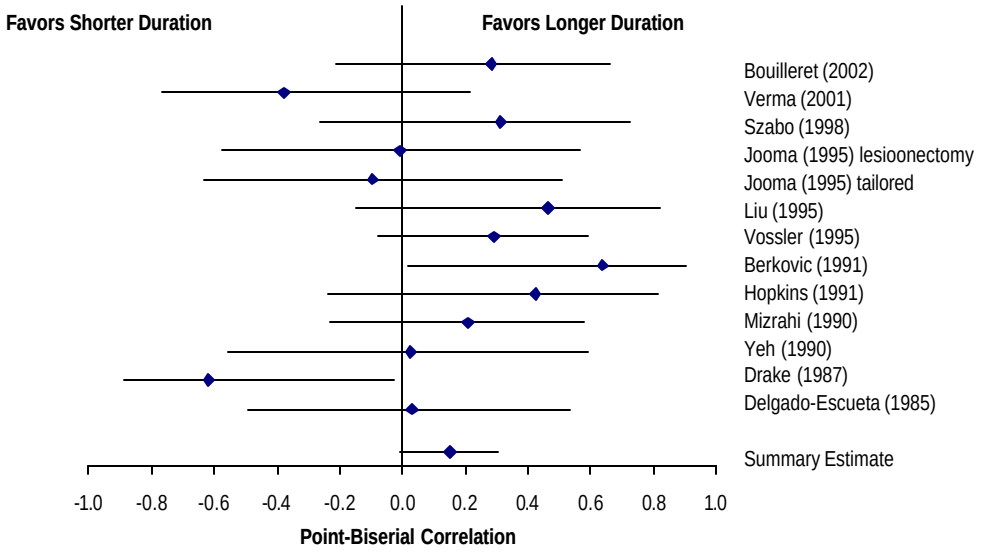


Figure 46. Forest plot: temporal lobe surgery and male and female patients

Studies reported the success of surgery among male and female patients

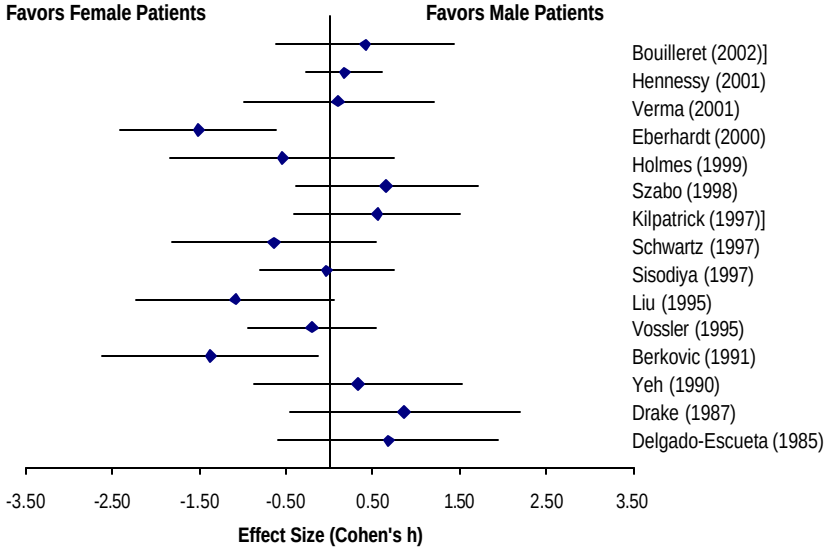


Figure 47. Meta-regression: temporal lobe surgery and male and female patients

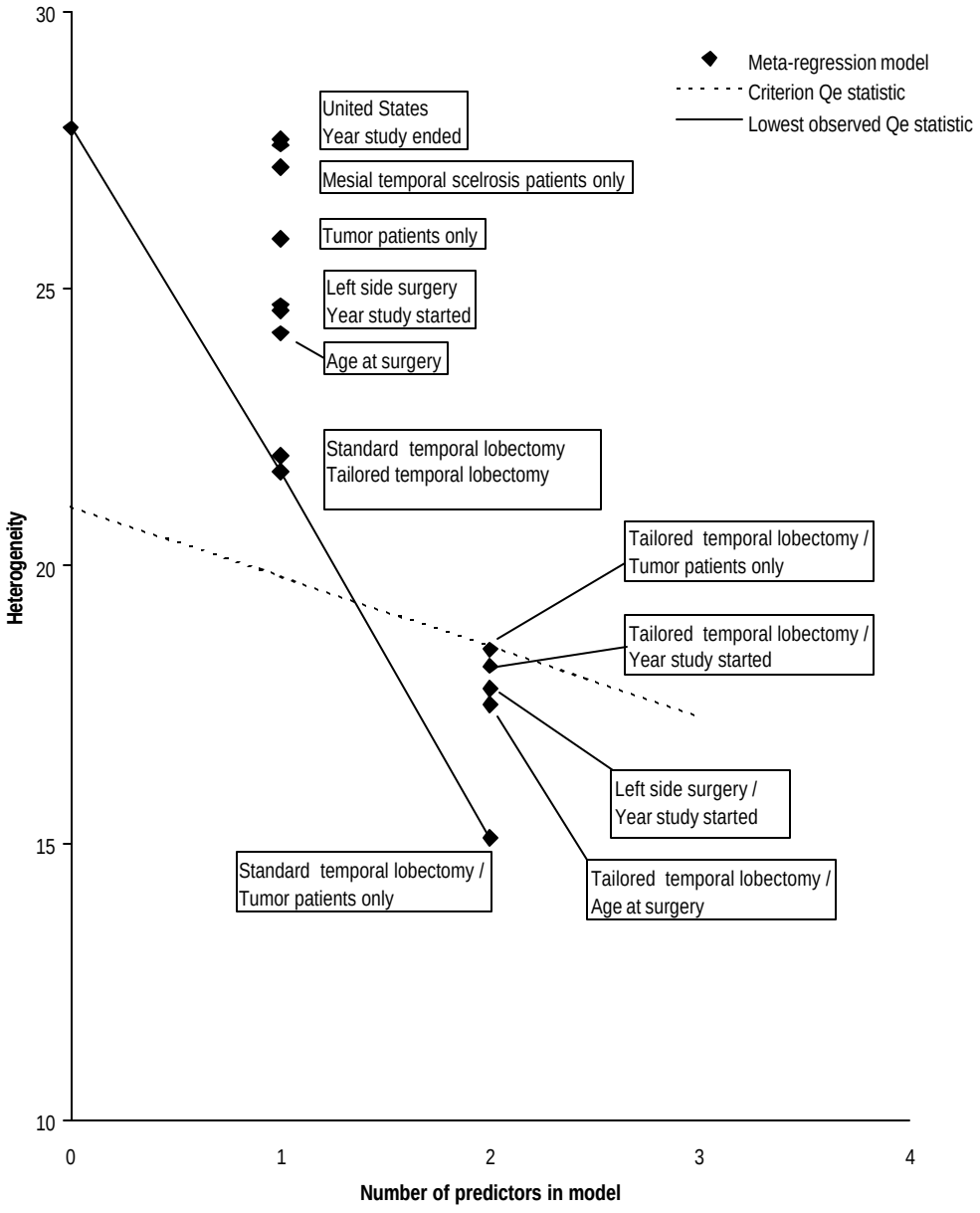


Figure 48. Forest plot: temporal lobe surgery and location of surgery

Studies reported the success of surgery among patients with left side and right side surgery

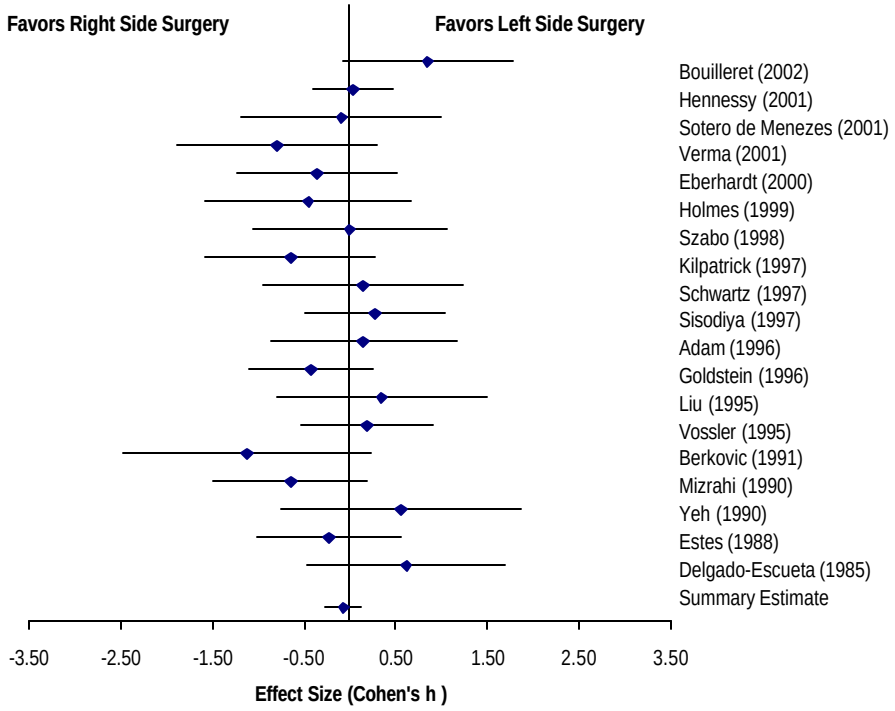


Figure 49. Forest plot: temporal lobe surgery and simple partial seizures

Studies reported the success of surgery in patients with and without simple partial seizures (SPS)

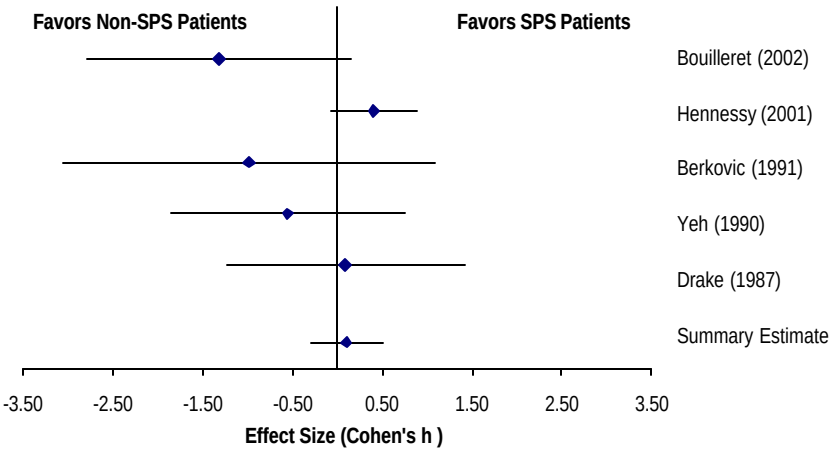
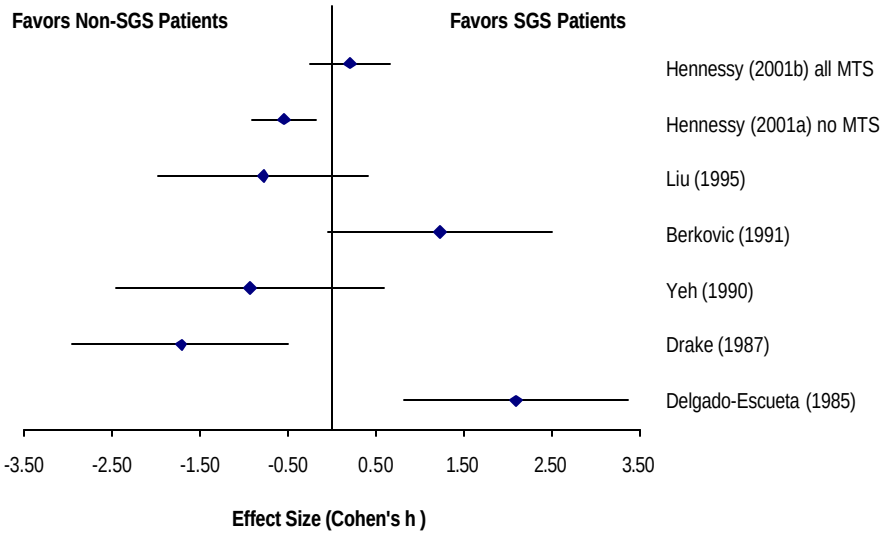


Figure 50. Forest plot: temporal lobe surgery and secondarily generalized seizures

Studies reported the success of surgery among patients with and without secondarily generalized seizures (SGS)



MTS = Patients with mesial temporal sclerosis

Figure 51. Meta-regression: temporal lobe surgery and secondarily generalized seizures

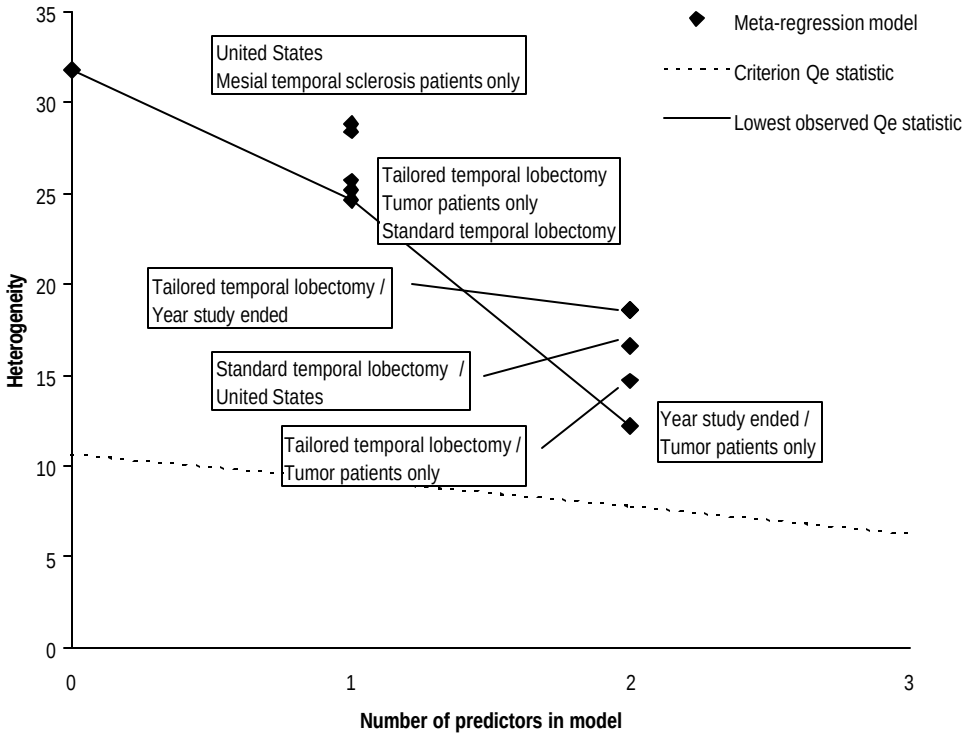
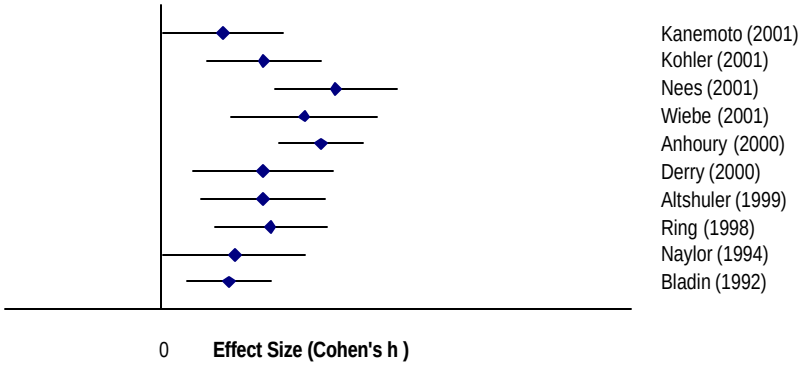


Figure 52. Forest plot: temporal lobe surgery and new cases of depression



A scale is not shown because the effect sizes were not calculated with actual control groups

Figure 53. Meta-regression: temporal lobe surgery and new cases of depression

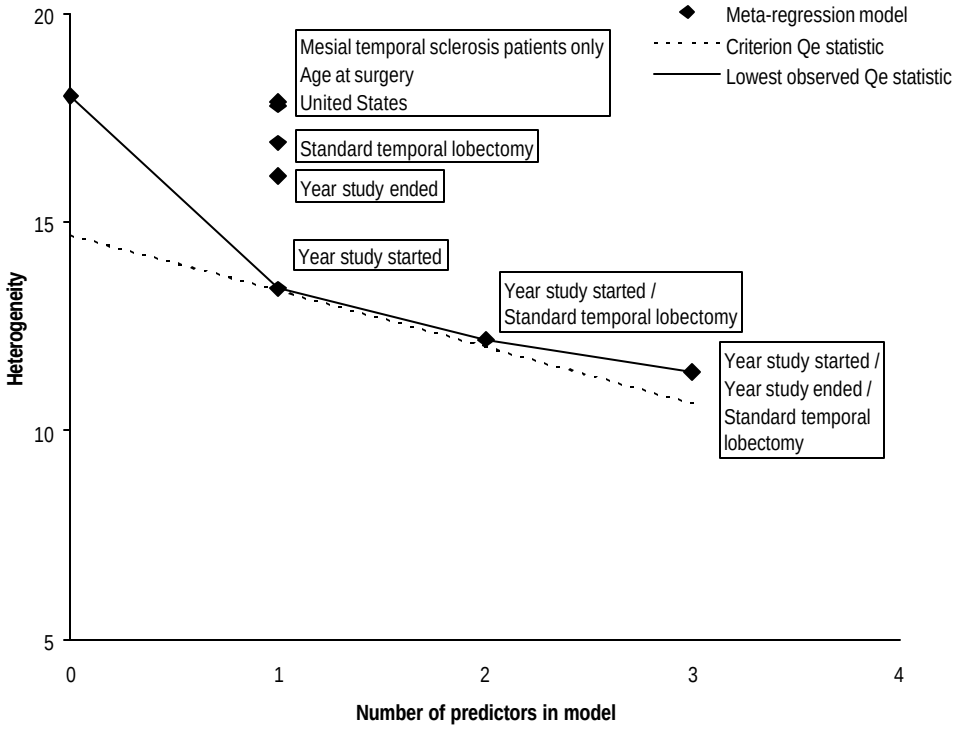
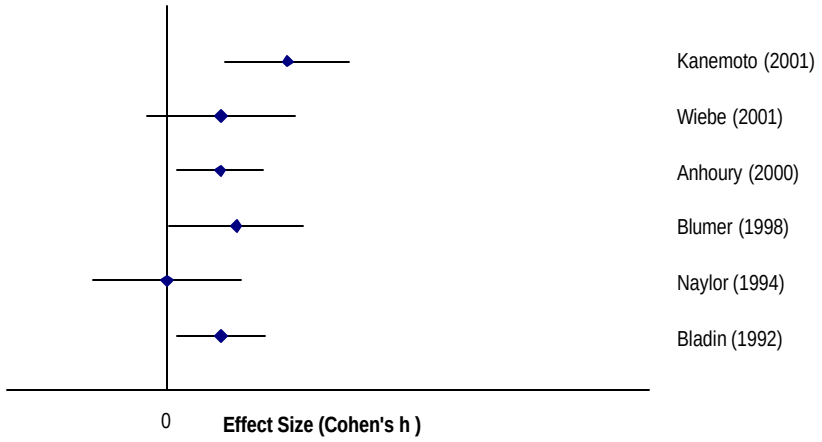


Figure 54. Forest plot: temporal lobe surgery and new cases of psychosis



A scale is not shown because the effect sizes were not calculated with actual control groups

Figure 55. Threshold analysis: temporal lobe surgery and new cases of psychosis

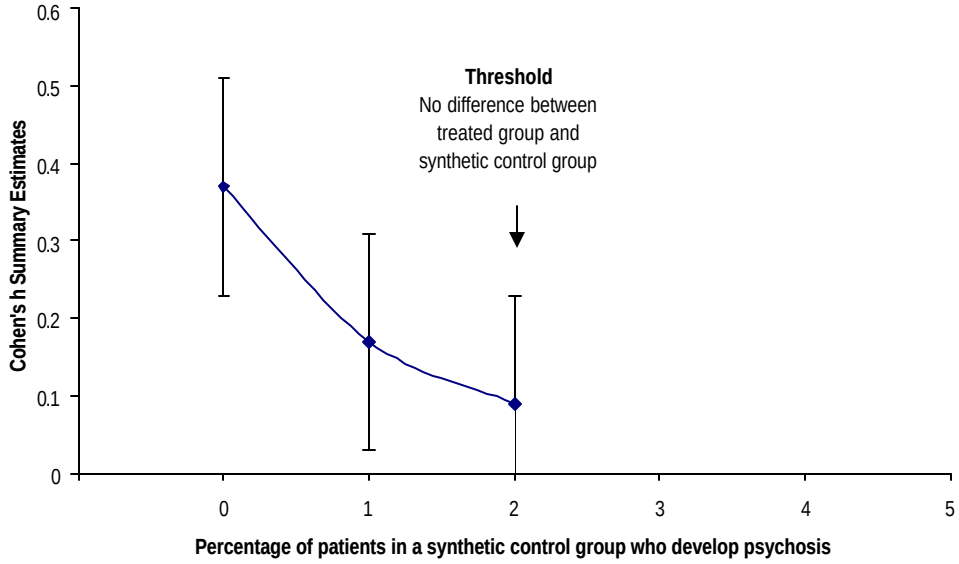
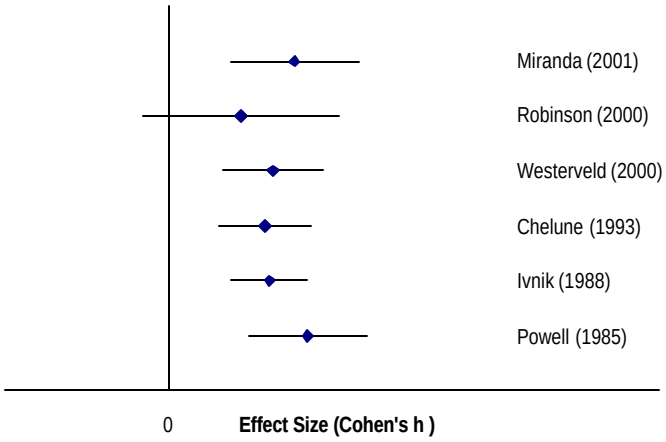


Figure 56. Forest plot: temporal lobe surgery and decreases in IQ after surgery

Studies reported individuals with significant decreases in IQ after surgery



A scale is not show n because the effect sizes were not calculated with actual control groups

Figure 57. Threshold analysis: temporal lobe surgery and decreases in IQ after surgery

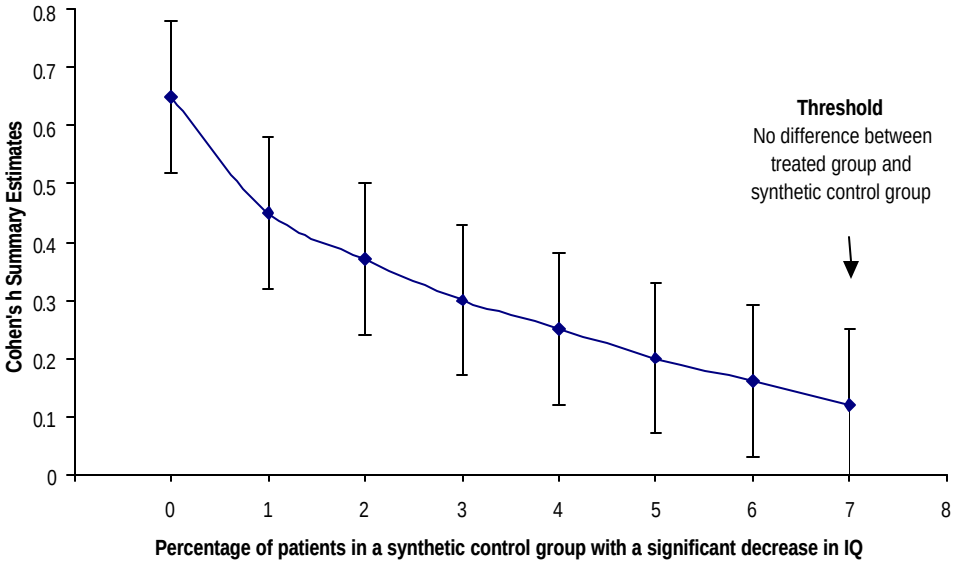
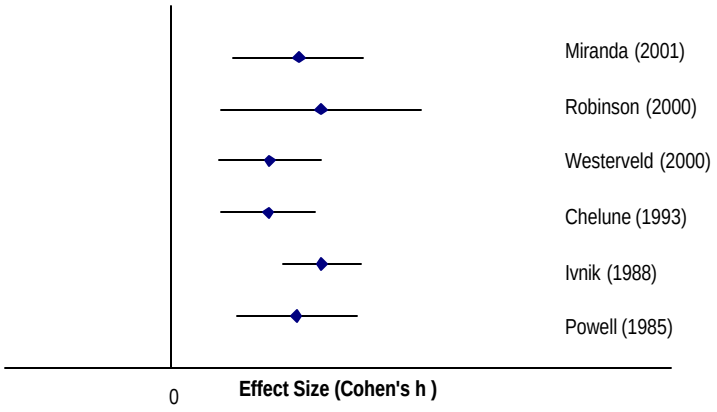


Figure 58. Forest plot: temporal lobe surgery and increases in IQ after surgery

Studies reported individuals with significant increases in IQ after surgery



A scale is not shown because the effect sizes were not calculated with actual control groups

Figure 59. Threshold analysis: temporal lobe surgery and increases in IQ after surgery

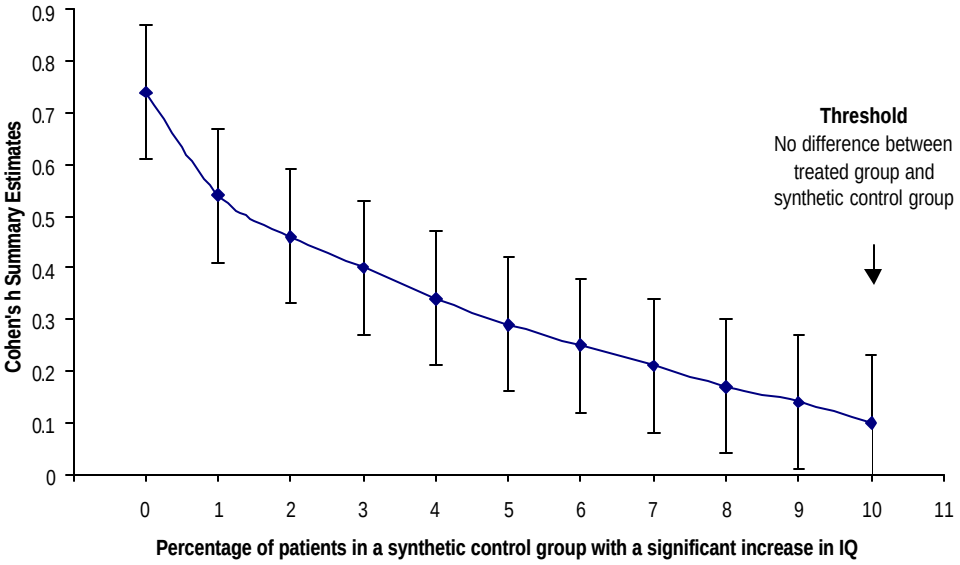


Figure 60. Forest plot: temporal lobe surgery and changes in mean IQ

Studies reported both presurgery and postsurgery mean IQ

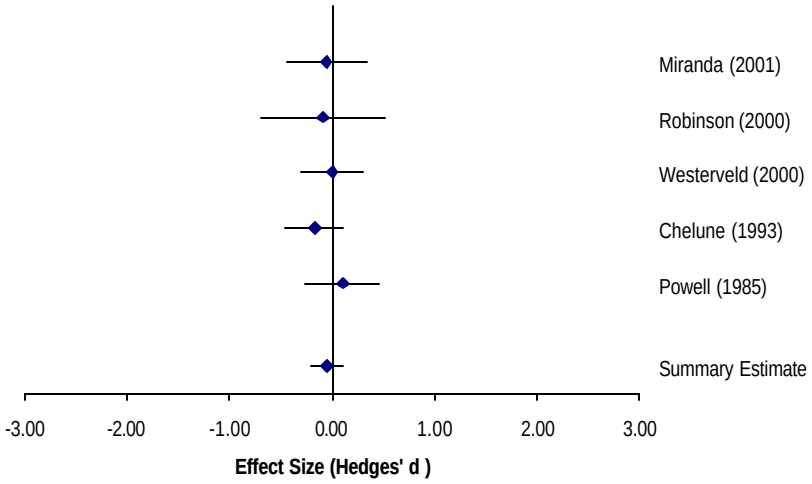


Figure 61. Temporal lobe surgery: changes in memory after surgery

Studies reported individuals with significant changes in memory after surgery

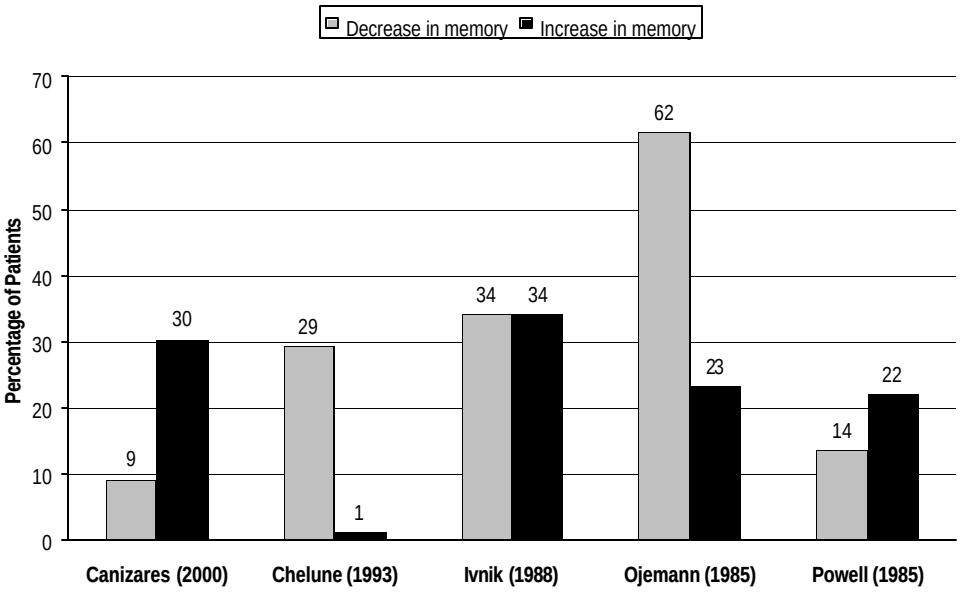
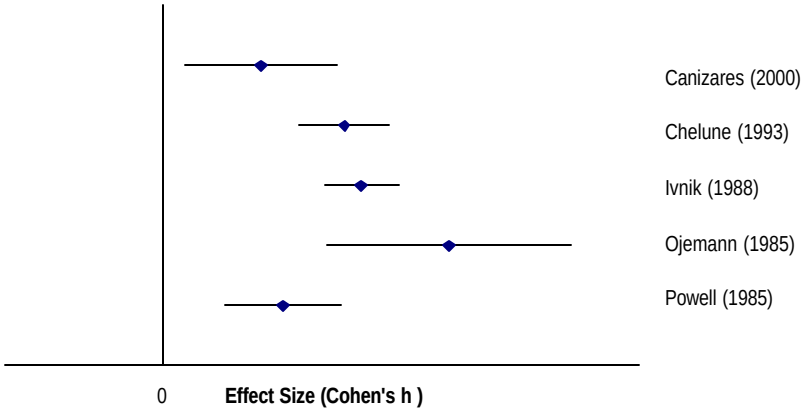


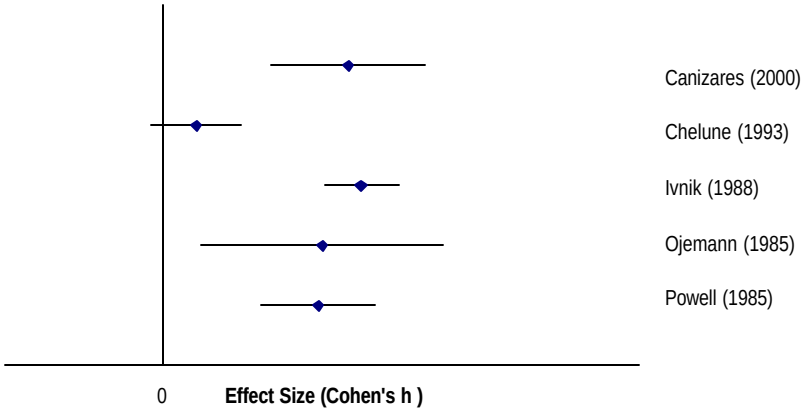
Figure 62. Forest plot: temporal lobe surgery and changes in memory

Decreases in memory scores



A scale is not shown because the effect sizes were not calculated with actual control groups

Increases in memory score.



A scale is not shown because the effect sizes were not calculated with actual control groups

Corpus Callosotomy

Resection of the corpus callosum is intended as a palliative procedure that reduces the frequency of seizures that could lead to injury or seriously interfere with life-style.[262-264] These patients typically have multifocal, unresectable, or unlocalized lesions.[263] Candidates for this procedure include both children and adult patients with atonic, tonic, and tonic-clonic seizures.[263] These patients typically have daily to weekly seizures of multiple types that occur despite therapeutic blood levels of AEDs for at least 2 years prior to surgery.[265]

Corpus callosotomy is not expected to eliminate all seizures. Under these circumstances, reduction in overall seizure frequency and in specific seizure frequencies are the most valuable outcome measurement for establishing if corpus callosotomy has been effective. Due to the complicated nature of the surgery, an assessment of surgical complications and deaths due to surgery is also necessary to judge the effectiveness of corpus callosum resection.

Percent Reduction in Overall Seizure Frequency

Excluded studies

We excluded one study of corpus callosotomy reporting seizure frequency outcome measures from the evidence base. This study and the reason for its exclusion are listed in Evidence Table 163.

Evidence base

Among the 26 studies of corpus callosotomy meeting our inclusion criteria, 12 reported an outcome measurement related to seizure frequency. Three hundred and forty-nine patients were examined in these studies. Table 29 presents a list of the seizure outcome categories used by each of the 12 studies. The most common means of reporting the effect of surgery on seizure occurrence was to classify patients into groups based on their percentage reduction in the frequency of all seizures or one particular seizure type. Eight studies considered the percentage reduction in all seizure types, while the remaining four studies measured only changes in the most disabling seizure, in disabling generalized seizures, or in drop attacks and generalized tonic-clonic seizures. Four studies reported the number of patients who were seizure-free for all seizure types.

Design and conduct of included studies

Internal validity

As noted for temporal lobe surgery, withholding surgery may be unethical, so the evidence base for corpus callosotomy consists mainly of uncontrolled trials. Indeed, none of the studies in the evidence base for corpus callosotomy employed a control group. Rather, all studies were case series. Therefore, all of the 12 studies in the evidence base may have biases that reduce internal validity as previously discussed for temporal lobe surgery. However, these patients have daily to weekly seizures of multiple types that occur despite therapeutic blood levels of AEDs.[265] Therefore, given the severe nature of the seizure activity in individuals considering this type of surgery, explanations for seizure reduction other than the effect of surgery may be considered implausible.

External validity

As stated earlier, patients being considered for resection of the corpus callosum experience characteristics seizures due to the multifocal nature of the lesions responsible for the seizures. Therefore, the patients in published studies of corpus callosotomy should be representative of all patients receiving this surgery. However, differences may exist across studies with regard to age or pathology.

The specific characteristics of corpus callosotomy patients reported in each study are presented in Evidence Table 164. In two studies, the patients were all less than 20 years of age at the time of surgery. In the other 10 studies, the mean age at the time of surgery varied between 20 to 30 years of age, with patient ages ranging from a youngest of about 5 years to an oldest of about 50 years of age. The mean age of seizure onset was less than 10 years of age in the nine studies reporting this patient characteristic. The age of seizure onset ranged from birth to 26 years of age. The mean duration of epilepsy prior to surgery was less than 10 years in one study and between 13 and 21 years in another eight studies. The range for duration of epilepsy prior to surgery was less than a year to 50 years.

Based on the distribution of patient characteristics, this evidence base seems to be generalizable to corpus callosotomy patients in clinical practice.

Synthesis of study results

As with temporal lobe surgery, several categories are used to describe reductions in seizure frequency. Evidence Table 165 presents the data from 12 studies organized according to the percentage of seizure frequency reduction. Ten studies reported a category of 90 percent or greater reduction. Nine studies reported frequency data that could be organized into categories of seizure reductions of greater than or equal to 90 percent, 75 percent to 90 percent, 50 percent to 75 percent, less than 50 percent, and no change or worse. One study reported only the number of patients in the 90 percent reduction group. Two studies did not report a 90 percent reduction category; one study separated patients above and below a 50 percent reduction in seizure frequency and the other reported the number of patients to achieve better than a 75 percent reduction. This last study also reported the number patients with no change or who became worse, but the remaining patients could have been anywhere between 1 percent and 74 percent.

As previously mentioned, the types of seizures being evaluated are not the same across studies. In four of the 12 studies reporting a percentage reduction in seizure frequency, only a single specific type of seizure was considered (Table 29).

Meta-analytic threshold analysis of 90 percent reduction in seizure frequency

Five studies reported the number of patients with a greater than or equal to 90 percent reduction for all seizure types. Evidence Table 166 presents the actual patient counts, percentages, and calculated effect sizes for each study used in this analysis. The individual study effect sizes (Cohen's h) presented in the Evidence Table were based on a control group in which no patients experience a 90 percent reduction in seizure frequency. Figure 63 presents a forest plot of these effect sizes to show the extent of variation between studies, but no scale is provided because these effect sizes were not calculated using actual control groups.

The results of our threshold analysis appear in Figure 64. The summary estimate calculated at the 0 percent point (no patients in a synthetic control group showed a 90 percent reduction in seizure frequency) was 0.94 (CI: 0.70 to 1.18, $p < 0.000001$). This summary estimate

corresponded to 20 percent (CI: 12 percent to 31 percent)[w] of patients experiencing a 90 percent reduction in seizure frequency after surgery. The summary estimate became nonsignificant (no statistically significant difference between surgery and control) when the proportion of patients in the synthetic control group reached 15 percent. There was no statistically significant heterogeneity in the threshold analysis (Q = 2.9, p = 0.58).

Meta-analysis of no change or increase in seizure frequency

Seven studies reported the number of patients who experienced no change or became worse for all seizure types, and we performed a meta-analysis of these studies. Evidence Table 167 presents the actual patient counts, percentage, and calculated effect sizes (Cohen's h) for each study used in this analysis. We calculated each study's Cohen's h presented in the Evidence Table under the assumption that no surgical patients would have been included under the category of no change or became worse. Figure 65 presents a forest plot of the effect sizes. We did not conduct a threshold analysis because control patients are expected to experience the outcome we are meta-analyzing, no change or an increase in seizure frequency.

Our meta-analysis produced a statistically significant summary estimate (0.83, CI: 0.62 to 1.03, p <0.000001) that corresponds to 16 percent (CI: 9 percent to 24 percent) of patients with no change or an increase in seizure frequency after surgery. There was no statistically significant heterogeneity in the threshold analysis (Q = 9.0, p = 0.17).

Analysis of seizure-free outcome measurements

Four studies reported the number of patients who became completely seizure-free (no auras) after resection of the corpus callosum. Of the 85 patients examined in these four studies, only five patients became seizure-free (6 percent). The range among the individual studies was 0 percent to 14 percent. Evidence Table 168 presents the data from these studies.

Analysis of presurgery and postsurgery seizure frequency outcome measurements

Pre- and postsurgery seizure frequency data were reported in only three studies. Therefore, we did not perform a meta-analysis of these data. All three studies reported a reduction in mean seizure frequency after surgery. Mean presurgery seizure frequency ranged from 110 to 178 seizures per month. The mean postsurgery seizure frequency dropped to a range of 20 to 78 per month. Because each of these studies reported individual patient data for seizure frequency, we looked for significant changes in seizure frequency in each study using a paired t-test. Two of the three studies showed statistically significant reductions in seizure frequency after surgery (p = 0.014 and 0.015) and the third showed a reduction close to being statistically significant (p = 0.065). These studies suggest that corpus callosotomy can be effective in reducing absolute seizure frequency. Evidence Table 169 presents the data abstracted from these studies and the results of our paired t-test calculations.

Analysis of nested case-control studies

Four nested case-control studies of corpus callosotomy presented an evaluation of patient characteristics that could potentially influence surgical outcomes. Evidence Table 170 presents the findings reported by each of the nested case-control studies in our evidence base for seizure frequency outcome measurements. One of the four studies used multiple regression, but did not assess age at surgery, age at seizure onset, or duration of epilepsy prior to treatment.

[w] Computed from a back-transformation of Cohen's h.

Meta-analysis of patient characteristics

As mentioned previously in the section on temporal lobe surgery, each nested case-control study may have been too small (i.e., had too little power) to detect clinically meaningful correlations between patient characteristics and successful surgery. To address this, we performed meta-analyses that combined individual patient data across studies. At least five studies reported individual patient data for one or more of the following continuous variables: age at surgery, age at seizure onset, or duration of epilepsy prior to surgery. We calculated a point-biserial correlation for each study and combined these in a meta-analysis. The coefficient was calculated so that a positive correlation indicated that an older age or longer duration favored a successful outcome and a negative correlation indicated that a younger age or shorter duration favored a successful outcome. Table 30 presents a list of the studies of corpus callosotomy that reported characteristics for patients with successful and nonsuccessful surgery. All of these studies were included in the previous meta-analyses examining the efficacy of surgery.

Age at surgery. Our first meta-analysis looks at whether different outcomes were obtained in patients of different ages at the time they receive surgery. Individual ages at surgery for patients with successful and nonsuccessful surgery were reported in six studies with 120 patients. Evidence Table 171 presents the definition used for successful surgery and the point-biserial correlation calculated for each of the six studies. Figure 66 presents a forest plot of the effect sizes.

The meta-analysis produced a summary estimate that was not statistically significant (r_{pb} = 0.14, CI: –0.05 to 0.32, p = 0.16) suggesting that age at surgery had no influence on the success of surgery in these studies. The effect sizes in this meta-analysis were not heterogeneous (Q = 4.1, p = 0.54).

We performed a sensitivity analysis to determine whether a single study had excessive influence over the results of the analysis. The summary estimate and other statistics did not change markedly because of the sensitivity analysis. The correlation changed by no more than 0.04 due to removal of studies during the sensitivity analysis. The summary estimates remained nonsignificant. The results of the sensitivity analysis and the original meta-analysis are presented in Evidence Table 172.

Age at seizure onset. In our second meta-analysis, we sought to determine whether different outcomes were obtained in patients of different ages at seizure onset. Individual ages at seizure onset for patients with successful and nonsuccessful surgery were reported in five studies with 105 patients. Evidence Table 173 presents the definition used for successful surgery and the point-biserial correlations calculated for each of the five studies. Figure 67 presents a forest plot of the effect sizes. The meta-analysis produced a summary estimate that was not statistically significant (r_{pb} = 0.04, CI: -0.16 to 0.24, p = 0.70) suggesting that age at seizure onset had little or no influence on the success of surgery in these studies. The effect sizes in this meta-analysis were not heterogeneous (Q = 2.6, p = 0.64).

We performed a sensitivity analysis to ensure that a single study did not have excessive influence over the results of the analysis. The summary estimate and other statistics did not change markedly because of the sensitivity analysis. The point-biserial correlation varied between 0.0 and 0.12 due to removal of studies during the sensitivity analysis. The summary estimates remained nonsignificant. The results of the sensitivity analysis as well as the original meta-analysis are presented in Evidence Table 174.

Duration of epilepsy prior to surgery. In our third meta-analysis, we sought to determine whether different outcomes were obtained in patients with different durations of epilepsy prior to surgery. Individual durations of epilepsy prior to surgery for patients with successful and nonsuccessful surgery were reported in five studies with 105 patients. These same five studies reported individual patient age at onset of seizures. Evidence Table 175 presents the definitions used for successful surgery and the point-biserial correlations calculated for each of the five studies. Figure 68 presents a forest plot of the effect sizes. The meta-analysis produced a summary estimate that was not statistically significant (r_{pb} = 0.15, CI: -0.05 to 0.34, p = 0.15) suggesting that duration of epilepsy prior to surgery had no influence on the success of surgery in these studies. The effect sizes in this meta-analysis were not heterogeneous (Q = 6.2, p = 0.18).

We performed a sensitivity analysis to ensure that a single study did not have excessive influence over the results of the analysis. The summary estimate and other statistics did not change markedly because of the sensitivity analysis. The point-biserial correlation varied between 0.06 and 0.21 due to removal of studies during the sensitivity analysis. The summary estimates remained nonsignificant. The results of the sensitivity analysis as well as the original meta-analysis are presented in Evidence Table 176.

Changes in the Frequency of Specific Seizure Types

In the previous section, we analyzed data on overall seizure reduction and estimated that, after corpus callosotomy, only 20 percent of patients are likely to exhibit a 90 percent reduction in all seizure types. The benefits of corpus callosotomy may also be determined from surgery's effect on the most disabling seizures experienced by the patients, or from surgery's effect on specific types of seizures. The most disabling seizures are primarily generalized seizures that can result in falls and injuries. The specific types of seizures for which corpus resection may have a beneficial effect are generalized tonic/clonic seizures, atonic seizures, and tonic seizures.[266]

Evidence base

Nine studies presented data on the number of patients who became free of specific types of seizures after corpus callosotomy. These studies are presented in Table 31.

Design and conduct of included studies

These nine studies were among the studies considered previously for changes in overall seizure frequency.

Synthesis of study results

Meta-analytic threshold analysis of most disabling seizure types

Seven studies reported patients who became free of their most disabling seizures after surgery. A total of 165 patients were examined in these studies. Evidence Table 177 presents the actual patient counts, percentages, and calculated effect sizes for each study in this analysis. The individual study effect sizes (Cohen's h) presented in the Evidence Table were based on a control group in which no patients became free of their most disabling seizures. Figure 69 presents a forest plot of these effect sizes to show the extent of variation between studies, but no scale is provided because these effect sizes were not calculated using actual control groups.

The results of our threshold analysis appear in Figure 70. The summary estimate calculated at the 0 percent point (no patients in a synthetic control group were free of their most disabling seizures) was 1.07 (CI: 0.86 to 1.29, p <0.000001) and corresponded to 26 percent (CI: 17

percent to 36 percent) of patients becoming free of their most disabling seizures after surgery. The summary estimate became nonsignificant (no statistically significant difference between surgery and control) when the proportion of patients in the synthetic control group reached 20 percent. There was no statistically significant heterogeneity in the threshold analysis (Q = 3.7, p = 0.72).

Meta-analytic threshold analysis of generalized tonic-clonic seizures

We next performed a threshold analysis of the eight studies reporting patients who were free of generalized tonic-clonic seizures after surgery. Two hundred and sixty-one patients were examined in these studies. Evidence Table 178 presents the actual patient counts, percentages, and calculated effect sizes (Cohen's h) for each study in this analysis. We calculated each study's effect size presented in the Evidence Table under the assumption that no patients in a synthetic control group would become free of generalized tonic-clonic seizures. Figure 71 presents a forest plot of these effect sizes to show the extent of variation between studies, but no scale is provided because these effect sizes were not calculated using actual control groups.

Our threshold analysis found statistically significant heterogeneity among the effect sizes (Q = 21.5, p = 0.003). Therefore, the summary estimates in any threshold analysis of these data were not calculated. Rather, we sought to "explain" the source(s) of heterogeneity using meta-regression.

Meta-regression. To "explain" this heterogeneity, we performed 21 meta-regressions. All of the study and patient characteristics used in the meta-regressions are presented in Evidence Table 179. The results of the meta-regression are presented in Evidence Table 180 and Figure 72. Of the meta-regressions, the only one-predictor model to explain the heterogeneity was the year the study enrollment ended. The two studies with the earliest enrollment dates (Spencer, Spencer, Williamson, et al.[267] and Gates, Rosenfeld, Maxwell, et al.[265]) reported the highest percentages of patients who became free of generalized tonic-clonic seizures. The reasons for this are unclear.

The intercept of this model represents the percentage of patients who would become seizure-free in a study with an average end date of enrollment and corresponds to 40 percent (CI: 29 percent to 50 percent) of patients becoming free of generalized tonic-clonic seizures after surgery. We next performed a threshold analysis on the intercept of the preceding regression model. This analysis (Figure 73) found that the effect of corpus callosotomy on seizure-freedom became statistically nonsignificant when 30 percent of patients in the synthetic control group became free of generalized tonic-clonic seizures.

Meta-analytic threshold analysis of atonic seizures

We performed a threshold analysis of the six studies reporting patients who were free of atonic seizures or "drop attacks" after surgery. Two hundred and twenty-six patients were examined in these studies. Evidence Table 181 presents the actual patient counts, percentages, and calculated effect sizes for each study used in this analysis. The individual study effect sizes (Cohen's h) presented in the Evidence Table were based on a control group in which no patients became free of atonic seizures. Figure 74 presents a forest plot of these effect sizes to show the extent of variation between studies, but no scale is provided because these effect sizes were not calculated using actual control groups.

The results of our threshold analysis appear in Figure 75. The summary estimate calculated at the 0 percent point (no patients in a synthetic control group were free of atonic seizures) was 1.81 (CI: 1.58 to 2.04, p <0.000001) and corresponded to 62 percent (CI: 50 percent to 72 percent) of patients becoming free of atonic seizures after surgery. The summary estimate

became nonsignificant (no statistically significant difference between surgery and control) when the proportion of patients in the synthetic control group reached 55 percent. There was no statistically significant heterogeneity in the threshold analysis (Q = 8.6, p = 0.13).

Employment Outcome Measurements

Evidence base

Only one study reporting employment data in corpus callosotomy patients, Sakas and Phillips[268] met our inclusion criteria. This study enrolled 20 patients.

Synthesis of study results

Evidence Table 182 presents a summary of the findings reported in this study. The study examined patients whose age at surgery was 15 to 37 years. At the time of surgery, none of the patients was working or in training for employment. An average of 6.7 years after surgery, 16 of 20 (80 percent) patients were employed either full-time or in training. In this study, a 50 percent or better reduction in the frequency of drop attacks and generalized tonic-clonic seizures was also seen in sixteen of the 20 patients (80 percent).

Cognitive Function Outcome Measurements – IQ

Evidence base

Only one study reporting changes in IQ in corpus callosotomy patients, Cohen, Holmes, Campbell et al.,[269] met our inclusion criteria. This study enrolled 10 patients.

Synthesis of study results

Evidence Table 183 presents a summary of the findings reported in this study. The study was restricted to patients less than 18 years of age. A change of greater than or equal to 8 points in the test instruments used in the study (Wechsler Intelligence Scale for Children-Revised, Stanford Binet Intelligence Scale and the Vineland Adaptive Behavior Scale) was considered a significant change in IQ. Due to the small sample size and the use of several intelligence tests, the authors did not perform a statistical analysis. Out of 10 patients, two showed a significant increase in IQ and one showed a significant decrease in IQ. Mean IQ did not change. In this study, a 50 percent or better reduction in the overall seizure frequency was seen in 6 patients between 1 and 2 years after surgery The authors concluded that the majority of patients did not appear to suffer any loss of cognitive ability due to surgery.

Complications Due to Surgery

The following section evaluates studies that reported cases of serious permanent complications and mild or transient complications resulting from surgical resection of the corpus callosum.

Evidence base

Among the 26 studies of corpus callosotomy meeting our inclusion criteria, 20 studies reported on complications due to surgery. The 20 studies examined 661 patients (Table 32). We abstracted data on serious permanent complications only if the publication specifically reported such a complication or specifically reported that no such complications occurred. We

considered disconnection syndrome, infraction, hemiparesis, and clinically significant language impairment to be serious permanent complications. We abstracted data on mild or transient complications only from studies reporting data on serious permanent complications. One of the 20 studies reported serious permanent complications but did not report on the occurrence of mild or transient complications.

Design and conduct of included studies

Internal validity

The complications reported by these studies could only have occurred because of surgery, so the internal validity with regard to the cause and effect is not in question. However, some potential biases may still be present. Investigator reporting bias may have affected the reporting of mild or transient complications because they may not be regarded as important by some investigators. Attrition bias is not a concern because all patients were examined after surgery. Maturation bias is also not a concern when reporting complications.

External validity

The specific patient characteristics of corpus callosotomy patients reported in each study are presented in Evidence Table 184. Three studies enrolled patients who had a mean age at surgery of less than 15 years with no patient exceeding 20 years of age. Seventeen studies had a mean age at surgery of greater than 15 years with youngest and oldest ages that varied between 1 year and 60 years. All 12 studies reporting age at seizure onset has a mean of less than 10 years for this patient characteristic. The range of seizure onset varied from birth to 33 years of age. In one study, the range of seizure onset was less than 1 year for all patients. Two studies reported a mean duration of epilepsy of less than 10 years and 10 studies reported a mean duration of epilepsy of between 12 and 21 years. The range for duration of epilepsy varied from less than 1 year to 50 years. Of the 20 studies, none began enrolling patients prior to 1972 (Table 32). Based on the distribution of patient characteristics, this evidence base seems to be generalizable to corpus callosotomy patients in clinical practice.

Synthesis of study results

Evidence Table 185 presents a study-by-study list of the complications reported in each of the 20 studies in the evidence base. Among the 661 corpus callosotomy patients, 24 serious permanent complications were reported. This corresponds to 3.6 percent of the patients or 36 serious complications per 1,000 surgery patients. One hundred and twenty-seven mild or transient complications were reported among 597 patients, which corresponds to 22 percent or 220 complications per 1,000 surgery patients. The number of mild or transient complications may be underestimated by these data because of differences in reporting these complications across studies. Clinician judgment as to the importance of reporting various mild or transient complications will likely vary across studies, whereas, the occurrence of permanent paralysis will usually warrant reporting.

Surgery-related Mortality

Any surgical procedure may result in such serious complications that death results. The following section evaluates studies that reported deaths due to corpus callosotomy.

Evidence base

Among the 26 studies of corpus callosotomy meeting our inclusion criteria, 18 studies reported a death due to surgery or specifically reported that no deaths occurred due to surgery. The 18 studies examined 643 patients (Table 33). Only one of these studies was not included in the evidence base for complications due to corpus callosotomy.

We abstracted only deaths specifically reported to be caused by surgery. Deaths as a result of invasive presurgical diagnostic procedures were not included.

Design and conduct of included studies

Internal validity

The deaths reported here could only have occurred through surgery. Investigator reporting bias, attrition bias, and maturation bias are not a concern when reporting surgery-related mortality.

External validity

The specific characteristics of corpus callosotomy patients reported in each study are presented in Evidence Table 184. Since the evidence base for death due to surgery is almost identical to the evidence base for complications, the ages at surgery, ages at seizure onset, and duration of epilepsy prior to surgery described earlier depict these patient characteristics for the studies reporting surgery-related mortality. Of the 18 studies, none began enrolling patients prior to 1972 (Table 33). Based on the distribution of patient characteristics, this evidence base seems to be generalizable to corpus callosotomy patients in clinical practice.

Synthesis of study results

Among the 643 corpus callosotomy patients, six deaths were reported (0.93 percent or 9.3 deaths per 1,000 patients). All of the deaths were reported in four studies (Table 35). The deaths were not found in the largest studies, and were not associated with the years in which the studies were conducted. That deaths were not reported in larger studies suggests that this death rate is uncertain.

Table 29. Corpus callosotomy and seizure frequency outcomes

Method of reporting effect of surgery on seizure frequency

Reference	Types of Seizures Evaluated	Percent Reduction in Seizure Frequency	Actual Change in Mean Seizure Frequency	Seizure-free
Kwan (2001)[270]	All seizure types	✓		
Maehara (2001)[271]	Disabling generalized seizures	✓		
Matsuzaka (1999)[272]	Most disabling seizures	✓		✓
McInerney (1999)[273]	Most disabling seizures	✓		✓
Sakas (1996)[268]	Drop attacks and generalized tonic-clonic seizures	✓		✓
Claverie (1995)[274]	All seizure types	✓		
Reutens (1993)[275]	All seizure types	✓		
Marino (1990)[276]	All seizure types	✓		✓
Murro (1988)[277]	All seizure types	✓	✓	✓
Purves (1988)[278]	All seizure types	✓		
Spencer (1988)[267]	All seizure types	✓	✓	✓
Gates (1987)[265]	All seizure types	✓	✓	✓

Table 30. Corpus callosotomy: individual patient data

Studies of corpus callosotomy reporting individual patient data for patients with successful and nonsuccessful surgery

Reference	Age at Treatment	Age at Seizure Onset	Duration of Epilepsy Prior to Treatment
Sakas (1996)[268]	✓	✓	
Claverie (1995)[274]	✓		
Nordgren (1991)[279]	✓	✓	✓
Marino (1990)[276]	✓	✓	✓
Murro (1988)[277]	✓	✓	✓
Purves (1988)[278]	✓	✓	✓
Spencer (1988)[267]	✓	✓	✓

Table 31. Corpus callosotomy and specific seizure types

Studies reported patients who were free of specific seizure types after surgery

Reference	Specified Most Disabling Seizure	Patients with Generalized Tonic-Clonic Seizures	Patients with Atonic Seizures
Kwan (2001)[270]		✓	✓
Maehara (2001)[271]		✓	✓
Matsuzaka (1999)[272]	✓		
McInerney (1999)[273]	✓	✓	✓
Sakas (1996)[268]	✓	✓	✓
Marino (1990)[276]	✓	✓	✓
Murro (1988)[277]	✓	✓	
Spencer (1988)[267]	✓	✓	
Gates (1987)[265]	✓	✓	✓

Table 32. Corpus callosotomy and complications due to surgery

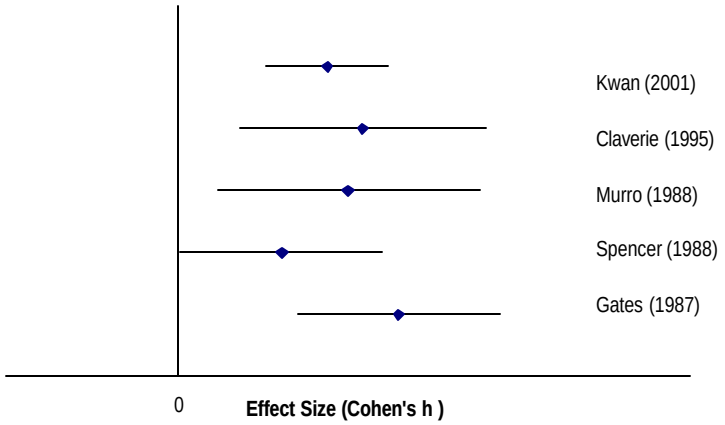
Reference	Number of Patients	Country	Years Study Conducted	Number of Permanent Complications
Hodaie (2001)[280]	17	Canada	1992-1999	0
Maehara (2001)[271]	52	Japan	1991-1998	0
Fandino-Franky (2000)[281]	97	Colombia	1989-1997	0
Pinard (1999)[282]	17	France	1989-1995	2
Carmant (1998)[283]	28	United States	1989-1993	0
Sorenson (1997)[284]	23	United States	1991-1994	1
Andersen (1996)[285]	20	Denmark	1988-1994	7
Rossi (1996)[286]	20	Italy	1988-1995	1
Sakas (1996)[268]	20	Ireland	1984-1993	0
Reutens (1993)[275]	64	Australia	1973-1991	4
Fuiks (1991)[287]	80	United States	1985-1990	0
Nordgren (1991)[279]	18	United States	1972-1987	1
Oguni (1991)[288]	43	Canada	1981-1989	0
Marino (1990)[276]	28	Brazil	1978-1985	1
Provinciali (1990)[289]	15	Italy	1987-1988	0
Sass (1990)[290]	32	United States	1985-1987	4
Murro (1988)[277]	25	United States	1980-1986	0
Purves (1988)[278]	24	Canada	1977-1987	0
Garcia-Flores (1987)[291]	14	Mexico	1980-1986	0
Gates (1987)[265]	24	United States	1979-1985	3

Table 33. Corpus callosotomy and surgery-related mortality

Reference	Number of Patients	Country	Years Study Conducted	Number of Deaths
Hodaie (2001)[280]	17	Canada	1992-1999	0
Maehara (2001)[271]	52	Japan	1991-1998	0
Fandino-Franky (2000)[281]	97	Colombia	1989-1997	0
McInerney (1999)[273]	47	United States	1972-1999	0
Carmant (1998)[283]	28	United States	1989-1993	0
Sorenson (1997)[284]	23	United States	1991-1994	0
Andersen (1996)[285]	20	Denmark	1988-1994	1
Sakas (1996)[268]	20	Ireland	1984-1993	0
Reutens (1993)[275]	64	Australia	1973-1991	0
Fuiks (1991)[287]	80	United States	1985-1990	2
Nordgren (1991)[279]	18	United States	1972-1987	1
Oguni (1991)[288]	43	Canada	1981-1989	0
Provinciali (1990)[289]	15	Italy	1987-1988	0
Sass (1990)[290]	32	United States	1985-1987	0
Murro (1988)[277]	25	United States	1980-1986	2
Purves (1988)[278]	24	Canada	1977-1987	0
Garcia-Flores (1987)[291]	14	Mexico	1980-1986	0
Gates (1987)[265]	24	United States	1979-1985	0

Figure 63. Forest plot: corpus callosotomy and reduction in seizure frequency

Studies reported patients with at least a 90 percent reduction in seizure frequency after surgery



A scale is not shown because the effect sizes were not calculated with actual control groups

Figure 64. Threshold analysis: corpus callosotomy and reduction in seizure frequency

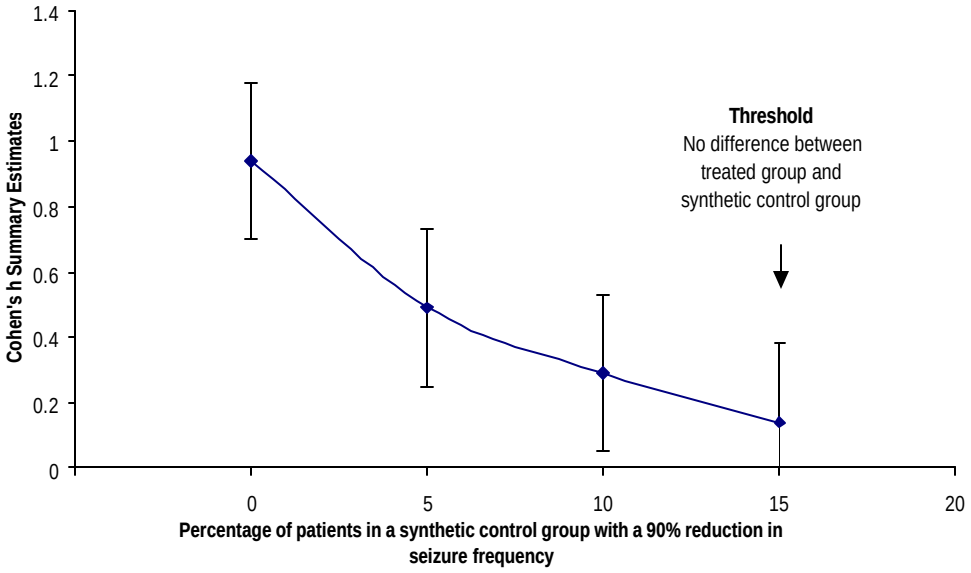


Figure 65. Forest plot: corpus callosotomy and no benefit from surgery

Studies reported patients who had no change or an increase in seizure frequency

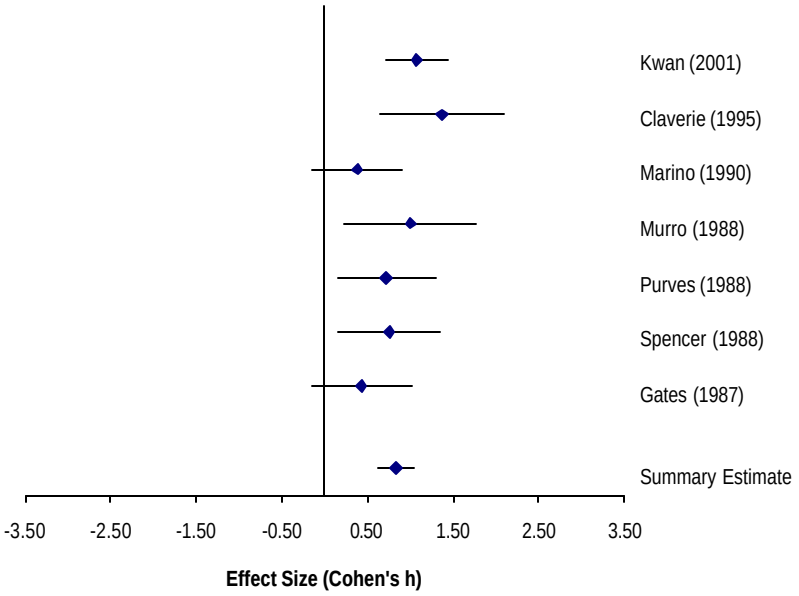


Figure 66. Forest plot: corpus callosotomy and patient age at surgery

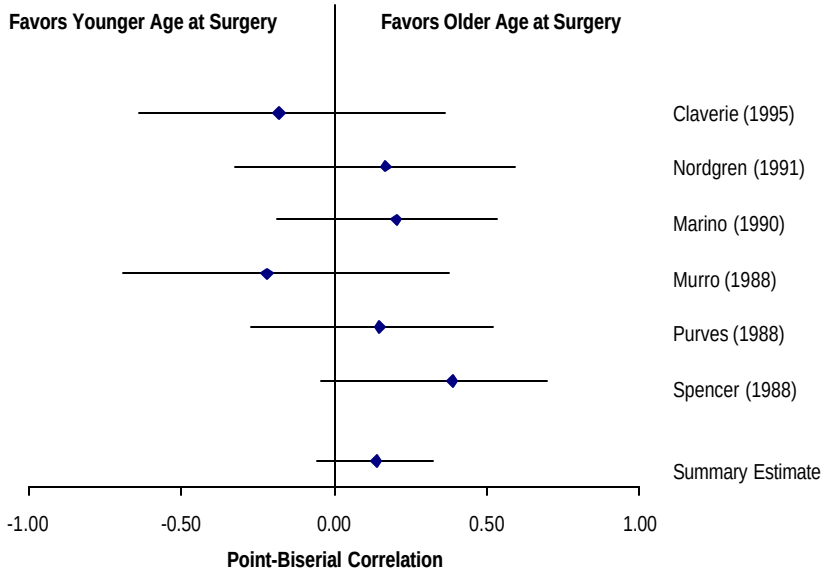


Figure 67. Forest plot: corpus callosotomy and patient age at onset of seizures

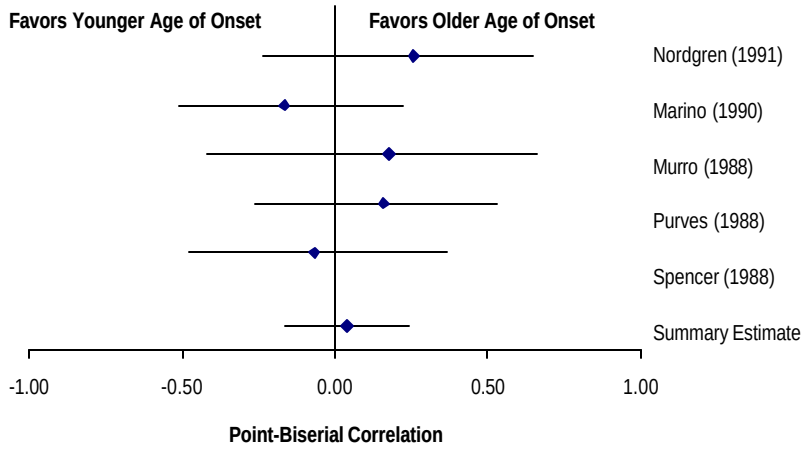


Figure 68. Forest plot: corpus callosotomy and duration of epilepsy prior to surgery

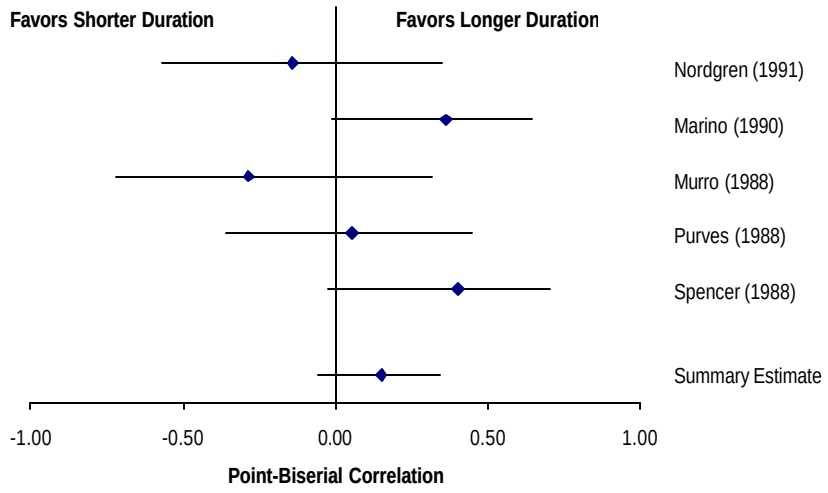
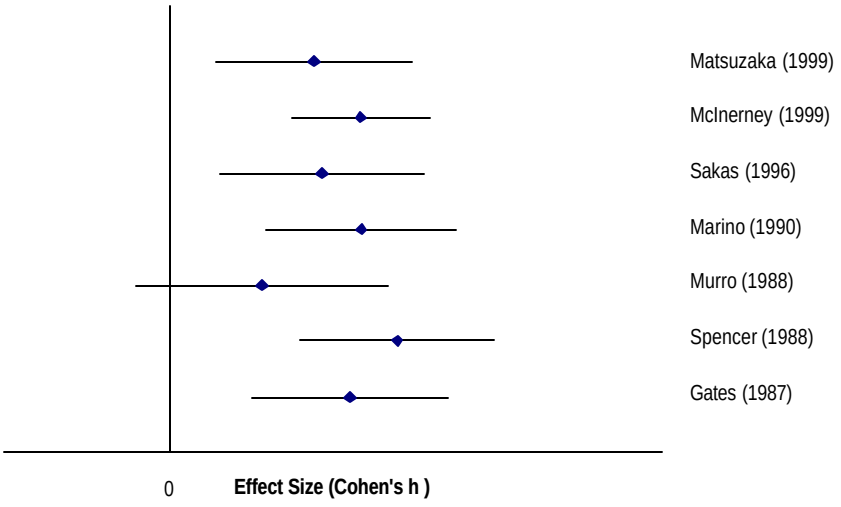


Figure 69. Forest plot: corpus callosotomy and most disabling seizures

Studies reported patients who were free of their most disabling seizures



A scale is not shown because the effect sizes were not calculated with actual control groups

Figure 70. Threshold analysis: corpus callosotomy and most disabling seizures

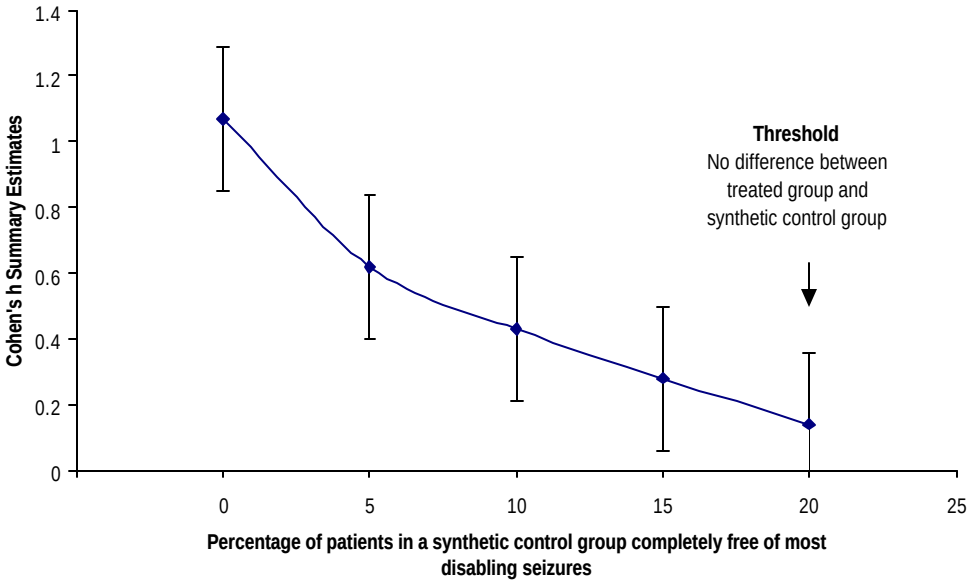
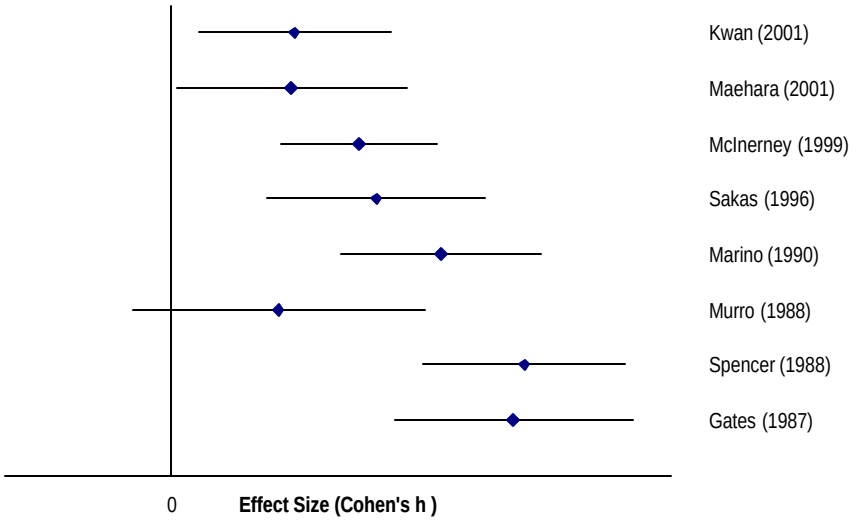


Figure 71. Forest plot: corpus callosotomy and generalized tonic-clonic seizures

Studies reported patients who were free of generalized tonic-clonic seizures



A scale is not shown because the effect sizes were not calculated with actual control groups

Figure 72. Meta-regression: corpus callosotomy and generalized tonic-clonic seizures

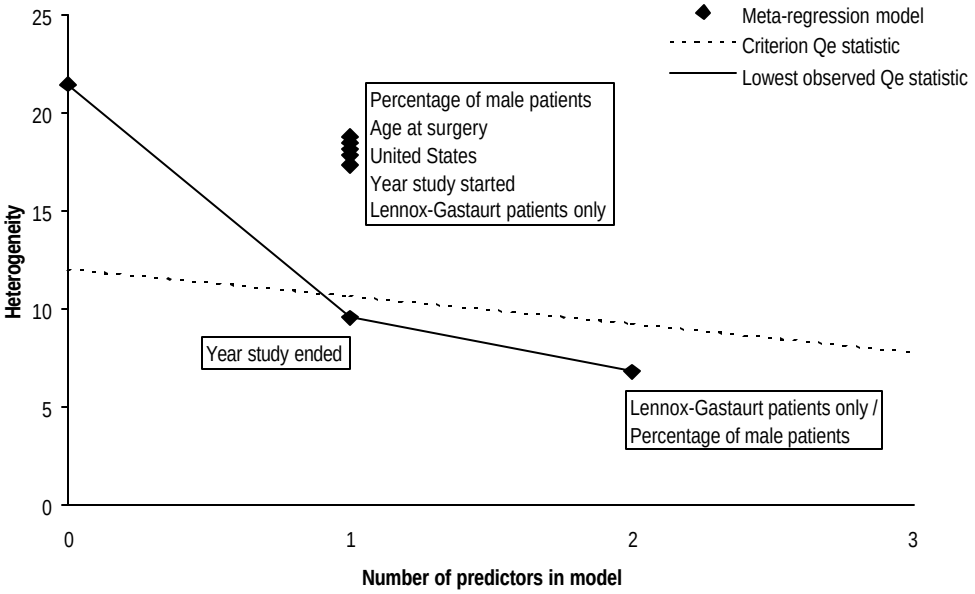


Figure 73. Threshold analysis: corpus callosotomy and generalized tonic-clonic seizures

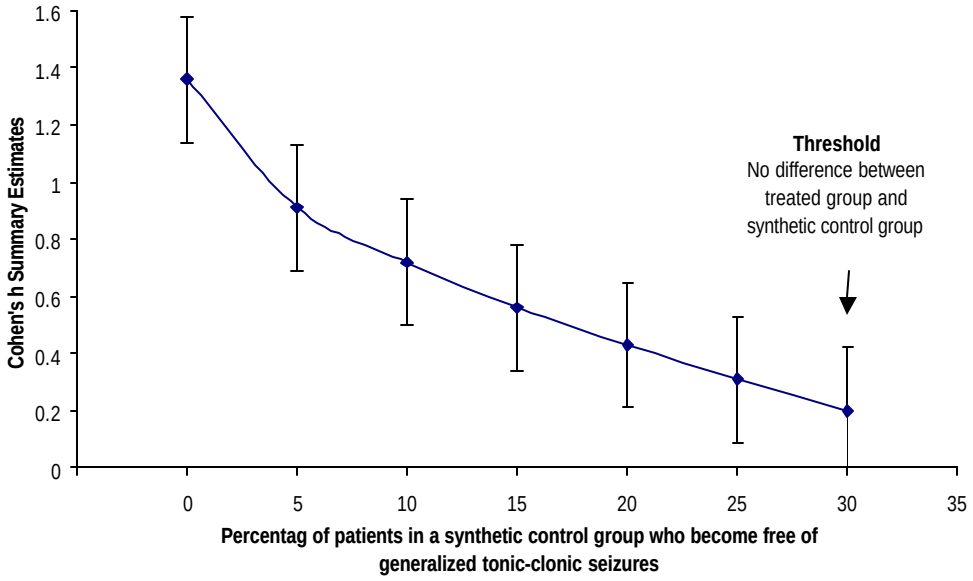
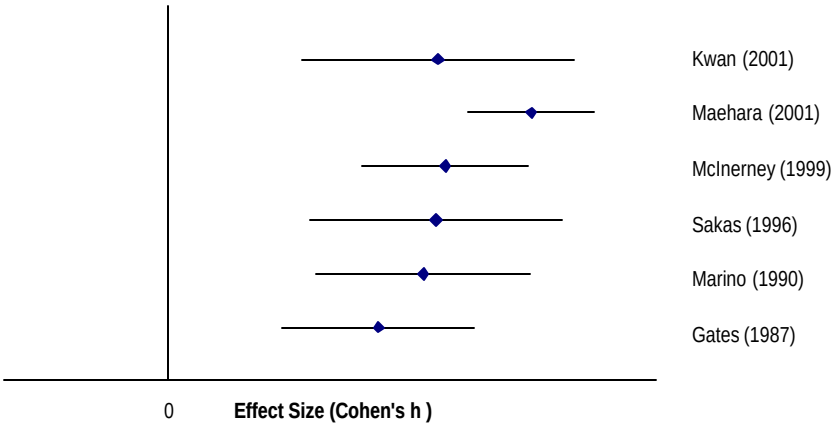


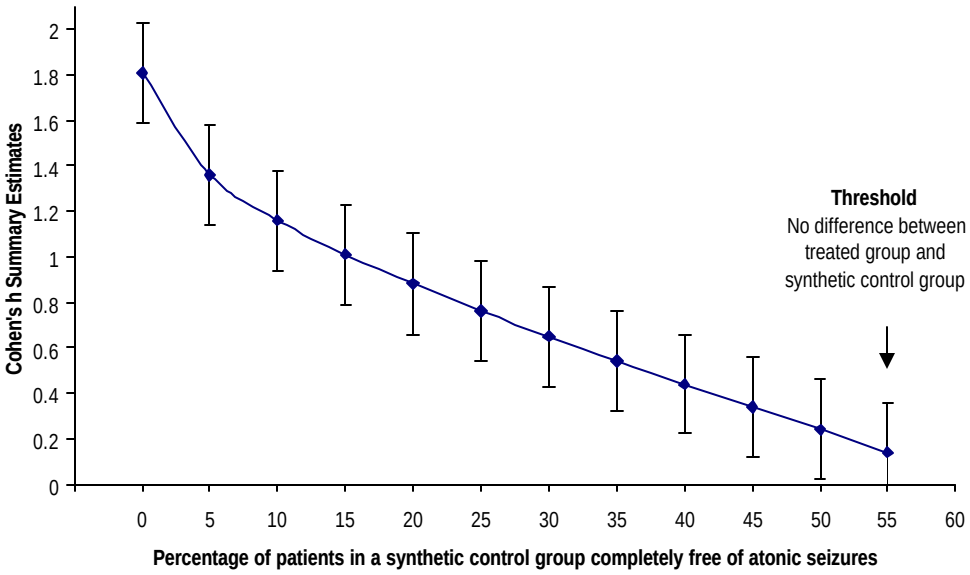
Figure 74. Forest plot: corpus callosotomy and atonic seizures

Studies reported patients who were free of atonic seizures



A scale is not shown because the effect sizes were not calculated with actual control groups

Figure 75. Threshold analysis: corpus callosotomy and atonic seizures



Frontal Lobe Surgery

Partial motor seizures on one side of the body are caused by lesions in the frontal lobe opposite to the side of the seizures.[1] The most common type of seizure with a frontal lobe origin begins with a turning movement of the head and eyes to the side opposite the lesion. The seizure often includes tonic contractions of the trunk and limbs. A generalized clonic seizure may then follow. A lesion in the frontal lobe may also result in generalized convulsive seizure without the initial turning of the head and eyes. Surgery is directed at resection of the lesion.

As with temporal lobe surgery, our basis for judging the success of frontal lobe surgery is the number of patients who are seizure-free after surgery. Surgical complications and deaths due to surgery will also be considered in determining the efficacy of this surgical procedure. No studies meeting our inclusion criteria reported data on the other outcome measurements listed in our inclusion criteria.

Seizure-free

Evidence base

Among the 18 studies of frontal lobe surgery meeting out inclusion criteria, 13 reported some sort of seizure-free outcome measurement. Studies classified patients who achieved freedom from seizures using one of four different definitions of "seizure-free" as previously discussed for temporal lobe surgery. Thus, some studies classified patients as seizure-free if they were completely seizure-free including auras, others classified patients as seizure-free regardless of whether patients experienced auras, other studies classified patients as seizure-free but did not state whether auras were considered in their determination, and still others classified patients as seizure-free if they were in Engel class I.

Table 34 presents a listing of the seizure-free categories used by each of the 13 studies reporting seizure-free outcome measurements after frontal lobe surgery. Two studies with a total of 33 patients reported seizure-free with no auras, two studies with a total of 37 patients reported seizure-free with auras, six studies with a total of 415 patients reported seizure-free but the presence of auras was undefined, and three studies with a total of 54 patients reported Engel class I. If a study reported separate outcome and patient information according to a specific age group, type of surgery, or pathology, these data are presented separately in the evidence tables. None of the studies reported more than one of the four seizure-free categories.

Design and conduct of included studies

Internal validity

None of the studies in the evidence base included data from a control group and all employed a case series design. Therefore, all of the 13 studies in the evidence base may have biases that reduce internal validity as previously discussed for temporal lobe surgery. However, these patients may have uncontrolled seizures for more than 4 years while using appropriate AEDs.[292] Therefore, given the occurrence of treatment-resistant seizure activity in these individuals, explanations for seizure reduction other than the effect of surgery may be considered implausible.

External validity

As described earlier, patients being considered for resection of the frontal lobe experience characteristic seizures due to the location of the lesion responsible for the seizures. Therefore, the patients in published studies of frontal lobe surgery should be representative of all patients receiving this surgery. However, differences may exist across studies with regard to age or pathology.

The specific patient characteristics of frontal lobe patients reported in each study are presented in Evidence Table 186. The mean age at the time of surgery varied between 10 to 30 years of age with patient ages ranging from a youngest of less than a year old to an oldest of 74 years of age. The mean age of seizure onset was between 1.5 to 28 years of age with a range of birth to 49 years. The mean duration of epilepsy prior to surgery was between 4 and 19 years. The range for duration of epilepsy prior to surgery was less than a year to 40 years.

Based on the distribution of patient characteristics, this evidence base seems to be generalizable to frontal lobe surgery patients in clinical practice.

Synthesis of study results

Meta-analytic threshold analysis of seizure-free undefined

Of the four "seizure-free" outcome measures, only seizure-free undefined was reported by five or more studies. Evidence Table 187 presents the actual patient counts, percentages, and calculated effect sizes (Cohen's h) for each study in this analysis. We calculated each effect size presented in the Evidence Tables under the assumption that no patients in a synthetic control group attained the seizure-free outcome measure being analyzed. Figure 76 presents a forest plot of these effect sizes to show the extent of variation between studies, but no scale is provided because these effect sizes were not calculated using actual control group.

Our threshold analysis of six studies reporting seizure-free undefined found statistically significant heterogeneity among the effect sizes ($Q = 43.2$, $p < 0.000001$). Therefore, we did not calculate the summary estimates required for a threshold analysis. Rather, we sought to "explain" the source(s) of heterogeneity using meta-regression.

Despite the heterogeneity, all six studies show significant effect sizes suggesting that frontal lobe surgery is effective in eliminating seizures. However, the percentage of patients considered seizure-free varied from 24 percent to 100 percent.

Meta-regression. To "explain" this heterogeneity, we performed 10 meta-regressions. Of these, no one-predictor model explained the heterogeneity, and a pair of two-predictor models did. Because there were only six studies, and we suspected over fitting of the models, we did not pursue a threshold analyses of the intercepts. All of the study and patient characteristics used in our meta-regression are presented in Evidence Table 188. The results of the meta-regression are presented in Evidence Table 189 and Figure 77.

Analysis of seizure-free with no auras, seizure-free with auras, and Engel class I

Evidence Table 190 presents the findings for the studies using seizure-free with no auras, seizure-free with auras, and Engel Class I to judge surgical success.[x] With fewer than five studies reporting each of these outcome measurements, we did not conduct a meta-analysis. All of the studies show that some proportion of patients do become seizure-free after surgery. However, the estimates varied greatly among studies. The three studies using Engel Class I reported success

[x] See the temporal lobe surgery section for a description of these seizure-free classificiations.

rates between 55 percent and 58 percent. These findings are similar to the two studies using seizure-free with no auras and reported success rates of 57 percent. However, studies using seizure-free with no auras might be expected to have a reduced success rate compared to studies using Engel Class I because seizure-free with no auras is a subgroup within Engel Class I. Studies using seizure-free with auras should be reporting higher percentages of patients than studies using seizure-free with no auras. However, the findings of the two studies using seizure-free with auras were considerably lower than the two studies using seizure-free with no auras. One study reported a 17 percent success rate, and the other study reported a 25 percent success rate for adult patients and a 31 percent success rate for children. An explanation for the large discrepancy between studies reporting seizure-free with no auras and seizure-free with auras is not apparent from the patient data reported in these studies.

Analysis of nested case-control studies

Two nested case-control studies of frontal lobe surgery presented an evaluation of patient characteristics that could potentially influence surgical outcomes. Evidence Table 191 presents the findings reported by each of the nested case-control studies in our evidence base for seizure frequency outcome measurements. Both studies used univariate methods. Only the study by Smith, Lee, King, et al.[293] looked at age at surgery, age at seizure onset, and duration of epilepsy prior to surgery. This study found no association between these patient characteristics and successful surgery.

Complications Due to Surgery

The following section evaluates studies that reported cases of serious permanent complications and mild transient complications resulting from frontal lobe surgery.

Evidence base

Among the 18 studies of frontal lobe surgery meeting our inclusion criteria, eight studies reported on complications due to surgery. The eight studies examined 369 patients (Table 35). We abstracted data on serious permanent complications only if the publication specifically reported such a complication or specifically reported that no such complications occurred. We considered disabling or spastic hemiparesis and worsening of the preoperative neurologic deficit to be serious permanent complications of surgery. We abstracted data on mild or transient complications only from studies reporting data on serious permanent complications.

Design and conduct of included studies

Internal validity

The complications reported by these studies could only have occurred because of surgery, so the internal validity with regard to the cause and effect is not in question. However, some potential biases may still be present. Investigator reporting bias may have affected the reporting of mild or transient complications because they may not be regarded as important by some investigators. Attrition bias is not a concern because all patients were examined after surgery. Maturation bias is also not a concern when reporting complications.

External validity

The specific patient characteristics of frontal lobe surgery patients reported in each study are presented in Evidence Table 192. The mean age at surgery was between 5 and 28 years of age

with a range that varied from a youngest patient of less than a year to an oldest patient of 49 years. Mean age at seizure onset was between 10 and 30 years of age. The range of seizure onset varied from less than a year to 59 years of age. Mean duration of epilepsy was from 4 to 16 years and the range varied from less than 1 year to 41 years.

Based on the distribution of patient characteristics, this evidence base seems to be generalizable to frontal lobe surgery patients in clinical practice.

Synthesis of study results

Evidence Table 193 presents a study-by-study list of the complications reported in each of the eight studies in the evidence base. Among the 369 frontal lobe surgery patients, 31 serious permanent complications were reported, mostly some form of partial paralysis. This corresponds to 8.4 percent of the patients or 84 serious complications per 1,000 surgery patients. One hundred and twenty mild or transient complications were reported among 337 patients, which correspond to 35.6 percent or 356 complications per 1,000 surgery patients. Most of the transient complications were also some form of partial paralysis.

Twenty-seven of the 31 serious permanent complications (87.1 percent) and 72 of the 120 mild or transient complications (60 percent) were reported in one study.[294] One year after surgery 27 out of 120 patients had spastic hemiparesis or pronounced worsening of their preoperative deficit that was considered a serious complication and 42 patients had hemiparesis or hemiplegia that was considered a transient complication. This study had the largest sample size and collected patient data back to the earliest year, 1964. The early start date may partially explain the high number of complications in this study. Only one other study reported serious complications.[293] In this study, 4 out of 53 patients had serious complications and two of these patients had disabling hemiparesis.

Surgery-related Mortality

Any surgical procedure may result in such serious complications that death results. The following section evaluates studies that reported deaths due to frontal lobe surgery.

Evidence base

Among the 18 studies of frontal lobe surgery meeting our inclusion criteria, three studies reported a death due to surgery or specifically reported that no deaths occurred due to surgery. The three studies examined 96 patients.[200,293,295] All three of these studies were included in the evidence base for complications due to frontal lobe surgery.

We abstracted only deaths specifically reported to be caused by surgery. Deaths as a result of invasive presurgical diagnostic procedures were not included.

Design and conduct of included studies

The specific patient characteristics of frontal lobe surgery patients reported in each study are presented in Evidence Table 194.

Synthesis of study results

Among the 96 frontal lobe surgery patients in the three studies reporting mortality data, one death was reported (1.0 percent or 10 deaths per 1,000 patients). The death was reported in a study that had 53 patients. The other two studies had 13 and 32 patients. Table 37 lists the studies

reporting mortality data. Given the small number of studies reporting mortality data, this estimate may be inaccurate.

Table 34. Frontal lobe surgery: seizure-free outcome reporting

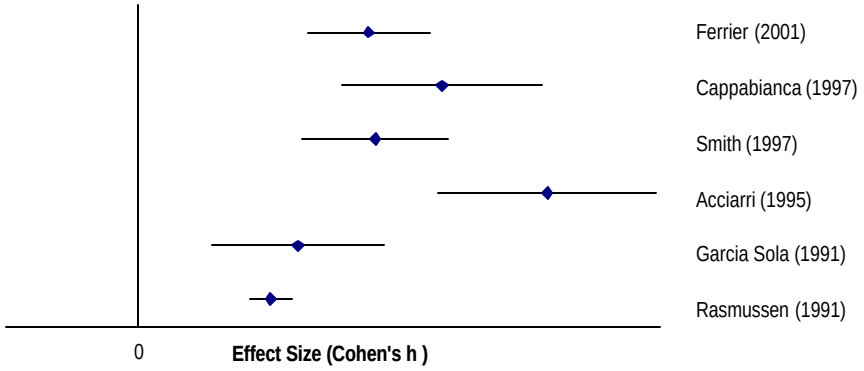
Reference	Seizure-Free Outcome Measurement			
	Undefined	No Auras	With Auras	Engel Class I
Ferrier (2001)[296]	✓			
Siegel (2001)[297]				✓
Hong (2000)[298]				✓
Eriksson (1999)[203]			✓	
Wennberg (1999)[299]				✓
Swartz (1998)[292]		✓		
Cappabianca (1997)[175]	✓			
Smith (1997)[293]	✓			
Acciarri (1995)[200]	✓			
Adler (1991)[300]		✓		
Garcia Sola (1991)[301]	✓			
Palmini (1991)[302]			✓	
Rasmussen (1991)[303]	✓			

Table 35. Frontal lobe surgery and complications due to surgery

Reference	Number of Patients	Country	Years Study Conducted	Number of Perrnanent Complications	Number of Surgery-related Deaths
Kral (2001)[295]	32	Germany	1989-2000	0	0
Mosewich (2000)[304]	68	United States	1987-1994	0	Not reported
Chassoux (1999)[294]	120	France	1964-1995	27	Not reported
Ferrier (1999)[305]	42	England	1975-1996	0	Not reported
Helmstaedter (1998)[306]	33	Germany	1995-1996	0	Not reported
Swartz (1998)[292]	19	United States	1986-1995	0	Not reported
Smith (1997)[293]	53	United States		4	1
Acciarri (1995)[200]	13	Italy	1975-1992	0	0

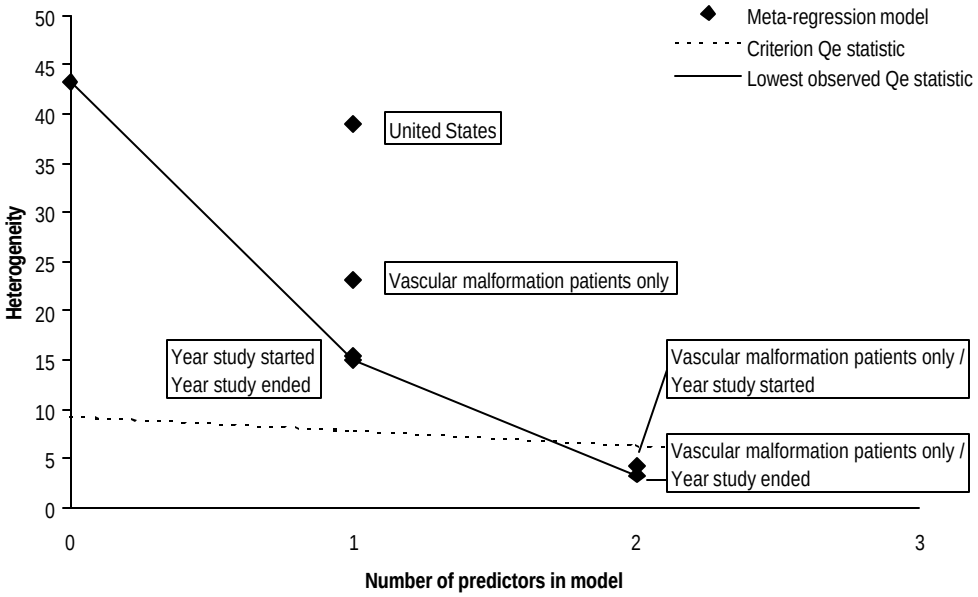
Figure 76. Forest plot: frontal lobe surgery and seizure-free (undefined)

Studies reported patients who were seizure-free undefined



A scale is not shown because the effect sizes were not calculated with actual control groups

Figure 77. Meta-regression: frontal lobe surgery and seizure-free (undefined)



Hemispherectomy

Hemispherectomy involves complete or partial removal of an entire cortical hemisphere of the brain including the motor and sensory cortex.[264] The intent of surgery is to eliminate seizures originating diffusely from a single cerebral hemisphere. The procedure is performed when smaller focal resections will not remove all of the epileptic region or when the progressive involvement of the remaining ipsilateral hemispheric cortex is inevitable.[307] Removal of the cortex of one hemisphere is used in patients with intractable unilateral, multifocal epilepsy associated with infantile hemiplegia or in some adults with severe cerebral disease and intractable unilateral motor seizures.[140,308] The etiological factors include injuries at birth, meningitis, acute and chronic encephalitis, head trauma, Rasmussen's syndrome, developmental dysplasia, and vascular problems.[307,309] The seizures experienced by these patients include partial motor seizures, unilateral tonic-clonic seizures, and drop attacks.[309]

Our basis for judging the success of hemispherectomy, which is governed by the available literature, is the number of patients who are seizure-free after surgery. Surgical complications and deaths due to surgery were also considered in determining the efficacy of this surgical procedure.

Seizure-free

Evidence base

Among the 11 studies of hemispherectomy meeting our inclusion criteria, three studies reported the number of patients that were seizure-free a minimum of 2 years after surgery. Table 36 presents a listing of these three studies, which enrolled 44 patients.

Design and conduct of included studies

Internal validity

None of the studies in the evidence base included data from a control group and all studies have case series design. Therefore, all three of the studies in the evidence base may have biases that reduce internal validity as previously discussed for temporal lobe surgery. However, these patients have a variety of etiologies for their seizure activity and they are not expected to improve without intervention.[307,309] Therefore, given the occurrence of treatment-resistant seizure activity in these individuals, explanations for seizure reduction other than the effect of surgery may be considered implausible.

External validity

As described earlier, patients being considered for hemispherectomy experience characteristic seizures due to damage or disease in a single hemisphere. The specific patient characteristics of hemispherectomy patients reported in each study are presented in Evidence Table 195. The mean age at the time of surgery varied between 2 to 14 years of age with patient ages ranging from a youngest of less than a year old to an oldest of 38 years of age. The mean age of seizure onset was about 5 years of age with a range of less than a year to 21 years. The mean duration of epilepsy prior to surgery was between 7 and 10 years. The range for duration of epilepsy prior to surgery was a year to 37 years.

Based on the distribution of patient characteristics, this evidence base seems to be generalizable to hemispherectomy patients in clinical practice.

Synthesis of study results

With only three studies in our evidence base reporting seizure-free outcome measures, we did not perform a meta-analysis. Evidence Table 196 presents the findings from these studies. From 40 percent to 70 percent of the patients in these studies were seizure-free a minimum of 2 years after surgery depending on the definition of seizure-free used in each study. Seizure-free with no auras was reported once, seizure-free undefined was reported twice, and Engel class I was reported twice among the three studies.

We calculated Cohen's h for each of the seizure-free outcomes reported in each of these studies to determine the magnitude of the effect and to determine if each result was significantly different from zero. The effect sizes were calculated using a synthetic control group in which none of the patients achieved the seizure-free outcome. These effect sizes are presented in Figure 78. The effect sizes were all statistically significantly different from zero.

Two studies also listed the number of patients classified in Engel class IV (no benefit). Both studies reported 1 out of 15 patients (6.7 percent) in this category.

Education Outcome Measurments

Evidence base

Only one study reporting data on education after hemispherectomy met our inclusion criteria. Lindsay, Ounsted, and Richards[310] examined 17 patients.

Design and conduct of included studies

This study had no control group. Therefore, all of the concerns with regard to internal validity for studies of case series design previously discussed with regard to seizure-free outcome measures apply to this study as well.

Synthesis of study results

Evidence Table 197 presents a summary of the findings reported in this study. Only one patient was in school or training prior to surgery, but eight patients were employed or attended school after a mean followup period of 14 years.

Cognitive Outcome Measurements - IQ

Evidence base

Only one study reporting data on changes in IQ after hemispherectomy met our inclusion criteria. Lindsay, Ounsted, and Richards[310] examined 17 patients.

Design and conduct of included studies

This study had no control group. Therefore, all of the concerns with regard to internal validity for studies of case series design previously discussed with regard to seizure-free outcome measures apply to this study as well.

Synthesis of study results

Evidence Table 198 presents a summary of the findings reported in this study. IQ increased by 10 points or more in six patients (35 percent) and decreased by 10 points or more in only two patients (12 percent). Mean IQ also increased by about seven points after surgery.

Complications Due to Surgery

Complications because of hemispherectomy are of particular concern.[140] The following section evaluates studies that reported cases of serious permanent complications and mild or transient complications resulting from hemispherectomy.

Evidence base

Among the 11 studies of hemispherectomy meeting our inclusion criteria, all 11 reported on complications due to surgery. These 11 studies examined 266 patients (Table 37). We abstracted data on serious permanent complications only if the publication specifically reported such a complication or specifically reported that no such complications occurred. We considered severe disability and coma to be serious permanent complications. We abstracted data on mild or transient complications only from studies reporting data on serious permanent complications.

Design and conduct of included studies

Internal validity

The complications reported by these studies could only have occurred because of surgery, so the internal validity with regard to the cause and effect is not in question. However, some potential biases may still be present. Investigator reporting bias may have affected the reporting of mild or transient complications because they may not be regarded as important by some investigators. Attrition bias is not a concern because all patients were examined after surgery. Maturation bias is also not a concern when reporting complications.

External validity

The specific patient characteristics of hemispherectomy patients reported in each study are presented in Evidence Table 199. The mean age at surgery was between 2 and 14 years of age. All studies had patients younger than 3 years of age while the oldest patient was 38 years of age. Mean age at seizure onset was between less than 1 year of age and 5 years of age. In one study, the age at onset for all patients was shortly after birth. The range of seizure onset varied from near birth to 38 years of age. Mean duration of epilepsy was between 4 and 11 years and the range varied between less than 1 year to 37 years.

Based on the distribution of patient characteristics, this evidence base seems to be generalizable to hemispherectomy patients in clinical practice.

Synthesis of study results

Evidence Tables 200 presents a study-by-study list of the complications reported in each of the 11 studies in the evidence base. Among the 266 patients, three serious permanent complications were reported, a severe disability due to bilateral brain swelling, hemosiderosis, and a coma. This corresponds to 1.1 percent of the patients or 11 serious complications per 1,000 surgery patients. Two of the complications were reported in studies that enrolled patients

prior to 1970.[307,310] Among the 193 patients in studies with enrollment after 1970, 0.5 percent of the patients experienced a serious complication (5 per 1,000 surgery patients).

Forty-one mild or transient complications were reported among 193 patients, which correspond to 21.4 percent or 214 complications per 1,000 surgery patients. Most of the transient complications were hydrocephalus requiring a shunt.

Surgery-related Mortality

Complications of hemispherectomy resulting in death are of particular concern.[140] The following section evaluates studies that reported deaths due to hemispherectomy.

Evidence base

All 11 studies of hemispherectomy reported whether there was a death due to surgery or specifically reported that no deaths occurred due to surgery. The 11 studies examined 266 patients (Table 39).

We abstracted only deaths specifically reported to be caused by surgery. Deaths as a result of invasive presurgical diagnostic procedures were not included.

Design and conduct of included studies

The specific patient characteristics of hemispherectomy patients reported in each study are presented in Evidence Table 199. Since the evidence base for death due to surgery contains the evidence base for complications, the ages at surgery , ages at seizure onset, and duration of epilepsy prior to surgery discussed previously applies to these patients as well.

Synthesis of study results

Among the 266 patients in the 11 studies reporting mortality data, seven deaths were reported (2.6 percent or 26 deaths per 1,000 patients). The deaths were reported in four of the 11 studies (Table 37). Four of the seven deaths were reported in studies that enrolled patients prior to 1970.[307,310] Among the 193 patients in studies with enrollment after 1970, 1.5 percent of the patients died as a result of surgery (15 per 1,000 surgery patients).

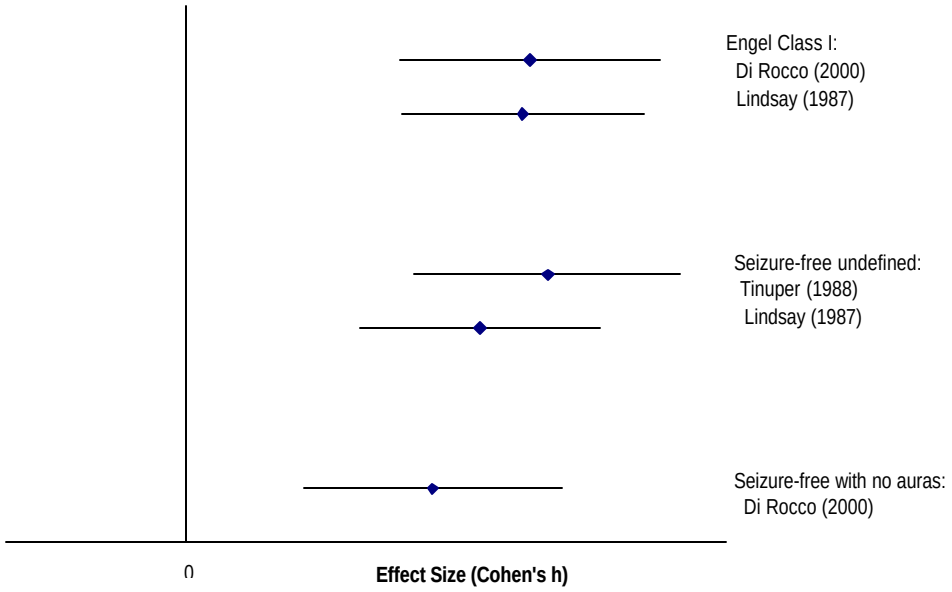
Table 36. Hemispherectomy: seizure-free outcome reporting

Reference	Number of Patients	Seizure-free with No Auras	Seizure-free Undefined	Engel Class I
Di Rocco (2000)[311]	15	✓		✓
Tinuper (1988)[309]	14		✓	
Lindsay (1987)[310]	15		✓	✓

Table 37. Hemispherectomy and complications and/or surgery-related mortality

Reference	Number of Patients	Country	Years Study Conducted	Number of Permanent Complications	Number of Surgery-related Deaths
Carreno (2001) [312]	13	United States	1992-1999	0	0
Schramm (2001)[313]	20	Germany	1991-1999	0	1
Di Rocco (2000)[311]	15	Italy	1985-1996	0	0
Shimizu (2000) [314]	34	Japan	1993-1999	1	0
Battaglia (1999)[315]	10	Italy	1987-1998	0	0
Wyllie (1998)[257]	16	United States	1990-1996	0	0
Vining (1997)[307]	58	United States	1968-1996	1	3
Peacock (1996)[316]	58	United States	1986-1995	0	2
Schramm (1995)[317]	13	Germany	1992-1994	0	0
Tinuper (1988)[309]	14	Canada	1974-1987	0	0
Lindsay (1987)[310]	17	England	1948-1986	1	1

Figure 78. Forest plot: hemispherectomy and seizure-free outcomes



A scale is not shown because the effect sizes were not calculated with actual control groups

Multiple Subpial Transection

Multiple subpial transection (MST) is intended for treatment-resistant patients whose epileptogenic lesion is located in cortical tissue controlling speech, movement, primary sensations, or memory.[139,318,319] The procedure is designed to horizontally sever interneuronal fibers longer than 5 mm while preserving neural elements and blood vessels that are vertically oriented. This procedure is relatively new compared to the other surgical procedures for epilepsy examined in this report. In 1989, Morrel, Whisler, and Bleck[318] published the first account of patients who received MST. Patients often have part of the temporal or frontal lobe resected in addition to the MST. The etiologies underlying the seizures include cortical dysplasia, Rasmussen's syndrome, gliosis, Landau-Kleffner syndrome, and tumors.[320-322]

The Expert Panel expressed a particular interest in the MST approach to the treatment of focal epilepsy. However, our inclusion criteria of a minimum 2-year followup period for all patients limited the size of our evidence base for this procedure. Therefore, we altered our inclusion criteria so that studies with a minimum followup period of 6 months could be included in our evidence base for assessing this procedure. In the absence of long-term followup data, patient improvement cannot be assumed permanent or long lasting. We included MST patients with or without other resections.

Our basis for judging the success of MST is the number of patients who are seizure-free after surgery and the number of patients showing a reduction in seizure frequency. Surgical complications and deaths due to surgery were also considered in determining the efficacy of this surgical procedure. No studies meeting our inclusion criteria reported data on the other outcome measurements listed in our inclusion criteria.

Seizure-free or Improvement After Surgery

Evidence base

Among the 10 studies of MST meeting out inclusion criteria, nine studies, totaling 212 patients, reported some sort of seizure-free or seizure improvement outcome measurement. Patients who achieved freedom from seizures were reported using one of three different outcome measures for "seizure-free" as previously discussed for temporal lobe surgery: seizure-free with auras, seizure-free undefined, and Engel class I.

Seizure-free with no auras (completely seizure-free) was not used by any of the studies in the MST evidence base. In addition to seizure-free outcome measurements, studies also reported the number of patients experiencing a 90 percent reduction in seizure frequency.

Table 38 presents a listing of the seizure outcome measurements used by each of the nine studies in the evidence base. Three studies reported seizure-free with auras, four studies reported seizure-free undefined, and four studies reported Engel class I. Ninety percent seizure frequency reduction was reported in four studies.

Design and conduct of included studies

Internal validity

None of the studies in the evidence base included data from a control group and all studies have retrospective case series design. Therefore, all nine studies in the evidence base may have biases that reduce internal validity as previously discussed for temporal lobe surgery. However,

these patients have a variety of etiologies for their seizure activity and they are not expected to improve without intervention.[320-322] Therefore, given the occurrence of treatment-resistant seizure activity in these individuals, explanations for seizure reduction other than the effect of surgery may be considered implausible.

External validity

Patients considered for MST experience characteristic seizures due to the location of the lesion responsible for the seizures. However, differences may exist across studies with regard to age or pathology.

The specific patient characteristics of MST patients reported in each study are presented in Evidence Table 201. The mean age at the time of surgery varied between 7 to 30 years of age with patient ages ranging from a youngest of less than a year old to an oldest of 54 years of age. The mean age of seizure onset was between 4 to 12 years of age with a range of birth to 39 years. The mean duration of epilepsy prior to surgery was between 3 and 17 years. The range for duration of epilepsy prior to surgery was less than a year to 42 years.

Based on the distribution of patient characteristics, this evidence base seems to be generalizable to MST patients in clinical practice.

Synthesis of study results

Analysis of seizure-free, seizure reduction, and Engel class I outcomes

Evidence Table 202 presents the findings for the studies reporting seizure-free with auras, seizure-free undefined, Engel class I, and a 90 percent reduction in seizure frequency. The studies vary widely in their estimate of the number of patients likely to become seizure-free after surgery. Among the three studies reporting seizure-free with auras, the percentage of patients in this category was between 37 percent and 57 percent. Within the four studies reporting seizure-free undefined, the percentage of patients achieving this category varied from 0 percent to 79 percent. The four studies using Engel Class I reported success rates between 20 percent and 57 percent. In the four studies reporting the number of patients with a 90 percent reduction in seizure frequency, the percentage of patients with this outcome measure was between 25 percent and 90 percent. Differences in how each outcome measure was recorded may account for the differences between studies. Patient age, pathology, the length of followup period, and the centers in which this new procedure was performed are also possible explanations for the variation in results.

We calculated Cohen's h for each of the seizure-free outcomes reported in each of these studies to determine the magnitude of the effect and to determine if each result was significantly different from zero. The effect sizes were calculated using a synthetic control group in which none of the patients achieved the seizure-free outcome. These effect sizes are presented in Figure 79.

Meta-analysis of patient characteristics

We performed separate meta-analyses that combined individual patient data across studies. Table 39 presents a list of the five studies of MST that provided individual patient data on age at surgery and successful surgery among male and female patients. All of these studies were included in the previous analysis examining seizure-free status after MST. Successful surgery was based on seizure-free undefined or a 90 percent reduction in seizure frequency.

For studies providing the age at surgery, we calculated a point-biserial correlation for each study and combined these data in a meta-analysis. The coefficient was calculated so that a positive correlation indicated that an older age favored a successful outcome and a negative correlation indicated that a younger age favored a successful outcome. For males versus females, we calculated each Cohen's h in the meta-analysis so that a positive effect size indicated that males had more successful surgery compared to females.

Although we are increasing the ability of our analysis to detect significant differences by increasing the sample size, our summary estimates are not adjusted for the influence of the other potentially important covariates in a study. An analysis using hierarchical modeling would be useful to search for factors that influence surgical outcomes by combining the patient-level data across studies, but such an analysis is beyond the scope of this report.

Age at surgery. Individual patient ages at surgery for patients with successful and nonsuccessful surgery were reported in five studies totaling 97 patients. Evidence Table 203 presents the definition used for successful surgery and the individual study point-biserial correlation calculated for each of the five studies. Figure 80 presents a forest plot of the effect sizes. The meta-analysis produced a summary estimate that was not statistically significant (r_{pb} = 0.14, CI: –0.07 to 0.34, p = 0.20) suggesting that age at surgery does not markedly influence surgical success. The effect sizes in this meta-analysis were not heterogeneous (Q = 3.3, p = 0.50).

We performed a sensitivity analysis to ensure that a single study did not have excessive influence over the results of the analysis. The summary estimate and other statistics did not change markedly because of the sensitivity analysis. Although, the point-biserial correlation varied from 0.07 to 0.24 as studies were removed during the sensitivity analysis, the summary estimates remained nonsignificant. The results of the sensitivity analysis as well as the original meta-analysis are presented in Evidence Table 204.

Gender. The percentage of male and female patients among patients with successful and nonsuccessful surgery was reported in five studies totaling 97 patients. Evidence Table 205 presents the individual numbers of male and female patients among the successful and nonsuccessful surgeries, the definition used for successful surgery, and the Cohen's h we calculated for each of these studies. Figure 81 presents a forest plot of the effect sizes.

The meta-analysis produced a summary estimate that was not statistically significant (0.24, CI: -0.19 to 0.66, p = 0.27), suggesting that gender has little or no influence on the success of surgery. The effect sizes in this meta-analysis were not heterogeneous (Q = 3.6, p = 0.46).

The summary estimate and other statistics did not change because of the sensitivity analysis. The difference between the percentage of male and female patients who achieved successful surgery varied between 0 percent and 3 percent as studies were removed during the sensitivity analysis. The sensitivity analysis and the original meta-analysis are presented in Evidence Table 206.

No Change or Increase in Seizure Frequency

Evidence base

Four studies, totaling 74 patients, reported the number of MST patients who had no change in seizure frequency or experienced an increase in seizure frequency (Table 38).

Design and conduct of the included studies

Internal validity

None of the studies in the evidence base included data from a control group and all studies have case series design. Therefore, all four studies in the evidence base may have biases that reduce internal validity as previously discussed for temporal lobe surgery.

External validity

The patient characteristics as described previously for the analysis of studies reporting seizure-free or improvement outcome measurements also apply to the studies reporting no benefit from surgery.

Synthesis of study results

A meta-analysis was not performed because less than five studies reported patients who had no change in seizure frequency or experienced an increase in seizure frequency. Evidence Table 207 presents the findings for these studies. The number of patients not benefiting from surgery varied greatly between studies. The percentage of patients with no change or an increase in seizure frequency ranged from 0 percent to 42 percent.

Complications Due to Surgery

The following section evaluates studies that reported cases of serious permanent complications resulting from MST.

Evidence base

Among the 10 studies of MST meeting our inclusion criteria, nine studies reported on complications due to surgery. The nine studies examined 236 patients (Table 40). We abstracted data on serious permanent complications only if the publication specifically reported such a complication or specifically reported that no such complications occurred. We considered hemiparesis and any form of aphasia as serious permanent complications. We abstracted data on mild or transient complications only from studies reporting data on serious permanent complications.

Design and conduct of included studies

Internal validity

The complications reported by these studies could only have occurred because of surgery, so the internal validity with regard to the cause and effect is not in question. However, some potential biases may still be present. Investigator reporting bias may have affected the reporting of mild or transient complications because they may not be regarded as important by some investigators. Attrition bias is not a concern because all patients were examined after surgery. Maturation bias is also not a concern when reporting complications.

External validity

The specific patient characteristics of MST patients reported in each study are presented in Evidence Table 201. The mean age at surgery was between 7 and 21 years of age with a range that varied from a youngest patient of less than a year to an oldest patient of 54 years. Mean age at seizure onset was between 4 and 13 years of age. The range of seizure onset varied from birth

to 39 years of age. Mean duration of epilepsy was between 3 and 17 years and the range varied from less than 1 year to 42 years.

Based on the distribution of patient characteristics, this evidence base seems to be generalizable to MST patients in clinical practice.

Synthesis of study results

Evidence Table 208 presents a study-by-study list of the complications reported in each of the nine studies in the evidence base. Among the 236 MST patients, 14 serious permanent complications were reported in four studies. Aphasia and dysphasia were reported in three of the studies. This corresponds to 5.9 percent of the patients or 59 serious complications per 1,000 surgery patients. Forty-five mild or transient complications were reported among 236 patients, which correspond to 19.1 percent or 191 complications per 1,000 surgery patients. Most of the transient complications were neurological deficits involving motor impairment. The mild and transient complications were reported in seven of the nine studies. Two studies reported no mild or transient complications.

Surgery-related Mortality

Any surgical procedure may result in such serious complications that death results. The following section evaluates studies that reported deaths due to MST.

Evidence base

Among the 10 studies of MST meeting our inclusion criteria, nine studies reported whether there was a death due to surgery or specifically reported that no deaths occurred due to surgery. These nine studies are the same studies reporting complications due to MST and examined 236 patients (Table 42).

Design and conduct of included studies

The specific patient characteristics of MST patients reported in each study are presented in Evidence Table 201.

Synthesis of study results

Among the 236 MST patients in the nine studies reporting mortality data, no deaths were reported (0 percent).

Table 38. Multiple subpial transection: seizure outcomes

Reference	Number of Patients	Seizure-Free With Auras	Seizure-Free Undefined	Engel Class I	90 Percent Reduction in Seizure Frequency	No Change or an Increase in Seizure Frequency
Mulligan (2001)[321]	12		✓		✓	✓
Orbach (2001)[320]	54	✓		✓	✓	
Shimizu (2000)[323]	25			✓		
Smith (1998)[324]	84			✓		
Hufnagel (1997)[325]	22	✓			✓	✓
Pacia (1997)[326]	21	✓		✓	✓	✓
Patil (1997)[327]	19		✓			✓
Morrell (1995)[328]	14		✓			
Sawhney (1995)[322]	21		✓			

Table 39. Multiple subpial transection: individual patient data

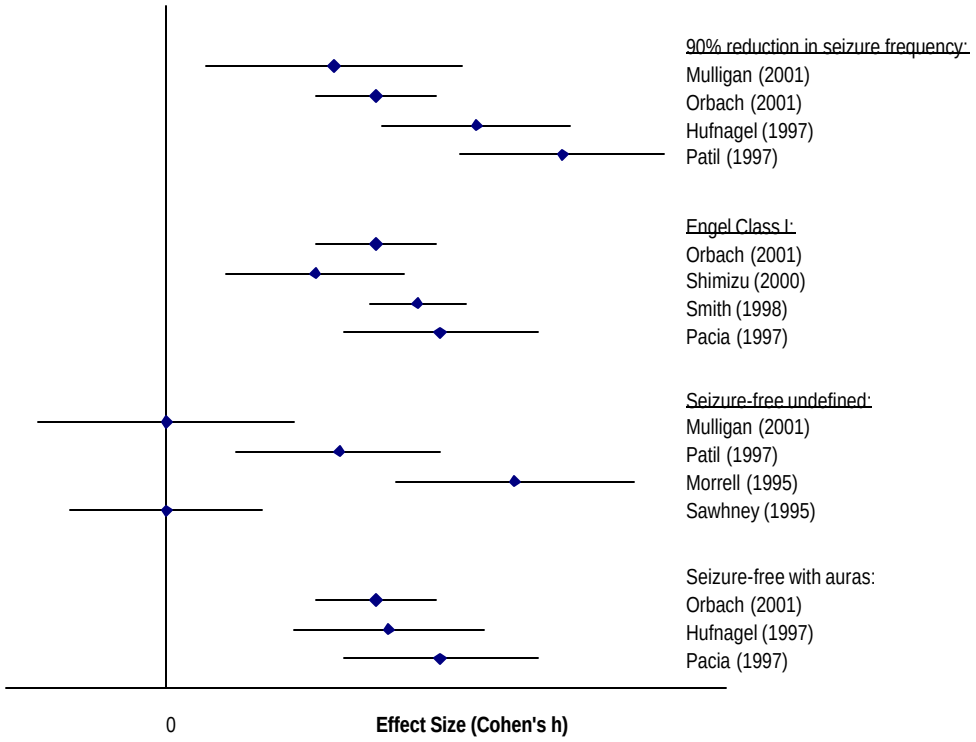
Studies of multiple subpial transection reporting individual patient data for patients with successful and nonsuccessful surgery

Reference	Number of Patients	Age at Surgery	Gender
Hufnagel (1997)[325]	22	✓	✓
Pacia (1997)[326]	21	✓	✓
Patil (1997)[327]	19	✓	✓
Morrell (1995)[328]	14	✓	✓
Sawhney (1995)[322]	21	✓	✓

Table 40. Multiple subpial transection and complications and/or surgery-related mortality

Reference	Number of Patients	Country	Years Conducted	Number of Permanent Complications	Number of Deaths
Mulligan (2001)[321]	12	United States	1990-1999	0	0
Shimizu (2000)[323]	31	Japan	1983-1998	0	0
Smith (1998)[324]	84	United States		7	0
Hufnagel (1997)[325]	22	Germany	1993-1996	4	0
Pacia (1997)[326]	21	United States	1992-1994	1	0
Patil (1997)[327]	19	United States	1991-1995	0	0
Morrell (1995)[328]	14	United States	1987-1994	0	0
Sawhney (1995)[322]	21	England	1989-1993	2	0
Shimizu (1991)[329]	12	Japan	1989-1990	0	0

Figure 79. Forest plot: multiple subpial transection and seizure-free outcomes



A scale is not shown because the effect sizes were not calculated with actual control groups

Figure 80. Forest plot: multiple subpial transection and patient age at surgery

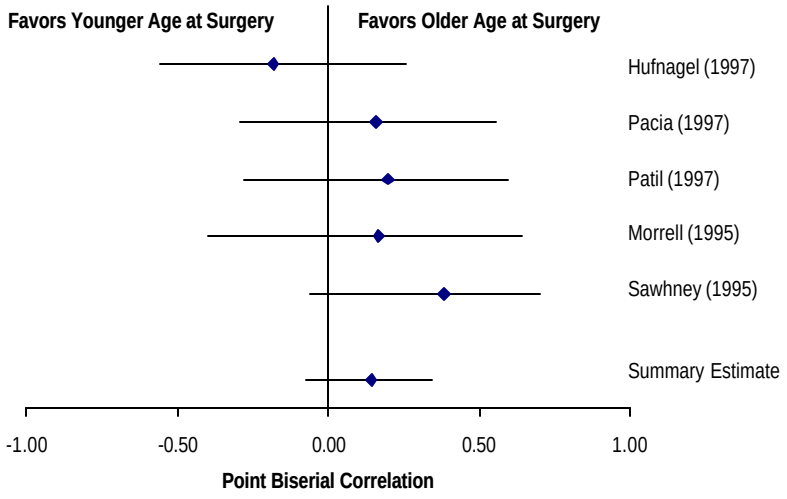
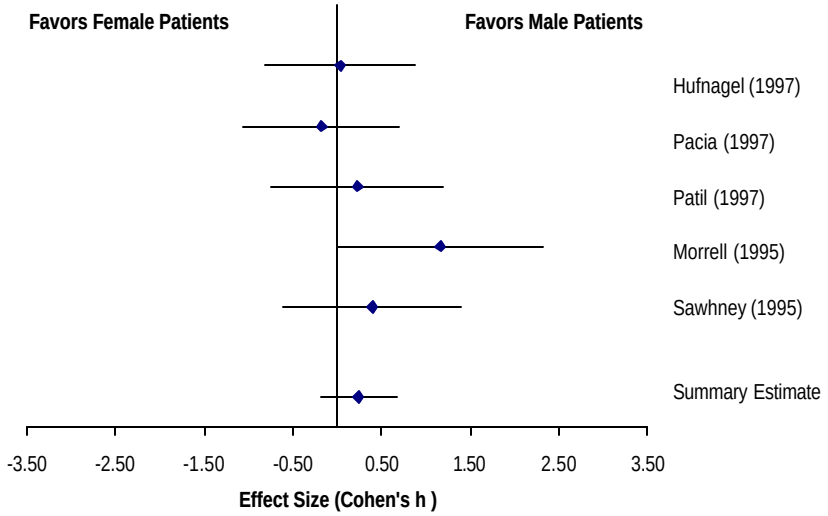


Figure 81. Forest plot: multiple subpial transection and male and female patients

Studies reported the success of surgery among male and female patients



Nondrug, Nonsurgical Treatments

In this subsection, we assess evidence pertaining to the effectiveness of several nondrug, nonsurgical treatments for patients with treatment-resistant epilepsy.

Question specific inclusion criteria

We included articles if they met both the general inclusion criteria presented in the Methodology section and if, as per discussions with the Expert Panel, Technical Experts, CDC, and SSA, the article reported on a study that evaluated the effectiveness and/or harms associated with one of the following nondrug, nonsurgical interventions:

- Vagal Nerve Stimulation (VNS)
- Ketogenic Diet
- Magnetic Therapy
- Vitamin B_6 Therapy
- Herbal Medicine
- Acupuncture
- Electrical Brain Stimulation
- Chiropractic Therapy
- Cranial Realignment
- Hyperbaric Oxygen Therapy

Number of articles addressing each intervention

The numbers of articles that address each of the interventions listed above are presented in Table 41. A full list of articles and the interventions that they address is presented in Evidence Table 209.

Only two treatment modalities, VNS and the ketogenic diet, were addressed by five or more studies. As per the general inclusion criteria specified in the Methodology section, data about treatments not addressed by at least five studies are not considered further in this section of the report. Consequently, we do not include further information about magnet therapy, vitamin B_6 therapy, herbal medicine, acupuncture, electrical brain stimulation, chiropractic therapy, cranial realignment, and hyperbaric oxygen therapy.

Vagal Nerve Stimulation

VNS is considered an adjunct therapy for patients with treatment-resistant epilepsy who are not candidates for surgery or who have undergone unsuccessful surgical intervention. Thus, patients considered for VNS tend to have the most severe forms of treatment-resistant epilepsy. VNS is presumed to elicit its antiseizure effects through the repeated stimulation of the left vagus nerve by an implanted electrode. Despite studies in both animals and man, the mode of action of VNS remains unknown.[330]

Excluded articles

Not all of the articles that met the general and subquestion specific inclusion criteria were included in the evidence base for this intervention. We list the single study that we excluded for reasons of quality in Evidence Table 210 along with an explanation as to why it was excluded.

Evidence base

Following the exclusion of the single article, 14 included articles describing 14 separate studies remained. These articles, which described data collected from 565 patients, are listed in Table 42. Complete details of the study characteristics are presented in Evidence Tables 211 through 217.

Design and conduct of included studies

Of the 14 articles included in this evidence base, two described double-blinded, multi-center RCTs (n > 50 in each study arm).[331,332] The remaining 12 articles described nonblinded, longitudinal case series studies. Two of these case series[31,333] were comprised of patients from the two RCTs. Patients who entered each of the original RCTs were, on their completion, entered into a separate long-term followup study. In the followup study, all patients (including those randomized to the control arm) were treated with VNS and followed for an extended period. The original RCTs followed patients for approximately 3 months, and both long-term followup studies followed patients for a further 12 months. This was possible because all patients included in both RCTs received a VNS device at the onset of the study. Those patients who were randomized to the control group had their device activated, but it was set to a level considered by the investigators to have minimal therapeutic effects (a so-called "active-control" group). At the end of the study, the VNS device parameters in all those patients in the active-control group were reset to therapeutic levels and all patients were then followed. In order to avoid double counting of patients, we have, when appropriate, analyzed data from the RCTs independently of data from the corresponding followup studies.

The remainder of this section presents the findings of our systematic assessment of the quality of the evidence for Question 5B. This systematic assessment consisted of an appraisal of each study's internal and external validity.

Internal validity

Sample selection bias, patient reporting bias, and measurement bias potentially affected all of the studies in the evidence base. The eight case series studies in which patients were followed for more than 12 months were all potentially affected by maturation bias. Investigator reporting bias, regression to the mean, and extraneous event bias were not present in the two RCTs conducted in the United States but may have affected all of the other studies. Selection bias was also not present in the two RCTs. Eight studies did not report on their patient recruiting methods and may be prone to sampling bias. Further details on the potentially biases present in the studies addressing this question are presented in Appendix B.

External validity

Complete details of the characteristics of the patients enrolled in the studies in the current evidence base are presented in Evidence Tables 218 and 219. In all of the studies, patients who received VNS were either not considered to be candidates for epilepsy surgery or, had undergone surgery that was unsuccessful in controlling their seizures.

Twelve of the 14 studies enrolled patients because they were considered representative of a specific subpopulation of patients with treatment-resistant epilepsy (Evidence Table 220). Nine of the studies, including the two RCTs, recruited patients because they had a particular seizure type or syndrome.[31,331-340] Three studies recruited patients because they were specifically interested in the assessing the effectiveness of VNS in children only.[337,340,341] The remaining two studies did not restrict their patient sample by age or seizure type.[342] Thus, the findings of the latter two studies may be considered as being more generalizable to the population of patients of interest than those reported by the former 12 studies. Five studies reported individual patient data and are presented in Evidence Tables 221 to 225.[343]

Synthesis of study results

Not all of the outcomes of interest to the Expert Panel and the Technical Experts were reported by all of the articles included for this subquestion. Those outcomes that were reported, and the articles that reported them, are listed in Table 43.

Seizure frequency-based outcome measures

As shown in Table 43, the following outcome measures related to seizure frequency were reported in the articles that comprise the present evidence base:

- Percentage reduction in seizure frequency from baseline.
- Change in absolute seizure frequency from baseline.
- Proportion of seizure-free patients.
- Proportion of patients with a greater than 50 percent reduction in seizure frequency.

Percentage change in seizure frequency from baseline. Twelve of the articles included in the present evidence base presented data on this outcome. Both of the RCTs reported this outcome and these data are presented in Evidence Table 226. Because data from only two RCTs met the criteria for inclusion in this report, we did not perform a meta-analysis.

Both RCTs found that patients in the treatment group experienced statistically significant reductions in seizure frequency when compared to patients in the active-control group. The between groups differences in mean improvement were small (12.7 percent, CI: 2.4 percent to 23.1 percent in clinical trial EO5; 18.4 percent, CI: 4.5 percent to 32.3 percent in clinical trial EO3). Therefore, the clinical importance of this difference is unclear.

However, patients in the active-control groups also demonstrated reductions in seizure frequency from baseline. This improvement in the "active-control" group may have been due to a placebo effect or regression to the mean, but the possibility remains that it was the result of VNS. Although the stimulation levels used in this patient group were minimal, the study authors did not establish that these stimulation levels were subtherapeutic. Indeed, in their discussion of Clinical Trial EO5, Handforth, DeGiorgio, Schachter, et al.[332] stated that they did not assume that low stimulation was ineffective. If the level of VNS stimulation experienced by the patients in the two active-control groups were therapeutic, then the effects of treatment demonstrated by the two RCTs may underestimate the true effectiveness of VNS when applied at maximum tolerable stimulation levels. On the other hand, the integrity of the blinding of the two RCTs cannot be assumed and[344,345] we cannot discount the possibility that the between-groups difference is actually an overestimate of treatment effectiveness.

Ten of the 12 case series reported on the percentage change in seizure frequency from baseline. These data are presented in Evidence Table 226. Only four studies reported means and standard deviations and therefore we calculated effect sizes for these four only. We then combined these data in a meta-analysis with the percentage change in seizure frequency from baseline effect size data calculated for the treatment arms of the two RCTs. Our threshold analysis (Evidence Table 227) revealed the presence of statistically significant heterogeneity ($Q = 13.12$; $p = 0.022$).

We next examined this heterogeneity using least-squares meta-regression. Predictor variables used in these analyses included: sample size; whether the data originated from a RCT; attrition rate (percent); followup time (months); patient age at surgical implant of the VNS device (years); the proportion of patients who were male (percent); the proportion of patients with partial seizures; the proportion of patients with generalized seizures; and the proportion of patients with Lennox-Gastaut syndrome.

The results of our analysis of this heterogeneity are presented in Evidence Tables 228 and Figure 82. They show that data originating from an RCT was a statistically significant predictor of the magnitude of the effect of VNS. The percentage reduction in seizure frequency from baseline from both of the RCTs combined (summary mean percentage reduction = 21.9, CI: 13.7 percent to 30.0 percent) was approximately 53 percent lower than that found in the case series (summary mean percentage reduction = 46.9, CI: 35.5 percent to 59.3 percent). This suggests that the case series included in the present analysis may overestimate the true magnitude of treatment effectiveness. Although the reason for this difference cannot be determined with certainty, one plausible explanation is that the double blinding in the two RCTs was successful. Because none of the case series were blinded, these studies may have been affected by both investigator and patient biases that caused an overestimation of the effectiveness of VNS by approximately two-fold.

An alternative explanation is that the effects of VNS may increase over time. Thus, the difference between data from the two RCTs, which had limited followup time (3 to 4 months) would be expected to demonstrate lower treatment effectiveness than the case series which had, on average, longer followup times (range: 3 months to over 2 years). Our least squares meta-regression analyses of the available data, however, did not suggest that VNS effectiveness increases with increasing followup time. Consequently, the difference in treatment effectiveness estimated from the RCTs and the case series is not likely to be explained by differences in followup time.

In light of the findings described above, we conclude that the case series studies in the present evidence base overestimate of the effectiveness of VNS. Although the precise cause of this overestimate cannot be determined, the most plausible explanation relates to the blinding status of the studies.

Because the quantitative evidence suggests that the case series are biased, we have not further considered data from these studies. Our conclusions regarding the effectiveness of VNS are thus based solely on the findings of the two RCTs. Consequently, our evidence-based conclusions regarding the effectiveness of VNS are only generalizable to patients with similar characteristics to those included in the two RCTs. These characteristics are patients in the age range between 12 and 60 years of age with partial seizures who were not considered candidates for surgery. Evidence-based conclusions about the effectiveness of VNS in other patient populations cannot be made with the available data.

Change in seizure frequency in absolute terms. One of the two RCTs (Clinical Trial EO3[331]) presented data describing changes in absolute seizure frequency from baseline. These data are presented in Evidence Table 229. Although the study investigators presented median pre- and posttreatment seizure frequency data, they did not present any dispersion data. Thus, we were precluded from performing any secondary analyses of this data. The study investigators reported a statistically significant difference between groups ($p = 0.02$), with patients who received VNS at maximum tolerable levels demonstrating the greatest reduction in seizure frequency. Because data were only available from a single RCT, the magnitude of the measured treatment effect may be unreliable. Therefore, conclusions cannot be drawn about the effectiveness of VNS, where effectiveness is defined by a reduction in absolute seizure frequency from baseline. However, the available evidence does strongly suggest that VNS may offer an effective adjunctive treatment option when treatment effectiveness is gauged by reductions in absolute seizure frequency from baseline levels.

Proportion of patients seizure-free. No RCTs reported this outcome. Thus, we have not considered it further. Evidence Table 230 presents data from a single case series study that reported this outcome.

Proportion of patients with ³50 percent reduction in seizure frequency. Both of the RCTs reported the difference in the proportion of patients who demonstrated a greater than 50 percent reduction in seizure frequency when compared to baseline levels in the treatment and control groups. These data are presented in Evidence Table 231.

Because data from only two RCTs were available, we did not subject them to a meta-analysis. Thus, our conclusions are based on a semi-quantitative analysis. Data from both RCTs suggest that VNS, when applied at maximal tolerable levels, reduces seizure frequency by greater than 50 percent in a statistically significantly higher proportion of patients than does VNS applied at just perceptible levels.

Non-seizure frequency-based outcome measures

Data for three different nonseizure frequency-related outcome measures were presented by one or both of the RCTs included in the present evidence base. These outcome measures were quality of life (one RCT), adverse events (both RCTs), and mortality (both RCTs).

Quality of life. Quality of life data from Clinical Trial EO3[333] are presented in Evidence Table 232. This study used two validated measurement instruments, Quality of Life in Epilepsy-31 (QOLIE-31) and Short-Form-36 (SF-36)). Quality of life data were not collected from all patients enrolled in the RCT (78 of 95 patients in treatment group and 82 of 103 patients in active-control group). Rather, they were collected from only 82 percent (160 of 195) of patients in the study. The authors reported that complete and usable test information was not available from the remaining patients because of mental retardation too severe to permit testing, postictal confusion following a recent seizure, or scheduling problems.

Quality of life data collected using SF-36 indicate that patients treated with VNS at maximum tolerable stimulation levels exhibited greater improvements compared to patients in the active-control group. However, no such benefit was found when quality of life was measured using the QOLIE-31. Since the QOLIE-31 was designed specifically for patients with epilepsy, this instrument might be expected to be more sensitive to improvements in quality of life occurring in this patient population. Given these mixed findings, firm evidence-based conclusions cannot be made about the influence of VNS on quality of life.

Adverse events. Both of the RCTs included in the evidence base reported adverse events. These data, which compare adverse event rates experienced by those patients in the treatment and active-control groups, are presented in Evidence Table 233.

The vagus nerve innervates the thoracic and abdominal organs as well as the larynx, pharynx, and palate, and contains motor fibers involved in swallowing, speech, and the gag reflex. In addition, its afferent component carries sensory information from the heart, lungs, digestive tract, and carotid artery. Therefore, adverse events related to these areas were investigated by both of the included studies.

Other than one case of sepsis that led to the death of a patient in the long-term followup of clinical trial EO5 described by DeGiorgio, Schachter, Handforth, et al.,[333] adverse events tended to be minor and reversible. The most common adverse events experienced by patients treated with VNS were hoarseness and throat irritation.

Mortality. No deaths were reported during the study of the two RCTs included in the evidence base. However, four deaths occurred among the patients in these RCTs during the long-term followup periods reported in the case series. The number of patients that died in each study and the reported causes of death are presented in Evidence Table 234.

Of the four deaths documented by the two long-term followup studies of the RCTs, one death could be directly attributable to VNS. This death resulted from untreated sepsis resulting from implantation of the VNS device. Thus, treatment-related mortality is rare among patients treated with VNS.

The Ketogenic Diet

The ketogenic diet as a means of seizure control in patients with epilepsy was first introduced in 1921 and was based on observations that fasting led to reductions in seizure frequency. This diet, which provides about 87 percent of its energy as fat, is primarily used in the treatment of children with treatment-resistant epilepsy.

Excluded articles

Four of the eight studies meeting our inclusion criteria were excluded for reasons of quality (Evidence Table 235). As per the general inclusion criteria specified in the Methodology section, data about treatments not addressed by at least five included studies (or at least one RCT with more than 50 patients in each study arm) are not considered further in this section of the report. Consequently, we do not further assess this intervention in this report.

Table 41. Articles addressing nondrug, nonsurgery interventions

Intervention	Total Number of Studies	Intervention	Total Number of Studies
Vagal Nerve Stimulation	15	Magnetic Therapy	0
Ketogenic Diet	8	Vitamin B6 Therapy	0
Herbal Medicine	4	Chiropractic Therapy	0
Electrical Brain Stimulation	4	Cranial Realignment	0
Acupuncture	2	Hyperbaric Oxygen Therapy	0

Table 42. Vagal nerve stimulation: included articles

Reference	Number of Patients	Study Design	Country	Multicenter?	Number of Centers	Industry Funded?
Aldenkamp (2001)[337]	16	Case Series	Holland	No	1	Not reported
Chayasirisobhon (2001)[346]	24	Case Series	United States	No	1	Yes
Ergene (2001)[335]	17	Case Series	United States	No	1	Not reported
Hoppe(2001)[339]	36	Case Series	Germany	No	1	Yes
DiGiorgio (2000)[333] Followup of Clinical Trial EO5	199	Case Series	United States	Yes	20	Yes
Hosain (2000)[336]	13	Case Series	United States	No	1	Not reported
Ben-Menachem (1999)[343]	64	Case Series	Sweden	No	1	No
Boon (1999)[338]	25	Case Series	Belgium	No	1	No
Clinical Trial EO4 Labar (1999)[334]	25	Case Series	United States	Yes	Not reported	Yes
Parker (1999)[340]	16	Case Series	United Kingdom	No	1	No
Clinicial Study EO5 Handforth (1998)[332]	199	Randomized Controlled Trial	United States	Yes	20	Yes
Lundgren (1998)[341]	16	Case Series	Sweden	No	1	Not reported
Salinski (1996)[31] Followup of Clinical Trial EO3	114	Case Series	Multinational[a]	Yes	17	Yes
Clinical Trial EO3 The VNS Group (1995)[331]	114	Randomized Controlled Trial	Multinational[a]	Yes	17	Yes

[a] United States / Germany / Sweden / Canada / Holland
NA Not applicable

Table 43. Outcome reporting in studies of vagal nerve stimulation

Reference	Reported Seizure Outcomes			
	Percent of Patients with >50 percent Reduction in Seizure Rate	Percent of Patients Sezure -free	Difference in Absolute Seizure Frequency	Percent Reduction in Seizure Frequency[a]
RCTs performed in United States				
Clinical Study EO5 (1998)[332]	✓[b]			✓[c]
Clinical Trial EO3 (1995)[331]	✓		✓	✓
Long-term followup of RCTs performed in United States				
DiGiorgio (2000)[333] Followup of Clinical Trial EO5	✓			✓
Salinski (1996)[31] Followup of Clinical Trial EO3	✓			✓
Case series performed in United States				
Chayasirisobhon (2001)[346]	✓		✓	✓
Ergene (2001)[335]				
Hosain (2000)[336]	✓			✓
Clinical Trial EO4 Labar (1999)[334]	✓			✓
Case series performed outside United States				
Aldenkamp (2001)[337]	✓		✓	✓
Hoppe (2001)[339]	✓		✓	✓
Ben-Menachem (1999)[343]	✓			
Boon (1999)[338]	✓	✓	✓	✓
Parker (1999)[340]	✓			✓
Lundgren (1998)[341]	✓		✓	✓
Totals	**13**	**1**	**6**	**12**

[a] Reduction from baseline unless otherwise indicated
[b] This data comes from article by Handforth, DeGiorgio, Schachter, et al.[332]
[c] This data comes from article by Dodrill and Morris[347]

Table 43. Outcome reporting in studies of vagal nerve stimulation (continued)

Reference	Reported Non-Seizure Outcomes							
	Quality of Life	Mood	Cognition	Ability to Return, Obtain, or Remain at Work	Ability to Return, Obtain, or Remain in School	Ability to Hold Driver's Licence	Adverse Event	Mortality
RCTs performed in United States								
Clinical Study EO5 (1998)[332]	✓[c]						✓[b]	✓[b]
Clinical Trial EO3 (1995)[331]							✓	✓
Long-term followup of RCTs performed in United States								
DiGiorgio (2000)[333] Followup of Clinical Trial EO5							✓	✓
Salinski (1996)[31] Followup of Clinical Trial EO3							✓	✓
Case series performed in United States								
Chayasirisobhon (2001)[346]	✓[d]						✓	✓
Ergene (2001)[335]	✓						✓	✓
Hosain (2000)[336]							✓	✓
Clinical Trial EO4 Labar. (1999)[334]							✓	✓
Case series performed outside United States								
Aldenkamp (2001)[337]	✓[e]							✓
Hoppe (2001)[339]								✓
Ben-Menachem (1999)[343]							✓	✓
Boon (1999)[338]							✓	✓
Parker (1999)[340]	✓[g]							✓
Lundgren (1998)[341]	✓[f]						✓	✓
Totals	**2[h]**	**0**	**0**	**0**	**0**	**0**	**11**	**14**

[b] This data comes from article by Handforth, DeGiorgio, Schachter, et al.[332]
[c] This data comes from article by Dodrill and Morris[347]

Footnotes continue on the next page.

[d] Authors reported that they measured quality of life and discuss findings in conclusions section. However, they do not present any data

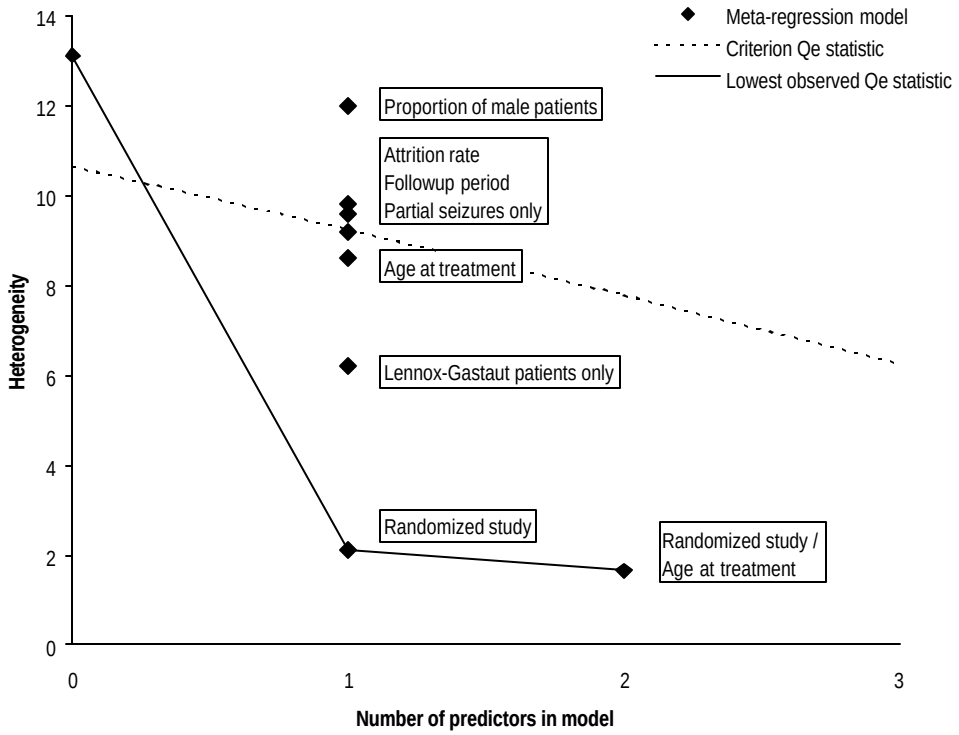
[e] Authors stated that they measured quality of life but did not use a validated quality of life instrument. Instead they used four instruments that measured four domains that the authors claim provide a estimate of quality of life. We were unable to confirm that these four domains are related to overall quality of life and thus do not consider this data further

[f] Authors used a nonvalidated quality of life instrument. These data are not considered further

[g] Authors used a validated instrument (Welcome Quality of Life Assessment). However, they only report p-values and did not present data that could be validated

[f] Excludes quality of life data from four articles (see footnotes d and g)

Figure 82. Meta-regression: vagal nerve stimulation and percentage change in seizure frequency



Nonmedical Treatments

In this section of the Evidence Report, we addressed Key Question #6: *Which social, psychological or psychiatric services for treatment-resistant epilepsy lead to, or can be expected to lead to improved patient outcomes?*

In the present question, we address whether nonmedical treatments have an effect on patients with treatment-resistant epilepsy. Nonmedical treatments include education and training in skills that may help prevent seizures or enable the patient to better adapt to seizures. For the purpose of this question, we separately consider each treatment or group of related treatments. Thus, groups of studies of each treatment are assessed, first for their quality, then for their outcomes. No attempt is made to compare different treatments.

Question specific inclusion criteria

In addition to the general inclusion criteria described in the Methodology section, we included studies for this question if:

1. They examined a social, psychological or psychiatric intervention. Studies reporting on the effects of a drug, device or surgical procedure were excluded.
2. They classified patients according to any classification system (studies were not required to classify patients according to the International League Against Epilepsy).

Number of studies addressing each intervention

Applying these criteria gave us 25 studies of 12 interventions. They are listed in Table 44. A full list of articles and the interventions that they address are presented in Evidence Table 236.

Only two treatment modalities, multidisciplinary neurobehavioral treatment and EEG biofeedback, were addressed by five or more studies. As per the general inclusion criteria specified in the Methodology section, data about treatments not addressed by at least five studies are not considered further in this section of the report.

Multidisciplinary Neurobehavioral Treatments

Multidisciplinary neurobehavioral treatments are comprised of epilepsy programs. Patients are taught to identify and avoid situations that may precipitate seizures. This may include stressful situations, loud noises, flashing lights, and other individualized environmental cues. Recognizing auras and applying techniques such as relaxation might help to prevent the aura from developing into a full seizure. Some programs also include individualized counseling, relaxation training, or EEG biofeedback.

Evidence base

Six studies describing 231 patients utilized a multidisciplinary neurobehavioral approach to treatment (Evidence Table 236).

Excluded studies

Two studies of neurobehavioral treatments were excluded for reasons of quality. In both studies, patients changed their AED regimens during treatment. This external event obscures any

association between the observed outcomes and the treatment. This left four studies remaining. Because fewer than five included studies examined this intervention, we did not proceed with a further analysis.

EEG Biofeedback

EEG biofeedback uses auditory or visual signals to train patients to control their EEG. By altering their EEG patterns to increase waveforms thought to discourage seizures and decrease waveforms believed to promote seizures, patients may be able to reduce or eliminate seizures. The precise waveforms promoted or repressed may vary somewhat among studies depending on the theoretical underpinnings of the treatment.

Evidence base

Six studies of EEG biofeedback, describing 143 patients, met the inclusion criteria for this question (Evidence Table 236).

Excluded studies

Two studies of EEG biofeedback were excluded for reasons of quality.[348,349] In one study, patients received behavioral therapy in addition to EEG biofeedback, and therefore neither therapy could be associated with the studies outcomes. In the other, some patients received EEG biofeedback while others received end-tidal CO_2 biofeedback. Outcomes for the two groups were not individually reported, again preventing any interpretation of the effect of EEG biofeedback. This left four studies remaining. Because fewer than five included studies reported on this intervention, we did not proceed with further analyses. A complete list of studies, and the reason for their exclusion, is presented in Evidence Table 236.

Table 44. Interventions for nonmedical treatments

Intervention	Number of Studies Reporting
Multidisciplinary Neurobehavioral Treatments	6
EEG Biofeedback	6[a]
Medical Resonance Therapy Music	1
Sahaja Yoga	1
Meditation	1
Physical Exercise	3
Self-Help Group (Group Therapy)	1
Counseling	1
Progressive Muscle Relaxation	2
End-Tidal CO_2 Biofeedback	2[a]
Vocational Services	1
Systematic Desensitization	1
Epilepsy Education	0

[a] One study reported on both of these interventions.

Employment and School

In this section of the Evidence Report, we addressed Key Question #7: *What characteristics of treatment-resistant epilepsy interfere with ability to obtain and maintain employment, or attend and perform well in school?*

In addressing this question, we consider whether published literature suggests that patient or disease characteristics can predict or correlate with poor performance or difficulty at work or school.

Question specific inclusion criteria

In addition to the general inclusion criteria described in the Methodology section, we included studies for this question if:

1. They attempted to identify relevant patient characteristics using regression techniques, or
2. They compared outcomes in different groups of patients with different characteristics

Because randomizing patients to groups with different employability or school attendance is not possible, the above two criteria offer the only realistic way of addressing this question.

We did not require that the study exclusively enroll patients with treatment-resistant epilepsy. An analysis of the effect of seizure frequency on patient employment or academic performance may be aided by including some patients with epilepsy whose seizure frequency is relatively low because it avoids range restriction. However, we did have two reasons to exclude studies as being beyond the scope of this Evidence Report. Studies of newly diagnosed patients (for whom the effect of epilepsy on employment or school attendance may not have been established) and studies in which the number of patients with nontreatment-resistant epilepsy (who are not the subject of this Evidence Report) exceeded 25 percent of the study population were excluded. This relaxation of our general exclusion criteria enabled us to expand the number of studies included and increased the possibility that we would include a sufficient number of studies to run an analysis.

Excluded Studies

Five studies met our inclusion criteria. Two were subsequently excluded for reasons of quality. These studies and the reasons for their exclusion are listed in Evidence Table 237.

Evidence Base

Three studies met all inclusion criteria; Seidenberg,[350] Sturniolo and Galletti,[351] and Bulteau, Jambaque, Viguier, et al.[352] Because fewer than five studies were available, we did not proceed with an analysis.

Mortality Rate

In this section of the Evidence Report, we addressed Key Question #8: *What is the mortality rate in patients with treatment-resistant epilepsy?*

In this question, we examine mortality rates in persons with treatment-resistant epilepsy. We consider mortality from a number of causes, including overall (all-cause) mortality, sudden unexpected death in epilepsy (SUDEP)-related mortality, and mortality from other causes. Whenever possible, we compared mortality in studies from a given country to the mortality in a reference population from the same country. For reference data, we searched national databases from different countries that contained mortality rates for general populations.

If not reported in the study, we have (wherever possible) calculated a standardized mortality ratio (SMR), which is based on the number of observed deaths divided by the number of expected deaths. The latter number is the number of deaths expected given the age distribution of the study population and the age-specific death rates in the general population (from the country where the study was conducted, if possible). SMR is the primary measure of interest because it provides a comparison to a reference population that is at least standardized by age. In contrast, drawing conclusions from nonstandardized (crude) mortality ratios (CMR) is difficult because mortality rates vary depending on the age of the population, and CMRs do not adjust for this. Comparison of CMRs from two populations with differing age distributions may therefore be misleading.

As an additional source of reference data suggested by the Technical Experts, we examined mortality rates and SMRs among patients newly diagnosed with epilepsy as reported in epidemiological studies and large clinical trials from institutions that conducted studies of patients with treatment-resistant epilepsy. Patients newly diagnosed will include those who have treatment-resistant epilepsy and those who do not. Therefore, mortality among patients with newly diagnosed disease could be lower than that in patients with treatment-resistant epilepsy.

Question specific inclusion criteria

For this question, we included mortality data from epidemiological studies, published clinical trials of epilepsy treatments, and the Physician's Desk Reference. We included studies if they:

1. Reported mortality rates (or enough information to allow independent calculation) or SMRs.
2. Reported overall mortality, SUDEP, or mortality from other causes.
3. The goal of the study was to evaluate mortality or adverse events (including mortality). This criterion ensures that we included only studies with an adequate number of person-years of followup to allow unbiased mortality data. Studies with too few person-years of followup are likely to have no deaths (or very few deaths), resulting in large fluctuations in reported mortality rates among smaller studies. Two analysts reviewed decisions as to which studies met this inclusion criterion.

Excluded studies

As discussed in the Methodology section, we retrieved articles identified by our literature searches according to certain *a priori* criteria. Of those that did meet these criteria, some were excluded for reasons of quality and some were excluded because they contained data that was

published in other included studies. Evidence Table 238 lists the studies we excluded and the reason for their exclusion.

Evidence base

After the above exclusions, there were 10 studies with 22,462 patients available to address this question.

Design and conduct of included studies

The ideal study design for addressing this question is a prospective cohort study that follows patients with treatment-resistant epilepsy for several years, records all-cause mortality and mortality from specific causes, and reports SMRs using age-specific mortality rates from a national reference population. Only one of the studies was prospective, a cohort study of mortality among surgical patients. Evidence Table 239 presents data relevant to study quality. The remaining studies were retrospective cohort studies and case series that evaluated mortality among surgical patients or used databases of information collected about patients who received a variety of AEDs.

Internal validity

We evaluated each study's potential for certain biases as discussed in the Methodology section. Three of 10 studies were vulnerable to mortality ratio bias for overall mortality, and all studies were vulnerable to mortality ratio bias for cause-specific mortality. Nine of 10 studies were vulnerable to sampling bias, and all studies were vulnerable to sample specification bias. More detailed information regarding the internal validity of these studies is presented in Appendix B.

External validity

Whether a study is an epidemiological study or clinical trial can affect its generalizability to the larger target population of patients (Table 47). Epidemiological studies tend to have the widest inclusion criteria, and therefore tend to have greater generalizability. The evidence base for this question included one study that examined all patients who received a specific AED at five United Kingdom centers,[353] and another study that examined all patients in long-term residential care at an epilepsy center.[354] Two more studies mixed aspects of epidemiological studies and clinical trials. An AED study of patients in a database contained a mix of clinical trial participants and "compassionate use" patients,[355] while a study of all patients who received VNS contained clinical trial participants and open market patients.[356]

Clinical trials of AEDs generally exclude patients with the most severe (life-threatening) epilepsy and often exclude patients with comorbidities (e.g. cardiovascular disease),[357] a practice that limits their generalizability. The remaining six studies that addressed this question evaluated either clinical trial patients from AED development databases,[357-359] or patients who received surgical treatment at a single institution.[360-362]

Patient treatment history is another variable that could affect the generalizability of these studies. Because the range of treatments given in these studies did not span the full spectrum of treatments, none of the studies may be fully generalizable to the overall population of patients with treatment-resistant epilepsy. For example, many patients become seizure-free following surgery. If these patients are included in the followup mortality analysis, the mortality rate may be lower than what might have been observed if all patients still experienced seizures. Two of

three surgical studies did not separate seizure-free patients from nonseizure-free patients in their analysis of mortality.[361,362] Therefore, although the reported mortality rates may accurately reflect the mortality rates for subgroups of patients who receive different treatments, they may not reflect the expected rate for the overall population of patients with treatment-resistant epilepsy. Studies of AED recipients that did not exclude less healthy patients may be the most generalizable, because this is the largest subgroup of patients with treatment-resistant epilepsy. Six studies reported AEDs as the primary treatment,[353-355,357-359] three studies evaluated surgical patients,[360-362] and one study evaluated patients who received VNS from an implanted device.[356] These studies are presented in Evidence Table 239.

Mean patient age (or age at death) in different studies may affect generalizability if SMRs vary among different age groups. Only five of 10 studies reported either mean patient age or age range (two reported both,[353,354] one reported only mean age,[360] and two reported only the range[357,359]). Therefore, determining whether the mean age or even the age ranges of the patients in these studies are typical of the overall population of patients with treatment-resistant epilepsy is difficult. One study specifically focused on adult patients (18 years or older).[354] The remaining studies included both pediatric and adult patients, with the possible exception of one study that did not report enough information to confirm this assumption.[360] Five of 10 studies reported either mean age at death or range of age at death,[354-356,360,362] and five of 10 studies reported seizure types of the patients in the respective study populations.[353,355,357,359,363] Therefore, determining whether the patients in these studies are representative of the overall population of patients with treatment-resistant epilepsy is difficult. The study by Annegers, Coan, Hauser, et al.[356] reported SMRs for different age subgroups, while the study by Racoosin, Feeney, Burkhart, et al.[357] reported enough information to allow us to independently calculate approximate SMRs for different age subgroups. Thus, these two studies present the most useful information related to mortality and age.

Synthesis of study results

The studies included for this question reported several different types of mortality (Table 46). Mortality specifically caused by epilepsy is sometimes difficult to determine; reported definitions include sudden unexpected death, accidents, and aspiration. Because not all accidents and sudden deaths are necessarily epilepsy-related, we have addressed each of these mortality rates separately. We addressed treatment-related mortality in Questions 4 and 5. The relationship between seizure type and frequency and sudden unexpected death in epilepsy is addressed in Question 9.

Overall Mortality

Overall mortality rates were obtained from six studies (three from the United States and three from the United Kingdom) of patients with treatment-resistant epilepsy that either calculated SMRs, or from which we could independently calculate SMRs (Evidence Table 240). Because SMRs from different studies are based on different standards (each study is standardized according to its own age distribution), they are inherently noncomparable.[364] For this reason, we have not combined individual study SMRs in a meta-analysis.

The SMRs from these six studies suggest that the overall mortality rate for patients with treatment-resistant epilepsy is approximately 1.9 to 10.4 times greater compared to that observed in general reference populations from the United States and United Kingdom (Figure 83 and Evidence Table 240). SMRs from the United States studies ranged from 3.6 to 4.7.[356,357,360] Two United States studies that separated results for males and females found that the increased

mortality rate was independent of gender.[356,357] Only two studies presented results for different age groups.[356,357] Although the studies did not use identical age subgroups, in both studies the highest SMRs (16.4 and 11.4) occurred in the youngest group and the lowest SMRs (2.2 and 1.8) occurred in the oldest group (Figure 84). As mentioned earlier, we independently calculated the SMRs from the study by Racoosin, Feeney, Burkhart, et al.[357] These approximate SMRs are less precise than those reported in the study by Annegers, Coan, Hauser, et al.[356] However, Figure 84 shows good agreement between the approximate SMRs we calculated and the SMRs reported by Annegers, Coan, Hauser, et al.[356] This increases our confidence that the independently calculated SMRs closely approximate the true numbers. Since the SMR seems to vary considerably depending on patient age group, summary SMRs derived from a study group with a large age range may not accurately reflect the SMRs for more specific age subgroups.

As an indirect comparison, a United Kingdom study of newly diagnosed patients with epilepsy found a two-fold higher mortality rate compared to the general reference population.[365] However, the patient age distributions in the treatment-resistant epilepsy studies were not identical to those in the newly diagnosed patient study, so caution is required when considering these comparisons.

Sudden Unexpected Death in Epilepsy (SUDEP)

Mortality resulting from SUDEP was reported in nine studies of patients with treatment-resistant epilepsy (Evidence Table 241). Neither SMRs nor even CMRs for SUDEP could be calculated because of the differing definitions of sudden unexpected death in patients with epilepsy and in the general population. A classification of sudden unexpected death in patients with epilepsy generally requires that the death be unexplained (no obvious cause appears on autopsy). On the other hand, most cases of sudden unexpected death in the general population have a definable cause upon autopsy (most frequently cardiac disease).[366] The implication is that the rate of sudden unexpected death as defined in the general population increases with age due to the large percentage of cases with cardiovascular causes.[367] In contrast, some studies have suggested that sudden unexpected death rates begin to decrease after middle age among patients with epilepsy.[368,369] Due to these differing definitions, sudden unexpected death rates among patients with treatment-resistant epilepsy and the general population are inherently noncomparable. Therefore, we report the SUDEP rates per 1,000 person-years among patients with treatment-resistant epilepsy. The rates ranged from 2.1 to 7.6 per 1,000 person-years, and they represented 6 percent to 55 percent of the total deaths reported in the individual studies.

Nine studies presented data concerning sudden unexpected death among patients with treatment-resistant epilepsy. Four were from the United States, two were multi-country studies that included United States patients, and three were from the United Kingdom. United States studies yielded SUDEP rates ranging from 3.8 to 7.5 per 1,000 person-years.[357,358,360,370] In contrast to other reports in the literature, the one study that met our inclusion criteria and reported SUDEP rates for four different age groups of patients did not find a decrease in SUDEP rates among the oldest age group (age 55-72).[357] However, since none of these cases was autopsied, the relatively high rate among older patients in this study could have resulted partly from cardiac causes that might have been identified by autopsy.[368,369]

Drowning

Drowning was reported in four studies examining patients with treatment-resistant epilepsy (Evidence Table 242).[354,355,361,362] One of the studies exclusively contained patients from the United States, and one international study contained some United States patients. CMRs

calculated using the average drown ing rate across all ages suggested a higher drowning rate among patients with treatment-resistant epilepsy compared to a general reference population. Even when we used the highest age-specific rate for drowning (for men age ≥ 85) from the general population to calculate CMRs, all but one of the CMRs were statistically significant (lower CI >1). This conservative analysis increases the confidence that the drowning rate is truly higher among patients with treatment-resistant epilepsy. However, better quality evidence is needed to determine the true magnitude of the mortality difference. CMRs comparing drowning rates from a study of newly-diagnosed epilepsy patients with a general reference population showed a trend toward a higher rate among patients with epilepsy, but it was not statistically significant.[365]

Accident-related mortality

Accident-related mortality combines death from all types of accidents (including drowning and automobile accidents which are also addressed separately). This outcome was reported in six studies of patients with treatment-resistant epilepsy (Evidence Table 243). Due to a lack of information, SMRs could not be calculated. Instead, we have calculated CMRs in an exploratory analysis with the caveat that these numbers may be imprecise since they were not adjusted for the age of the study populations. Two studies were from the United States, two were multi-country studies that included United States patients, and two were from the United Kingdom. We compared mortality rates from United States studies and international studies that included patients from the United States to the age-adjusted accident-related mortality rate reported in the U.S. Census Bureau Statistical Abstract of the United States for 2000.[371] For United Kingdom studies, we compared mortality rates to crude accident-related mortality rates reported in Mortality Statistics (England and Wales, 1999).[372]

Three out of six studies showed a significantly higher accident mortality rate (lower CI above 1.0) among patients with treatment-resistant epilepsy compared to a general reference population. One was a United States study,[357] one was an international study that included United States patients,[359] and one was a United Kingdom study.[362] The remaining three studies showed a trend in the same direction that was not statistically significant. To test the robustness of these findings, we conducted a sensitivity analysis by calculating CMRs using the highest general population age-specific mortality rate that matched the age of patients in each individual study (Evidence Table 243). Two of the three studies that had shown a statistically significant CMR in the initial analysis became nonsignificant in the sensitivity analysis (one just by a slight margin, however); the remaining study remained statistically significant. Thus, although one study still suggests that overall accident rates are elevated among patients with treatment-resistant epilepsy, the other two studies do not. Under these circumstances, the inherent inaccuracy in these crude ratios precludes determining with certainty whether the accident rate differs between these populations.

One United States study of newly diagnosed patients showed a trend toward higher accident-related mortality among newly diagnosed patients compared to the general reference population, but the trend was not statistically significant.[357]

Automobile accident-related mortality

Only one study reported a death resulting from a motor vehicle accident (other studies may have subsumed automobile accidents in the broader category of accidents).[360] Only a CMR could be calculated from the reported information. Although there was a trend toward a higher rate among treatment-resistant patients, it was not statistically significant (Evidence Table 244).

Since this is only one study and an SMR could not be calculated, no evidence-based conclusions can be reached.

Aspiration-related mortality

Aspiration-related mortality refers to death because of accidental inhalation of food or fluid that blocks respiration. This outcome was reported in four studies of patients with treatment-resistant epilepsy (Evidence Table 245). Because our searches did not locate any references containing aspiration-related mortality rates among the general population or even among patients newly-diagnosed with epilepsy, no mortality ratios could be calculated. Therefore, the tabled mortality rates may not be comparable to rates in other populations.

Mortality from pneumonia

Pneumonia-related mortality was reported in three studies of patients with treatment-resistant epilepsy (Evidence Table 246). For the two studies of United States patients,[355,360] we calculated CMRs using pneumonia death rates from the United States population age-adjusted to the 1940 standard (which is closer to the age distribution of the patients in these studies).[371] For the remaining study (from the United Kingdom),[354] the crude pneumonia mortality rate from the United Kingdom population was used as a reference standard (because the study group had a similar age distribution to the current United Kingdom population).[373] The CMRs varied considerably, and only one was statistically significant. This is possibly due to the effect of age on pneumonia susceptibility, as the study with the oldest mean patient age had the highest pneumonia mortality rate. A sensitivity analysis using the mortality rate of the oldest general population reference group that matched the age of patients in each individual study, overturned the statistically significant finding in the one study that showed a difference, indicating that the original finding is not robust.

One long-term study following newly diagnosed patients reported an SMR of 5.9 (CI: 4.1-8.0), suggesting a higher pneumonia mortality rate among these patients.[365] This study also had, on average, an older patient population. Since the populations in different studies have different age distributions, determining whether the pneumonia mortality rates differ between patients with treatment-resistant epilepsy and the general population of patients with epilepsy was not possible.

Cardiovascular mortality

Although cardiovascular mortality was reported in three studies of patients with treatment-resistant epilepsy, we do not present the results of two of these trials due to bias regarding this particular outcome. The two studies in question evaluated mortality reported in AED databases containing predominantly patients involved in clinical trials.[355,357] Clinical trials of AEDs generally exclude patients with cardiovascular disease, meaning that cardiovascular mortality would be underrepresented in this group of patients. This left one United Kingdom study of surgical patients that did not have this bias (Evidence Table 247). The CMR we calculated from this study was not significantly different compared to the general population,[362] but the inability to calculate an SMR and the low number of studies prevents drawing firm conclusions.

Another United Kingdom long-term study of newly diagnosed patients with epilepsy reported an SMR of 1.1,[365] suggesting that cardiovascular mortality rates did not differ between patients with epilepsy and the general population. Determining whether rates differed between patients with treatment-resistant epilepsy and newly diagnosed epilepsy was not possible.

Cerebrovascular mortality

Cerebrovascular mortality was reported in three studies of patients with treatment-resistant epilepsy (Evidence Table 248).[354,355,357] Again, SMRs could not be calculated for these studies. An exploratory analysis using CMRs did not show any statistically significant differences in cerebrovascular mortality rate among treatment-resistant patients compared to general reference populations in any of the studies. Therefore, we did not perform any additional sensitivity analysis on these results. One United Kingdom study of newly diagnosed patients with epilepsy reported an SMR of 3.2 (CI: 2.2-4.4), suggesting a significantly higher cerebrovascular mortality rate among patients with epilepsy.[365] However, the lack of comparable mortality ratios precludes any firm conclusion concerning relative mortality rates.

Cancer mortality

One United Kingdom study reported SMRs for overall cancer mortality and specific types of cancer (Evidence Table 249).[354] Therefore, we did not attempt to calculate CMRs from other studies because higher quality data was available. This study reported an SMR of 2.0 (CI: 1.3-2.9), indicating a significantly greater cancer mortality rate among patients with treatment-resistant epilepsy. Among specific types of cancer, the highest SMRs were observed for hepatobiliary cancers (17.6, CI: 3.6-51.5) and pancreatic cancer (6.2, CI: 1.7-15.8). In addition, a United Kingdom long-term study of newly diagnosed patients reported an SMR of 2.6 (CI: 1.9-3.4) for overall cancer mortality.[365] These two studies had, on average, the oldest patient populations. An elevated cancer mortality rate may possibly exist among older treatment-resistant patient populations compared to the general population, but more evidence (in the form of studies reporting SMRs) is needed to confirm this trend. There is not enough evidence to determine whether a difference exists between patients with treatment-resistant epilepsy and the overall population of patients with epilepsy.

Suicide

Suicide rates were reported in three studies of patients with treatment-resistant epilepsy (Evidence Table 250).[355,360,362] Because SMRs could not be calculated, we conducted an exploratory analysis using CMRs. The CMRs suggested a trend toward a higher suicide rate in patients with treatment-resistant epilepsy compared to the United States general population. However, the CIs overlapped 1.0 (indicating no statistical significance), and the general population rates could not be age-adjusted to any of the studies. Because none of the studies showed a statistically significant between-population difference in suicide rates, we did not perform a sensitivity analysis.

In summary, the present evidence is insufficient to determine whether suicide rates among patients with intractable epilepsy are higher than expected in the general population. A CMR derived from a United States study of newly diagnosed patients also showed a nonsignificant trend toward higher suicide rates among these patients.[357] Again, there is not enough evidence to make any conclusions regarding the relative suicide rates of any of these populations.

Table 45. External validity of studies of mortality rate

Reference	Country	Number of Patients (Person-Years of Followup)	Type of Study	Treatment	Mean Age (Range)	Mean Age at Death (Range)
Physician's desk reference Gabapentin trial data (2001)[358]	United States	2203 (2103)	Clinical trial	AEDs	Not reported	Not reported
Racoosin (2001)[357]	United States	9144 (13617)	Clinical trial	AEDs	(1-72)	Not reported
Wong (2001)[353]	United Kingdom	1050 (2294)	Epidemiologic	AEDs	31 (7-77)	Not reported
Annegers (2000)[356]	United States	1819 (3176)	Epidemiologic with some clinical trial patients	VNS	Not reported	(6-52)
Hennessy (1999)[362]	United Kingdom	305 (2729)	Clinical trial	Surgery	Not reported	34 (19-54)
Sperling (1999)[360]	United States	194 (801.5)	Clinical trial	Surgery	33.4	34.6 (22.5-42)
Vickrey (1997)[361]	United States	248 (1488)	Clinical trial	Surgery or nonsurgical treatment (not described)	Not reported	Not reported
Leestma (1997)[355]	United States, United Kingdom, Europe, Australia, South Africa	4700 (5747)	Epidemiologic with some clinical trial patients	AEDs	Not reported	36 (0.5-74)
Leppik (1995)[359]	United States, Europe, Australia	2600 (1810)	Clinical trial	AEDs	(12-77)	Not reported
Klenerman (1993)[354]	United Kingdom	Not reported (3392)	Epidemiologic	AEDs	52 (18-91)	64 (23-91)

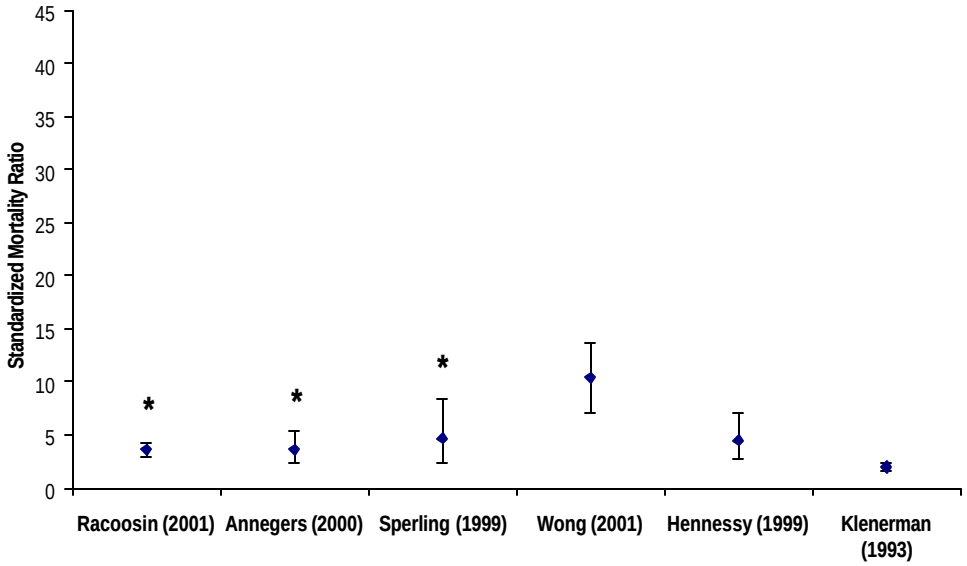
Table 45. External validity of studies of mortality rate (continued)

Reference	Percent Female Patients (all patients)	Percent Female Patients (Deaths)	Patients Recruited Because of Seizure Type	Seizure Types in Study Group
Physician's desk reference, Gabapentin trial data (2001)[358]	Not reported	Not reported	Not reported	Not reported
Racoosin (2001)[357]	45	Not reported	No	Partial, generalized tonic-clonic, Lennox-Gastaut syndrome
Wong (2001)[353]	50.4	Not reported	No	Partial, generalized
Annegers (2000)[356]	Not reported	44	No	Not reported
Hennessy (1999)[362]	Not reported	60	No	Not reported
Sperling (1999)[360]	48.1	45.5	No	Tonic-clonic, simple partial, complex partial, others not reported
Vickrey (1997)[361]	Not reported	Not reported	No	Not reported
Leestma (1997)[355]	Not reported	37.8	Yes	Partial, partial with secondary generalization, generalized
Leppik (1995)[359]	Not reported	Not reported	No	Partial, other uncontrolled seizures (not described)
Klenerman (1993)[354]	33.3	Not reported	No	Not reported

Table 46. Types of reported mortality

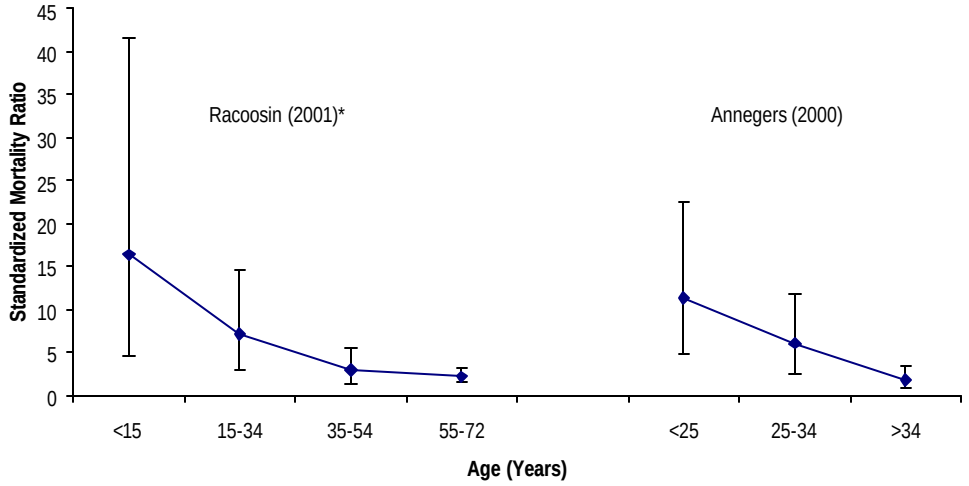
Reference	Overall Mortality	Sudden Unexpected Death (SUDEP)	Epilepsy-Related Mortality	Aspiration	Drowning	Accident	Suicide	Pneumonia	Cardiovascular Disease	Cerebrovascular Disease	Cancer	Treatment Complication
Physician's desk reference, Gabapentin trial data (2001)[358]		✓										
Racoosin (2001)[357]	✓	✓		✓		✓			✓	✓	✓	
Wong (2001)[353]	✓	✓										
Annegers (2000)[370]	✓	✓										
Hennessy (1999)[362]	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓
Sperling (1999)[360]	✓	✓				✓	✓	✓			✓	
Leestma (1997)[355]	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	
Vickrey (1997)[361]	✓											✓
Leppik (1995)[359]	✓	✓				✓					✓	
Klenerman (1993)[354]	✓	✓	✓	✓	✓	✓		✓		✓	✓	

Figure 83. Standardized mortality ratios for overall mortality



*Studies conducted in the United States

Figure 84. Standardized mortality ratios for age-specific mortality



*Approximate SMRs for this study calculated by ECRI

Frequency and Type of Seizure and Sudden Death

In this section of the Evidence Report, we addressed Key Question #9: *Is there a correlation between the number and/or type of seizure and sudden death?*

In the present question, we address whether persons who experience SUDEP and persons with epilepsy who do not experience SUDEP have a history of different seizure types and/or a history of different numbers of seizures.

The present question is not amenable to study by a randomized or nonrandomized clinical trial. Consequently, we address this question using information from observational studies. In the present context, studies of this design cannot be considered to be of lower quality compared to those of RCTs or nonrandomized-controlled trials. This is because the present question requires a comparison of outcomes (in this case, sudden death) in groups of patients that are different from each other. This stands in sharp contrast to the optimal situation in the study of interventions, where having identical groups is desirable.

Specifically, we address the present question using data from case-control studies, where the "cases" are individuals who experienced SUDEP, and the "controls" are individuals with epilepsy who did not experience such a death. More specifically, the "controls" are comprised of patients with epilepsy who were still alive at the time of the study or who had died of causes not related to epilepsy. In addressing the present question, we do not consider studies in which the controls are persons without epilepsy. By its nature, this question requires a group of controls that are vulnerable to seizures. Thus, the controls must have epilepsy. A presentation of sudden death mortality rates among patients with treatment-resistant epilepsy is part of Question 8.

Question specific inclusion criteria

In addition to the general inclusion criteria described in the Methodology section, we included studies for this question if they:

1. Enrolled at least some patients with treatment-resistant epilepsy. We did not require all patients in a study to have this form of epilepsy and, consequently, we did not exclude studies if they enrolled some patients with medically controlled epilepsy. Inclusion of studies that evaluated only treatment-resistant patients would have introduced a "range restriction" in seizure frequency (i.e. no patients could have had zero seizures). This would have made detecting a potential correlation between sudden death and seizure frequency more difficult in these studies. This criterion allowed us to include studies that performed multiple regression and that had no range restriction.
2. Compared seizure rates and/or types in persons who experienced SUDEP (cases) to rates and/or types in persons with epilepsy who did not experience SUDEP (controls). Controls could be living patients or patients who died from other causes.
3. Included patients receiving any type of standard treatment for epilepsy (including surgery).

Excluded studies

All of the studies that met our inclusion criteria were included in our analysis. No studies were excluded for reasons of quality.

Evidence base

Nine studies with 8018 patients addressed this question. These studies are listed in Table 47.

Design and conduct of included studies

Internal validity

We evaluated each study's potential for certain biases as discussed in the Methodology section. At least four of nine studies were vulnerable to cause validation bias (in three studies this could not be determined). Seven of nine studies were vulnerable to sampling bias, while seven of nine studies were also vulnerable to statistical control bias. No studies accounted for the effects of all possible confounding variables. More detailed information regarding internal validity is presented in Appendix B. Additional details on design and conduct are provided in Evidence Table 251.

External validity

Knowledge of characteristics of patients in study groups is important for determining the degree of generalizability of a given study. Patient characteristics are shown in Table 47. Eight of nine studies provided at least some information on the seizure types in their respective study groups, although the terminology used to characterize seizures varied somewhat among these studies. Seven of nine studies evaluated some patients with generalized seizures, while five studies evaluated some patients with partial seizures.

The studies that evaluated patients with different seizure types as well as receiving AEDs are probably the most generalizable. By this criterion, the most generalizable studies were conducted by Walczak, Leppik, D'Amelio, et al.[369] and Nilsson, Farahmand, Persson, et al.[374] In addition, these studies provided data concerning both seizure type and seizure frequency. Four additional studies evaluated patients who were receiving AEDs, one study evaluated surgical patients, and two studies did not report treatment information.

Synthesis of study results

Sudden unexpected death and seizure frequency

To address this question, we first looked for any evidence in the literature that suggested a correlation between SUDEP and seizure frequency. If a correlation was found in any study, we then looked for evidence that other variables were correlated with SUDEP. If so, we asked whether their effects were adjusted for in a multiple regression. Eight studies reported information concerning seizure frequency (Table 48).

The two studies that used multiple regression to evaluate the potential relationship between SUDEP and seizure frequency are shown in Evidence Table 252. Figure 85 shows the studies that presented an odds ratio (or allowed independent calculation of an odds ratio or relative risk) for the risk of SUDEP with increasing seizure frequency. Walczak, Leppik, D'Amelio, et al.[369] adjusted for the potential influence of the frequency of tonic-clonic seizures and the number of AEDs used. After these adjustments, the odds ratio for the relationship between SUDEP and overall seizure frequency was reduced to a statistically nonsignificant level (OR 1.1, CI: 0.3-4.0), while the frequency of tonic-clonic seizures remained statistically significant. However, the authors did not adjust for duration of epilepsy or low IQ, two variables that also showed a significant association with SUDEP in linear regression models. Nilsson, Farahmand, Persson, et al.[374] adjusted for the potential effects of epilepsy type, age at epilepsy onset, number of AEDs, and changes in AED dose per year. The relative risk for SUDEP with increasing seizure

frequency was still statistically significant (Evidence Table 252). A univariate analysis stratified by gender suggested that the increased risk was higher among males. However, the authors did not evaluate the potential effect of frequency of tonic-clonic seizures on SUDEP.

The data reported by the studies that did not statistically adjust for differences between patients tended to find no statistically significant relationship between seizure frequency and SUDEP (Evidence Table 253). This was true in the four studies that performed statistical calculations as well as the two studies where we performed independent calculations. However, one study showed a statistically significant difference in seizure frequency between SUDEP patients and living patients with epilepsy, and two other studies showed trends toward higher seizure frequency among SUDEP cases that were not statistically significant. We mention this because these were mostly small studies that had low statistical power (meaning that the effect size may have been statistically significant with a larger study group). In particular, the study by Sperling, Feldman, Kinman, et al.[360] would have shown a statistically significant odds ratio if the study had been slightly larger (the odds ratio for seizures vs. no seizures was 13.76; the study had enough power to detect a minimal difference of 15.1). However, even if a statistically significant relationship was present, whether this would remain if the authors had adjusted for the effects of other variables cannot be determined. The available data are insufficient to provide strong support for a relationship between seizure frequency and SUDEP. However, enough data suggests such a correlation that it cannot be ruled out at this time.

Sudden unexpected death and seizure type

The results of the two studies that used multiple regression to evaluate the potential correlation between SUDEP and seizure type are presented in Evidence Table 254.[369,374] Walczak, Leppik, D'Amelio, et al.[369] presented odds ratios from a multiple regression model that adjusted for the potential effects of overall seizure frequency and number of AEDs. The adjusted odds ratio for SUDEP with increasing frequency of tonic-clonic seizures was statistically significant (OR 7.0, CI: 2.0-24.2), suggesting that an increased frequency of tonic-clonic seizures was associated with an increased risk of sudden death. A univariate analysis stratified by gender suggested that the increased risk was most pronounced among females. A potential weakness of this study was that half of the SUDEP cases were not diagnosed by autopsy. Furthermore, the authors did not adjust for the effect of duration of epilepsy or low IQ, which also showed a significant association with SUDEP in linear regression models.

Nilsson, Farahmand, Persson, et al.[374] presented relative risks for epilepsy type in SUDEP cases vs. controls from a multiple regression model that adjusted for several other variables (seizure frequency, age at epilepsy onset, number of AEDs, and changes in AED dose per year). They did not analyze tonic-clonic seizures as a separate group, but did divide seizure type into generalized idiopathic, partial symptomatic, and partial cryptogenic. These authors found no increased risk of SUDEP for any of these seizure types.

The remaining studies did not adjust for the effects of possible confounding variables (Evidence Table 255). Although the results are less reliable compared those of the above studies, they are presented as additional lower level evidence that may support the results of the higher quality studies.

Two of these studies supported the results of Walczak, Leppik, D'Amelio, et al.[369] Sperling, Feldman, Kinman, et al.[360] reported frequency of tonic-clonic seizures among sudden death cases and controls, although no statistical analysis was performed. Our calculation of odds ratios showed that the presence of tonic-clonic seizures had a statistically significant association with sudden death. However, there was also a trend (though not statistically significant) toward higher

seizure frequency among SUDEP cases. In the absence of multiple regression, which of these factors had the strongest relationship cannot be determined. The relative odds ratios for this study and Walczak, Leppik, D'Amelio, et al.[369] are shown in Figure 86. Timmings[375] reported a statistically significant relationship between SUDEP and idiopathic generalized tonic-clonic seizures (chi-square, $p < 0.05$), while no statistically significant relationship was found between duration of epilepsy or seizure frequency.

The remaining four studies tended to report seizure types as generalized and partial, with occasional subdivisions of these two categories. One study reported p-values and two studies did not perform any statistical analysis. We performed independent calculations of odds ratios for each study. A statistically significant odds ratio was found in only one study.[376] It suggested an association between generalized cryptogenic/symptomatic seizures and sudden death, but there was insufficient data to allow adjustment for the effects of potential confounding variables. Two other studies showed a moderate but nonsignificant trend toward generalized seizures (primary and/or secondary) among SUDEP cases. Figure 87 shows studies that reported odds ratios (or that allowed independent calculation of odds ratios) for generalized seizures and SUDEP.

Table 47. External validity in studies of mortality related to seizure type and frequency

Reference	Country	Number of Patients	Mean Age (Range)	Percent Female	Seizure Types in Study Group	Treatment
Studies conducted in the United States						
Walczak (2001)[369]	United States	4578	Not reported (but contained children and adults up to age 80+)	SUDEP: 60 Controls: 45	Generalized tonic-clonic, others not described	AEDs
McKee (2000)[377]	United States	180	20.3	46	Not reported	AEDs
Sperling (1999)[360]	United States	393	32.7	48.1	Tonic-clonic, complex partial	Surgery
Jick (1992)[378]	United States	3280 (only 31 were relevant to this question)	(15-49)	SUDEP: 36 Controls: 45	Primary generalized, primary partial, unknown	AEDs
Birnbach (1991)[379]	United States	108	SUDEP: 31.6 Non-SUDEP: 28.7 Living: 29.4	SUDEP: 24 Non-SUDEP: 34.8 Living: 29.4	Generalized convulsive, others not described	Not reported
Studies conducted in the United Kingdom						
Nilsson (1999)[374]	United Kingdom	228	SUDEP: 44 Non-SUDEP: 44.7	SUDEP: 40.4 Non-SUDEP: 40.4	Generalized idiopathic, partial symptomatic, partial cryptogenic, undetermined	AEDs
Nashef (1995)[376]	United Kingdom	601	32.5 (10-80)	Not reported	Partial cryptogenic/ symptomatic, generalized idiopathic, generalized cryptogenic/ symptomatic, undetermined	Not reported

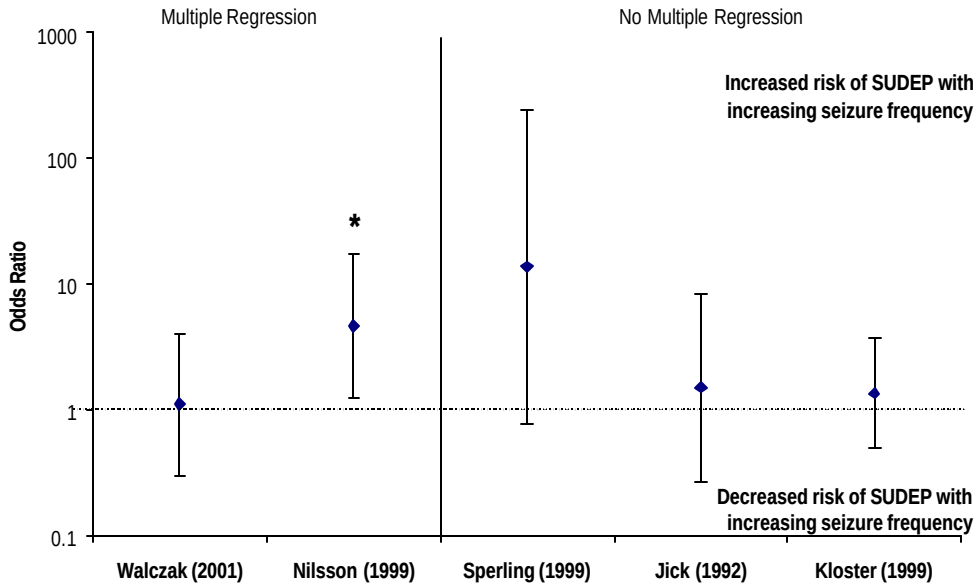
Table 47. External validity in studies of mortality related to seizure type and frequency (continued)

Reference	Country	Number of Patients	Mean Age (Range)	Percent Female	Seizure Types in Study Group	Treatment
Timmings (1993)[375]	United Kingdom	1820	SUDEP: 35 (20-69)	SUDEP: 35.7 Controls: 53	Idiopathic generalized tonic clonic, partial seizures (with or without secondary generalization)	AEDs
Studies conducted in other countries						
Kloster (1999)[380]	Norway	79	SUDEP: 27.9 Non-SUDEP: 32.6	SUDEP: 38 Non-SUDEP: 49	Generalized motor seizures, partial seizures	AEDs

Table 48. Reporting of seizure type and seizure frequency in studies of mortality

Reference	Seizure Type	Seizure Frequency
Walczak (2001)[369]	✓	✓
McKee (2000)[377]		✓
Kloster (1999)[380]	✓	✓
Nilsson (1999)[374]	✓	✓
Sperling (1999)[360]	✓	✓
Nashef (1995)[376]	✓	
Timmings (1993)[375]	✓	✓
Jick (1992)[378]	✓	✓
Birnbach (1991)[379]	✓	✓

Figure 85. Risk of SUDEP with increasing seizure frequency



*The study by Nilsson, Farahmand, Persson et al.[374] reported relative risks rather than odds ratios.

Figure 86. Risk of SUDEP in patients with tonic-clonic seizures

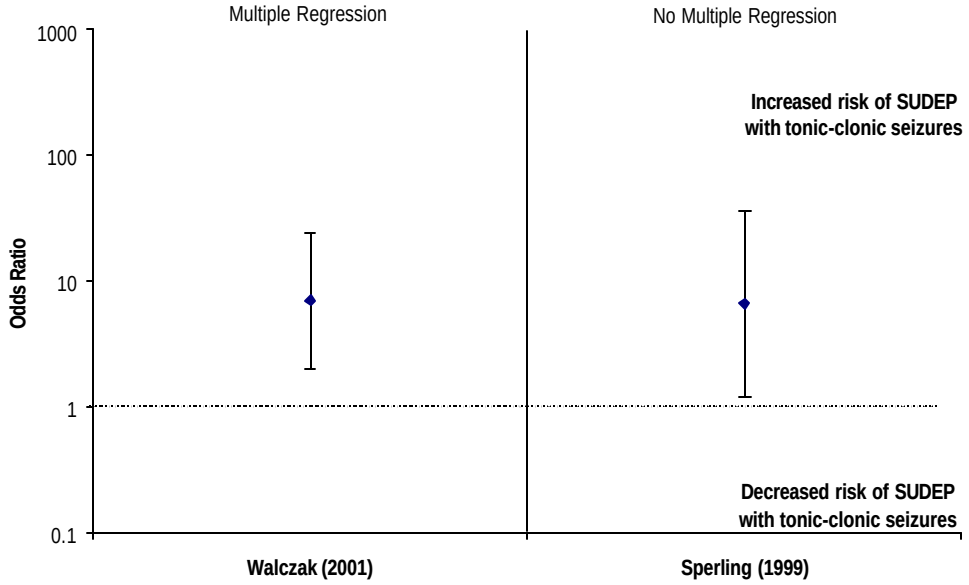
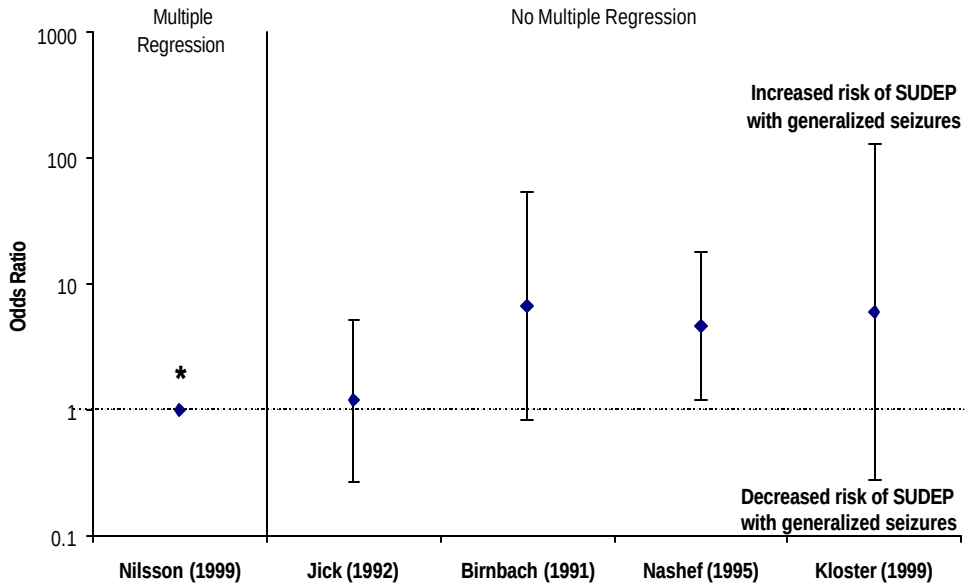


Figure 87. Risk of SUDEP in patients with generalized seizures (primary and/or secondary)



*The study by Nilsson, Farahmand, Persson et al.[374] reported relative risks rather than odds ratios.