Outcome assessment in epilepsy: available rating scales for adults and methodological issues pertaining to the development of scales for childhood epilepsy

Hans A. Carpay *, Willem F.M. Arts

Westeinde Hospital, Department of Child Neurology, Den Haag, The Netherlands

Received 11 October 1995; revised 2 February 1996; accepted 2 February 1996

Abstract

During the past decade, several scales have been developed to improve the assessment of outcome in epilepsy. These scales were developed for adults and their reliability, validity and usefulness have been established. However, there is also a need for alternative measures of outcome in childhood epilepsy, especially a measure of seizure severity (SS) and measures pertaining to quality of life (QoL). Four of these adult scales are reviewed and compared to examine their applicability in childhood epilepsy. Two important methodological differences between them are discussed: (a) patient self-report vs. physician-based scales and (b) generic vs. disease-specific instruments. QoL in epilepsy is briefly reviewed. Severity of seizures and severity of side-effects are relatively neglected areas of importance to QoL in epilepsy. The existing instruments for adults are not appropriate for children in their present form. Some specific methodological issues, which are relevant for the development of scales for children with epilepsy, are subsequently discussed. New scales pertaining to physical and psychosocial aspects of QoL in childhood epilepsy are being developed. In the near future, data on their reliability, validity and usefulness will become available. A combination of scales focusing on specific aspects of QoL, including SS and severity of adverse effects, and more traditional clinical data may provide a more complete assessment of outcome in childhood epilepsy.

Keywords: Quality of life; Outcome measure; Seizure severity; Adverse effect; Intractable epilepsy; Children

1. Introduction

Traditionally, outcome assessment in epilepsy has pertained to seizure frequency (SF). The limitations of this approach are obvious. Clearly, there are more variables that may be relevant for patients with epilepsy. In the past decade, several health measurement instruments have been developed to improve assessment of outcome in epilepsy [11,23,28,71]. These alternative outcome measures were developed for adults with epilepsy. As such, they are probably inappropriate for children [59]. As epilepsy in childhood is relatively common and not always easy to control, the need for improvement of traditional measures of outcome for children with epilepsy is well...
recognized [19,25]. We studied the relevant literature on outcome measures in epilepsy, as a starting point for the development of scales for childhood epilepsy. Furthermore, we explored the specific problems relating to the development of outcome measures for childhood epilepsy. Although a wide spectrum of clinical and psychosocial problems applies to patients with epilepsy, this paper will only address the clinician’s primary concerns in epilepsy treatment: seizures and adverse effects of medication.

Drug treatment will succeed in controlling seizures within a short period of time in ~ 70% of new onset patients [3,4,14,16]. In the remaining 30%, reasonable regimens fail to bring seizures under complete control. In 10–15%, chances of remission are extremely poor [14,41]; these patients are often referred to as ‘intractable’ [15]. In clinical practice, the aim in such cases is to find the optimal balance between seizure suppression and adverse effects of medication. Thus, a careful and complete assessment of the frequency and severity of seizures, as well as the prevalence and severity of side-effects, is necessary.

2. Scales to measure the clinical severity of epilepsy in adults

A comparison of four scales that have been developed to measure the severity of seizures and adverse effects in adult epilepsy demonstrates quite clearly the different strategies in the approach to outcome assessment in epilepsy. We selected these scales because they are well described in the literature. (1) Two physician-based scales will be described. (2) Two patient-based scales will be introduced which have been incorporated in quality of life (QoL) studies. The first three scales predominantly contain items which are specific for epilepsy, the last is a more generic scale. These differences in methodology will be discussed in subsequent sections.

2.1. Veterans administration (VA) rating scales for seizure type and frequency, neurotoxicity and systemic toxicity (VA scale)

The VA scale provides a composite score intended to represent ‘the overall effect of seizures and toxicity from medication on the QoL of a patient’ [23]. It relates to SF, seizure type and severity, and severity of antiepileptic drug (AED) toxicity. This physician-based scale, developed in 1983 by Cramer et al., was used in an influential comparative trial of four AEDs [51]. A slightly modified scale has been validated in a Dutch population with epilepsy by Wijsman et al. [74]. This modified scale has been used for audit studies in adult epilepsy [45,75].

The design of the VA scale for seizure type and frequency reflects the idea that the three most common seizure types in adults are not equally severe. Separate subscales provide ratings for generalized tonic-clonic seizures, complex partial seizures and simple partial seizures. The combination of type and frequency gives the basic score, with subsequent modifications for the presence of a useful aura, an avoidable precipitating factor, ‘subtherapeutic’ AED levels and specific patterns of seizures. Since seizure classification is a major factor in the determination of seizure severity (SS) in this scale, it has certain limitations, as discussed below (see Section 4.1). The VA scale has no rating for such generalized seizure types as absence and atonic seizures, which are rare in adults but not in children.

The VA scale takes the perspective of the medical-professional and does not measure patient perceptions. This causes concern about the relevance of the scale and the validity of the complex scoring system. For example, in the systemic toxicity scale the same score is given to ‘reduced platelet count (< 75,000)’ and to ‘frequent vomiting’. As far as the patient is concerned, these are problems with an entirely different impact. As such, the VA scale is not adequate for assessments of QoL as we would define it today (see Section 3.1).

2.2. Chalfont SS scale

This scale was designed by Duncan and Sander in 1991 and measures only SS [28]. It is completed by the physician, preferably during an interview with the patient and an eye-witness of the seizures, as detailed information about the seizures is requested. The physician has to make a medical-professional distinction between different seizure types in a patient. As in the VA scale, separate columns provide separate severity scores per seizure type. Individual item scores are modified according to the relative
frequency of occurrence in a seizure type (e.g. when incontinence occurs in 25% of seizures, the incontinence score is quartered).

The Chalfont scale aims to measure aspects of SS that can be recorded objectively. Although the scale and weightings were based on interviews with both patients and their close relatives, it does not necessarily reflect the individual patient’s opinion.

To our knowledge, the Chalfont scale has not been validated in children.

2.3. Liverpool SS scale

This epilepsy-specific scale was developed to measure the (‘subjective’) severity of seizures as perceived by the patient, although it includes some ‘objective’ items [11]. The instrument consists of two subscales: perception of control and ictal and postictal events. Most items use a simple subjective four-point response scale, ranging from 1 (the patient perceived that aspect of the seizure to be no problem) to 4 (a severe problem). The patient may choose to complete different columns for his/her ‘major’ and ‘minor’ seizures. There is increasing evidence that this scale is reliable, valid and useful as an outcome measure in a clinical trial [10,68]. It has been included in a multidimensional model for QoL in epilepsy developed by Baker and colleagues [9,10]. Recently, a similar patient-based adverse event profile has been developed (Baker et al., submitted for publication).

Although these scales reflect the individual patient’s perception, both the Liverpool SS scale and the adverse events profile agree well with the clinician’s paradigm of the severity of epilepsy [37]. As these are self-report scales for adults, they would have to be modified to be applicable in a childhood population.

2.4. Epilepsy surgery inventory (ESI)-55

The ESI-55 was developed by Vickrey and colleagues in 1992 [71]. Like the Liverpool scales, the ESI-55 is a subjective patient-based scale. This self-report scale was developed for studies assessing the outcome of epilepsy surgery in adults. It consists of a generic core (the RAND 36-Item Health Survey) with additional items, providing an epilepsy-specific supplement. The items about physical function and pain do not specifically address the consequences of seizures or antiepileptic medication. When compared to the Liverpool scales, the ESI-55 is a more generic measure of QoL and its relation to epilepsy-specific clinical variables, like SS and severity of adverse effects, is not as direct as in the Liverpool scales. The ESI-55 has been applied to evaluate seizure-based outcome systems from studies addressing the outcome of epilepsy surgery [72].

Because the ESI-55 was designed for patients with a severe form of epilepsy who are candidates for epilepsy surgery, it may not be valid for less severe forms of epilepsy. For this reason, the ESI-55 was expanded for broader application as the QoL in epilepsy (QOLIE) scale [57]. In their present form, as self-report scales for adults, they are not appropriate for children with epilepsy.

3. Definitions and basic differences in methodology

The previous section has shown that there are important differences between existing health measurement scales for seizures and adverse effects of AED. Not surprisingly, none of these scales seems directly suitable for use in childhood epilepsy. However, they illustrate some important differences in methodology.

First, it seems appropriate to define two terms associated with assessment of outcome of chronic disorders: QoL and clinimetrics. Two basic methodological differences in the presented sample of scales are subsequently discussed.

3.1. QoL

Health-related QoL (or perceived health status) refers to a scientific analysis of the functional outcome of a disease and its treatment in a patient [63]. The goal is to quantify patient’s perceptions as valid and reliable data [9]. According to Schipper et al., the questions included in the assessment may be drawn from the experience of patients, relatives and health care providers, but they should be answered from the patient’s perspective [63]. Some authors, however, take a more restrictive – patient-centred –

view [34]. Karnofsky and colleagues were the first in the literature to demonstrate the importance of assessment of the effects of disease on a patient’s functional status [42]. The formal use of the term QoL developed later and its definition is still – to some extent – a matter of debate.

QoL is essentially a multidimensional concept, including physical, psychological, social and economic domains [63]. It is widely recognized as an important outcome measure in chronic disease [1]. The major drawback in the use of QoL assessments, is the lack of consensus on how it should be measured [34,35] and the complexity of most attempts to measure it.

3.1.1. QoL and epilepsy

Epilepsy is a relative latecomer in the field of formal QoL studies [9,39]. A great deal of past psychosocial research on the impact of epilepsy might be appropriately termed ‘QoL-research’ [39,62]. QoL in epilepsy has been reviewed elsewhere [39,52,62,67]. From these reviews it becomes clear that the physical domain has been neglected as compared to the psychosocial domain. Many studies claiming to report QoL have only addressed psychosocial issues. Hermann identifies symptoms and functional status as underinvestigated areas of QoL in epilepsy [39].

As QoL is the ultimate outcome of medical treatment of any chronic disorder, there is no doubt that it is relevant in epilepsy and should be incorporated in trials comparing treatment regimes [24,25,52,76].

3.2. Clinimetrics

Clinimetrics refers to the use of rating scales which transform clinical data into a score. This allows the summation of different medical variables and, thus, facilitates statistical comparisons of a patient’s status and the assessment of change after treatment [31]. This does not imply that the clinician is the only source of information. Clinimetric tools usually relate to severity of disease in medical terms, but can be a part of QoL research, when they quantify clinical data from the patient’s perspective: a clinimetric approach to assessing QoL in epilepsy [22].

3.3. Methodological issues

A comparison of the scales in Section 2 indicates that there are important methodological differences between them. At least, we can make a distinction along two axes.

(1) Physician-based vs. patient-based scales (2.1 and 2.2 vs. 2.3 and 2.4).

(2) Epilepsy-specific vs. generic scales (2.1 to 2.3 vs. 2.4).

Clearly, two questions are of importance:

– Who determines the clinical severity of disease: the doctor or the patient?

– When should we use disease-specific scales and when a (more well known) generic profile?

3.3.1. Who determines disease severity?

Many authors agree that the patient is the most important authority regarding the effects of disease on his/her life [24,69]. At least, doctors and nurses are not able to assess overall QoL of their patients in a meaningful and reliable way [66]. This favours the assessment of subjective disease severity as perceived by the patient and QoL assessments are based on this principle. The openly subjective nature of QoL assessments, however, is a source of unease among some investigators [63]. In the opinion of some authors, physician-based scales are more ‘scientific’ than scales reflecting patient’s perceptions [13].

Perhaps, a sharp division between ‘patient-based’ and ‘physician-based’ scales is too artificial. Any health measurement scale reflects patients’ as well as doctors’ opinions both in the development and selection of the scale as well as in the completion of the scale in an individual patient’s case.

3.3.2. Generic and specific scales

Disease severity can be assessed with generic scales, which are broadly applicable across different diseases and populations, or with disease- or population-specific instruments [32,36,39,56]. A clear advantage of many generic scales is that they are well validated. This is especially important, since for most health measurement scales there is no ‘gold standard’ to assess their validity in a simple way.

However, using such generic scales, one can eas-
ily miss important problems that relate only to a specific condition. If one focuses more accurately on a specific disease or population, this will result in increased responsiveness of an scale [36]. As epilepsy is not an ‘average’ disease, we agree with others that at least a fair amount of epilepsy-specific items should be included in any scale developed to assess treatment effects in clinical trials concerning epilepsy [67].

4. Epilepsy-specific scales: seizures and adverse effects

As the severity of uncontrolled epilepsy depends to a large extent on the frequency and severity of seizures and side-effects of AEDs, these variables should be addressed adequately in epilepsy-specific scales.

4.1. SF and type

Most clinical trials in intractable epilepsy are based on an assessment of change in SF. However, there is no consensus on how a change in SF should be measured or reported. A large number of different outcome systems, typically classifying patients into categories relating to SF, is used [5,23,72]. For some seizure types, reports of SF may be unreliable as they are difficult to quantify without sophisticated techniques [30].

The outcome systems used in most studies fail to recognize that the consequences of seizures determine the number of attacks which an individual patient will tolerate. At least for patients with several ‘minor seizures’ a day, the absolute quantity of seizures may not be clinically relevant, nor would a reduction in SF of 50% be very significant [15,60,65,72].

4.2. Seizure type and severity

SF is quite meaningless without at least information about seizure type. Clinicians will have some idea of the average severity of different seizure types. For example, seizure duration, an obvious indicator of SS, is related to seizure type [47,73]. Loss of consciousness during a seizure influences both classification and SS. However, the ILAE seizure classification [20] is not a complete and sensitive indicator of SS.

- It lacks content validity for this purpose: many factors that intuitively are relevant to SS, are not included: incontinence, injuries, postictal dysfunction etc.

- It lacks discriminative ability: within one seizure type, factors relating to severity may vary, e.g. a complex partial seizure lasting 15 s is less severe than a complex partial seizure lasting 8 min, despite identical classification.

- It is not adequately sensitive to change in SS: a change of SS can only be monitored if it results in a change in classification.

- It is not as practical as it might seem. Especially, in the more complex childhood epilepsies seizure classification is not a simple matter [73].

In conclusion, for studies of uncontrolled epilepsy, a combination of a measure of SF and seizure classification is not sufficient. The development of a complete seizure-based outcome system should include a specific measure of SS.

In Section 2, we reviewed two scales, providing such a measure of SS: the Chalfont and the Liverpool SS scales (Section 2.2, Section 2.3, respectively). Both scales include a large extent the same variables, as shown in Table 1.

According to Cramer, the major problems inherent in assessing SS are dependence on patient recall and reporting, and observer documentation [22].

4.3. Timing, predictability and control of seizures

The VA and Chalfont scales include modifying factors for timing and/or predictability. The Liverpool SS scale has a subscale for this construct (Table 1), although it is less reliable and sensitive than the ictal subscale [68]. The common idea is that some degree of predictability and control makes seizures less likely to cause injuries and perhaps less severe.

4.4. Adverse effects of antiepileptic drug treatment

At present, in most studies, reports of adverse effects are descriptive and lack quantification [46,50]. Assessment of adverse effects is of great importance in epilepsy treatment [10,33,41,48]. Considerations
Table 1
Comparison of variables included in two scales measuring SS

<table>
<thead>
<tr>
<th>Liverpool SS scale</th>
<th>Chalfont SS scale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Perception subscale</strong></td>
<td></td>
</tr>
<tr>
<td>Timing</td>
<td>Warning before loss of awareness</td>
</tr>
<tr>
<td>Aura</td>
<td></td>
</tr>
<tr>
<td>Predictability</td>
<td></td>
</tr>
<tr>
<td>Prevention</td>
<td></td>
</tr>
<tr>
<td>Clustering</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td></td>
</tr>
<tr>
<td>Day-to-day activity</td>
<td></td>
</tr>
<tr>
<td>Occurrence in sleep</td>
<td>Nocturnal seizures only; divide score by 2</td>
</tr>
<tr>
<td><strong>Ictal and postictal subscale</strong></td>
<td></td>
</tr>
<tr>
<td>Overall severity of seizures</td>
<td></td>
</tr>
<tr>
<td>Loss of consciousness</td>
<td>Duration of awareness</td>
</tr>
<tr>
<td>Duration of loss of consciousness</td>
<td></td>
</tr>
<tr>
<td>Lip-smacking/fidgeting</td>
<td>Automatisms</td>
</tr>
<tr>
<td>Postictal confusion</td>
<td>Time to return to normal from onset</td>
</tr>
<tr>
<td>Duration of postictal confusion</td>
<td></td>
</tr>
<tr>
<td>Falling</td>
<td>Fall to ground</td>
</tr>
<tr>
<td>Headache</td>
<td></td>
</tr>
<tr>
<td>Incontinence</td>
<td></td>
</tr>
<tr>
<td>Tongue-biting</td>
<td></td>
</tr>
<tr>
<td>Incontinence</td>
<td></td>
</tr>
<tr>
<td>Injury, including tongue-biting</td>
<td></td>
</tr>
<tr>
<td>Drop, spill a held object</td>
<td></td>
</tr>
<tr>
<td>Convulsion (clonic jerking of limbs)</td>
<td></td>
</tr>
</tbody>
</table>

regarding adverse effects influence the selection of an AED regimen [18,38,48,51,70]. Adverse effects are associated with high doses and the use of multiple AEDs [48,58]. Thus, patients with uncontrolled seizures are at risk for adverse effects, as they are often prescribed high doses or a combination of AEDs [64].

It is possible to classify adverse effects along different axes, such as relationship to duration of therapy (early vs. late adverse effects) or relationship to dose (idiosyncratic vs. dose-dependent adverse effects) or by organ system [26]. The most frequent adverse effects are dose-related, not dangerous and quite subjective [48]. They may come and go when the dose is changed or when the patient’s tolerance to the drug changes [29]. Therefore, an individual, subjective measure for these ‘subtle’ adverse effects seems useful. Clinically severe adverse effects are a rare and clear medical-professional problem and the patient’s perception in such cases is less relevant. Behavioural and cognitive adverse effects are usually evaluated with standardized psychological tests [2,53]. As these psychological tests are not easily administered, especially in children, a more simple (screening) tool may be useful [21].

5. Health measurement scales in childhood epilepsy

It is not surprising that health measurement scales developed for an adult population are seldom appropriate for children [59]. We concluded that the reviewed scales for adult epilepsy patients in their present form are inappropriate for childhood epilepsy. Let us look at the reasons for this conclusion and at some specific issues relating to scales for children and childhood epilepsy.

5.1. Who determines disease severity in children?

Children are a problematic group for self-report scales. Self-report scales for children require separate scales at least for different age groups. Probably, sex groups should be analysed separately at different developmental stages, as during the school age period, girls may have relatively superior language and social skills [43]. All this makes it difficult to compare results.

Second, any researcher developing a self-report scale for children should be aware of their structurally different way of thinking about disease [44]. A model of severity of epilepsy that makes sense to adults, may not agree with the ideas that children have.

Third, for children who develop epilepsy, the parents play a crucial role in relation to rationalization and, as suggested by Scambler, the opinions of young children about their epilepsy may in fact be very similar to their parents’ opinions [61]. This raises the question: Is it worthwhile developing a childhood self-report scale if the same information can be collected more easily by questioning the parents?

In most studies on disease severity in children, information comes from the parents, even when the
aim is to assess the children’s QoL [59]. Apparently, many researchers believe that the parents are so close to their child that their perception of disease severity approaches the child’s perspective. Furthermore, such an approach rightly recognizes parents as experts on their own children [54]. Self-report scales in childhood epilepsy are usually aimed at adolescents. Ratings by adolescents and their parents may well disagree [7].

When subjective assessments of disease severity are used, one must consider the possibility of biased assessments. For example, when parents are asked to quantify the severity of specific aspects of a child’s epilepsy, they may be biased by the overall disease severity and may have difficulties to focus on the specific issue addressed in a scale. Anxiety and coping difficulties may bias ratings by the parents in severe cases of epilepsy. In some cases, clinicians may feel the parents have produced ‘idiosyncratic’ scores, which are totally out of line with the clinician’s view. Health measurement scales are to be used to supplement the standard clinical information and results should always be interpreted in this context. Nevertheless, it is not always within the physician’s competence to judge at what point ratings by the parents become exaggerated and such a conclusion should, therefore, be reached with caution.

5.2. Specific characteristics of childhood epilepsy

The need for a childhood epilepsy-specific scale is further emphasized by the following specific characteristics of childhood epilepsy.

Childhood epilepsy is often more complex than adult epilepsy. Children with chronic epilepsy frequently have polymorphous seizure disorders.

Predictability and control of seizures may also be relevant issues in childhood seizures. However, a warning signal for a seizure may not be as useful in children as in adult patients, as children, especially those who are retarded, may be less able to take appropriate action.

Many children with chronic epilepsy are mentally retarded or have other impairments [12,41]. In polyhandicapped patients, it is more difficult to determine to what extent the epilepsy contributes to the overall impairment [27].

The incidence and nature of adverse effects reported in children are different from those in adults [49,70]. A scale to assess the presence and severity of adverse effects of AEDs in children should be adjusted to the specific characteristics of adverse effects in this population.

Daily life and responsibilities in children are clearly different from those of adults and, thus, the consequences of seizures are different [55]. Hence, the extent to which epilepsy is a handicapping disorder or causes disabilities [77] differs in adults and children.

6. Development of improved outcome measures for childhood epilepsy: on-going projects

Most research into improvement of outcome assessment and QoL in childhood epilepsy is of very recent date and only preliminary reports are available. Studies addressing QoL are focused on psychosocial, rather than physical or clinical factors. For this reason, a discussion of projects addressing physical aspects of QoL in childhood epilepsy has to be brief.

Baker et al. proposed a multidimensional model to assess QoL in childhood epilepsy, including measures of SS, severity of adverse effects, mood and behaviour, intellectual function and physical coordination [9]. A questionnaire based on this model was used in a study of lamotrigine in children with severe epilepsy and learning difficulties (Baker, written communication).

In 1993, the Dutch Study Group of Epilepsy in Childhood has embarked on a project to improve outcome assessment in childhood epilepsy. It was decided to focus on the physical domain of QoL, specifically on SS and severity of adverse effects. Subjective parent-completed scales were developed [78]. Furthermore, a scale was developed to measure the severity of disabilities due to restrictions in activities of daily life [17].

Austin et al. have compared the QoL of children with asthma and children with epilepsy, using more generic and predominantly psychosocial measures [8]. Children with epilepsy had a more compromised QoL in the psychological, social and school domain, and children with asthma in the physical domain. By Austin and Austin, social activities and parental su-
 pervision of children with new-onset epilepsy were studied [6]. Reduced child activity was found to be related to SF.

Hoare and Russel published a pilot validation study using a newly developed parental QoL questionnaire measuring the impact of illness on children with epilepsy and their families [40]. No data to support the scale’s reliability were given. Their scale reflects a generic approach, as it was intended to be valid for children with other disabilities as well.

In conclusion, the perfect assessment of outcome in childhood epilepsy will perhaps never be achieved, but in our opinion a combination of well validated and reliable measures pertaining to specific physical and psychosocial aspects of QoL, with medical-professional outcome variables in epilepsy, would seem an important step forward. In the near future, such new measures of outcome will become available. However, the final proof of their usefulness must come from clinical trials.

References


