Epilepsia Open

Not all that glitters is gold: A guide to the critical interpretation of drug trials in epilepsy

*†Emilio Perucca and ‡§Samuel Wiebe

SUMMARY

Clinical trials represent the best source of evidence on which to base treatment decisions. For such evidence to be utilized meaningfully, however, it is essential that results are interpreted correctly. This requires a good understanding of strengths and weaknesses of the adopted design, the clinical relevance of the outcome measures, and the many factors that could affect such outcomes. As a general rule, uncontrolled studies tend to provide misleading evidence as a result of the impact of confounders such as regression to the mean, patient-related bias, and observer bias. On the other hand, although randomized controlled trials (RCTs) are qualitatively superior, aspects of their execution may still decrease their validity. Bias and decreased validity in RCTs may occur by chance alone (for example, treatment groups may not necessarily be balanced for important variables despite randomization) or because of specific features of the trial design. In the case of industry-driven studies, bias often influences the outcome in favor of the sponsor's product. Factors that need to be carefully scrutinized include (1) the purpose for which the trial is conducted; (2) potential bias due to unblinding or lack of blinding; (3) the appropriateness of the control group; (4) the power of the study in detecting clinically relevant differences; (5) the extent to which eligibility criteria could affect outcomes and be representative of routine clinical practice; (6) whether the treatments being compared are used optimally in terms of dosing, duration of treatment, and other variables; (7) the appropriateness of the statistical comparisons; (8) the clinical relevance of the outcome measures and whether all key outcome information is reported (for example, responder rates in completers); and (9) potential bias in the way results are presented and discussed. This article discusses each of these aspects and illustrates the discussion with examples taken from published antiepileptic drug trials.
KEY WORDS: Clinical trials, Randomized trials, Epilepsy, Antiepileptic drugs, Interpretation.

Prof. Emilio Perucca
is President of ILAE
and Scientific Director,
C Mondino National
Neurological Institute,
Pavia.

Following the gradual introduction of second-generation antiepileptic drugs (AEDs) over the last 25 years, the pharmacological armamentarium against epilepsy now includes more than 30 different medications.[1] This is a welcome development, because it provides unprecedented opportunities to tailor treatment choices to the characteristics and needs of the individual. On the other hand, the availability of so many medications, most of which have overlapping indications, challenges the skills of the busy physician, who must be able to select the most appropriate treatment based on sound evidence about the comparative value of each therapeutic option.

Whereas all relevant information, including results from observational studies, case reports, and personal clinical

KEY POINTS

- Uncontrolled trials typically provide misleading estimates of efficacy because of the influence of regression to the mean and other confounders
- Randomized controlled trials (RCTs) are the best source of evidence on which to base treatment decisions but must be interpreted critically
- The validity of RCTs can be decreased by weaknesses and bias in the study design, which often favor the sponsor's product

experience, should be taken into consideration when selecting an AED, there is no doubt that the best source of evidence is the randomized controlled trial (RCT). For such evidence to be optimally utilized, however, it is essential to interpret trial results correctly with regard to their validity and applicability. This requires a good understanding of the strengths and weaknesses of the adopted methodology, including the many potential bias that could affect the findings and, not uncommonly, invalidate the study's conclusions as reported in the final publication. The purpose of the present article is to highlight key issues that need to be considered when interpreting the results of AED trials.

ASSESSING THE RISK OF BIAS AND APPLICABILITY OF STUDY RESULTS

Clinicians seeking to apply the results of studies about therapeutic interventions in clinical practice need to assess three key aspects: (1) how serious is the risk of bias (i.e., are there systematic flaws that favor one group over the other)? (2) What are the results (i.e., can we make sense of them)? and (3) Can we apply the results to patient care?[2] Table 1 presents a simple checklist to assess these three generic aspects of study validity and usability that can be applied to clinical studies about AEDs.

Following these generic measures of validity and applicability of clinical trial results, we touch on methodological aspects that are often problematic in trials of AEDs and use examples from the literature to illustrate these problems.

UNCONTROLLED STUDIES

Uncontrolled trials represent the overwhelming majority of AED studies. For example, a recent systematic review showed that, of all studies evaluating AEDs as initial treatment in patients with juvenile myoclonic epilepsy, only one had a randomized design.[3] Likewise, of 32 studies of levetiracetam monotherapy in children identified by a 2015 systematic review, only 4 were randomized, 18 were

uncontrolled, and 10 were case reports.[4] Although most clinicians recognize that uncontrolled studies are not optimal for evaluating the effect of a therapeutic intervention, not all appreciate the extent to which uncontrolled studies can lead to misleading conclusions. Evidence from many fields of medicine indicates that uncontrolled studies can grossly overestimate treatment benefits as compared with RCTs. For example, in an early review of trials performed in patients with acute myocardial infarction, favorable treatment effects were found in 56% of nonrandomized studies but in only 30% of blinded RCTs.[5]

A common misperception when interpreting results of uncontrolled studies of AEDs is that any improvement in seizure frequency recorded after introducing a therapeutic intervention (e.g., the addition of a new AED) is largely related to the effect of the treatment. In reality, many other factors could explain such improvement, and in many situations the therapeutic effect of the drug (if any) is quantitatively the least important among them. This is illustrated by a systematic review of all clinical studies using lamotrigine in patients with Lennox-Gastaut syndrome up to 1997, when the only placebo-controlled RCT of lamotrigine for Lennox-Gastaut was published (Table 2). In uncontrolled studies,[6–12] 70% of patients were considered to be "responders," defined as having at least a 50% reduction in seizure frequency compared with baseline. In the RCT, however, the proportion of responders in the group randomized to lamotrigine was only 33%, compared with 16% in the group randomized to placebo.[13] Thus, the actual gain in responder rate associated with active treatment versus placebo was actually 17% (33% minus 16%). Although statistically significant, this was far less impressive than the apparent "response" reported in uncontrolled studies.

Uncontrolled studies provide misleading estimates not only for efficacy but also for adverse effects. In a recent meta-analysis of RCTs in focal epilepsy, 60.3% of patients allocated to placebo groups reported treatment-emergent adverse events, and in 3.9% of placebo-treated patients those adverse events were so severe as to cause withdrawal from the trial.[14]

Where does the "placebo response" in epilepsy come from?

In addition to drug effects, many factors contribute to improvement in seizure frequency after a treatment is started or changed. These include emotional and psychological influences on seizure susceptibility, patient-related bias (natural tendency to please caregivers), and observers' bias, that is, the unconscious tendency to find what one expects or hopes for.[15]

For patients with chronic uncontrolled epilepsy, probably the most important determinant of improvement in seizure frequency is regression to the mean.[16–18] Seizures are unpredictable events with considerable fluctuations over time. Patients going through a period of seizure frequency

Table 1. Assessing the validity of antiepileptic drug trials and the applicability of the results	
	Comments
(1) What is the trial about?	
(a) What are the PICO elements of the trial?	The acronym *PICO* stands for P = patients, I = interventions, C = comparison, O = outcomes. This basic information will help us determine whether the patients in the trial are similar to our patients, whether the interventions under comparison are of interest and feasible in our setting, whether the comparator used is adequate, and whether the outcome chosen for the analyses is clinically relevant
(2) Is the study valid and can we believe the results?	
(a) What are the baseline characteristics of the patients?	Of particular importance is to assess whether patients in the treatment and control groups were similar at the outset with regard to important prognostic factors that could influence the results in favor of one group. Randomization does not guarantee balance
(b) Was concealed randomization used?	This applies only to randomized studies. The one important goal of randomization is to ensure that the treatment and control groups are balanced with respect to factors that may influence outcome. *Concealed randomization* refers to the fact that neither researchers nor patients can guess the sequence of randomization, i.e., they cannot guess which treatment the next patient will receive
(c) Was blinding of study participants appropriate?	The purpose of blinding is to avoid bias in outcome assessment. Patients are blinded to avoid being influenced by knowledge of what treatment they have been allocated to. Clinicians in the trial are blinded to ensure that patients in various groups are treated and assessed similarly. Data collectors and analysts are blinded to avoid bias in their respective roles in the trial
(d) Was follow-up complete and sufficiently long?	Patients who do not finish and those who do finish the trial often have different outcomes. The statistical maneuvers used to deal with patients who do not finish the trial are extremely important in epilepsy. For example, the "last observation carried forward" analysis has important potential for biases. Readers of trials need to be able to judge both the proportion of patients not completing the trial and the assumptions made in the analyses of these incomplete data. Outcomes should be reported separately for all patients and for completers. Also, the shorter the trial, the harder it is to extrapolate to long-term or infrequently occurring outcomes
(e) Was there an intention-to-treat analysis?	This refers to counting and analyzing patients within the group they were randomized to, regardless of whether they received or completed the assigned treatment. The reason is that patients who "cross-over" or do not adhere to the assigned treatment arm may do so for a reason, including having prognostic factors that influence the outcome
(3) What are the results and can they be applied to our patients?	
(a) Is the main result stated clearly and meaningfully?	There should be a clear measure of the main outcome of the trial (e.g., absolute or relative risk reduction, hazard ratio), accompanied by an estimate of precision, such as 95% confidence intervals, not just a p value. This should allow readers to judge whether the magnitude of the effect is clinically important, not just statistically significant, and to understand the uncertainty around the estimates. We should be able to assess net benefits (such as number needed to treat) and have an idea of impact on costs and quality of life
(b) Are the most important clinical outcomes addressed?	If the study fails to address the outcomes that are of importance clinically, the results may be irrelevant regardless of their statistical significance. Also, if the study focuses on surrogate or intermediate outcomes (e.g., improvement in EEG features instead of seizures), or if important results are buried in composite outcomes (the combination of outcomes of varying relevance into a single score), we may not be able to judge actual efficacy. Surrogate and composite outcomes are sometimes used because they are more likely to yield a statistically significant outcome and to contribute to a "positive" trial, not because they are clinically important or relevant

higher than average are more likely to seek medical attention and to be asked (or to meet eligibility criteria) to be included in a drug trial as compared with patients who, at the same time, are experiencing relatively good seizure control. This implies that, irrespective of treatment efficacy, seizure frequency after enrollment in a trial is expected to show a spontaneous amelioration, reflecting a natural tendency to regress toward the frequency that is average for that patient. This also explains why response to active treatment is generally greater in uncontrolled studies than in RCTs (Table 2). Unlike uncontrolled studies, for which baseline seizure frequency is usually assessed retrospectively, RCTs typically involve a prospective run-in period during which the baseline (pretreatment) seizure frequency is established. Consequently, in RCTs regression to the mean generally starts during the prospective baseline, and its impact during the subsequent treatment phase is attenuated (in some trials, this phenomenon may be partly

Table 2. Responder rate (percentage of patients showing at least 50% reduction in seizure frequency compared with baseline) in all the studies of lamotrigine in Lennox-Gastaut syndrome published up to 1997, the year of publication of the first RCT in this indication. Overall responder rate across uncontrolled studies was 70% (56/80)

References	Study design	Number of patients who received lamotrigine	Responder rate (%)
Timmings and Richens[6]	Uncontrolled, open-label	11	91
Schlumberger et al.[7]	Uncontrolled, open label	10	60
Suárez et al.[8]	Uncontrolled, open label	10	90
Buchanan[9]	Uncontrolled, open label	14	57
Donaldson et al.[10]	Uncontrolled, open label	15	53
Farrell et al.[11]	Uncontrolled, open label	15	73
Yen et al.[12]	Uncontrolled, open label	5	80
Motte et al.[13]	Randomized controlled, double-blind	79	33[a]

[a]Responder rate in the group randomized to placebo was 16% (14/90).

counterbalanced by the fact that patients whose seizure frequency falls below a minimal level during baseline are not permitted to enter the treatment phase). Awareness of the phenomenon of regression to the mean helps not only to interpret clinical trials but also to evaluate treatment outcomes in clinical practice. Even a marked improvement in seizure frequency after a treatment change could be explained by spontaneous fluctuation in seizure control.[15]

The magnitude of the placebo response varies depending on many factors, including the year in which the trial was conducted (more recent studies are associated with larger placebo responses), characteristics of the enrolled population (lower placebo responder rates are associated with prior exposure to a high number of AEDs or prior epilepsy surgery, a high baseline seizure frequency, and adult age and older age at diagnosis), type of trial design and statistical analysis, and geographical area in which the trial was conducted.[14,19–27] The mechanisms by which some of these factors affect seizure outcomes are not fully understood.

Could historical clinical data be used to validate the results of uncontrolled studies?

If the magnitude of the placebo response were similar and consistent over time, across trials and interventions, and in all geographic settings, one could argue that there is no need to include a control group in any drug trial, because the outcome observed in the AED group could be compared with the outcome reported for historical controls. Unfortunately, even after controlling for some of the variables discussed in the previous section, responder rates in untreated (or placebo-treated) groups vary markedly and unpredictably from one trial to another. For example, a systematic review of all RCTs conducted in adults with focal epilepsy between 1960 and 2009 found that the proportion of responders among groups assigned to placebo ranged from as small as <5% to close to 40%.[22]

Without an adequate internal control group, it is virtually impossible to determine whether any change in seizure frequency observed after introducing an AED is due to the pharmacological effect of the drug or to the influence of confounders, such as regression to the mean or imbalance of prognostic factors. This concept is illustrated nicely by a comparison of two adjunctive-therapy double-blind RCTs of lamotrigine and topiramate in patients with primarily generalized tonic-clonic seizures (Fig. 1). The two trials had a similar design and were coordinated by the same investigator.[28,29] Although both lamotrigine and topiramate were significantly superior to placebo in reducing generalized tonic-clonic seizure frequency, the responder rate in the placebo groups differed widely. Specifically, the responder rate on placebo (39%) in the lamotrigine trial was almost twice as high as the responder rate on placebo (20%) in the other trial. When calculated over the maintenance period, the responder rate on placebo in the lamotrigine trial was as high as 49%, compared with an even higher responder rate (72%) for lamotrigine. Without an internal placebo control,
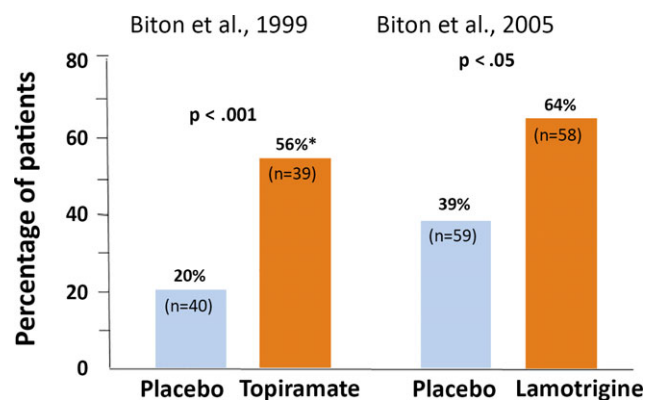
Figure 1.
Responder rates (proportions of patients with at least 50% reduction in primary generalized tonic-clonic seizure frequency compared with baseline) in two placebo-controlled, adjunctive-therapy RCTs of topiramate and lamotrigine in patients with primarily generalized tonic-clonic seizures.[28,29] Despite use of a very similar design in both trials, there was a prominent difference in responder rates in the groups assigned to placebo treatment.
Epilepsia Open © ILAE

one might have concluded that lamotrigine was slightly more efficacious than topiramate—in fact, the actual gain in responder rate over placebo for lamotrigine was considerably less than that reported for topiramate. These data demonstrate that without appropriate internal controls, it may be impossible to draw valid conclusions about the real efficacy of therapeutic interventions.

The choice of an appropriate control group is a most critical consideration when designing a clinical trial. Although placebo provides a highly informative comparator in studies designed to establish the efficacy and tolerability of an investigational compound, the use of placebo in epilepsy trials may raise serious ethical concerns.[26] Specifically, because of the risks associated with uncontrolled seizures, as a general rule the use of placebo as sole treatment (monotherapy) is ethically unacceptable in patients with active seizure disorders. Concerns have also been raised about the risks associated with prolonged placebo treatment in adjunctive-therapy trials. To address these concerns, innovative trial designs that minimize duration of placebo exposure have recently been proposed.[26]

Are uncontrolled studies of any value in specific situations?

A number of considerations are often put forward to justify uncontrolled studies. A commonly used argument is that, compared with RCTs, they reproduce more closely routine clinical conditions and therefore provide information that is more directly applicable to the everyday practice. This argument is wrong on at least two counts. First, it is feasible to design pragmatic RCTs that mimic clinical practice equally well. Second, no objective can justify a study whose results cannot be meaningfully interpreted.[15] Another common proposition is that uncontrolled studies are the only feasible option to investigate therapeutic interventions in rare syndromes for which it would be impossible to enroll a population large enough to conduct an RCT. This argument is incorrect because methodologically sound RCTs of specific types can be conducted in small patient groups,[30] and even in individual patients,[31,32] provided that certain conditions are satisfied.

Because results of uncontrolled studies can be seriously misleading, they represent a much weaker class of evidence. As provocatively stated by David Sackett, the father of evidence-based medicine, "If you find that a study was not randomized, we'd suggest that you stop reading it and go on to the next article."[33] However, this does not mean that uncontrolled trials serve no purpose at all. In some cases, only uncontrolled studies may be available, and clinicians must judiciously use this weaker level of evidence to inform their decisions. In fact, uncontrolled studies provide useful information in many situations. They can be valuable to identify risk factors and prognostic indicators when RCTs are unfeasible or unethical. In addition, if carefully designed and executed, uncontrolled observational studies can yield estimates of treatment effects that are similar to those derived from RCTs.[34] Long-term observational studies are also valuable to assess the course of illness and to identify rare or delayed adverse effects. Uncontrolled studies can also be of value in the early phases of development of potential AEDs to evaluate pharmacokinetics and drug interactions and to obtain preliminary estimates of tolerability and potential efficacy in specific seizure types and syndromes. The primary purpose of the latter studies, however, is to identify signals and to generate hypotheses that need to be confirmed in appropriately designed RCTs. For example, the early positive findings of uncontrolled studies with lamotrigine in Lennox-Gastaut syndrome (Table 2) provided valuable information that led to a well-designed placebo-controlled RCT that confirmed the efficacy of the drug.[6] Importantly, preliminary positive signals generated by uncontrolled studies are not always confirmed when tested in RCTs. For example, cinromide also yielded promising early results in patients with Lennox-Gastaut syndrome, but it was no different from placebo when tested in a large, well-designed RCT.[35]

Uncontrolled studies as a tool for drug promotion

Confounding and imbalance in prognostic variables (Table 1) account for a large proportion of the improvement in seizure frequency typically observed in uncontrolled studies in chronic refractory epilepsy (Table 2). A common statement among biostatisticians is that the best way to improve the outcome of a therapeutic trial is to leave out controls.

These considerations explain why uncontrolled trials have been widely used as a tool for drug promotion, particularly for the implementation of seeding trials. The latter are defined as postmarketing studies of little or no scientific value that are intended to fulfill one or more of the following objectives: (1) generate inflated "efficacy" estimates of the therapeutic value of a specific product; (2) familiarize clinicians with its use; (3) increase long-term prescriptions through short-term enrollment of patients in the "study"; and (4) provide misleading efficacy "evidence" to support use of the product for nonapproved indications. The improper utilization of clinical trials as marketing tools is extensively discussed in a landmark open-access article by Kessler et al.,[36] which every clinician should read.

RANDOMIZED CONTROLLED TRIALS

RCTs are widely recognized as the best tool to assess the comparative effectiveness of therapeutic interventions. However, these trials may vary substantially in their methodological quality and their applicability to clinical practice, according to current standards (Table 1). For example, in a recent systematic review[3] of 34 RCTs comparing the effectiveness of AEDs in adults with newly diagnosed focal epilepsy, only 4 were rated as Class 1 (the

Table 3. Advantages and disadvantages of regulatory randomized controlled trials in epilepsy	
Advantages	Disadvantages
Typically double-blind, which minimizes probability of results being biased	Question being addressed often differs from what clinicians need to know (see text)
Inclusion of a placebo control (commonly included in add-on trials) permits unequivocal interpretation of efficacy and tolerability findings	Strict exclusion criteria typically result in a trial population that is poorly representative of routine clinical practice
Well-standardized methodology, based on high scientific standards	Dosing regimens do not usually allow the flexibility required to achieve optimal outcomes
	The number of patients is relatively small and inadequate to identify uncommon but potentially important adverse effects
	Duration of treatment is generally short, which may not allow detection of chronic or delayed adverse effects

highest quality rating), 1 was rated as Class 2 (intermediate quality), and 29 were rated as Class 3 (the lowest rating). Of 19 RCTs in children with focal seizures, only 1 was rated as Class 1, and 18 were rated as Class 3. All 19 RCTs in adults with primarily generalized tonic-clonic seizures were rated as Class 3. Common methodological flaws included inadequate power, suboptimal trial duration, and choice of an inadequate comparator.

Although randomization is essential to reduce the risk of bias, randomization alone, even if appropriately done, is not the only factor required to avoid the risk of bias (Table 1).[15,37,38] Bias in RCTs may occur by chance (for example, despite randomization, treatment groups may not necessarily be balanced for important variables) or because specific features of the trial design, in the case of industry-driven studies, influence the outcome in favor of the sponsor's product. Factors that need to be considered include (1) the purpose for which the trial is conducted; (2) whether key aspects related to trial design, execution, and analysis are methodologically sound and adequate to achieve the stated objective; and (3) whether the results are described and interpreted correctly (Table 1). Relevant aspects of this process are discussed in the sections below.

What was the purpose for conducting the trial?

Because RCTs are a prerequisite to obtain evidence for marketing approval, it is no surprise that most RCTs of AEDs have been conducted for regulatory purposes. According to current regulations, marketing approval requires demonstration that a product is efficacious and safe. With respect to efficacy, it is sufficient to show that the product is "better than nothing," that is, that it is superior to placebo or to a suboptimally used active treatment in reducing seizure frequency.[26] Therefore, the question asked in these trials is of modest relevance to practicing clinicians whose primary interest is how a new drug compares with previously established treatment options in terms of efficacy and tolerability, not whether a new drug is better than nothing. Unfortunately, regulatory trials generally do not

provide this information (Table 3). The only exceptions are a few trials designed to obtain a monotherapy indication in Europe. These trials required a comparison with an optimally used active comparator, although they have also been criticized because of concerns with assay sensitivity.[26,39,40] Specifically, all the latter trials used a noninferiority design, and the finding of noninferiority (or equivalence) cannot exclude the possibility that both treatments might have been equally ineffective in the specific population and under the conditions in which the studies were done.[39,41] Other limitations of regulatory trials are the inclusion of highly selected populations, a short duration of assessment in a disease that is chronic and fluctuating, and the use of predetermined, nonflexible dosing schemes. This limits considerably the applicability of study results to routine clinical practice.[42]

In contrast with regulatory trials, nonregulatory RCTs are usually conducted to assess the comparative effectiveness of different treatment options, and, in this respect, their results may be of greater relevance for clinicians. Many large-scale double-blind RCTs funded by nonprofit organizations fall within this category[43–46] and provide highly valuable evidence that can be applied to treatment decisions. The fact that a study was supported by a nonprofit organization, however, does not necessarily imply that it is unbiased or that it does not have weaknesses in terms of generalizability or external validity. Clinicians should also be aware that some postmarketing nonregulatory RCTs have been conducted with the nondisclosed objective of generating supportive material for drug promotion.[15,36,38,47] Not surprisingly, the design of such studies often tends to favor the sponsor's product, and clinicians are advised to apply the principles listed in Table 1 and discussed in these sections to assess the validity and applicability of all AED trials.

The sections below provide examples of how even apparently minor aspects of study design or implementation can influence outcomes to a clinically important extent.

Could selection of the study population have biased the results?

Close scrutiny of the study populations is essential for a correct interpretation of the results of a clinical trial. One important consideration is whether there are any imbalances between treatment arms for variables that could have influenced the outcomes. As noted in Table 1, randomization is the best defense against such imbalance, but imbalances can still occur by chance.

A second consideration is whether selection bias, which can occur even if treatment arms are balanced, limits the applicability of the results. In particular, regulatory trials typically have a long list of exclusion criteria that are intended to select a homogeneous population and thereby minimize the probability of confounders obscuring treatment effects. However, this can result in trials whose population is not representative of that encountered in everyday practice.[15] This could limit the applicability of the trial results in several ways, and clinicians are well advised to carefully assess the characteristics of the patients included and the list of exclusion criteria. For example, if one of the two treatments being compared has a high propensity to produce adverse psychiatric effects in patients with a history of behavioral disturbances, exclusion of patients with a positive psychiatric history could prevent identification of an important difference in tolerability. In fact, patients at special risk such as those with comorbidities, those with a history of severe drug reactions, those who are elderly, and women of childbearing age are typically underrepresented or even excluded in RCTs.

At the other end of the spectrum, excessive heterogeneity can invalidate the primary outcome of a trial. A typical example is provided by the many RCTs in newly diagnosed epilepsy that enrolled a mixed population of patients with focal seizures and primarily generalized tonic-clonic seizures and that used as primary outcome the proportion of the pooled population of patients who remain on treatment.[3,45] These trials tend to bias the results in favor of broad-spectrum AEDs compared with AEDs that are preferentially efficacious against focal seizures.

When the size of one of the subgroups is small, however, results can lead to the misleading conclusion that the assessed treatments have comparable efficacy against both seizure types when in fact the reverse is true. This was the case for a double-blind RCT comparing lamotrigine and gabapentin in a mixed population of patients with focal and generalized epilepsy that concluded that the two drugs were "similarly effective…in terms of seizure control and tolerability in patients with partial seizures with or without secondary generalization or primary generalised tonic-clonic seizures."[48] In fact, of the 309 patients included in the study, only 58 had primarily generalized seizures. Of those, none of the 27 patients randomized to lamotrigine met exit criteria (failed therapy), whereas 5 out of 31 in the gabapentin group exited, a clear signal that the two treatments might

not have been "similarly effective" in this subpopulation. The trial was underpowered to detect potentially important differences in outcomes between patient subgroups.

Apparently minor aspects in inclusion criteria could substantially influence study results. For example, the landmark RCTs conducted by the US Veterans Administration (VA) collaborative networks to assess the comparative effectiveness of different AED monotherapies enrolled not only patients with previously untreated epilepsy but also patients who were previously "undertreated," that is, exposed to "subtherapeutic AED doses and blood levels."[43–45] However, it is now known that newly diagnosed patients often achieve seizure freedom at low AED doses[49] or at serum AED levels well below the reference range.[50] Inclusion of previously "undertreated" patients, therefore, may have weighted the patient population toward those who had failed to respond to a potentially effective dose of one or more of the treatments being compared. In the trial that compared carbamazepine with valproate, 26% of patients were receiving an AED at the time of enrollment, but the proportion of those who had been exposed to one of the study drugs was not reported.[44] In the RCT that compared lamotrigine, gabapentin, and carbamazepine in epilepsy with onset in old age, 36.9% of the enrolled population had not responded to low doses or serum levels of phenytoin, a drug with a mechanism of action very similar to carbamazepine, potentially biasing the results in favor of the other AEDs included in the comparison.[46]

These examples illustrate the importance of carefully scrutinizing the population under study, and the interpretation of the results, to determine the risk of bias and whether the results are applicable in practice. Such scrutiny should include a careful review of the criteria used to determine eligibility of patients and to ensure that study participants do have epilepsy and their seizure types are classified correctly. Comparison of outcomes becomes meaningless if the correct diagnosis of study participants is in doubt.

Was treatment in the reference (control) group chosen and used appropriately?

The most important information that clinicians require is how a new treatment compares with the gold standard used for the same indication. In some comparative monotherapy trials, however, the choice of the reference (control) treatment has been questionable. For example, in the only large double-blind RCT that evaluated the efficacy and tolerability of oxcarbazepine in children with newly diagnosed (mostly) focal epilepsy,[51] the AED selected for comparison was phenytoin, which is usually not a first choice in children because of side effects, including hirsutism and gum hyperplasia. Not surprisingly, oxcarbazepine showed superior tolerability in that trial, but a comparison with carbamazepine would have been more meaningful clinically.

Even if the most appropriate comparator was selected, it is essential to determine whether the treatments being

compared were used optimally in terms of dose titration, dosing frequency, and individualization of dose.[37] Even minor differences in these parameters can affect clinical outcomes.[47] For example, in an early RCT that compared the effectiveness of carbamazepine and vigabatrin in newly diagnosed epilepsy, carbamazepine could be adjusted on the basis of clinical response up to the highest tolerated dose, whereas the dose of vigabatrin could not exceed a predetermined ceiling irrespective of seizure control and adverse effects.[52] Although carbamazepine was an appropriate comparator, adopting different criteria for optimizing the dose could have biased the study results, which favored vigabatrin with regard to tolerability and carbamazepine with regard to seizure freedom.

Even the type of formulation selected for AED trials can be important.[53] Most monotherapy trials of new AEDs have used as comparator an immediate-release formulation of carbamazepine given on a twice daily dosing schedule,[3,45] which is clearly not an optimal regimen, particularly for patients who require relatively high doses.[54] The suboptimal use of carbamazepine in these trials is likely to have biased outcomes in favor of the sponsor's product.[15] A possible example is provided by the results of two double-blind RCTs that compared lamotrigine with carbamazepine in patients with older-onset epilepsy (Fig. 2).[55,56] Both trials used a very similar design and dosing schedules, but the formulation of carbamazepine differed. Whereas retention in the trial (primary outcome) in the lamotrigine arm was almost identical in the two trials, for carbamazepine the outcome was markedly better when the controlled-release formulation was used, mainly because of improved tolerability. Although other factors might have contributed to the differences in outcome, the findings are consistent with evidence that, when used on a twice daily schedule, controlled-release carbamazepine is better tolerated than the immediate-release dose form.[54]

In some situations, no meaningful interpretation of data is possible without inclusion of a placebo control. This is the case for adjunctive-therapy AED trials, where the investigational treatment is generally not expected to show greater efficacy than an optimally chosen active comparator.[26] According to international guidelines, an equivalence or noninferiority design can be applied meaningfully only when evidence exists that, under the specified study conditions, effective treatments can be consistently differentiated from less effective or ineffective treatments, and sufficient data exist to allow an estimate across studies of the magnitude of difference in outcome between the reference treatment and placebo.[39] This is clearly not the case for adjunctive-therapy AED trials. For example, although most placebo-controlled studies have shown that add-on levetiracetam is superior to placebo in reducing seizure frequency in patients with focal epilepsy,[57] in some instances the drug could not be differentiated from placebo in this indication, despite use of a standard trial design.[58] This
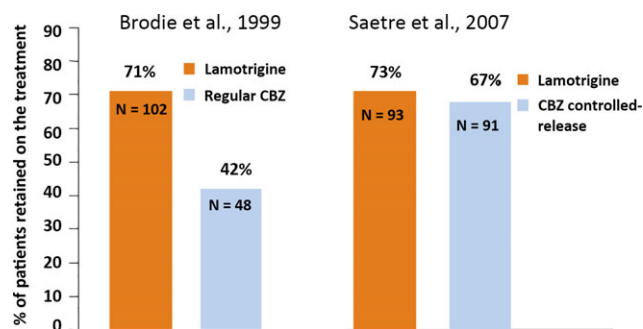
Figure 2.
Retention in the trial (a combined measure of efficacy and tolerability) in two double-blind RCTs comparing the outcome of treatment with lamotrigine and carbamazepine (CBZ) monotherapy in patients aged 65 years and over with newly diagnosed epilepsy with onset in old age.[55,56] Outcome with lamotrigine was similar in the two trials, whereas outcome on carbamazepine was better in the trial that used the controlled-release formulation. Both trials enrolled very similar populations and used identical dosing schemes. Duration of follow-up was longer for the trial that used controlled-release carbamazepine (40 vs. 24 weeks). Reproduced from Perucca[53] with permission.
Epilepsia Open © ILAE

observation can undermine the conclusions of a recent large double-blind RCT in which improvements in seizure frequency in patients with focal epilepsy were not different with pregabalin or levetiracetam.[59] In the absence of a placebo control, the possibility exists that the two AEDs were equally ineffective, given the trial design.

Including a placebo arm is also essential when interventions are tested for indications for which no established efficacious treatments exist. Failure to recognize this concept could lead to faulty trial designs and consequent misinterpretation of the results. For example, no AED has ever been demonstrated to be effective in preventing late unprovoked seizures when used prophylactically after craniotomy.[60] Yet, there have been studies in which different AEDs were compared in this indication, and lack of differences in seizure outcome was interpreted as a demonstration that both treatments were equally efficacious.[61] In fact, both treatments might have been equally nonefficacious in those trials.

Was the duration of assessment adequate?

The duration of assessment has a critical impact on the ability of a trial to detect differences in efficacy and tolerability between AEDs. Short-term assessments cannot identify differences in the occurrence of chronic, late-onset adverse effects, and they can also lead to misleading estimates of effectiveness. For example, in the RCT comparing the effectiveness of gabapentin and lamotrigine monotherapy,[48] eligibility criteria allowed enrollment of patients who had as few as one seizure over the previous 12 months, and the duration of assessment was limited to only 30 weeks,

including the titration period and time allowed to adjust dose if needed. Although retention in the trial was similar for the two drugs, the duration of follow-up was insufficient to detect potential differences in effectiveness between treatments.

Many AED trials suffer from similar shortcomings in duration of assessment. In a recent multicenter RCT comparing levetiracetam with lamotrigine in patients with newly diagnosed focal or generalized epilepsy, defined by either two or more unprovoked seizures or one first seizure with high risk for recurrence, no difference in seizure outcomes between the two treatments was identified, but follow-up treatment was limited to only 26 weeks.[62] Interestingly, the primary endpoint in the latter trial was the proportion of seizure-free patients in the first 6 weeks, which was intended to show superiority of levetiracetam (uptitrated over 22 days) over lamotrigine (uptitrated over 71 days). Presumably, baseline seizure frequency in most patients included in the trial was too low for treatment effects to be differentiated, and possibly even detected, over a period as short as 6 weeks.

Insufficient duration of assessment may also affect the validity of trials reporting superiority of one treatment over another. In a double-blind RCT involving 660 patients with newly diagnosed focal seizures, 6-month seizure freedom rates (the primary endpoint) were significantly higher for patients randomized to lamotrigine than for those randomized to pregabalin.[63] As correctly pointed out by the authors, however, the efficacy of pregabalin in this trial might have been underestimated as a result of selection of a suboptimally low initial maintenance dose coupled with a duration of follow-up that was insufficient to assess seizure outcomes following optimization of dose.

The statistical power to detect an event is related to the duration of assessment. A useful statistical concept pertaining to duration of assessment is the "rule of 3." This refers to the fact that the probability of observing an event with 95% certainty requires an observation period three times longer than the usual interval between events. For example, to be 95% certain of observing a seizure in a patient who has one seizure every 6 months on average would require a minimum period of assessment of 18 months. This concept, used by the International League Against Epilepsy (ILAE) to establish the minimum period of observation required in individual patients to determine whether epilepsy is drug-resistant,[64] is not formally used in the design or analysis of RCTs of AEDs.

Is the trial adequately powered?

The power of a study to detect small but clinically important differences is directly related to the sample size and inversely related to the variability of the outcome of interest. Although variability of outcomes among patients in an RCT cannot be controlled, sample size is under the direct control of the researcher. The smaller the sample size, the higher the chance of a false-negative trial. The majority of comparative monotherapy AED trials conducted to date failed to identify differences in efficacy or effectiveness between treatment arms. However, systematic reviews of these trials have demonstrated that very few enrolled a number of patients sufficient to exclude clinically important differences in outcome.[3,45] For example, sample size for more than half of 19 double-blind RCTs in adults with focal epilepsy was not large enough to exclude that one treatment was >30% less effective or less efficacious than the reference treatment with which it was compared.[45] The proportion of adequately powered studies is even smaller if we consider those that were able to detect outcome differences between treatments that are smaller than 30% but still clinically important. For trials enrolling populations heterogeneous in terms of seizure type or epilepsy syndrome, statistical power is almost invariably insufficient to obtain meaningful comparison among these subgroups.

Open-label versus blinded trials: Does it matter?

Because most seizures are events with outward manifestations that can be observed, there is a common belief that counting them is not greatly affected by patients' or observers' biases. In fact, that assumption has not been adequately verified and experience from other areas of medicine has shown that lack of blinding can introduce major bias in efficacy estimates.[5] Additionally, it is important to remember that, for a treatment to be effective in practice, it has to be tolerated. Specifically, if patients discontinue treatment prematurely because of an adverse drug reaction, they will never be able to achieve sustained seizure freedom on that treatment.

Patients and doctors must be informed about the specific side effects of the treatments being administered, and such knowledge inevitably influences reporting of adverse experiences. For example, patients started on a drug known to cause life-threatening skin reactions will pay special attention to any cutaneous event, and their doctors may be more likely to discontinue that drug should a minor and potentially transient skin rash occur. If the same trial involves an unblinded comparison with another medication known to be generally devoid of serious skin reactions, the probability of the same event going unrecognized or not acted upon would obviously be greater. Ultimately, these biases significantly influence retention in the trial, sustained seizure freedom rates, and frequency and types of adverse events.

In a double-blind trial, patients and doctors do not know which treatment is being administered, and therefore "fears" about potential adverse effects will not bias outcomes for or against any of the treatments being compared. Current standards recommend that to avoid bias and to maintain balance between groups in clinical management, for assessment of outcomes and interpretation and analyses of data in an RCT five groups of individuals should ideally be blinded: the patients, the clinicians looking after the patients, the data

collectors, the clinicians adjudicating outcomes (for example, to determine seizures or adverse events), and those in charge of data analysis (Table 1).

Is the choice of the primary endpoint appropriate for the purpose of the trial?

The choice of the primary endpoint has a major influence on the outcome of a clinical trial. Ideally, the endpoint should be the one with the greatest clinical significance, but there are many instances in which appropriateness of the choice of primary endpoint can be questioned. For example, in a large multicenter trial designed to investigate the efficacy and safety of lamotrigine monotherapy in comparison with carbamazepine or valproic acid in adolescents and adults with newly diagnosed epilepsy, the primary efficacy endpoint was the percentage of patients seizure-free between treatment weeks 17 and 24.[65] Such a short and seemingly arbitrary period is of little relevance in a population for which the goal of therapy is long-term, sustained seizure freedom. Additionally, the short duration of the trial (24 weeks) was inadequate not only to assess meaningfully seizure outcome but also to allow optimization of dose.

In some situations, endpoints of modest clinical relevance can be justified, depending on the purpose of the trial. In particular, although there is universal agreement that sustained seizure freedom should be the primary objective of AED treatment, in adjunctive-therapy trials only few patients achieve this goal, and therefore demonstration of efficacy has to rely on the use of other endpoints such as percent reduction in seizure frequency or responder rate (proportion of patients with at least 50% reduction in seizure frequency compared with baseline).[26] Although use of the latter endpoints is acceptable for regulatory purposes, it is clear that being a responder according to these criteria does not necessarily imply having a clinically significant benefit. Indeed, many studies have shown that quality of life (QOL) in people with epilepsy is primarily conditional on achieving complete seizure freedom without intolerable adverse effects and is little affected by a simple reduction in the frequency of seizures.[66] This concept is well illustrated by the findings of an RCT of vigabatrin in patients with focal seizures.[67] As illustrated in Fig. 3, an improvement in QOL could be demonstrated only for patients who became seizure free, and even a 75–99% reduction in seizure frequency had little or no impact on QOL.

Another example of an endpoint that is of limited relevance to everyday practice is time to exit according to predetermined exit criteria, as used in conversion-to-monotherapy trials conducted to obtain approval of a monotherapy indication in the U.S.A.[39] In these trials, patients are randomized to be converted to monotherapy over a short period, and they are forced to exit the study when a predetermined number (or type) of seizures occur. The design of these trials has evolved over the years because of safety concerns,[68] but their justification remains
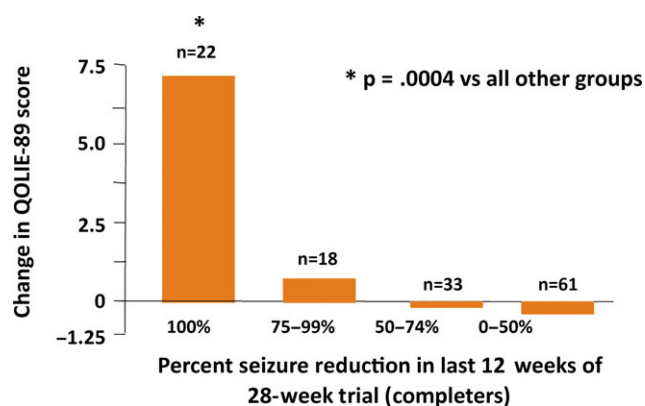
Figure 3.
Relationship between percent seizure reduction (vs. baseline) during the last 12 weeks of a 28-week placebo-controlled adjunctive-therapy trial and mean change in health-related quality of life (QOLIE-89 score).[67] A significant improvement in quality of life was found only for patients who achieved complete freedom from seizures.
*Epilepsia Open* © ILAE

questionable on ethical and scientific grounds and because of the difficulty in extrapolating results to the routine clinical setting.[40]

Is all important information reported? The case of LOCF versus completers analysis

As a rule, changes in seizure frequency or responder rates in adjunctive therapy are calculated by applying the last-observation-carried-forward (LOCF) analysis.[23] In this type of analysis, when patients withdraw prematurely from the trial as a result of adverse events or other reasons, efficacy estimates are calculated by using the seizure outcomes recorded up to the time of exit. For example, if in a 16-week trial a patient exits at 1 week because of intolerable adverse events and no seizures occurred over the preceding 7 days, that patient would be considered seizure-free for the entire duration of the trial, for the purpose of the final analysis.[69] This explains why, paradoxically, in some studies the responder rates in one or more treatment arms have been greater than the proportion of patients who were able to complete the trial (Fig. 4).[70] Although there are some justifications for applying a LOCF analysis (mainly to obtain an estimate of efficacy not confounded by other variables such as failure to tolerate the treatment), every clinician will agree that a much more meaningful estimate of treatment effects is represented by the number of responders who were able to complete the trial.

Remarkably, such essential information is very seldom reported. For example, in a recent 15-week placebo-controlled adjunctive-therapy RCT of clobazam in Lennox-Gastaut syndrome, an impressive 24.5% of patients allocated to the highest dose were reported to have become completely free from drop attacks.[71] However, no information was given
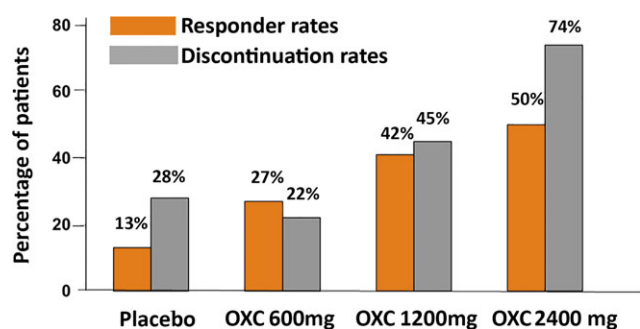
Figure 4.
Responder rates (proportion of patients with >50% decrease in seizure frequency compared with baseline) and rates of premature discontinuation from the trial in a placebo-controlled adjunctive-therapy trial of oxcarbazepine (OXC) in a total of 694 patients with focal seizures.[70] For oxcarbazepine-treated groups, most premature discontinuations were due to adverse events. Because of the application of last-observation-carried-forward analysis, the proportion of responders at the 2,400-mg dose was lower than the proportion of patients who discontinued prematurely owing to adverse events.

Epilepsia Open © ILAE

on whether those patients were seizure-free for a short time before exiting owing to intolerability or they were seizure-free for the entire duration of the trial.

Patients who do not finish and those who do finish a trial often have different outcomes and clinical characteristics, that is, discontinuation does not necessarily occur at random. Furthermore, a recently published open-access systematic review found that up to 30% of positive trials published in high-impact journals would lose statistical significance if patients who were not followed until the completion of the trial had a range of plausible outcomes different from those assumed by researchers in the analyses (Table 1).[72] Importantly, some authors use survival analysis as a method to remedy incomplete follow-up in studies of dichotomous outcomes (e.g., seizure free vs. not seizure free) because, with this methodology, patients are censored at their last visit, that is, they contribute information until their last visit. However, survival analysis is no remedy for a high proportion of patients not completing the study.

A systematic review of all adjunctive-therapy RCTs conducted in adults with focal epilepsy during the period 1967–2009 found that only 3 out of 63 trials (<5%) had reported the proportion of responders who were able to complete the trial.[22] Clearly, LOCF can lead to gross overestimates of the actual efficacy of an AED, and journal editors and reviewers should ensure that responder rates for completers also be reported in publications describing the results of clinical trials. Needless to say, clinicians need to carefully scrutinize these results and be aware that the methodology used for data analysis as well as the incomplete disclosure of trial data have important implications for the interpretation of the results.

Did the trial control for the confounding effect of drug interactions?

In adjunctive-therapy trials, drug-drug interactions can have a profound impact on the results. If an added drug causes a fall in the serum concentration of underlying AEDs, its efficacy could be underestimated because of the attenuated antiseizure effect of the comedications. Conversely, when an added drug increases the concentration of concomitant AEDs, it may be difficult to determine whether any improvement in seizure control is due to an independent antiseizure effect or to an increased effect of the comedications. An example of the interpretative difficulties related to the latter situation is provided by the trials conducted with stiripentol, a drug licensed in Europe for use "in conjunction with clobazam and valproate as adjunctive therapy of refractory generalized tonic-clonic seizures in patients with Dravet syndrome whose seizures are not adequately controlled with clobazam and valproate."[73] In the RCTs that led to regulatory approval, stiripentol was found to cause an approximately two- to 3-fold increase in serum concentration of clobazam and a 5-fold increase in serum concentration of the active clobazam metabolite, norclobazam.[73,74] On the basis of these data, it is difficult to conclude whether stiripentol has efficacy on its own or its therapeutic effects, as well as some of its adverse effects, are related to the increased serum levels of clobazam and its metabolite.

Effects of pharmacodynamic interactions are more difficult to ascertain in clinical trials, but they can also confound the assessment of efficacy and tolerability. For example, a pooled analysis of placebo-controlled adjunctive-therapy RCTs provided evidence that lacosamide shows better antiseizure effects and better tolerability when it is combined with AEDs that act through mechanisms that do not involve blockade of voltage-dependent sodium channels.[75]

CONCLUSION

RCTs provide the most valid and relevant information on the efficacy and safety of existing treatment options. However, for a correct application of this information to clinical practice, it is essential that the results of these studies are interpreted critically. This requires good understanding of the strengths and weaknesses of the applied methodology and careful assessment of potential bias associated with trial design and implementation and with the analysis and reporting of the data. An overall evaluation of available evidence reveals that the quality of most AED trials conducted to date is less than optimal and that with careful scrutiny of the data the conclusions reached by the authors at times are not fully supported. Hopefully, an improved awareness of these issues will assist clinicians not only to apply a critical approach to the interpretation of AED trials but also to become more actively engaged in developing high-quality outcomes research in epilepsy.

Acknowledgments

Conflict Of Interest

References

1. Shorvon SD, Perucca E, Engel J Jr (Eds) *The treatment of epilepsy*. 4th ed. Oxford: Wiley-Blackwell; 2016.
2. Walsh M, Perkovic V, Manns BJ, et al. Therapy (Randomized Trials). In Guyatt G, Rennie D, Meade MO, Cook DJ (Eds) *Users' guides to the medical literature: a manual for evidence-based clinical practice.* 3rd ed. New York: McGraw-Hill, 2015. Available at: http://jama evidence.mhmedical.com.ezproxy.lib.ucalgary.ca/content.aspx?bookid =847&Sectionid=69031464. Accessed February 12, 2016.
3. Glauser T, Ben-Menachem E, Bourgeois B, et al. Updated ILAE evidence review of antiepileptic drug efficacy and effectiveness as initial monotherapy for epileptic seizures and syndromes. *Epilepsia* 2013;54:551–563.
4. Weijenberg A, Brouwer OF, Callenbach PM. Levetiracetam monotherapy in children with epilepsy: a systematic review. *CNS Drugs* 2015;29:371–382.
5. Chalmers TC, Celano P, Sacks HS, et al. Bias in treatment assignment in controlled clinical trials. *N Engl J Med* 1983;309:1358–1361.
6. Timmings PL, Richens A. Lamotrigine as an add-on drug in the management of Lennox-Gastaut syndrome. *Eur Neurol* 1992;32:305–307.
7. Schlumberger E, Chavez F, Palacios L, et al. Lamotrigine in treatment of 120 children with epilepsy. *Epilepsia* 1994;35:359–367.
8. Suárez P, Vadillo J, Suárez C, et al. Adding lamotrigine to the treatment of epilepsy which is difficult to control (article in Spanish). *Rev Neurol* 1995;23:1190–1192.
9. Buchanan N. Lamotrigine: clinical experience in 200 patients with epilepsy with follow-up to four years. *Seizure* 1996;5:209–214.
10. Donaldson JA, Glauser TA, Olberding LS. Lamotrigine adjunctive therapy in childhood epileptic encephalopathy (the Lennox Gastaut syndrome). *Epilepsia* 1997;38:68–73.
11. Farrell K, Connolly MB, Munn R, et al. Prospective, open-label, add-on study of lamotrigine in 56 children with intractable generalized epilepsy. *Pediatr Neurol* 1997;16:201–205.
12. Yen DJ, Yiu CH, Kwan SY, et al. Lamotrigine as add-on therapy in adult patients with refractory epilepsy. *Zhonghua Yi Xue Za Zhi (Taipei)* 1997;59:303–307.
13. Motte J, Trevathan E, Arvidsson JF, et al. Lamotrigine for generalized seizures associated with the Lennox-Gastaut syndrome. Lamictal Lennox-Gastaut Study Group. *N Engl J Med* 1997;337:1807–1812.
14. Zaccara G, Giovannelli F, Schmidt D. Placebo and nocebo responses in drug trials of epilepsy. *Epilepsy Behav* 2015;43:128–134.
15. Perucca E. What can we learn from clinical trials of anticonvulsant drugs in epilepsy? *Eur J Pain* 2002;6(Suppl. A):35–44.
16. Spilker B, Segreti A. Validation of the phenomenon of regression of seizure frequency in epilepsy. *Epilepsia* 1984;25:443–449.
17. Barnett AG, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. *Int J Epidemiol* 2005;34:215–220.
18. Da Costa BR, Juni P. Systematic reviews and meta-analyses of randomized trials: principles and pitfalls. *Eur Heart J* 2014;35:3336–3345.
19. Rheims S, Cucherat M, Arzimanoglou A, et al. Greater response to placebo in children than in adults: a systematic review and meta-analysis in drug-resistant partial epilepsy. *PLoS Med* 2008;5:e166.
20. Baulac M, Leon T, O'Brien TJ, et al. A comparison of pregabalin, lamotrigine, and placebo as adjunctive therapy in patients with refractory partial-onset seizures. *Epilepsy Res* 2010;91:10–19.
21. Guekht AB, Korczyn AD, Bondareva IB, et al. Placebo responses in randomized trials of antiepileptic drugs. *Epilepsy Behav* 2010;17:64–69.
22. Rheims S, Perucca E, Cucherat M, et al. Factors determining response to antiepileptic drugs in randomized controlled trials. A systematic review and meta-analysis. *Epilepsia* 2011;52:219–233.
23. French JA, Krauss GL, Biton V, et al. Adjunctive perampanel for refractory partial-onset seizures: randomized phase III study 304. *Neurology* 2012;79:589–596.
24. French JA, Baroldi P, Brittain ST, et al. Efficacy and safety of extended-release oxcarbazepine (Oxtellar XR™) as adjunctive therapy in patients with refractory partial-onset seizures: a randomized controlled trial. *Acta Neurol Scand* 2014;129:143–153.
25. Goldenholz DM, Goldenholz SR. Response to placebo in clinical epilepsy trials—old ideas and new insights. *Epilepsy Res* 2016;122:15–25.
26. Franco V, French JA, Perucca E. Challenges in the clinical development of new antiepileptic drugs. *Pharmacol Res* 2016;103:95–104.
27. Tachibana Y, Narukawa M. Investigation of influencing factors on higher placebo response in East Asian versus Western clinical trials for partial epilepsy: a meta-analysis. *Clin Drug Investig* 2013;33:315–324.
28. Biton V, Montouris GD, Ritter F, et al. A randomized, placebo-controlled study of topiramate in primary generalized tonic-clonic seizures. Topiramate YTC Study Group. *Neurology* 1999;52:1330–1337.
29. Biton V, Sackellares JC, Vuong A, et al. Double-blind, placebo-controlled study of lamotrigine in primary generalized tonic-clonic seizures. *Neurology* 2005;65:1737–1743.
30. Cornu C, Kassai B, Fisch R, et al. Experimental designs for small randomised clinical trials: an algorithm for choice. *Orphanet J Rare Dis* 2013;8:48.
31. Privitera MD, Treiman DM, Pledger GW, et al. Dezinamide for partial seizures: results of an n-of-1 design trial. *Neurology* 1994;44:1453–1458.
32. Willoughby JO, Pope KJ, Eaton V. Nicotine as an antiepileptic agent in ADNFLE: an N-of-one study. *Epilepsia* 2003;44:1238–1240.
33. Sackett DL, Richardson WS, Rosenberg W, et al. *Evidence-based medicine: how to practice and teach EBM*. New York, NY: Churchill Livingstone; 1997.
34. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med* 2000;342:1878–1886.
35. Group for the Evaluation of Cinromide in the Lennox-Gastaut Syndrome. Double-blind, placebo-controlled evaluation of cinromide in patients with the Lennox-Gastaut syndrome. *Epilepsia* 1989;30:422–429.
36. Kessler DA, Rose JL, Temple RJ, et al. Therapeutic-class wars—drug promotion in a competitive marketplace. *N Engl J Med* 1994;331:1350–1353.
37. Perucca E, Tomson T. Monotherapy trials with the new antiepileptic drugs: study designs, practical relevance and ethical implications. *Epilepsy Res* 1999;33:247–262.
38. Every-Palmer S, Howick J. How evidence-based medicine is failing due to biased trials and selective publication. *J Eval Clin Pract* 2014;20:908–914.
39. Perucca E. Designing clinical trials to assess antiepileptic drugs as monotherapy: difficulties and solutions. *CNS Drugs* 2008;22:917–938.
40. Mintzer S, French JA, Perucca E, et al. Is a separate monotherapy indication warranted for antiepileptic drugs? *Lancet Neurol* 2015;14:1229–1240.
41. Leber PD. Hazards of inference: the active control investigation. *Epilepsia* 1989;30(Suppl. 1):S57–S63.
42. Walker MC, Sander JW. Difficulties in extrapolating from clinical trial data to clinical practice: the case of antiepileptic drugs. *Neurology* 1997;49:333–337.

43. Mattson RH, Cramer JA, Collins JF, et al. Comparison of carbamazepine, phenobarbital, phenytoin, and primidone in partial and secondarily generalized tonic-clonic seizures. *N Engl J Med* 1985;313:145–151.

44. Mattson RH, Cramer JA, Collins JF, et al. A comparison of valproate with carbamazepine for the treatment of complex partial seizures and secondarily generalized tonic-clonic seizures in adults. *N Engl J Med* 1992;327:765–771.

45. Glauser T, Ben-Menachem E, Bourgeois B, et al. ILAE treatment guidelines: evidence-based analysis of antiepileptic drug efficacy and effectiveness as initial monotherapy for epileptic seizures and syndromes. *Epilepsia* 2006;47:1094–1120.

46. Rowan AJ, Ramsay RE, Collins JF, et al. New onset geriatric epilepsy: a randomized study of gabapentin, lamotrigine, and carbamazepine. *Neurology* 2005;64:1868–1873.

47. Lexchin J, Bero LA, Djulbegovic B, et al. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003;326:1167–1170.

48. Brodie MJ, Chadwick DW, Anhut H, et al. Gabapentin versus lamotrigine monotherapy: a double-blind comparison in newly diagnosed epilepsy. *Epilepsia* 2002;43:993–1000.

49. Kwan P, Brodie MJ. Effectiveness of first antiepileptic drug. *Epilepsia* 2001;42:1255–1260.

50. Patsalos PN, Berry DJ, Bourgeois BF, et al. Antiepileptic drugs—best practice guidelines for therapeutic drug monitoring: a position paper by the subcommission on therapeutic drug monitoring. ILAE Commission on Therapeutic Strategies. *Epilepsia* 2008;49:1239–1276.

51. Guerreiro MM, Vigonius U, Pohlmann H, et al. A double-blind controlled clinical trial of oxcarbazepine versus phenytoin in children and adolescents with epilepsy. *Epilepsy Res* 1997;27:205–213.

52. Kälviäinen R, Aikiä M, Saukkonen AM, et al. Vigabatrin vs carbamazepine monotherapy in patients with newly diagnosed epilepsy. A randomized, controlled study. *Arch Neurol* 1995;52:989–996.

53. Perucca E. Extended-release formulations of antiepileptic drugs: rationale and comparative value. *Epilepsy Curr* 2009;9:153–157.

54. Canger R, Altamura AC, Belvedere O, et al. Conventional vs controlled-release carbamazepine: a multicentre, double-blind, cross-over study. *Acta Neurol Scand* 1990;82:9–13.

55. Brodie MJ, Overstall PW, Giorgi L, et al. Multicentre, double-blind, randomised comparison between lamotrigine and carbamazepine in elderly patients with newly diagnosed epilepsy. *Epilepsy Res* 1999;37:81–87.

56. Saetre E, Perucca E, Isojärvi J, et al. An international multicenter randomized double-blind controlled trial of lamotrigine and sustained-release carbamazepine in the treatment of newly diagnosed epilepsy in the elderly. *Epilepsia* 2007;48:1292–1302.

57. Mbizvo GK, Dixon P, Hutton JL, et al. Levetiracetam add-on for drug-resistant focal epilepsy: an updated Cochrane Review. *Cochrane Database Syst Rev* 2012;9:CD001901.

58. Xiao Z, Li JM, Wang XF, et al. Efficacy and safety of levetiracetam (3,000 mg/day) as an adjunctive therapy in Chinese patients with refractory partial seizures. *Eur Neurol* 2009;61:233–239.

59. Zaccara G, Almas M, Pitman V, et al. Efficacy and safety of pregabalin versus levetiracetam as adjunctive therapy in patients with partial seizures: a randomized, double-blind, noninferiority trial. *Epilepsia* 2014;55:1048–1057.

60. Weston J, Greenhalgh J, Marson AG. Antiepileptic drugs as prophylaxis for post-craniotomy seizures. *Cochrane Database Syst Rev* 2015;3:CD007286.

61. Beenen LF, Lindeboom J, Kasteleijn-Nolst Trenité DG, et al. Comparative double blind clinical trial of phenytoin and sodium valproate as anticonvulsant prophylaxis after craniotomy: efficacy, tolerability, and cognitive effects. *J Neurol Neurosurg Psychiatry* 1999;67:474–480.

62. Rosenow F, Schade-Brittinger C, Burchardi N, et al. The LaLiMo Trial: lamotrigine compared with levetiracetam in the initial 26 weeks of monotherapy for focal and generalised epilepsy—an open-label, prospective, randomised controlled multicenter study. *J Neurol Neurosurg Psychiatry* 2012;83:1093–1098.

63. Kwan P, Brodie MH, Kälviäinen R, et al. Efficacy and safety of pregabalin versus lamotrigine in patients with newly diagnosed partial seizures: a phase 3, double-blind, randomised, parallel-group trial. *Lancet Neurol* 2011;10:881–890.

64. Kwan P, Arzimanoglou A, Berg AT, et al. Definition of drug resistant epilepsy: consensus proposal by the ad hoc Task Force of the ILAE Commission on Therapeutic Strategies. *Epilepsia* 2010;51:1069–1077.

65. Steinhoff BJ, Ueberall MA, Siemes H, et al. The LAM-SAFE study: lamotrigine versus carbamazepine or valproic acid in newly diagnosed focal and generalised epilepsies in adolescents and adults. *Seizure* 2005;14:597–605.

66. Luoni C, Bisulli F, Canevini MP, et al. Determinants of health-related quality of life in pharmacoresistant epilepsy: results from a large multicenter study of consecutively enrolled patients using validated quantitative assessments. *Epilepsia* 2011;52:2181–2191.

67. Birbeck GL, Hays RD, Cui X, et al. Seizure reduction and quality of life improvements in people with epilepsy. *Epilepsia* 2002;43:535–538.

68. Perucca E. When clinical trials make history: demonstrating efficacy of new antiepileptic drugs as monotherapy. *Epilepsia* 2010;51:1933–1935.

69. Gazzola DM, Balcer LJ, French JA. Seizure-free outcome in randomized add-on trials of the new antiepileptic drugs. *Epilepsia* 2007;48:1303–1307.

70. Barcs G, Walker EB, Elger CE, et al. Oxcarbazepine placebo-controlled, dose-ranging trial in refractory partial epilepsy. *Epilepsia* 2000;41:1597–1607.

71. Ng YT, Conry JA, Drummond R, et al. Randomized, phase III study results of clobazam in Lennox-Gastaut syndrome. *Neurology* 2011;77:1473–1481.

72. Akl EA, Briel AM, You JJ, et al. Potential impact on estimated treatment effects of information lost to follow-up in randomised controlled trials (LOST-IT): systematic review. *BMJ* 2012;344:e2809.

73. European Medicines Agency. Diacomit. Summary of product characteristics. Available at:http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000664/WC500036518.pdf. Accessed March 14, 2016.

74. Chiron C, Marchand MC, Tran A, et al. Stiripentol in severe myoclonic epilepsy in infancy: a randomised placebo-controlled syndrome-dedicated trial. STICLO study group. *Lancet* 2000;356:1638–1642.

75. Sake JK, Hebert D, Isojärvi J, et al. A pooled analysis of lacosamide clinical trial data grouped by mechanism of action of concomitant antiepileptic drugs. *CNS Drugs* 2010;24:1055–1068.