Epilepsy Research 139 (2018) 43–50

Contents lists available at ScienceDirect

Epilepsy Research

journal homepage: www.elsevier.com/locate/epilepsyres

Do neurologists around the world agree when diagnosing epilepsy? – Results of an international EpiNet study

Peter S. Bergin[a,*], Ettore Beghi[b], Lynette G. Sadleir[c], Manjari Tripathi[d], Mark P. Richardson[e], Elisa Bianchi[b], Wendyl J. D'Souza[f], on behalf of the EpiNet Study Group[1]

[a] Department of Neurology, Auckland City Hospital, Grafton, Auckland, New Zealand
[b] IRCCS – Istituto di Ricerche Farmacologiche Mario Negri, Milan, Italy
[c] Department of Paediatrics, University of Otago, Wellington, New Zealand
[d] Department of Neurology, All India Institute of Medical Sciences, New Delhi, India
[e] Division of Neuroscience, King's College London, London, UK
[f] Department of Medicine, St. Vincent's Hospital, The University of Melbourne, Australia

ARTICLE INFO

Keywords:
Epilepsy
Seizures
Diagnostic agreement
Multicentre collaboration
EpiNet

ABSTRACT

Objective: Previous studies have shown moderate agreement between physicians when diagnosing epilepsy, but have included small numbers. The EpiNet study group was established to undertake multicentre clinical trials in epilepsy. Before commencing trials, we wanted to determine levels of agreement between physicians from different countries and different health systems when diagnosing epilepsy, specific seizure types and etiologies.
Methods: 30 Case scenarios describing six children and 24 adults with paroxysmal events (21 epileptic seizures, nine non-epileptic attacks) were presented to physicians with an interest in epilepsy. Physicians were asked how likely was a diagnosis of epilepsy; if seizures were generalised or focal; and the likely etiology. For 23 cases, clinical information was presented in Step 1, and investigations in Step 2.
Results: 189 Participants from 36 countries completed the 30 cases. Levels of agreement were determined for 154 participants who provided details regarding their clinical experience. There was substantial agreement for diagnosis of epilepsy (kappa = 0.61); agreement was fair to moderate for seizure type(s) (kappa = 0.40) and etiology (kappa = 0.41). For 23 cases with two steps, agreement increased from step 1 to step 2 for diagnosis of epilepsy (kappa 0.56–0.70), seizure type(s) (kappa 0.38–0.52), and etiology (kappa 0.38–0.47). Agreement was better for 53 epileptologists (diagnosis of epilepsy, kappa = 0.66) than 56 neurologists with a special interest in epilepsy (kappa = 0.58). Levels of agreement differed slightly between physicians practicing in different parts of the world, between child and adult neurologists, and according to one's experience with epilepsy.
Conclusion: Although there is substantial agreement when epileptologists diagnose epilepsy, there is less agreement for diagnoses of seizure types and etiology. Further education of physicians regarding semiology of different seizure types is required. Differences in approach to diagnosis, both between physicians and between countries, could impact negatively on clinical trials of anti-epileptic drugs.

1. Introduction

There have been few studies to determine how consistently or accurately doctors diagnose epilepsy. Agreement when two to six physicians have independently determined if patients have epilepsy has been fair to moderate (Hoefnagels et al., 1992; Rinaldi et al., 2000; Stroink et al., 2004; van Donselaar et al., 1989). Agreement regarding type of seizure (Bodensteiner et al., 1988; van Campen et al., 2013; van Donselaar et al., 1990) and epilepsy syndrome (Ottman et al., 1990; Rinaldi et al., 2000) has generally been poorer with occasional exceptions (Berg et al., 1999; Kellinghaus et al., 2004). No studies have compared diagnostic agreement between neurologists from multiple countries.

Most clinical trials in epilepsy have presumed investigators can make an accurate diagnosis, and distinguish generalised from focal

Abreviations: AEDs, Anti epileptic drug; AVM, arterio-venous malformation; CAE, Childhood absence epilepsy; EEG, Electroencephalogram; ILAE, International league against epilepsy; JAE, Juvenile absence epilepsy; JME, Juvenile myoclonic epilepsy; NOS, not otherwise specified; REM, Rapid eye movement; TC , Tonic clonic

* Corresponding author.
E-mail addresses: pbergin@adhb.govt.nz (P.S. Bergin), ettore.beghi@marionegri.it (E. Beghi), lynette.sadleir@otago.ac.nz (L.G. Sadleir), mark.richardson@kcl.ac.uk (M.P. Richardson), elisa.bianchi@marionegri.it (E. Bianchi), wendyl@unimelb.edu.au (W.J. D'Souza).

[1] Members of the EpiNet study group who contributed to this study are listed in an Appendix A after the Acknowledgements section

https://doi.org/10.1016/j.eplepsyres.2017.10.014
Received 2 January 2017; Received in revised form 17 April 2017; Accepted 20 October 2017
Available online 26 October 2017
0920-1211/

seizures. This is not necessarily the case. Trials including large numbers of centres from multiple countries tend to have lower effect sizes than trials conducted in fewer centres or a single country (Friedman and French, 2012). Some trials may have obtained negative results because patients have been included who do not have epilepsy, or the seizure type under investigation (Friedman and French, 2012).

The EpiNet study group was established to undertake investigator-led clinical research in epilepsy. Initially formed by members of the New Zealand League against Epilepsy, it now comprises neurologists and epileptologists from many countries (www.epinet.co.nz) (Bergin et al., 2007; Bergin et al., 2012). The EpiNet study group has recently commenced trials in newly diagnosed patients with epilepsy (Bergin et al., 2015). The EpiNet steering committee undertook a Validation study to determine the diagnostic accuracy and consistency of investigators prior to undertaking these trials.

It was decided to invite any epileptologist, neurologist or paediatrician with an interest in epilepsy to participate in this study.

The EpiNet Validation study had two aims:

1) to accredit investigators for the EpiNet-First trials.
2) to determine how much variability there is between neurologists and epileptologists in diagnoses.

A separate paper looking at the process to accredit investigators for the EpiNet-First trials is in press (Bergin et al., 2017).

This paper addresses the second of these aims – assessing the levels of agreement between neurologists and epileptologists from different countries around the world.

2. Methods

The processes undertaken for the EpiNet Validation study have been described elsewhere (Bergin et al., 2017). In summary, the EpiNet steering committee prepared 30 case scenarios describing patients with various paroxysmal attacks. Case histories described real patients, nearly all of whom had been seen by members of the steering committee; all identifying details were removed. Not all patients had epileptic seizures; nine patients had attacks often confused with epilepsy (e.g. syncope, psychogenic non-epileptic episodes). The 30 cases were chosen by consensus by the EpiNet steering committee from an initial pool of 40 cases that had been prepared by PB. Most of the cases represented patients with new-onset epilepsy or patients who were seen in first-seizure clinics, since the study was also being undertaken to accredit investigators for the EpiNet-First trials (Bergin et al., 2017). Many patients had typical histories, but cases where the diagnosis was less clear were also included. Some scenarios described attacks in detail, but others contained limited information, as it is recognised that physicians are sometimes required to make decisions regarding treatment when information is incomplete.

The 30 case reports comprised six children (6–15 years) and 24 adults (18–90 years); 21 had epilepsy, and nine had an alternative diagnosis (Table 1). Examples of cases have been published elsewhere (Bergin et al., 2017).

1) how likely was it that the patient had experienced epileptic seizures;
2) the type of seizure, using the ILAE 2010 classification schema (Berg et al., 2010);
3) the etiology of the epilepsy, in broad categories, using the ILAE 2010 classification schema.

For the analysis performed here, there were three broad categories for diagnosis of epilepsy: Epilepsy; Possible Epilepsy; and Not Epilepsy.

Five broad categories were presented for classification of the attacks: Generalised seizures; Focal seizures; Epileptic seizures, but uncertain if focal or generalised; Turns/Attacks, possibly epileptic; Attacks not epileptic.

Etiology was classified according to the three major categories proposed in the ILAE 2010 report (genetic (presumed); structural/metabolic; unknown) (Berg et al., 2010) and Not Epilepsy.

In 23 cases, information was presented in two steps, with responses required after each step. Step 1 consisted of clinical information. Step 2 consisted of results of investigations (neuroimaging studies, EEGs and occasionally video monitoring). Investigators had the option of changing their responses from step 1 to step 2.

Inter-rater agreement was determined using the Kappa statistic to adjust for chance agreement. For the diagnosis of epilepsy (level of confidence that a patient had epilepsy) the overall agreement was calculated using the mean of all pairwise weighted kappa values. Pairwise weighted kappa values were obtained using Fleiss-Cohen weights (Fleiss and Cohen, 1973), giving the following values to the level of confidence that a patient had epilepsy: epilepsy = 2, possible epilepsy = 1, not epilepsy = 0. For seizure types and etiology, inter-rater agreement was calculated using Fleiss' kappa (Fleiss, 1971). Only the broad categories for seizure type and etiology were considered. Kappa values with 95% confidence intervals (CI) were calculated separately (and for each step separately) in the entire sample and in subgroups from differing geographic areas (Europe, North America, Latin America, Asia, and Oceania (Australia and New Zealand)), epileptologists vs. neurologists with a special interest in epilepsy, adult vs. child neurologists, subgroups with different levels of experience (< 15 vs. ≥ 15 years of experience) and investigators from countries where English was listed as an official language vs. those from other countries. Results for the steering committee were excluded. Kappa values were classified as poor (less than chance – kappa below 0), slight (0.01–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), almost perfect (0.81–1.00) (Landis and Koch, 1977). No formal statistical tests were performed. Differences between subgroups of investigators should be considered as statistically significant if 95% CI do not overlap.

The study commenced in December 2013, and continued until the end of 2014. Epileptologists, neurologists and paediatricians with an interest in epilepsy were invited to participate. The study was advertised at international meetings, via the ILAE website and through the ILAE chapters.

The Northern B New Zealand Ethics committee approved this study.

Participants from outside New Zealand who completed the set of 30 cases before the end of March 2014 entered a draw to win a holiday in New Zealand.

3. Results

189 participants from 36 countries completed all 30 cases. 159 of those who completed the study (83%) provided information about their professional role and experience with epilepsy (See Table 2). One medical student, one neurosurgeon, one primary-care physician and two physicians (not otherwise specified) completed the study, but their results have been removed from this analysis. The results from the remaining 154 participants comprise the data set for the analysis reported here. Fifty-three of the 154 participants described themselves as epileptologists and 56 as neurologists with a special interest in epilepsy. Other roles are shown in Table 2.

Kappa values for the sample of 154 investigators and for subgroups of investigators are shown in Table 3. The overall kappa value for the 154 investigators for a diagnosis of epilepsy at step 1 was 0.61 (*95% CI 0.609–0.613*); seizure type was 0.40 (*0.399–0.403*); and etiology was 0.41 *(0.406–0.410)*.

For the 23 cases with two steps, kappa values increased between steps from 0.56 *(0.559–0.563)* to 0.70 *(0.692–0.700)* for diagnosis of epilepsy; from 0.38 *(0.379–0.383)* to 0.52 *(0.514–0.518)* for seizure type; and from 0.38 *(0.380–0.384)* to 0.47 *(0.464–0.468)* for etiology.

Agreement was better for epileptologists kappa for diagnosis of epilepsy = 0.66, *(0.647-0.663)* than neurologists with a special interest

Table 1
Summary of cases.

Case	Age/Sex	Final Consensus Diagnosis/Etiology (EpiNet steering committee)	Seizure type (EpiNet steering committee)
1	28, M	Genetic Generalised Epilepsy, NOS	Generalised (Tonic-Clonic)
3	90, M	Focal seizure/Structural Metabolic − metastasis	Focal (Focal Motor)
4	20, F	Genetic Generalised Epilepsy (JME)	Generalised (Tonic-Clonic, Myoclonic)
5	20, F	Frontal lobe epilepsy/Unknown	Focal seizure (Frontal lobe seizure)
8	7, F	Genetic Generalised Epilepsy (CAE)	Generalised (Absence)
9	30, M	Occipital lobe epilepsy/Unknown	Focal (Occipital Lobe seizure)
10	36, F	Focal seizure/Structural Metabolic − AVM	Focal (Focal Motor)
11	17, F	Genetic Generalised Epilepsy (JME)	Generalised (Tonic-Clonic)
13	6, F	Focal seizure/Unknown	Focal (Rolandic Seizure)
14	22, M	Epilepsy-Nature Uncertain/Unknown	Epileptic seizure-Nature Uncertain/
15	16, M	Genetic Generalised Epilepsy (JME)	Generalised (Tonic-Clonic, Myoclonic)
17	48, M	Temporal lobe epilepsy/Unknown	Focal (Temporal lobe, evolving to TC)
18	21, M	Temporal lobe epilepsy/Unknown	Focal (Temporal lobe, evolving to TC)
20	13, M	Occipital lobe epilepsy/Unknown	Focal (Occipital Lobe seizure)
22	11, F	Genetic Generalised Epilepsy (eyelid myoclonia)	Generalised (eyelid myoclonia)
23	15, F	Generalised seizures/Unknown	Generalised (Tonic-Clonic)
24	50, F	Focal seizure/Structural Metabolic − Glioma	Focal (temporo-occipital seizure)
26	15, F	Genetic Generalised Epilepsy (JAE)	Generalised (Absence)
28	8, F	Focal seizure/Unknown	Focal (Rolandic Seizure)
29	85, M	Focal seizure/Structural Metabolic −Stroke	Focal seizure
30	32, F	Genetic Generalised Epilepsy (NOS)	Generalised (Tonic-Clonic)
2	51, M	Syncope	Not epilepsy
6	42, F	Psychogenic non-epileptic seizures	Not epilepsy
7	80, M	Syncope	Not epilepsy
12	23, F	Paroxysmal kinesogenic dyskinesia	Not epilepsy
16	43, F	Syncope	Not epilepsy
19	53, M	Sleep apnoea	Not epilepsy
21	51, F	Syncope	Not epilepsy
25	50, M	Psychogenic non-epileptic seizures	Not epilepsy
27	72, M	Sleep myoclonus, REM Sleep behaviour disorder	Not epilepsy

The seizure types, final consensus diagnoses and aetiologies as determined by the steering committee after Step 2 are listed here. Only the broad categories for seizure type and etiology were considered when calculating kappa values; more precise classification of seizure types or epilepsy syndromes (shown in brackets) were not considered when calculating kappa values.
Abbreviations;:AVM = arterio-venous malformation; CAE = Childhood absence epilepsy; JAE = Juvenile absence epilepsy; JME = Juvenile Myoclonic Epilepsy; NOS = not otherwise specified; REM = Rapid Eye Movement; TC = Tonic clonic

in epilepsy (0.58, (*0.568–0.584*)).

Investigators with less than 15 years of experience had a slightly better agreement (0.62, *(0.618–0.626)*) than investigators with longer experience (0.60, *(0.597–0.609)*); this difference was more pronounced for the cases where information was presented in two steps (0.72, *(0.711–0.723)* vs 0.67, *(0.666–0.682)*).

The kappa value for investigators from countries with English as an official language was greater (0.64, *(0.633–0.641)*) than for investigators from countries where English is not an official language (0.58, (*0.569–0.585*)).

Agreement on the diagnosis of epilepsy was greatest in participants from Oceania (0.67, (*0.664–0.680*)).

The kappa value for epilepsy diagnosis for the 88 investigators who exclusively looked after adults was 0.67 *(0.666–0.674)* while the kappa value for the 44 investigators who only looked after children was 0.56 *(0.550–0.566)*.

4. Discussion

4.1. Diagnosis of epilepsy

The 'gold standard' for diagnosing epilepsy is video-EEG recording; however, it is not practical to perform video-EEG recording in patients with infrequent episodes, and relatively few patients have this test before a diagnosis of epilepsy is made. In practice, the diagnosis of epilepsy is usually made on clinical grounds. Two aspects regarding diagnosis need to be considered; is the diagnosis correct? (accuracy, validity), and how consistently do different observers make the diagnosis? (interobserver variability, reliability) (van Donselaar et al., 2006). This paper solely addresses the levels of agreement between physicians (interobserver variability).

To make a clinical diagnosis, physicians need to obtain an accurate history from the patient, and, if possible a witness (D'Souza et al., 2010). Our study did not address clinicians' varying abilities in teasing out important aspects of the history. Instead, it was a pragmatic attempt to assess diagnostic variability between physicians from different countries and health systems, and everyone was therefore provided with ***exactly the same*** information. Many participants in this study would almost certainly have obtained more information from interviewing the patients than from the case scenarios, but some may have obtained less information. We deliberately included some 'grey' cases in this study because we wanted to see how investigators dealt with diagnostic uncertainty. Some of the patients in this study had experienced few attacks, and had normal or no investigations. These patients were included because the EpiNet study group is undertaking trials in newly diagnosed patients, and patients sometimes get started on AEDs before any tests are performed; we therefore wanted to see how much variability there was in the diagnosis of epilepsy in this group.

4.2. Factors determining levels of agreement

Levels of agreement improved substantially once results of investigations were provided to investigators. This is to be expected. Kappa values for diagnosis of epilepsy increased between steps 1 and 2 from 0.56 (moderate) to 0.70 (substantial), while kappa values for seizure type increased from 0.38 (fair) to 0.52 (moderate); kappa values increased from 0.38 (fair) to 0.47 (moderate) for etiology.

The magnitude of the kappa coefficient represents the level of agreement greater than expected by chance. The interpretation of this coefficient is not straightforward, as factors that can influence its magnitude include the prevalence index, i.e. the frequency of a given code in the entire diagnostic spectrum (Sim and Wright, 2005). If a

Table 2
Demographic details of Participants (N = 154).

	No	Percent
Gender		
Male	89	57.8
Female	65	42.2
Age		
< 30	3	2.0
30–40	46	29.9
40–50	61	39.6
50–60	31	20.1
> 60	13	8.4
Role		
Epileptologist	53	34.4
Neurologist with special interest in epilepsy	56	36.4
General neurologist	22	14.3
Paeditrician	6	3.9
Epilepsy fellow	6	3.9
Postgraduate neurology trainee	11	7.1
Number patients with epilepsy seen per week		
0–10	60	38.9
10–20	48	31.2
> 20	46	29.9
Years of experience with epileptic patients		
< 15	89	58.2
≥15	64	41.8
NA	1	
Nature of practice		
Outpatients with epilepsy	149	96.8
Inpatients with epilepsy	118	76.6
Patients being assessed for surgery	70	45.5
Clinical research in epilepsy	60	38.9
Laboratory research in epilepsy	10	6.5
Patients looked after		
Adults	110	71.4
Children	66	42.9
Infants	50	32.5
Region in which participants practice		
Oceania	51	33.1
Europe	47	30.5
North America	28	18.2
Asia	17	11.1
Latin America	9	5.8
Africa	2	1.3
Language		
English as official language	102	66.6
English not official language	52	33.3

Demographic details of participants whose results were included for calculation of kappa values.

condition is common (i.e. the prevalence index is high), it may be difficult to obtain an agreement substantially better than chance; in these circumstances, the kappa value may appear surprisingly low, even though actual agreement is reasonable. Since two thirds of the patients in this study had epilepsy, and there were only three diagnostic categories, one would expect moderate agreement simply by chance. Any kappa value greater than zero indicates agreement greater than can be accounted for by chance.

It is likely that some of the contribution to the lower kappa scores for seizure types and etiology relates to differing levels of confidence investigators require before making a specific diagnosis. Some investigators were prepared to commit to a specific etiology or seizure type after Step 1, while others waited for results of investigations. The inclusion of 'Uncertain' or 'Possible' options for all three categories may have produced lower kappa scores than if investigators had been given a forced-choice task, in which 'Uncertain' was not an option.

4.3. Varying seizure classifications

We asked investigators to classify patients' seizure types and etiology using the classification schema proposed by the ILAE Classification Commission in 2010 (Berg et al., 2010). This had been in use for only three years when this study was undertaken, and it had not been universally accepted (Avanzini, 2010; Commission International Classification, 1981; Ferrie, 2010; Fisher, 2010; Guerrini, 2010; Wolf, 2010). Since this study was conducted, the ILAE Classification Commission has further developed the classification (Fisher et al., 2016). For practical purposes, although there has been a change in terminology, there has been relatively little change in the classification of major seizure types, and we do not think the change in nomenclature of seizures will have had a significant effect on levels of agreement. Once consensus is finally reached regarding terminology, it would be of interest to undertake a further similar study looking at variability in diagnoses using the definitive classification.

4.4. Limitations

The study was biased towards adult neurologists as only six of the cases were children. Non-native English speakers were also at a disadvantage, as the cases were presented in English. These factors are not, in fact, independent of one another, as 45% of non-native English speakers in this study only look after children, whereas only 28% of those from English-speaking countries look after children exclusively. Unfortunately, it is not possible to do multivariate analysis to adjust kappa values for confounding factors, and we are not able to establish which factors are more or less important in determining observed differences between subgroups. We found a slight difference in levels of agreement according to years of experience, and this, too, might be explained by the higher prevalence of doctors from English speaking countries among those with less than 15 years of experience.

No definitive conclusions can be drawn regarding levels of agreement in different parts of the world, as, apart from New Zealand, relatively few specialists from any country participated in this study. The highest level of agreement was found with investigators from Australia and New Zealand, where subspecialty training in epilepsy is reasonably standardised under a single college. There is also a cultural bias to the entire study. The EpiNet website has been developed in New Zealand, and most of the cases were seen in the New Zealand public health service.

This is by far the largest study of its type ever undertaken. Despite this, the conclusions should not be generalised. The neurologists and epileptologists who participated are not necessarily representative of other physicians working in their countries or in their particular subspecialties, as they were self-selected. We think it is likely that those who did participate will have had higher levels of agreement than their colleagues who did not participate, though we cannot be certain of this.

Some physicians participated in this study because they hoped to win a holiday in New Zealand. We do not think this invalidates our findings. While the chance to win a holiday may have acted as an incentive, there is no reason to think that physicians' responses to the case scenarios will have been altered by the chance to win this prize.

4.5. Results in context of previous research

Our findings are generally consistent with those previously reported.

Three neurologists obtained moderate agreement when determining whether a first paroxysmal event was epileptic in nature (kappa = 0.58) (van Donselaar et al., 1989), while six doctors obtained kappa values between 0.40 and 0.63 when determining if episodes of collapse were syncope or seizures (Hoefnagels et al., 1992). Twenty two epileptologists obtained kappa values of 0.56 when making diagnoses (epilepsy, psychogenic non-epileptic seizures or physiologic non-epileptic events) by reviewing 22 video-recorded episodes (Benbadis et al., 2009).

There is a wider range of levels of agreement for specific seizure types. Six neurologists obtained a kappa score of 0.26 when determining whether seizures were focal or generalised from onset (van

Table 3
Kappa values.

Investigators	Form	All cases (30)		Cases with 2 steps (23) Step 1		Cases with 2 steps (23) Step 2	
		K	95% CI	K	95% CI	K	95% CI
All (N = 154)							
	Diagnosis	0.611	0.609–0.613	0.561	0.559–0.563	0.696	0.692–0.700
	Seizure type	0.401	0.399–0.403	0.381	0.379–0.383	0.516	0.514–0.518
	Etiology	0.408	0.406–0.410	0.382	0.380–0.384	0.466	0.464–0.468
Epileptologists (N = 53)							
	Diagnosis	0.655	0.647–0.663	0.615	0.607–0.623	0.738	0.730–0.746
	Seizure type	0.424	0.418–0.430	0.402	0.396–0.408	0.548	0.542–0.554
	Etiology	0.454	0.448–0.460	0.427	0.421–0.433	0.527	0.521–0.533
Neurologists with a special interest in epilepsy (N = 56)							
	Diagnosis	0.576	0.568–0.584	0.503	0.495–0.511	0.644	0.636–0.652
	Seizure type	0.419	0.413–0.425	0.392	0.386–0.398	0.537	0.531–0.543
	Etiology	0.409	0.403–0.415	0.374	0.368–0.380	0.459	0.453–0.465
< 15 years of experience with epileptic patients (N = 89)							
	Diagnosis	0.622	0.618–0.626	0.570	0.564–0.576	0.717	0.711–0.723
	Seizure type	0.396	0.392–0.400	0.376	0.372–0.380	0.507	0.503–0.511
	Etiology	0.413	0.409–0.417	0.387	0.383–0.391	0.467	0.463–0.471
≥15 years of experience with epileptic patients (N = 64)							
	Diagnosis	0.603	0.597–0.609	0.557	0.549–0.565	0.674	0.666–0.682
	Seizure type	0.411	0.407–0.415	0.392	0.386–0.398	0.530	0.524–0.536
	Etiology	0.402	0.398–0.406	0.376	0.370–0.382	0.464	0.458–0.470
Look after adults only (N = 88)							
	Diagnosis	0.670	0.666–0.674	0.635	0.629–0.641	0.765	0.759–0.771
	Seizure type	0.429	0.425–0.433	0.409	0.405–0.413	0.549	0.545–0.553
	Etiology	0.447	0.443–0.451	0.430	0.426–0.434	0.499	0.495–0.503
Look after children/infants only (N = 44)							
	Diagnosis	0.558	0.550–0.566	0.485	0.475–0.495	0.629	0.619–0.639
	Seizure type	0.377	0.371–0.383	0.359	0.351–0.367	0.493	0.485–0.501
	Etiology	0.367	0.361–0.373	0.329	0.321–0.337	0.429	0.421–0.437
Oceania (N = 51)							
	Diagnosis	0.672	0.664–0.680	0.650	0.642–0.658	0.785	0.775–0.795
	Seizure type	0.440	0.434–0.446	0.427	0.421–0.433	0.575	0.567–0.583
	Etiology	0.429	0.423–0.435	0.416	0.410–0.422	0.481	0.475–0.487
Europe (N = 47)							
	Diagnosis	0.617	0.609–0.625	0.559	0.549–0.569	0.735	0.725–0.745
	Seizure type	0.416	0.410–0.422	0.394	0.388–0.400	0.567	0.559–0.575
	Etiology	0.421	0.415–0.427	0.396	0.388–0.404	0.500	0.492–0.508
North America (N-28)							
	Diagnosis	0.640	0.626–0.654	0.592	0.576–0.608	0.688	0.672–0.704
	Seizure type	0.367	0.357–0.377	0.344	0.334–0.354	0.438	0.426–0.450
	Etiology	0.416	0.406–0.426	0.384	0.372–0.396	0.474	0.462–0.486
Asia (N = 17)							
	Diagnosis	0.581	0.556–0.606	0.499	0.472–0.526	0.602	0.573–0.631
	Seizure type	0.406	0.390–0.422	0.387	0.367–0.407	0.488	0.468–0.508
	Etiology	0.382	0.364–0.400	0.332	0.312–0.352	0.417	0.397–0.437
Latin America (N = 9)							
	Diagnosis	0.595	0.550–0.640	0.503	0.452–0.554	0.588	0.535–0.641
	Seizure type	0.417	0.382–0.452	0.395	0.356–0.434	0.509	0.468–0.550
	Etiology	0.387	0.352–0.422	0.370	0.331–0.409	0.421	0.382–0.460
English-speaking countries (N = 102)							
	Diagnosis	0.637	0.633–0.641	0.599	0.595–0.603	0.710	0.706–0.714
	Seizure type	0.406	0.404–0.408	0.387	0.383–0.391	0.506	0.502–0.510
	Etiology	0.419	0.417–0.421	0.393	0.389–0.397	0.469	0.465–0.473
Non English-speaking countries (N = 52)							
	Diagnosis	0.577	0.569–0.585	0.506	0.498–0.514	0.674	0.664–0.684
	Seizure type	0.395	0.389–0.401	0.372	0.366–0.378	0.535	0.529–0.541
	Etiology	0.386	0.380–0.392	0.361	0.355–0.367	0.459	0.453–0.465

Kappa values have been calculated for the 154 participants who provided details of their clinical experience, and subgroups.

Donselaar et al., 1990). One senior neurologist and three neurology residents obtained kappa scores of 0.24–0.38 when determining seizure types from verbatim descriptions of seizures (Bodensteiner et al., 1988). Five neurologists obtained kappa scores of 0.69 when determining seizure types by review of case histories for 21 of 27 patients admitted to an epilepsy monitoring unit (Beniczky et al., 2012).

Three paediatric neurologists obtained kappa scores of 0.31 when determining seizure types using the 2010 seizure classification (van Campen et al., 2013).

4.6. Implications

Drug trials often recruit patients with particular seizure types – e.g. focal seizures – but may not state how diagnoses should be made. Differences in diagnostic accuracy could be an explanation for negative results in some clinical trials of anti-epileptic drugs. Recent trials of perampanel (French et al., 2012), carisbamate (Sperling et al., 2010), and brivaracetam (Biton et al., 2014; Ryvlin et al., 2014), all found differing effect size in trials of similar design conducted in different parts of the world. If trials conducted in some centres enrolled patients who did not actually have focal seizures, this would have diluted the

effect size and reduced statistical power. The possibility that physicians from different parts of the world use different diagnostic criteria needs to be acknowledged. Formal processes should be established to ensure that appropriate patients are recruited into trials. In addition, the ILAE needs to continue efforts to educate physicians regarding diagnostic criteria for seizure types and aetiologies.

4.7. Future directions

This study is also being used to accredit investigators for the EpiNet-First studies, which are being conducted in people with newly diagnosed epilepsy (www.epinet.co.nz) (Bergin et al., 2015) Details regarding the accreditation process are described in a separate paper (Bergin et al., 2017) Many of the investigators who participated in this study have been accredited, on the basis of the results from this validation study. We would suggest that other collaborative study groups also conduct studies to determine the levels of diagnostic agreement between participants in multicentre studies.

5. Conclusion

There is substantial agreement regarding the diagnosis of epilepsy when physicians who have an interest in epilepsy interpret the same clinical data. Agreement regarding seizure type and etiology is only fair to moderate when based on history alone; however, agreement improves significantly when results of investigations are considered.

Funding

This study was funded by the Neurological Foundation of New Zealand and the New Zealand Health Research Council. The prize of a self-drive holiday in New Zealand was offered by Fine Tours New Zealand, with travel to New Zealand provided by the Julius Brendel trust.

Disclosures

None of the authors has any conflict of interest to disclose.

Dr PS Bergin received grants for the development of the EpiNet platform from the Health Research Council (NZ), the Neurological Foundation of New Zealand, the A + Charitable trust, the Julius Brendel trust, CSL Biotherapeutics, Sanofi, Enigma Publishing, and Cyberonics; and funding for a trip from Sanofi.

Dr E Beghi received personal fees from Viropharma to participate in an Advisory Board, research grants from UCB-Pharma, EISAI, the Italian Drug Agency (AIFA), the Italian Ministry of Health, the American ALS Association, and the Borgonovo Foundation, outside the submitted work.

Dr LG Sadleir reports no disclosures.

Ass Professor M Tripathi reports no disclosures.

Professor MP Richardson reports no disclosures.

Dr E Bianchi reports no disclosures.

Ass Professor WJ D'Souza has received travel, investigator-initiated, and speaker honoraria from UCB Pharma, educational grants from Novartis Pharmaceuticals and Pfizer Pharmaceuticals, educational, travel and fellowship grants from GSK Neurology Australia, and honoraria from SciGen Pharmaceuticals.

Acknowledgements

This study was funded by the Neurological Foundation of New Zealand and the New Zealand Health Research Council. The prize of a self-drive holiday in New Zealand was offered by Fine Tours New Zealand, with travel to New Zealand provided by the Julius Brendel trust.

Appendix A

Members of the EpiNet Study group who completed the 30 cases, listed by country and then alphabetically by family name.

Argentina
Analía Calle
Australia
Anita Cairns
Patrick Carney
Donald Craig
Deepak Gill
Sachin Gupta
Cecilie Lander
Hanka Laue-Gizzi
Michelle Kiley
Nicholas Lawn
Elizabeth Reyneke
Kate Riney
Meng Tan
Michael Tan
Mark Thieban
Emma Whitham
Chong Wong
Belgium
Benjamin Legros
Michel Ossemann
Germainvan Rijckevorsel
Brazil
Ana Gabriela Ferrari Strang
Angela Gifoni
Linden Helio
Bruno Monnerat
Canada
Paula Brna
Elizabeth Donner
Stephanie Jacques
Nathalie Jette
Richard McLachlan
Ismail Mohamed
Thi Phuoc Yen Tran
China
Song Fan
Yang Guang
Ming Li
Kang Wang
Shouwen Zhang
Colombia
Lady Ladino
Denmark
Margarethe Sophie Kölmel
Finland
Annukka Uusitalo
Paivi Vieira
France
StephaneAuvin
Georgia
Maia Alkhidze
Tamar Ediberidze
Nino Gogatishvili
Germany
Dieter Dennig
Anja Grimmer
Rosa Michaelis
Susanne Schubert-Bast
Caspar Stephani
Stefan Stodieck

Martin Vollbrandt
Andreas Zellner
Greece
DimitriosZafeiriou
Hungary
Andras Fogarasi
Peter Halasz
India
Satish Jain
Raj Nair
Pravar Passi
Surekha Rajadhyaksha
Sita Jayalakshmi Sattaluri
Kavita Srivastava
Vrajesh Udani
Ireland
Daniel Costello
Italy
Umberto Aguglia
Paolo Benna
Simone Beretta
Edoardo Ferlazzo
Alberto Spalice
Pasquale Striano
Alberto Verrotti
Clara Zanchi
Jamaica
Amza Ali
Malaysia
Kheng Seang Lim
Hui Jan Tan
Mexico
Alfredo Ramirez
Ildefonso Rodriguez-Leyva
New Zealand
Neil Anderson
Alan Barber
Pietro Cariga
James Cleland
Nicholas Child
Suzanne Davis
Viswas Dayal
Cameron Dickson
Roderick Duncan
Richard Frith
Pratima Giri
Michael Herd
David Hutchinson
Ivan Iniesta
Jayaganth Jayabal
Bethany Jones
Dean Kilfoyle
Nicole McGrath
John Mottershead
Colette Muir
Melinda Nolan
Jennifer Pereira
Anna Ranta
Ian Rosemergy
Sneha Sadani
Mark Simpson
Claire Spooner
Paul Timmings
Elizabeth Walker
Diana Wei
Ernest Willoughby
Edward Wong
Teddy Wu
Nigeria
Birinus Ezeala-Adikaibe
Talabi Olusola
Pakistan
Hiba Mahmud
Zarine Mogul
Peru
Julio Espinoza
Jose Hernandez Vizarreta
Portugal
Rute Teotónio
Serbia
Bosanka Jocic-Jakubi
Stevo Lukic
Slovenia
Marko Korošec
Tomaz Zgur
Spain
María Gómez Eguílaz
Sweden
Fredrik Asztely
Thailand
Pasiri Sithinamsuwan
United Kingdom
Joseph Anderson
Pauls Auce
Archana Desurkar
Andrew Kelso
Violeta Sanchez
Aurangzeb Sidra
Phil Smith
Tim Wehner
Gavin Winston
United States of America
Edgard Andrade
Michelle Boudreau
Kevin Chapman
Geetha Chari
Brian Droker
Mirret El-Hagrassy
Dawn Eliashiv
Christi Heck
Arif Kabir
Dmitriy Kolesnik
Alice Lam
Jonathan Lopez
Tammaa Maamoon
Rama Maganti
Chinasa Nwankwo
Kristen Park
Simona Proteasa
Evan Sandok
Syndi Seinfield
Julia Toub
Vietnam
Truong Tran Thien

References

Avanzini, G., 2010. A sound conceptual framework for an epilepsy classification is still lacking. Epilepsia 51, 720–722.
Benbadis, S.R., LaFrance Jr., W.C., Papandonatos, G.D., Korabathina, K., Lin, K., Kraemer, H.C., Treatment Workshop, N.E.S., 2009. Interrater reliability of EEG-video monitoring. Neurology 73, 843–846.
Beniczky, S.A., Fogarasi, A., Neufeld, M., Andersen, N.B., Wolf, P., van Emde Boas, W.,

Beniczky, S., 2012. Seizure semiology inferred from clinical descriptions and from video recordings. How accurate are they? Epilepsy Behav. 24, 213–215.
Berg, A.T., Levy, S.R., Testa, F.M., Shinnar, S., 1999. Classification of childhood epilepsy syndromes in newly diagnosed epilepsy: interrater agreement and reasons for disagreement. Epilepsia 40, 439–444.
Berg, A.T., Berkovic, S.F., Brodie, M.J., Buchhalter, J., Cross, J.H., van Emde Boas, W., Scheffer, I.E., 2010. Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005–2009. Epilepsia 51, 676–685.
Bergin, P.S., Frith, R., Walker, E., Timmings, P., 2007. How to get the answer to nearly everything: using the internet for epilepsy research. Epilepsia 48, 1415–1417 discussion 1417–1424.
Bergin, P.S., Sadleir, L., Legros, B., Mogal, Z., Tripathi, M., Dang, N., et al., 2012. for the EpiNet Study Group.,. An international pilot study of an Internet-based platform to facilitate clinical research in epilepsy: the EpiNet project. Epilepsia 53, 1829–1835.
Bergin, P.S., Beghi, E., Tripathi, M., Smith, P., D'Souza, W.J., 2015. Invitation to participate in the EpiNet-First trials. Epilepsia 56, 807.
Bergin, P.S., Beghi, E., Sadleir, L., Brockington, A., Tripathi, M., Richardson, M., for the EpiNet Study Group, et al., 2017. EpiNet as a way of involving more physicians and patients in epilepsy research; validation study and accreditation process. Epilepsia Open 2, 20–31.
Biton, V., Berkovic, S.F., Abou-Khalil, B., Sperling, M.R., Johnson, M.E., Lu, S., 2014. Brivaracetam as adjunctive treatment for uncontrolled partial epilepsy in adults: a phase III randomized, double-blind, placebo-controlled trial. Epilepsia 55, 57–66.
Bodensteiner, J.B., Brownsworth, R.D., Knapik, J.R., Kanter, M.C., Cowan, L.D., Leviton, A., 1988. Interobserver variability in the ILAE classification of seizures in childhood. Epilepsia 29, 123–128.
Commission International Classification, 1981. Proposal for revised clinical and electroencephalographic classification of epileptic seizures. Commission on Classification and Terminology of the International League Against Epilepsy. Epilepsia 22, 489–501.
D'Souza, W.J., Stankovich, J., O'Brien, T.J., Bower, S., Pearce, N., Cook, M.J., 2010. The use of computer-assisted-telephone-interviewing to diagnose seizures, epilepsy and idiopathic generalized epilepsy. Epilepsy Res. 9, 20–27.
Ferrie, C.D., 2010. Terminology and organization of seizures and epilepsies: radical changes not justified by new evidence. Epilepsia 51, 713–714.
Fisher, R., Cross, J., French, J., Higurashi, N., Hirsch, E., Jansen, F., et al., 2016. Operational Classification of Seizure Types by the International League Against Epilepsy. http://www.ilae.org/Visitors/Centre/documents/ClassificationSeizureILAE-2016.pdf.
Fisher, R.S., 2010. What is a classification essay? Epilepsia 51, 714–715.
Fleiss, J.L., Cohen, J., 1973. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ. Psychol. Meas. 33, 613–619.
Fleiss, J.L., 1971. Measuring nominal scale agreement among many raters. Psychol. Bull. 76, 378–382.
French, J.A., Krauss, G.L., Biton, V., Squillacote, D., Yang, H., Laurenza, A., et al., 2012. Adjunctive perampanel for refractory partial-onset seizures: randomized phase III study 304. Neurology 79 (6), 589–596.
Friedman, D., French, J.A., 2012. Clinical trials for therapeutic assessment of antiepileptic drugs in the 21 st century: obstacles and solutions. Lancet Neurol. 11, 827–834.
Guerrini, R., 2010. Classification concepts and terminology: is clinical description assertive and laboratory testing objective? Epilepsia 51, 718–720.
Hoefnagels, W.A., Padberg, G.W., Overweg, J., Roos, R.A., 1992. Syncope or seizure?: A matter of opinion. Clin. Neurol. Neurosurg. 94, 153–156.
Kellinghaus, C., Loddenkemper, T., Najm, I.M., Wyllie, E., Lineweaver, T., Nair, D.R., Luders, H.O., 2004. Specific epileptic syndromes are rare even in tertiary epilepsy centers: a patient-oriented approach to epilepsy classification. Epilepsia 45, 268–275.
Landis, J.R., Koch, G.G., 1977. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 33, 363–374.
Ottman, R., Hauser, W.A., Stallone, L., 1990. Semistructured interview for seizure classification: agreement with physicians' diagnoses. Epilepsia 31, 110–115.
Rinaldi, G., Zarrelli, M.M., Beghi, E., Apollo, F., Germano, M., Viesti, P.D., Simone, P., 2000. The international classification of the epilepsies and epileptic syndromes. An algorithm for its use in clinical practice. Epilepsy Res. 41, 223–234.
Ryvlin, P., Werhahn, K.J., Blaszczyk, B., Johnson, M.E., Lu, S., 2014. Adjunctive brivaracetam in adults with uncontrolled focal epilepsy: results from a double-blind, randomized, placebo-controlled trial. Epilepsia 55, 47–56.
Sim, J., Wright, C.C., 2005. The kappa statistic in reliability studies: use, interpretation, and size requirements. Phys. Ther. 85, 257–268.
Sperling, M.R., Greenspan, A., Cramer, J.A., Kwan, P., Kalviainen, R., Halford, J.J., et al., 2010. Carisbamate as adjunctive treatment of partial onset seizures in adults in two randomized, placebo-controlled trials. Epilepsia 51, 333–343.
Stroink, H., van Donselaar, C.A., Geerts, A.T., Peters, A.C., Brouwer, O.F., van Nieuwenhuizen, for the Dutch Study of Epilepsy in Childhood, et al., 2004. Interrater agreement of the diagnosis and classification of a first seizure in childhood. The Dutch Study of Epilepsy in Childhood. J. Neurol. Neurosurg. Psychiatry 75, 241–245.
van Campen, J.S., Jansen, F.E., Brouwer, O.F., Nicolai, J., Braun, K.P., 2013. Interobserver agreement of the old and the newly proposed ILAE epilepsy classification in children. Epilepsia 54, 726–732.
van Donselaar, C.A., Geerts, A.T., Meulstee, J., Habbema, J.D., Staal, A., 1989. Reliability of the diagnosis of a first seizure. Neurology 39, 267–271.
van Donselaar, C.A., Stroink, H., Arts, W.F., for the Dutch Study of Epilepsy in Childhood, 2006. How confident are we of the diagnosis of epilepsy? Epilepsia 47 (Suppl. 1), 9–13.
Wolf, P., 2010. Much ado about nothing? Epilepsia 51 (4), 717–718.