

Interrater agreement of the diagnosis and classification of a first seizure in childhood. The Dutch Study of Epilepsy in Childhood

H Stroink, C A van Donselaar, A T Geerts, A C B Peters, O F Brouwer, O van Nieuwenhuizen, R F M de Coo, H Geesink, W F M Arts

J Neural Neurosurg Psychiatry 2004;**75**:241–245. doi: 10.1136/jnnp.2003.015826

Objective: To assess the interrater agreement of the diagnosis and the classification of a first paroxysmal event in childhood.

Methods: The descriptions of 100 first paroxysmal events were submitted to two panels each consisting of three experienced paediatric neurologists. Each observer independently made a diagnosis based on clinical judgment and thereafter a diagnosis based on predefined descriptive criteria. Then, the observers discussed all patients within their panel. The agreement between the six individual observers was assessed before discussion within each panel and after that, between the two panels.

Results: Using their clinical judgement, the individual observers reached only fair to moderate agreement on the diagnosis of a first seizure (mean (SE) kappa 0.41 (0.03)). With use of defined descriptive criteria the mean (SE) kappa was 0.45 (0.03). The kappa for agreement between both panels after intra-panel discussion increased to 0.60 (0.06). The mean (SE) kappa for the seizure classification by individual observers was 0.46 (0.02) for clinical judgment and 0.57 (0.03) with use of criteria. After discussion within each panel the kappa between the panels was 0.69 (0.06). In 24 out of 51 children considered to have had a seizure, agreement was reached between the panels on a syndrome diagnosis. However, the epileptic syndromes were in most cases only broadly defined.

Conclusions: The interrater agreement on the diagnosis of a first seizure in childhood is just moderate. This phenomenon hampers the interpretation of studies on first seizures in which the diagnosis is only made by one observer. The use of a panel increased the interrater agreement considerably. This approach is recommended at least for research purposes. Classification into clinically relevant syndromes is possible only in a very small minority of children with a single seizure.

See end of article for authors' affiliations

Correspondence to:
H Stroink, Department of
Neurology, St Elisabeth
Hospital, PO Box 90151,
5000 LC Tilburg, The
Netherlands;
H.Stroink@Elisabeth.nl

Received 31 March 2003
In revised form
22 June 2003
Accepted 28 June 2003

The diagnosis and classification of a first seizure in childhood may be difficult. The differential diagnosis of a single paroxysmal event is extensive, particularly in young children. The consequences of the diagnosis of a first seizure are far reaching: it causes an emotional shock in the family and leads to restriction of activities. The subsequent classification may have consequences for the prognosis. According to the recent practice parameter, treatment with anti-epileptic drugs does not prevent the development of epilepsy, and treatment should be considered only in special circumstances.¹ Nevertheless, many children are at present still treated with anti-epileptic drugs after a first unprovoked seizure.² An objective test to confirm or refute the diagnosis of first seizure is missing. Epileptiform discharges on EEG recordings are not rare in children without epilepsy,^{3–5} whereas as many as 41% of patients with epilepsy and 56% of children with a first seizure have no epileptiform discharges on their standard EEG.^{6,7} The very low diagnostic value of EEG in children with single events of disputable origin was shown in an earlier study.^{7,8} Therefore, the diagnosis has to be based on the description of the episode given by an eyewitness, or sometimes by the child itself if he or she is old enough. For these reasons it is difficult to assess the accuracy^{8,9} of the diagnosis and classification of a first paroxysmal event, and little is known about the reliability (consistency, interrater and intrarater agreement) of the diagnosis. Earlier studies on children with single seizures did not mention these diagnostic problems.^{10–19} A study in adult patients showed that the use of diagnostic criteria formulated

in simple descriptive terms and discussion between neurologists improved the diagnostic agreement.²⁰

In a prospective hospital based multicentre study (Dutch Study of Epilepsy in Childhood, DSEC), we enrolled all children with suspected single seizures⁷ or epilepsy.^{21,22} We used previously defined descriptive criteria to diagnose seizures. In this part of the study under experimental conditions we evaluated the interrater agreement on the diagnosis and classification of a first paroxysmal event in childhood, and compared the results with the original diagnosis. We assessed whether the use of predefined criteria and discussion of the available data in a panel improved the interrater agreement.

PATIENTS AND METHODS

Two hundred and thirty three children, aged one month to 16 years, were included in the DSEC after a single unprovoked paroxysmal episode. This episode was considered as either a seizure or an unclear event by the paediatric neurologist of one of the four participating hospitals.⁷ Children with a clear diagnosis other than epileptic seizure were not referred systematically. The mean age was 6.2 years, median 6.0 years (25th percentile 2.0; 75th percentile 9.0); 110 were boys. The paediatric neurologist made a description of the event, and

Abbreviations: DSEC, Dutch Study of Epilepsy in Childhood; ILAE, International League Against Epilepsy

completed an extensive questionnaire on the episode, previous medical history, and findings on physical examination. All children were discussed in the original panel of the four paediatric neurologists participating in the DSEC (HS, AP, OB, WA) to assess the diagnosis according to predefined diagnostic criteria (table 1). This list contained descriptions of all possible seizure types, but in table 2 of this paper we only mention seizures which may present as a single event. The events were classified as epileptic seizure (170), other diagnosis (9), or unclear event (54). The study on the prognosis and prognostic determinants of these children was published in 1998.⁷

One year after the intake for children with a single event into the DSEC had been closed, two of the authors (HS, CD) selected 100 events from the diagnostic categories mentioned above. The intake panel of the DSEC considered 51 children to have had an epileptic seizure, nine an event with a clear other diagnosis (like breath holding spell or syncope), and 40 an unclear event. The number of children with an unclear event was set proportionally higher than in the original cohort of the DSEC to encourage discussion on their diagnosis and to diminish agreement due to chance. The mean age of the children was 5.6 years, median 6.0 years (25th percentile 2.0; 75th percentile 9.0); 52 were boys.

Two new panels were formed. Panel A consisted of three of the four paediatric neurologists from the original panel, each with at least 10 years of experience in paediatric epilepsy and in working in such an interactive way (AP, OB, WA). Together with HS, they started the DSEC in 1988. Panel B consisted of three experienced senior paediatric neurologists at that time working in other hospitals. One (HG) was attached to an epilepsy clinic, one (ON) to a university centre for epilepsy surgery, and one (RC) to a university hospital for children.

The members of both panels received anonymous descriptions of the 100 events, as given in the letter to the family physician. This included possible provoking factors and postictal signs, the previous medical history, the results of

Table 1 Descriptive criteria for the diagnosis of a single seizure. (Seizures that never present as a single event, like absences, are not described in this table.)

A	Episode with loss of consciousness
1	and jerks in one or more extremities or the face with or without tongue bite (generalised seizure)
2	and stiffness in one or more extremities with or without tongue bite (generalised seizure)
3	with only tongue bite (generalised seizure)
4	preceded by a classical aura (generalised seizure)
B	Episode of lowered consciousness without reaction to external stimuli
1	staring spell without automatisms, but preceded by a classical aura (complex partial seizure)
2	episode lasting more than 30 seconds with automatisms like smacking, swallowing, blinking, grimacing, fumbling, or making continuous stereotyped movements, not remembered by the patient (complex partial seizure or absence status)
C	Episode with normal consciousness and
1	one sided myoclonic jerks or stiffness in the face, jaw and/or tongue
2	one sided myoclonic jerks of the extremities
3	involuntary noises from the throat
4	paresthesias or numbness in the face, mouth, throat, tongue or extremities (one sided)
5	hypersalivation
6	anarthria
	(1-6: simple partial seizure)
	The event should not have been provoked by fever (temperature above 38°C, febrile convulsion); pain, acute emotional upset, or other known reason (collapse, blue or pale breath holding spell, Sandifer syndrome etc), nor should it have a reasonable explanation other than epileptic.

Table 2 Questionnaires

Diagnosis based on your own judgment	
1	If you were the attending physician, would you consider this episode to be an epileptic seizure, irrespective of the criteria presented in table 1? Yes/no/unclear
2	If you answered no, your diagnosis of the event is:
3	If you answered yes, how would you classify the seizure? (a) simple partial (b) complex partial (c) generalised with partial onset (d) generalised without partial onset (e) equivocal generalised or partial
Diagnosis with use of previously defined criteria	
4	Do you think the history of the event satisfies the descriptive criteria listed in table 1? Yes/no/unclear
5	If you answered yes, which category is applicable? (answer one or more) A1 / A2 / A3 / A4 B1 / B2 C1 / C2 / C3 / C4 / C5 / C6
6	Do you consider the seizure had a partial onset/partial characteristics? Yes/no/unclear
Diagnosis as a panel	
After your diagnosis and classification, your panel will discuss the case to reach a final common decision about the questions 4, 5, 6. After this, you will receive the results of the physical examination, EEG, imaging study, and other additional investigations.	
7	If your panel diagnosed the event as an epileptic seizure, the aetiology according to this panel is: (a) idiopathic (b) remote symptomatic (c) acute symptomatic (d) associated with mental retardation (e) equivocal
8	Are you able to classify the type of epilepsy according to the ILAE classification, knowing the results of the additional investigations?

the physical examination, and an assessment of the mental development. They were distributed in random order and did not include the results of additional investigations (EEG, imaging, etc). The paediatric neurologists were not aware of the stratification policy. Firstly, each member decided independently on the question "Was it a seizure?" and, if applicable, on seizure classification according to his personal judgment. Subsequently, they independently repeated this process using the predefined descriptive criteria (table 2). Then the observers discussed all patients within their own panel until they reached consensus on the diagnosis and, if applicable, classification of the event according to the predefined descriptive criteria. Next the panels received information on the results of the EEG, imaging study, and possible other relevant information. With this new information, both panels were independently asked again to classify the seizure in an epileptic syndrome, according to the classification of the International League Against Epilepsy (ILAE), despite the single occurrence of the event.²³ The panels were forced to reach consensus in all cases.

We evaluated the interrater agreement between the individual paediatric neurologists, between both panels after discussion between their members, and between both panels and the original panel deciding on inclusion in the DSEC. As part of the observed agreement can be attributed to chance, we used kappa statistics to assess the interrater agreement for each pair of observers and between the panels. The kappa is the ratio of the observed agreement beyond chance to the maximal potential agreement beyond chance. A kappa of 0.0 indicates that the observed agreement can be attributed completely to chance. A kappa of 1.0 means the observed agreement is maximal, a kappa of -1.0 means the observers totally disagree.²⁴ For intermediate values, Landis and Koch suggested the following interpretations: below 0.0, poor;

Table 3 Agreement (kappa) among observers concerning the question "Was it an epileptic seizure?" without (upper part table) and with (lower part) use of descriptive criteria (question 1 and question 4 of table 2). A1–3 are members of panel A, and B1–3 of panel B.

Diagnosis of a seizure with use of descriptive criteria	Diagnosis of a seizure without use of descriptive criteria					
	A1	A2	A3	B1	B2	B3
A1		0.52 (0.07)	0.43 (0.07)	0.39 (0.07)	0.31 (0.07)	0.48 (0.07)
A2	0.53 (0.07)		0.50 (0.07)	0.34 (0.08)	0.37 (0.07)	0.53 (0.07)
A3	0.47 (0.07)	0.56 (0.07)		0.25 (0.08)	0.60 (0.07)	0.60 (0.07)
B1	0.34 (0.07)	0.33 (0.07)	0.35 (0.07)		0.19 (0.08)	0.27 (0.08)
B2	0.37 (0.07)	0.48 (0.07)	0.60 (0.07)	0.23 (0.07)		0.44 (0.08)
B3	0.49 (0.07)	0.48 (0.07)	0.68 (0.07)	0.39 (0.07)	0.49 (0.07)	

Standard error (SE) shown in parentheses.

0.00–0.20, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; 0.81–1.00, almost perfect.²⁵ All 17 categories of the question concerning seizure classification using predefined criteria were collapsed into three new categories (A1 to A4, B1 to B2, and C1 to C6; table 1). Simplification of the complex ILAE syndrome classification was reached by grouping together the categories 4, 5, and 6; 7 and 8; 9 and 10 (table 5).

RESULTS

The kappa for pairs of individual observers for the question "Was it an epileptic seizure?" according to personal judgment varied between 0.19 and 0.60, median kappa 0.43, mean kappa 0.41 (SE 0.03). The use of the diagnostic criteria resulted in kappa values for pairs of individual observers between 0.23–0.68, median kappa 0.48, mean kappa 0.45 (0.03) (table 3). Both panels succeeded in all cases to reach consensus on the diagnosis after discussion. The kappa between the panels was 0.60 (0.06).

The kappas for the agreement between the diagnoses made by the panels participating in this experiment and the original panel deciding on entry into the DSEC were 0.72 (SE 0.07) for the experienced panel and 0.66 (0.08) for the inexperienced panel. Conspicuously, the experimental panels agreed on the epileptic nature of the event in 61 children, whereas the paediatric neurologists deciding on entry into the DSEC considered the event to be epileptic in only 51 cases.

For seizure classification, the kappas for pairs of individual observers without use of descriptive criteria varied between 0.29 and 0.60 (median 0.46, mean 0.46, SE 0.02). The use of the predefined criteria resulted in kappas of 0.34–0.74 (median 0.62, mean 0.57, SE 0.03; table 4). The kappa

between the panels after discussion within each panel was 0.69 (0.06).

Finally, after the results of the electroencephalograms and imaging study had been made available, each panel was asked to classify the epilepsy syndrome for the children diagnosed with an epileptic seizure. In 24 of the 61 children in whom both teams agreed there had been an epileptic seizure, the panels reached consensus on the syndrome diagnosis (table 5).

DISCUSSION

Our study shows that the agreement (mean kappa 0.41) between paediatric neurologists on the diagnosis of a first event as an epileptic seizure without use of criteria and without discussion is below the usual level of agreement in making a clinical diagnosis.^{9, 25} The agreement between the members of panel A, experienced in making a diagnosis in an interactive session using predefined criteria, was slightly but not significantly better than for panel B, whose members did not have this experience (tables 3 and 4). Agreement could be improved only slightly by the use of descriptive criteria, but discussion led to a better agreement (kappa 0.60). We also found substantial, but not perfect, agreement between both experimental panels and the panel originally deciding on the diagnosis at the moment of inclusion in the DSEC. The experienced panel did slightly better than the panel whose members were not used to working in such a collaborative way. This was probably not because of the effect of memory, despite the fact that the experienced panel contained three of the four members of the original panel. All case descriptions had been anonymised and the time elapsed between inclusion in the DSEC and the experiment described here varied between one and six years. On the contrary, it is

Table 4 Agreement (kappa) among observers concerning the classification of the seizure type without (upper part table) and once with (lower part) use of descriptive criteria (question 3 and question 5 of table 2). A1–3 are members of panel A and B1–3 of panel B

Classification of a seizure with use of descriptive criteria	Classification of a seizure without use of descriptive criteria					
	A1	A2	A3	B1	B2	B3
A1		0.60 (0.06)	0.47 (0.06)	0.46 (0.06)	0.47 (0.06)	0.42 (0.06)
A2	0.69 (0.07)		0.58 (0.06)	0.43 (0.06)	0.51 (0.06)	0.51 (0.06)
A3	0.64 (0.07)	0.64 (0.07)		0.38 (0.07)	0.59 (0.07)	0.44 (0.06)
B1	0.40 (0.08)	0.44 (0.08)	0.40 (0.08)		0.39 (0.06)	0.29 (0.06)
B2	0.62 (0.08)	0.61 (0.07)	0.74 (0.07)	0.34 (0.07)		0.42 (0.06)
B3	0.63 (0.08)	0.69 (0.08)	0.69 (0.07)	0.42 (0.07)	0.56 (0.07)	

Standard error (SE) shown in parentheses.

Table 5 Inter-panel agreement on syndrome classification

Panel B	Panel A								Total
	No epileptic seizure	1	2	3	4, 5, 6	7, 8	9, 10	11	
No epileptic seizure	25					1		8	34
1		4		1					5
2			5						5
3	2	2		7	1			7	19
4, 5, 6	2				5		1	22	30
7, 8			1			7			2
9, 10						1		1	2
11	1							2	3
Total	30	6	6	8	6	3	1	40	100

Classification according to the ILAE:

1. Benign childhood epilepsy with centro-temporal spikes;
2. Other symptomatic localisation-related epilepsies and syndromes;
3. Cryptogenic localisation-related epilepsies and syndromes;
4. Benign myoclonic epilepsy in infants;
5. Other generalised idiopathic epilepsies not defined above;
6. Generalised idiopathic epilepsies characterized by specific modes of precipitation;
7. Other symptomatic generalised epilepsies not defined above;
8. Specific symptomatic generalised syndromes;
9. Other undetermined epilepsies not defined above;
10. Epilepsies and syndromes without unequivocal generalised or focal features;
11. Isolated seizures or isolated status epilepticus.

Agreement between both panels is indicated in the table by bold italic numbers.
Syndromes not mentioned in the table were not diagnosed.

surprising that the agreement between the opinions of this panel at entry into the DSEC and some years later was not better than 0.72. Both the interrater and the “intra-panel” disagreement illustrated here suggest that the diagnosis of an isolated epileptic seizure may be extremely difficult and should always be looked at with some suspicion, especially in the context of a clinical research study.

For this study we had deliberately selected 100 cases with a 1:1 ratio between epileptic seizures on the one hand and unclear or non-epileptic events on the other. Although kappa statistics take into account the agreement due to chance, the results are influenced by the distribution of the possible diagnoses. In case of an askew distribution, kappa statistics will be lower than in case of a 1:1 ratio between the various diagnoses.²⁶ Therefore, the fair to moderate agreement rates found before discussion cannot be explained by an askew distribution between epileptic seizures versus unclear or non-epileptic events. One may even argue that our findings are biased towards a higher agreement.

Only in 24 cases was consensus reached on the syndrome diagnosis (table 5). A conspicuous discrepancy existed between the ways the panels used the ILAE classification. Panel B often tried to reach consensus on such classifications as “cryptogenic localisation-related epilepsy” or “idiopathic generalised epilepsy not otherwise defined”. Panel A classified most of these seizures as “isolated seizure or isolated status”. These classifications of both panels are safe when evidence concerning the nature of the seizure is lacking or inconclusive, but they do not contribute to our knowledge on the causal diagnosis of the child, the prognosis, or the way in which he or she should be treated. Only in the four children with benign childhood epilepsy with rolandic spikes did agreement exist on a syndrome with prognostic significance. King *et al* stated that syndrome diagnosis is possible in most patients presenting with only one seizure.²⁷ However, 45% of the patients in their study had suffered more than one seizure.²⁸ These patients were carefully excluded in our study by using standardised questionnaires.⁶ Moreover, most patients in the study of King *et al* were adults over 30 years old, so many had probably suffered a remote symptomatic partial seizure.^{28–30} More recently, the CAROLE group also made a syndrome classification in patients with newly

diagnosed epilepsy or only a single seizure.² In this study a panel made the diagnosis. However, the patients with a single seizure were much older than in our study (mean age 19 years) and most classifications were broadly defined as well. In our study, classification in a clinically relevant syndrome was possible only in a very small minority of children.

We know of only one study in which the reliability of the diagnosis of a first seizure was studied by assessing interrater agreement. This study was done in adults.²⁰ Other studies in epilepsy, in which kappa statistics were used, concerned generalised or partial seizure onset in adults,³¹ and seizure classification³² and syndrome classification in children,³³ all patients with multiple seizures. The study on seizure classification in children³² showed poor interrater correlations, and suggested that specific criteria for the categorisation of symptoms could reduce the interrater variability. Combined with our results, this suggests that the best agreement could be obtained if the seizures were not only classified according to pre-defined criteria (like in our study), but also the symptoms categorised according to a standardised questionnaire to the patient and any witnesses of the seizure.

The study on interrater agreement on classification of childhood epilepsy syndromes³³ showed excellent agreement using the ILAE classification of epilepsy syndromes, although a substantial proportion of children were classified into relatively non-specific syndromes.²³ However, in this study only children with newly diagnosed epilepsy were classified, not children with a single seizure.

Even experienced paediatric neurologists frequently disagree about the diagnosis and classification of a first seizure in children. In this study the diagnosis was based on a careful written description made by experts with the aim of an extensive questionnaire. The agreement among neurologists may be even lower when they have to listen to the actual histories from the parents themselves.

Our results may at least partly explain the widely discrepant recurrence risks reported in first seizure studies.^{10–19} In this study, the use of a panel was the best means to increase the interrater agreement. We recommend such an approach for research purposes, although even then

in many cases the diagnosis will remain uncertain. However, better ways to diagnose first seizures are not currently available.

ACKNOWLEDGEMENTS

This study was financially supported by the Dutch National Epilepsy Fund (grant numbers A 72 and A 85) and by the Princess Irene Fund, Arnhem, the Netherlands.

Authors' affiliations

H Stroink, Department of Neurology, St Elisabeth Hospital and TweeSteden Hospital, the Netherlands

C A van Donselaar, Rudolf Magnus Institute of Neuroscience, Epilepsy Section, University Medical Centre, Utrecht, the Netherlands

A T Geerts, **R F M de Coo**, **W F M Arts**, Department of Paediatric Neurology, Erasmus Medical Centre/Sophia Children's Hospital, Rotterdam, the Netherlands

A C B Peters, **O van Nieuvenhuizen**, Department of Paediatric Neurology, University Medical Centre, Utrecht, the Netherlands

O F Brouwer, Department of Paediatric Neurology, University Hospital, Groningen, the Netherlands

H Geesink, Stichting Epilepsie Instellingen, Heemstede, the Netherlands

REFERENCES

- Hirtz D, Berg A, Bettis D, et al. Practice parameter: Treatment of the child with a first unprovoked seizure: Report of the Quality Standards Subcommittee of the American Academy of Neurology and the Practice Committee of the Child Neurology Society. *Neurology* 2003;**60**:166–75.
- Jallon P, Loiseau P, Loiseau J, et al. Newly diagnosed unprovoked epileptic seizures: presentation at diagnosis in Carole Study. *Epilepsia* 2001;**42**:464–75.
- Cavazzuti GB, Cappella L, Nalin A. Longitudinal study of epileptiform EEG patterns in normal children. *Epilepsia* 1980;**21**:43–55.
- Binnie CD. What's the use of the EEG in epilepsy? *Br J Hosp Med* 1988;**39**:99–102.
- Gibbs J, Appleton RE. False diagnosis of epilepsy in children. *Seizure* 1992;**1**:15–8.
- Carpay JA, de Weerd AW, Schimsheier RJ, et al. The diagnostic yield of a second EEG after partial sleep deprivation: a prospective study in children with newly diagnosed seizures. *Epilepsia* 1997;**38**:595–9.
- Stroink H, Brouwer OF, Arts WFM, et al. The first unprovoked, untreated seizure in childhood: a hospital-based study of the accuracy of the diagnosis, rate of recurrence, and long-term outcome after recurrence. Dutch study of epilepsy in childhood. *J Neural Neurosurg Psychiatry* 1998;**64**:595–600.
- Stroink H, van Donselaar CA, Geerts AT, et al. The accuracy of the diagnosis of paroxysmal events in children. *Neurology* 2003;**60**:979–82.
- Sackett DL, Haynes RB, Tugwell P. *Clinical epidemiology, a basic science for clinical medicine*, 2nd edn. Boston, Toronto, London: Little, Brown and Company, 1991.
- Camfield PR, Camfield CS, Dooley JM, et al. Epilepsy after a first unprovoked seizure in childhood. *Neurology* 1985;**35**:1657–60.
- Boulouche J, Leloup P, Mallet E, et al. Risk of recurrence after a single, unprovoked, generalised tonic-clonic seizure. *Dev Med Child Neurol* 1989;**31**:626–32.
- Camfield P, Camfield C, Dooley J, et al. A randomised study of carbamazepine versus no medication after a first unprovoked seizure in childhood. *Neurology* 1989;**39**:851–2.
- Hauser WA, Rich SS, Annegers JF, et al. Seizure recurrence after a 1st unprovoked seizure: an extended follow-up. *Neurology* 1990;**40**:1163–70.
- Shinnar S, Berg AT, Moshe SL, et al. Risk of seizure recurrence following a first unprovoked seizure in childhood: an extended follow-up. *Paediatrics* 1996;**98**:216–25.
- First Seizure Trial Group. Randomised clinical trial on the efficacy of antiepileptic drugs in reducing the risk of relapse after a first unprovoked tonic-clonic seizure. *Neurology* 1993;**43**:478–83.
- Berg AT, Shinnar S. The risk of seizure recurrence following a first unprovoked seizure: a quantitative review. *Neurology* 1991;**41**:965–72.
- Hopkins A, Garman A, Clarke C. The first seizure in adult life: value of clinical features, electroencephalography and computerized tomographic scanning in prediction of seizure recurrence. *Lancet* 1988;**1**:721–6.
- Elwes RDC, Chesterman P, Reynolds EH. Prognosis after a first untreated tonic-clonic seizure. *Lancet* 1985;**2**:752–3.
- Hauser WA, Anderson VE, Loewenson RB, et al. Seizure recurrence after a first unprovoked seizure. *N Engl J Med* 1982;**26**:522–8.
- Van Donselaar CA, Geerts AT, Meulstee J, et al. Reliability of the diagnosis of a first seizure. *Neurology* 1989;**39**:267–71.
- Arts WFM, Geerts AT, Brouwer OF, et al. The early prognosis of epilepsy in childhood: the prediction of a poor outcome. The Dutch Study of Epilepsy in Childhood. *Epilepsia* 1999;**40**:726–34.
- Carpay JA, Arts WFM, Geerts AT, et al. Epilepsy in childhood: an audit of clinical practice. *Arch Neurol* 1998;**55**:668–73.
- Commission on Classification and Terminology of the International League Against Epilepsy. Proposal for revised classification of epilepsies and epileptic syndromes. *Epilepsia* 1989;**30**:389–99.
- Cohen JA. Coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 1960;**20**:37–46.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74.
- Kraemer HC. Ramifications of a population model for Kappa as a coefficient of reliability. *Psychometrika* 1979;**44**:461–72.
- King MA, Newton MR, Jackson GD, et al. Epileptology of the first-seizure presentation: a clinical, electroencephalographic, and magnetic resonance imaging study of 300 consecutive patients. *Lancet* 1998;**352**:1007–11.
- Lhato SD, Shorvon SD. Relevance of the first seizure. *Lancet* 1998;**352**:1003–4.
- Chadwick D, Smith D. Comment on epileptology of the first-seizure presentation. *Lancet* 1998;**35**:1855.
- Addy DP. Comment on epileptology of the first-seizure presentation. *Lancet* 1998;**35**:1857.
- Ottman R, Lee JH, Hauser WA, et al. Reliability of seizure classification using a semi-structured interview. *Neurology* 1993;**43**:2526–30.
- Bodensteiner JB, Brownsworth RD, Knapik JR, et al. Interobserver variability in the ILAE classification of seizures in childhood. *Epilepsia* 1988;**29**:123–8.
- Berg AT, Levy SR, Testa FM, et al. Classification of childhood epilepsy syndromes in newly diagnosed epilepsy: inter-rater agreement and reasons for disagreement. *Epilepsia* 1999;**40**:439–44.