Diagnostic reference frames for seizures: a validation study

J Neurol (2006) 253 : 372–376
DOI 10.1007/s00415-005-0011-0

ORIGINAL COMMUNICATION

J. F. van Ast
W. O. Renier
J. L. Talmon
J. M. A. Roos
A. Hasman

Diagnostic Reference Frames for Seizures:
A validation study

Received: 23 January 2005
Received in revised form: 9 May 2005
Accepted: 18 May 2005
Published online: 14 November 2005

J. F. van Ast (✉) · J. L. Talmon · J. M. A. Roos
University of Maastricht
Research Institute Caphri
Department of Medical Informatics
P. O. Box 616
6200 MD Maastricht, The Netherlands
Tel.: +31 (0) 433882240
Fax: +31 (0) 433884170
E-Mail: w.vanast@mi.unimaas.nl

W. O. Renier
Department of Child Neurology
University Medical Center St. Radboud
P. O. Box 9101
CZZO 326
6500 HB, Nijmegen, The Netherlands

A. Hasman
Department of Medical Informatics
Academic Medical Center Amsterdam
Meibergdreef 15
1105 AZ
Amsterdam Z.O, The Netherlands

Abstract *Introduction* We developed structured descriptions of signs and symptoms for specific seizure types (called Diagnostic Reference Frames – DRFs – by us) that can serve as a frame of reference in the process of classifying patients with epileptic seizures. In this study the validity of the DRFs for clinical use is evaluated and described. *Material and methods* In this study we use a decision support system based on the DRFs and using Bayes's rule for the validation of the DRFs. Patient's manifestations are entered in the decision support system and by successively applying Bayes's rule posterior probabilities are calculated. The DRFs with the highest posterior probability gives an indication of the classification of the seizure. The validation of the DRFs was performed by comparing the seizure type with the highest posterior probability with the classification of experienced epileptologists on a series of test cases with known epileptic seizures. In this way we assessed the accuracy of the DRFs in classifying patients with epileptic seizures. *Results* We included sixty-six patients in this efficacy study. The patients and/or their relatives described the manifestations occurring during a seizure. Sixty cases (91 %) were correctly classified using the decision support system. *Discussion* The accuracy of 91 % indicates that the knowledge encoded in the DRFs for the included seizure types is valid. The next step is to test the DRFs in a clinical setting to evaluate the applicability in daily practice.

Key words epilepsy · clinical features · computer-assisted diagnosis · classification · seizure

Introduction

Since the 1960s the International League Against Epilepsy (ILAE) has successfully developed several classification systems for epileptic seizures [1, 2]. The most recent accepted International Classification of Epileptic Seizures (ICES) dates from 1981 [1]. This classification provides rather general descriptions of the seizure types, which has led to variability in its use [3–6]. It has been shown that the use of well defined and clearly agreed upon definitions of seizures reduces the inter-observer variability [4].

We developed structured descriptions of epileptic seizures, called Diagnostic Reference Frames – DRFs – by us. A DRF consists of a list of manifestations. For each manifestation a conditional probability is given. A high probability e. g. 0.9 indicates that a manifestation is frequently occurring. A low probability e. g. 0.05 indicates that a manifestation is seldom occurring in the seizure type. The same list of manifestations is used across all included seizure types. Neurologists/epileptologists

provided estimates of these frequencies of occurrence on a 5-point scale from which the probabilities for the DRFs were derived. DRFs were developed for the most common seizure types: simple partial seizures, complex partial seizures, myoclonic seizures, tonic-clonic seizures and typical absence seizures. A DRF can serve as a frame of reference in the process of classifying patients with epileptic seizures.

The knowledge included in the DRFs needs to be verified and validated before it can be used in routine clinical practice. Verification refers to the internal consistency of the data. It can be performed without test cases [7]. It has already been shown, using the intraclass correlation coefficient (ICC) that the estimated frequencies by the experts for each seizure type are consistent [8] and that the DRFs are face-valid.

Validation is the "act of comparing properties of an object with the stated goal as a frame of reference concluding on degree of fulfillment" [9]. In our case the validation of the DRFs is performed by comparing the classifications of a decision support system based on the DRFs with the classification of experienced epileptologists using a series of test cases with known epileptic seizures.

This article describes the results of our validation study. In the following section we first introduce the decision support system in more detail. Next we describe the study design and the results. In the discussion we reflect on the observed accuracy of the DSS and the limitations of our study and the used DSS.

In the following sections the term (seizure) manifestations will be used to indicate both signs and symptoms.

Material and methods

Material

Diagnostic Seizure Support

As mentioned in the introduction we developed a decision support system, which we call Support in Classification of Seizures (SICS). SICS is a stand-alone system that provides assistance in classifying patients with epileptic seizures. The current system is discriminating between five seizure types: simple partial seizures (SPS), complex partial seizures (CPS), typical absence seizures (TAS), myoclonic seizures (MS) or tonic-clonic seizures (TCS).

In SICS the user fills out the administrative patient data: name, gender, date of birth and consultation date. Then the seizure manifestations are presented to the user in seven clusters: 'seizure characteristics', 'pre-ictal phenomena', 'motor signs and symptoms', 'non-motor symptoms', 'consciousness', 'somatic reactions', and 'post-ictal phenomena'. The user can report the manifestations in the order he/she wants.

The role of SICS is to provide a classification of a seizure based on the patient's manifestations. The system uses Bayes' rule to compute posterior probabilities for the included seizure types based on the a priori probabilities and the conditional probabilities for the manifestations reported. We used the normalized incidence rates of the seizure types as a priori probabilities. The system assumes independence among the manifestations. For each seizure type the system provides a probability and the seizure type with the highest probability is considered to be the correct classification.

Each time new information is entered the system updates the posterior probabilities for the different seizure types using Bayes' rule. If the patient does not know the value of a manifestation, the user can report this manifestation as 'unknown' and the posterior probabilities are not updated.

There is also a function available that determines the most informative manifestation that, on average, changes the probability distribution over the included seizure types. This function takes the information gathered so far into account (advice-function). This allows a sequential elicitation of the most informative manifestations [10, 11]. Further, the user has the possibility of examining and printing the reported manifestations.

Protocol for systematic registration of manifestations in SICS

For a systematic registration of seizure manifestations we developed a protocol. After the registration of the patient data the user reports the status of four key manifestations: 'onset of seizure', 'duration of seizure', 'movements' and 'consciousness' (step 1). Then the system is asked to generate the most informative manifestation and the user reports the answer of the patient or his/her relative(s) for this manifestation. Then the next informative manifestation is determined and presented to the user. Ten times in a row the most informative manifestation is determined and the value entered into the system (step 2). Finally the user can enter manifestations mentioned by the patient or his/her relative(s) and not suggested by the system in the order he/she wants (step 3) so as to have a more complete registration of the seizure.

Patient cases

The patients were selected by three experienced epileptologists from their own patient population.

To include the patients the following criteria were used: 1) the patient was at least 12 years old; 2) the patient was still under treatment of the epileptologists; 3) the epileptologist diagnosed the seizure type and 4) the patient was willing to cooperate.

When a patient was willing to participate the researcher (WvA) interviewed the patient or his/her relative(s) about the manifestations occurring during a seizure following the protocol. This interview took place directly after the consultation with the epileptologist or at a later time by telephone. The researcher was blinded for the diagnosis of the patient prior to the interview except for the MS cases. One of the epileptologists was asked to provide these cases and only such cases as to increase the number of cases with MS in our test set.

Methods

To determine the accuracy of the classification of the system the seizure type with the highest posterior probability was compared with the seizure type diagnosed by the epileptologist who treated the patient. The diagnosis made by the epileptologist was assumed to be correct and serves as a reference standard for the classifications provided by the system.

Further we assessed how the posterior probabilities changed after each step in the protocol.

When the system's classification differed from the reference standard the patient case was evaluated by one of the epileptologists (WR). The manifestations reported for these cases were presented to the epileptologist. He was asked to give his classification based on the reported manifestations.

Results

Description of patient cases

We included sixty-six patients. The patients and/or their relatives described the manifestations occurring during a seizure. Table 1 gives an overview of how many patients were included for each seizure type.

Accuracy of DSS

Sixty of the patient cases (91 %, 95 % confidence interval: 84 %–98 %) were correctly classified by the system after the registration of all elicited manifestations (Table 1). In 50 of the 60 correctly classified cases the posterior probability for the correct seizure type was 0.90 or higher (Table 2).

For six cases the system did not provide the correct classification. Based on the manifestations reported by the patients or his/her relatives the system provided classification other than that of the epileptologists. One epileptologist evaluated these six patient cases as well as one case that was correctly classified by the system. Based on the reported manifestations for these seven cases the epileptologist classified the same seizure type as the system in 5 cases. For the case which was correctly classified by the system, the epileptologist also provided the correct seizure type.

Evaluation of the protocol

Based on the four key manifestations, 46 (70 %) of the cases were classified correctly. By asking the ten most informative manifestations 8 more cases (82 %) were correctly classified (Table 1). Reporting additional manifestations resulted in a further increase in the number of correctly classified seizures: one MS case and five CPS cases. For TAS the number of correctly classified seizures did not change when manifestations additional to the key manifestations were reported.

Table 2 Distribution of final posterior probabilities for seizure types correctly classified by SICS

	Final posterior probabilities				
	> 0.5	> 0.6	> 0.7	> 0.8	> 0.9
SPS	5	5	5	5	5
CPS	9	8	5	4	4
MS	13	13	12	12	12
TCS	28	28	28	26	24
TAS	5	5	5	5	5
Total	60	59	56	52	50

Discussion

Study design

We validated the DRFs in an experimental setting. The patient was asked to tell his/her seizure manifestations to the researcher after the consultation with the treating physician. This does not reflect the normal consultation setting, but more closely resembles the situation in which a (specialist) nurse interviews the patient to elicit relevant clinical information before seeing the physician. The experimental setting in which the researcher is blinded for the classification of the seizure allows us to determine the accuracy of the system under optimal conditions (efficacy test) using the classification of the epileptologists as a reference [12].

Case selection was limited to two locations: an academic neurology department and a specialized epilepsy center. For logistic reasons, neurologists see their patients only on a specific day in the week and patients only visit the epileptologist once or twice a year, data collection took nearly one year. Including a larger population would have required a larger number of participating centers.

Accuracy of DSS

The performance of the system was 91 %. Compared with other decision support systems in neurology the

Table 1 Number of correctly classified seizure types after reporting key, most informative and additional manifestations

	Total # of patients included	Key manifestations	Informative manifestations	Additional manifestations	# of wrong diagnosis
SPS	7	3 (43 %)	5 (71 %)	5 (71 %)	2
CPS	10	4 (40 %)	4 (40 %)*	9 (90 %)	1
MS	15	9 (60 %)	12 (80 %)	13 (87 %)	2
TCS	28	25 (89 %)	28 (100 %)	28 (100 %)	0
TAS	6	5 (83 %)	5 (83 %)	5 (83 %)	1
Total	66	46 (70 %)	54 (82 %)	60 (91 %)	6

* For a discussion of the results for CPS see section Evaluation of the protocol

accuracy of SICS is good, even taking into account the relatively wide 95 % confidence interval due to the small number of cases. In Epilepsy Expert, a decision support system based on the International Classification of Epilepsies and Epileptic Syndromes [13], 18 of the 25 (72 %) patient cases were correctly diagnosed when the diagnosis was based on clinical manifestations solely [14]. Diagnoses given by MICROSTROKE, an expert system for stroke type diagnosis based on clinical information, have been tested for conformity with the final diagnosis of 250 cases in the Hamburg Stroke Data Bank and were found to be correct in 72.8 % of the 250 cases [15].

Although SICS performs better then the other two decision support systems in neurology it is not yet proven that an accuracy of 91 % is sufficient for clinical practice. One must also take into account the therapeutic consequences of a wrong classification. For the seizure types included in our system two groups of seizure types can be distinguished, each treated with a different therapy. The first group consists of SPS, CPS and TCS and the second group consists of TAS and MS. When the classification is wrong but falls within the same group, there are no therapeutic consequences. In our study there are two cases for which the classification of the system and the correct diagnosis fall in a different therapeutic group. This indicates that only for those two cases the classification would have therapeutic consequences.

In daily practice the physician also is not always making a final classification of the seizure during the first consultation. An evaluation of the manifestations and further diagnosis during the next consultations is always needed.

Evaluation of the protocol

Table 1 shows that the accuracy of SICS increased when *informative* manifestations are reported. Reporting *additional* manifestations seem to add less to the accuracy of SICS. This indicates that when key and *informative* manifestations are reported completely *additional* manifestations do not have to be reported any more. Only for CPS reporting *additional* manifestations did seem necessary. The results reported in this paper are based on a reanalysis of the data with the most recent version of SICS. We needed to update the conditional probabilities halfway through our study. For several CPS cases this resulted in different most *informative* manifestations. For many of these most *informative* manifestations we had to report 'unknown' because these manifestations were not asked about during the initial interview. Thus the posterior probabilities after the most *informative* manifestations were based on insufficient information. When all manifestations registered during the initial interview were reported as *additional* manifestations an increase of the posterior probabilities for CPS was observed. Therefore one should ignore the number of correct classifications after the *informative* manifestations in Table 1 for CPS. For the other seizure types this effect is not observed because the most *informative* manifestations remained largely the same although the sequence in which they were asked changed.

After each step in the protocol we determined a threshold for the maximum posterior probability such that no wrong classifications are made. A posterior probability for a seizure type > 0.97 after reporting key manifestations indicated a correct classification. When this threshold is not reached one should continue to elicit the most informative manifestations. After all informative manifestations are recorded a posterior probability for a seizure type > 0.91 indicates again a correct classification. For the remaining cases additional manifestations should be reported. As soon as a posterior probability reaches 0.91 one can stop the elicitation process. When there are no manifestations to be reported, e. g. because the patient or relative cannot provide a more detailed description and the probability is less than 0.91, the classification may not be correct. Again we conclude that additional manifestations seem to add less to the accuracy of DSS, unless the posterior probability of 0.91 is not reached after the informative manifestations.

The key and informative manifestations provide highly relevant information to obtain the correct classification. These manifestations should always be asked of the patient or relative in classifying seizures.

From the analysis of the cases that were classified incorrectly by the system and also incorrectly classified by the epileptologist based on the reported manifestations it seems that the patient's description of the manifestations was incomplete for these cases or that other information has provided evidence for the treating neurologist to make another diagnosis.

Limitations of SICS

Only five seizure types were included in the system. These five selected seizure types are the most frequently occurring seizure types in adolescents and adults.

Eleven patients were suffering from secondary generalized TCS according to the diagnoses of the epileptologists. Because for primary or secondary generalized tonic-clonic seizures all anti-epileptic drugs (AEDs) are in general acceptable, except ethosuccimide that is only indicated for absence seizures, the primary and secondary generalized TCS were put together in one category. A distinction between these types of seizures cannot be made by SICS since it is not able to reason with temporal ordering of the manifestations.

The system can only be used when a physician is certain that the patient has epilepsy but still has to determine the kind of seizure type the patient is suffering from. The system is not built for the differential diagnosis of epileptic-like phenomena.

Recommendations

The accuracy of the system in this efficacy study is quite good. This is a strong indication that the DRFs of the included five seizure types are valid. The next step is to test the system in a clinical setting e. g. a general neurology department/clinic.

It is also possible to use the system in educational settings. Medical students and/or neurology residents can be trained to ask and report manifestations in a systematic way by using simulated patients. When these students or residents are later confronted with epilepsy patients in daily practice they may have developed a strategy to diagnose the seizure type.

Conclusion

This efficacy study of the diagnostic reference frames showed an accuracy of 91 % of support in classification of seizures (SICS) for the most common seizure types. This indicates that the knowledge encoded in the DRFs is valid. The next step is to test the DRFs in a clinical setting to evaluate the applicability in every daily practice.

Acknowledgements We thank the epileptologists and patients for their cooperation. This study was financially supported by the Dutch Epilepsy Foundation.

References

1. Commission on Classification and Terminology of the International League Against Epilepsy (1981) Proposal for revised clinical and electroencephalographic classification of epileptic seizures. Epilepsia 22:489–501
2. Gastaut H (1970) Clinical and electroencephalographical classification of epileptic seizures. Epilepsia 11:102–113
3. Bodensteiner JB, Brownsworth RD, Knapik JR, Kanter MC, Cowan LD, Leviton A (1988) Interobserver variability in the ILAE classification of seizures in childhood. Epilepsia 29:123–128
4. Van Donselaar CA, Geerts AT, Meulstee J, Habbema JD, Staal A (1989) Reliability of the diagnosis of a first seizure. Neurology 39:267–271
5. Reutens DC, Howell RA, Gebert KE, Berkovic SF (1992) Validation of a questionnaire for clinical seizure diagnosis. Epilepsia 33:1065–1071
6. Sander JW, Shorvon SD (1987) Incidence and prevalence studies in epilepsy and their methodological problems: a review. J Neurol Neurosurg Psychiatry 50:829–839
7. Engelbrecht R, Rector A, Moser W (1995) Verification and validation. In: Van Gennip EMS, Talmon JL (eds) Assessment and evaluation of information technologies. Amsterdam: IOS Press; pp 51–66
8. Van Ast JF, Talmon JL, Renier WO, Hasman A (2004) An approach to knowledge base construction based on expert opinions. Methods Inf Med 43:427–432
9. Brender J, Talmon J, Nykanen P, McNair P, Demeester M, Beuscart R (1995) On the evaluation of system integration. In: Van Gennip EMS, Talmon JL (eds) Assessment and evaluation of information technologies. Amsterdam: IOS Press; pp 189–208
10. Rector AL, Ackerman E (1975) Rules for sequential diagnosis. Comput Biomed Res 8:143–155
11. Gorry GA, Barnett GO (1968) Sequential diagnosis by computer. JAMA 205:849–854
12. Nøhr C (1995) From assessment to decision-making. In: Van Gennip EMS, Talmon JL (eds) Assessment and evaluation of information technologies. Amsterdam: IOS Press; pp 117–126
13. Commission on Classification and Terminology of the International League Against Epilepsy (1989) Proposal for revised classification of epilepsies and epileptic syndromes. Epilepsia 30:389–399
14. Korpinen L, Pietila T, Peltola J, Nissila M, Keranen T, Touvinen T, Petranek ES, Frey H (1994) Evaluation of Epilepsy Expert-a decision support system. Comput Methods Programs Biomed 45:223–231
15. Spitzer K, Thie A, Caplan LR, Kunze K (1989) The MICROSTROKE expert system for stroke type diagnosis. Stroke 20:1353–1356