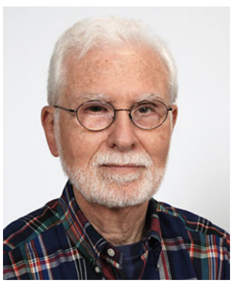
FULL-LENGTH ORIGINAL RESEARCH

The semiology of psychogenic nonepileptic seizures revisited: Can video alone predict the diagnosis? Preliminary data from a prospective feasibility study

*Giuseppe Erba, †Giorgia Giussani, *Adam Juersivich, ‡Adriana Magaudda, §Valentina Chiesa, ‡Angela Laganà, ¶Gabriella Di Rosa, †Elisa Bianchi, *John Langfitt, and †Ettore Beghi

Epilepsia, **(*):1–9, 2016
doi: 10.1111/epi.13351



Giuseppe Erba is professor emeritus of neurology at the University of Rochester and investigator at Ist. M. Negri.

Summary

Objective: **To investigate if, when, and to what extent visual information contained in a video-recorded event allows experienced epileptologists to predict the diagnosis of psychogenic nonepileptic seizures (PNES) without the aid of electroencephalography (EEG).**

Methods: **Five neurologists actively practicing in epilepsy centers in Italy and the United States were asked to review 23 videos capturing representative events of 21 unselected consecutive patients admitted for long-term video-EEG monitoring (VEM). Four raters were blind to EEG and clinical information; one rater was not. They were requested to (1) rate the videos for quality and content; (2) choose among four diagnoses: (a) epileptic seizures (ES); (b) PNES; (c) Other nonepileptic seizures (NES; (syncope, movement disorder, migraine, etc.); (d) "Cannot Say"; and (3) explain in their own words the main reasons leading to the diagnosis of choice.**

Results: **All raters predicted the diagnosis correctly in 7 of 23 videos (all ES or PNES) (30.4%), whereas all raters failed in 5 of 23 cases (three Other NES, one PNES, one Cannot Say) (21.7%). The conditions that facilitate, and those that interfere with, a confident diagnosis were predictable. Degree of accuracy among raters was not uniform and was consistently better in three raters. Two among the four blind raters were as accurate as the rater who was not blinded. Interrater agreement was "moderate" (k = 0.52) for the overall group; "moderate" for ES (k = 0.53); "substantial" for PNES (k = 0.63); "fair" for Other NES (k = 0.21)—similar to the results obtained in a previous study evaluating the reliability of combined video-EEG.**

Significance: **In about one third of cases, a confident diagnosis of PNES/ES can be established on clinical grounds based on video data alone. Our results benefit all affected patients, particularly those with no access to video-EEG monitoring units.**

KEY WORDS: **Video, Monitoring, Semiology, Psychogenic seizures, Epileptic seizures, Nonepileptic seizures.**

Accepted February 9, 2016.

*Department of Neurology, SEC, University of Rochester, Rochester, New York, U.S.A.; †Laboratory of Neurological Disorders, Department of Neuroscience, IRCCS-Pharmacological Research Institute "Mario Negri," Milan, Italy; ‡Department of Neuroscience, Epilepsy Center, University of Messina, Messina, Italy; §Neurology Unit 2, Epilepsy Center, San Paolo Hospital, Milan, Italy; and ¶Department of Pediatric, Gynecological, Microbiological and Biomedical Science, Unit of Infantile Neuropsychiatry, University of Messina, Messina, Italy

Address correspondence to Ettore Beghi, Laboratory of Neurological Disorders, IRCCS – Istituto di Ricerche Farmacologiche "Mario Negri," Via Giuseppe La Masa, 19, 20156 Milan, Italy. E-mail: ettore.beghi@marionegri.it

Wiley Periodicals, Inc.

Outside of epilepsy monitoring units the diagnosis of psychogenic nonepileptic seizures (PNES) constitutes a major challenge. Because the manifestations of PNES are greatly variable and often resemble those of epileptic seizures (ES),[1] no single feature has proved to be pathognomonic, although a recent study found that either diagnosis is associated with a distinct cluster of signs.[2] Based on current practice, the gold standard (GS) for a definitive diagnosis of PNES is the documentation of a normal electroencephalography (EEG) study during events with semiology and a patient's history consistent with the diagnosis of PNES. Thus, the GS implies accessibility to a monitoring unit with

Key Points

- Trained epileptologists can successfully distinguish ES from PNES with video alone in a subset of critical events with objective motor manifestations
- Events characterized by subjective sensory, nonmotor, autonomic manifestations require comprehensive diagnostic procedures, including VEM
- Clinical assessment on video alone can lead to misdiagnosis, with significant risk to patients, especially if the doctor has low level of training

specialized reviewers and services. Nonetheless, video-EEG monitoring (VEM) not infrequently fails to capture the events, despite induction attempts, and it will not differentiate certain types of frontal lobe ES from PNES. Moreover, the diagnosis of PNES, depends primarily on clinician judgment and, unlike epilepsy, its reliability cannot be objectively verified by pathology or treatment outcome.

In an attempt to assess the accuracy of VEM, 22 board-certified neurologists actively practicing in epilepsy centers from the United States and Europe were asked to predict the diagnosis in 22 consecutive patients with mixed seizure types (PNES, ES, and other nonepileptic events) based exclusively on video-EEG vignettes.[3] Interrater agreement was moderate across all three diagnostic categories ($k = 0.57$) and was moderate for PNES ($k = 0.57$), substantial for ES ($k = 0.69$), and low for other nonepileptic episodes. The conclusion was that the diagnosis of these disorders based on combined video-EEG data presents inherent difficulties, particularly it might be faulted by subjective components. Furthermore, adherence to the current standard of care will prevent or delay the diagnosis in many patients worldwide who have no access to a VEM facility.

In response to this problem, the International League Against Epilepsy (ILAE) Nonepileptic Seizures Task Force published recommendations indicating that in certain cases it is possible to reach the diagnosis of PNES on clinical grounds in the absence of concomitant EEG.[4] This can be done in stages, with different levels of diagnostic certainty, depending on the information available. For instance, a typical event observed by experienced clinicians in person or on video, especially if corroborated by signs of preserved consciousness and/or avoidance, provides a "probable" diagnosis of PNES. If a similar non-witnessed event were recorded separately, for instance during a routine EEG, and were accompanied by a normal EEG, the diagnosis is "clinically established." However, a diagnosis can be fully "documented" only through expert review of simultaneous video-EEG data. The task force also emphasized the importance of the patient's past history and comorbidities as corroborating evidence, and concluded that a diagnosis of PNES could be considered "definitive" only if supported by the patient's global assessment and full VEM documentation.

The task force also considered advantages and limitations of home-recorded videos but did not recommend their use because the diagnostic yield of typical events recorded by witnesses has not been systematically evaluated. Of the two modalities used in VEM, video recording is technically easier to obtain and less expensive than a video with simultaneous EEG. An earlier study demonstrated that, in a proportion of cases, neurologists could make a confident diagnosis of ES/PNES based entirely on videotapes recorded by hospital staff with a handheld camcorder.[5] More recent studies have confirmed that the differential diagnosis between ES and PNES based on video footage alone is possible but requires neurologic training.[6] Likewise, training of medical students through video-based modules of ES and PNES improves accuracy of seizure diagnosis.[7] Finally, it has been demonstrated that trained epileptologists, contrary to untrained eyewitnesses, could easily recognize key signs characteristic of either syndrome simply by analyzing seizures on video, and blinded to EEG.[2] Thus, there is evidence that in the hands of capable reviewers, video monitoring alone could represent a useful clinical tool.

The aim of this study was to investigate if, when, and to what extent the information conveyed by a paroxysmal event captured on video and interpreted by experts in epilepsy could contribute to the differential diagnosis without the aid of simultaneous EEG. This information would be particularly relevant when the simultaneous EEG is not available or in the event that clinicians may elect to screen the events of patients with mixed seizure disorders on video before referral to a monitoring unit.

Materials and Methods

This study represented the feasibility trial of a larger project currently in progress between the University of Rochester (UR) and three Italian Institutions: IRCCS Pharmacological Research Institute "Mario Negri," Milan; University of Messina; and Azienda Hospital San Paolo, Milan, Italy. The study protocol was reviewed and approved by the Research Subject Review Board (RSRB) of the UR where the patients were recruited and the video produced.

Patients

Patients 18 years and older consecutively admitted between July 1 and September 10, 2014, were asked to participate. Patients were excluded if clinical judgment indicated that they lacked the intellectual capacity to answer questionnaires designed for the larger project. We enrolled prospectively all patients who consented, irrespective of other clinical information. For each of them, at the time of discharge, a representative audio-video segment (or segments), deprived of the EEG tracing, was submitted to five designated independent raters for review and prediction of diagnosis. Where possible, each video included testing of patient's responsiveness by staff.

Raters

The five raters were board certified neurologists/child neurologists (or the Italian equivalent), all practicing full time in tertiary epilepsy centers. The four raters from the Italian institutions (R-1, R-2, R-3, and R-4), were completely blinded to the EEG findings, and to the patient's history and comorbidities, and were unaware of the final diagnosis established by the clinical team. The fifth rater (R-5) was a faculty member of the epilepsy unit at UR. Although not responsible for direct patient care during the admission, R-5 was not blind to the patient's condition and results of monitoring but was specifically instructed not to access history, laboratory results, or additional vignettes while reviewing the submitted video. This rater was included to investigate how awareness of ancillary clinical information would influence the rating by comparing R-5 score to the scores of the other four blinded raters.

Individual raters' profiles are reported in Table 1. All raters received their training through an epilepsy/clinical neurophysiology fellowship.

Procedure

Each rater was asked to review each video and render a diagnosis to the best of his or her capacity based only on audiovisual information. Raters were also asked if the video was technically satisfactory and "adequate," meaning it provided all the information necessary to formulate the diagnosis and, if not, why. Arbitrarily, we considered a video adequate for the task if at least three of five raters agreed that the video was sufficiently informative.

Raters were given four diagnostic options and required to choose one among the following: (1) ES, defined according to the ILAE classification[8]; (2) PNES, classified according to the six categories proposed by Seneviratne et al.:[9] 1. Rhythmic motor, 2. Hypermotor, 3. Complex motor, 4. Dialeptic, 5. Nonepileptic auras, and 6. Mixed; (3) Other nonepileptic seizures (NES), due to paroxysmal nonepileptic events other than psychogenic (syncope or other dysautonomic manifestations, migraine, movement disorder, panic attacks, etc.); (4) "Cannot Say."

In addition, raters had to specify the main reasons leading to the diagnosis of choice and describe any behavioral observations that most contributed to their diagnostic decision.

Each rater worked independently and filed the data directly into a database set up at the IRCCS-Pharmacological Research Institute "Mario Negri" in Milano, Italy, for statistical analysis.

We evaluated diagnostic accuracy as the ability of each individual rater to correctly predict the GS diagnosis, based on audiovisual evidence alone. The GS diagnosis was that established by the clinical team after a comprehensive evaluation of the patient's risk factors, comorbidities, psychosocial status, results of neurologic examination and neuroimaging, video semiology, EEG findings including purely electrical seizures, and the results of monitoring other physiologic parameters (ECG [electrocardiography], blood pressure, orthostatic testing, blood sugar, and so on) as appropriate. Raters' accuracy in predicting the GS diagnosis was presented as the proportion of raters that correctly predicted the GS.

We calculated interrater agreement among all raters, between pairs of raters, and between each rater and the GS using Fleiss' Kappa,[10] with 95% confidence intervals (CIs). The Kappa statistic is a measure of interrater agreement adjusted by the amount of the agreement expected to occur by chance alone. Kappa values were used to assess overall agreement across all diagnostic categories (PNES, ES, Other NES, Cannot Say), and agreement in differentiating between the diagnosis of ES, PNES, Other NES, and Cannot Say. Kappa values were classified as poor (<0.00), slight (0.01–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), or almost perfect (0.81–1.00).[11] This classification of the magnitude of Kappa values is only for descriptive aims and does not refer to statistical significance. Data were analyzed using the SAS statistical package (version 9.2; SAS Institute Inc, Cary, NC, U.S.A.).

For each seizure that was reviewed, we carefully surveyed each rater's comments and compiled a list of all individual signs or symptoms identified as significant and their relative frequency. In addition, we took note of any specific observations (the sequence of certain manifestations, patient's affect, and incongruous behavior) underpinning the raters' diagnostic reasoning.

Table 1. Individual profiles of raters

	R-1	R-2	R-3	R-4	R-5
Age (y)	56	48	39	39	32
Title	Neurologist, child neurologist	Neurologist	Child neurologist	Neurologist	Neurologist
Caring for patients with epilepsy (y)	30	15	9	9	2.5
Monthly hours for patients with epilepsy	50	60	30	150	100
Monthly visits for patients with seizures	25–50	>50	>50	>50	>50
Adults/children	90% adults	90% adults	100% children	99% adults	75% adults, 25% children
Blind to patients' info and EEG	Yes	Yes	Yes	Yes	No

doi: 10.1111/epi.13351

Results

A total of 21 patients were enrolled. Each had at least one typical event recorded on video during the course of the investigation. Twenty patients reported one type of event each represented in a single video. Case 3 reported three different types of events (3a, 3b, and 3c). Therefore, 23 videos were submitted to each rater for review.

Video quality

Raters considered "adequate" 10 (43%) of the 23 videos submitted, whereas 13 (57%) of 23 were, in their judgement "inadequate." This was due mainly to technical deficiencies or insufficient information (i.e., patient responsiveness not tested or incompletely tested, patient out of screen or poorly visible, defective audio).

Raters' accuracy in predicting the diagnosis

Table 2 shows the degree of concordance between raters' diagnostic predictions and GS diagnosis. This table includes the two cases (7 and 16) where the clinical team had reached no definite diagnosis (NDD).

All five raters were correct in predicting the diagnosis in 7 (30.4%) of 23 cases. Of these, three had ES and four PNES. On the contrary, none of the five raters was in agreement with the GS diagnosis in 5 (21.7%) of 23 cases: three had Other NES, one PNES, and one NDD.

The ability to predict the diagnosis in the remaining 11 cases was intermediate (see Table 2 for details).

Because raters were predicting the diagnosis based on the information contained in the video vignette, we reviewed each event in an attempt to identify the clues that could have influenced the raters' decision. It became apparent that success (5/5 raters correct) was related to events characterized predominantly by motor manifestations, whereas failure (0/5 raters correct) was related to events characterized mainly by subjective sensory symptoms with little or no motor manifestations. Figure 1 shows the correlation between different degrees of accuracy and type of events recorded on video (motor in red, nonmotor in blue). Success rate was higher for ES and PNES with motor manifestations and lower for Other NES, ES, or PNES with no motor manifestations. Moreover, the number of elementary signs identified as significant by the reviewer (shown in Table 3) seemed to correlate with accuracy. In the group of seven videos where all raters were correct, the total number was 56 (average eight signs/video, range 5–11) compared to 20 in the group of five videos, where all raters had failed in their prediction (average four signs/video, range 1–8).

Finally, to assess the influence of video quality on raters' accuracy, we mapped rate of success against video quality as perceived by the reviewers. Of the seven cases where all raters were correct, five videos were considered adequate and two not adequate. On the contrary, in the five cases where all raters had failed, the proportion was reversed with one video adequate against four not adequate.

Interrater agreement

The overall agreement among all five raters was moderate (k = 0.52, 95% CI 0.44–0.60). By diagnostic category, the agreement was moderate for ES diagnosis (k = 0.53, 95% CI 0.39–0.67), substantial for PNES (k = 0.63, 95% CI 0.49–0.77), and fair for Other NES (k = 0.21, 95% CI 0.07–0.35).

Table 4 shows interrater agreement within each pair of raters and between each individual rater and the GS. Combining all diagnoses (Overall), agreement between pairs of raters was higher between R-1 and R-2 (k = 0.73), R-4 and R-5 (k = 0.69), and R-1 and R-5 (k = 0.63). However, the pairs with the highest agreement varied when limiting the diagnosis to PNES and ES (Other NES excluded). Overall agreement between three of the above raters and GS was in the moderate range for raters R-5 (k = 0.58), R-4 (k = 0.56), and R-1 (k = 0.49). It was substantial or above for two of the diagnostic categories: respectively, 0.81, 0.63, and 0.82 for PNES; 0.90, 0.81, and 0.81 (almost perfect) for ES. It was remarkably lower for Other NES. In comparison, Kappa values for raters R-2 and R-3 were lower in all categories.

For the entire group of videos, agreement among the five raters was only slightly higher when inadequate videos were excluded (k = 0.59, 95% CI 0.45–0.63).

Video content and raters' strategy leading to the diagnosis of choice

Analysis of the reviewers' spontaneous comments and the terminology they used provided ground for assessing the diagnostic strategy. Table 3 lists in order of decreasing frequency and by diagnostic category, the specific terms used by raters to describe the signs that caught their attention. It also shows that the observable signs and symptoms reported for the ES and PNES categories, where raters were most accurate in predicting the diagnosis, were by far more numerous than for the Other NES and Cannot Say categories, where raters were most likely to fail.

The term "semiology" was the most frequently mentioned, either as a positive statement, to indicate that it was a key element to the diagnosis, mostly in the case of ES and PNES, or in a negative way, to state that it was not consistent with either ES or PNES or was not sufficiently definable, which was most likely the case of Other NES.

Discussion

This feasibility study offered an opportunity to better define how experts in the field of epileptology utilize visual clues to establish the diagnosis of epilepsy versus nonepilepsy. The approach is both analytical and global, based on

Table 2. Accuracy of raters in predicting the diagnosis

			Gold standard		Reviewer diagnosis					
Video	Adequacy of video	Semiology	Diagnosis	Seizure type	R-1	R-2	R-3	R-4	R-5	Accuracy
1	Yes	Motor	ES	Partial with sec. gen.	ES	ES	PNES	ES	ES	4/5
2	No	Nonmotor	Other NES	Dysautonomic	Cannot Say	Cannot Say	PNES	PNES	Cannot Say	0/5
3a	No	Motor	PNES	Complex motor	PNES	PNES	PNES	PNES	PNES	5/5
3b	No	Motor	PNES	Mixed	PNES	Es	PNES	PNES	PNES	4/5
3c	No	Nonmotor	PNES	Nonepileptic aura	Cannot Say	ES	Cannot Say	Cannot Say	Cannot Say	0/5
4	Yes	Motor	PNES	Rhythmic motor	PNES	PNES	PNES	PNES	PNES	5/5
5	No	Motor	ES	Partial with sec. gen.	ES	ES	PNES	ES	ES	4/5
6	No	Motor	PNES	Rhythmic motor	PNES	PNES	PNES	PNES	OTHER	4/5
7	Yes	Nonmotor	NDD	–	PNES	PNES	PNES	PNES	OTHER	0/5
8	Yes	Motor	ES	Partial with sec. gen.	ES	ES	ES	ES	ES	5/5
9	Yes	Motor	PNES	Rhythmic motor	PNES	PNES	PNES	Cannot Say	PNES	4/5
10	Yes	Motor	ES	Partial with sec. gen.	ES	ES	ES	ES	ES	5/5
11	Yes	Motor	ES	Partial with sec. gen.	ES	ES	ES	ES	ES	5/5
12	No	Nonmotor	Other NES	Dysautonomic	Cannot Say	ES	Cannot Say	ES	Cannot Say	0/5
13	No	Motor	ES	Partial with sec. gen.	ES	ES	OTHER	Cannot Say	Cannot Say	2/5
14	Yes	Nonmotor	ES	Simple partial	Cannot Say	Cannot Say	Cannot Say	ES	ES	2/5
15	Yes	Motor	ES	Complex partial	ES	PNES	PNES	ES	ES	3/5
16	No	Nonmotor	NDD	–	ES	ES	Cannot Say	Cannot Say	Cannot Say	3/5
17	No	Nonmotor	PNES	Nonepileptic aura	PNES	PNES	Cannot Say	PNES	PNES	4/5
18	No	Nonmotor	Other NES	Dysautonomic	Cannot Say	Cannot Say	OTHER	OTHER	OTHER	3/5
19	Yes	Motor	PNES	Complex motor	PNES	PNES	PNES	PNES	PNES	5/5
20	No	Nonmotor	Other NES	Dysautonomic	Cannot Say	Cannot Say	Cannot Say	Cannot Say	Cannot Say	0/5
21	No	Nonmotor	PNES	Dialeptic	PNES	PNES	PNES	PNES	PNES	5/5

ES, epileptic seizures; PNES, psychogenic nonepileptic seizures; NES, nonepileptic seizures.

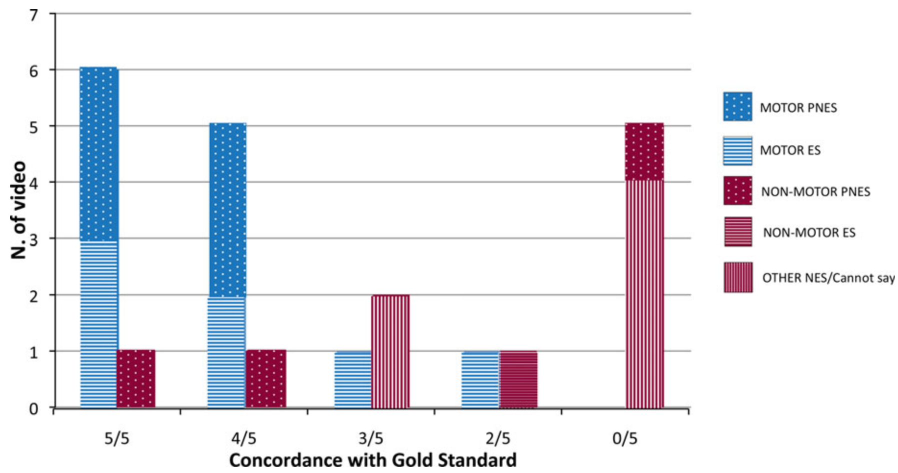


Figure 1.
Correlation between accuracy and type of events. Shows the association between degrees of accuracy (x axis) and type of events recorded on video (number of videos on y axis): motor PNES in blue with dots, motor ES in blue with horizontal lines, nonmotor PNES in red with dots, nonmotor ES in red with horizontal lines, and other NES/Cannot Say in red with vertical lines. ES, epileptic seizures; PNES, psychogenic nonepileptic seizures; NES, nonepileptic seizures.
Epilepsia © ILAE

Table 3. Signs and symptoms leading to the diagnosis

Signs	Compound frequency	ES	PNES	Other NES	Cannot Say
Semiology	28	1, 5, 10, 11, 14, 15	3a, 3b, 4, 6, 17, 19, 21	2, 18, 20	7
Wax/waning	21	1, 15	3a,3b, 4, 6, 9, 17, 19, 21		7
Subjective feelings	17		3c, 4, 17	12, 18, 20	
Automatisms	13	5, 8, 10, 13, 15			
Shaking	11	1, 15	3a, 4, 6, 9,		7
Long duration	10		3a, 3b, 4, 9, 17, 19	2	
Dystonic posturing	10	1, 5, 8, 10, 15	3b, 4, 19, 21		
Slow postictal recovery	10	5, 8, 10, 11, 14	3b		
Short duration	9	5, 10, 11, 15	3c	18	
Eyes close	8		3a, 4, 19, 21		
[Forced] head deviation	8	1, 5, 10, 11, 15			
Expression of emotionality	7		4, 17		
Modality of onset: out of sleep	7	1, 8, 13		2	
Eyes open	6	5, 8, 10			
Increased tone	6	1, 5, 10, 11	9		
Crying	6		21		7
Loss of consciousness	6	5, 10, 11, 15	3b, 3c		
Preserved consciousness	6		3a, 3b,	2	7, 16
Tremors	5		3a, 3b, 6		7
Slumping	5		3a, 19, 21		
Modality of onset: abrupt	5	5, 10	4, 9, 21		
Modality of onset: gradual	5	15	3a, 3b, 9		
Confusion	4	8, 10, 14			
Fast postictal recovery	4		3a, 3c, 4		
Asynchrony	3	1,5,10			
Tingling	2		3c, 21		
Arrhythmicity	1		19		
Internal sensation	1				7
Visual distortion	1				16
Hyperventilation	1				7
Synchrony	0	0			
Rhythmicity	0	0			

ES, epileptic seizures; PNES, psychogenic nonepileptic seizures; Cannot Say, no diagnosis possible.
Numbers in the table refer to the specific seizure video for which that particular sign was mentioned as being significant.

Table 4. Interrater agreement

	Overall		PNES		ES		Other NES	
Pair	Kappa	95% CI	Kappa	95% CI	Kappa	95% CI	Kappa	95% CI
R-5 vs. R-1	0.63	0.38–0.88	0.81	0.40–1.00	0.70	0.29–1.00	−0.07	−0.48–0.34
R-5 vs. R-2	0.37	0.12–0.62	0.62	0.21–1.00	0.35	−0.06–0.76	−0.07	−0.48–0.34
R-5 vs. R-3	0.44	0.19–0.69	0.37	−0.04–0.78	0.49	0.08–0.90	0.33	−0.08–0.74
R-5 vs. R-4	0.69	0.44–0.94	0.62	0.21–1.00	0.90	0.49–1.00	0.45	0.04–0.86
R-1 vs. R-2	0.73	0.44–1.00	0.82	0.41–1.00	0.63	0.22–1.00	–[a]	–
R-1 vs. R-3	0.48	0.21–0.75	0.56	0.15–0.97	0.40	−0.01–0.81	−0.05	−0.46–0.36
R-1 vs. R-4	0.54	0.27–0.81	0.82	0.41–1.00	0.62	0.21–1.00	−0.02	−0.43–0.39
R-2 vs. R-3	0.34	0.07–0.61	0.56	0.15–0.97	0.25	−0.16–0.66	−0.05	−0.46–0.36
R-2 vs. R-4	0.40	0.11–0.69	0.63	0.22–1.00	0.45	0.04–0.86	−0.02	−0.43–0.39
R-3 vs. R-4	0.48	0.23–0.73	0.56	0.15–0.97	0.40	−0.01–0.81	0.64	0.23–1.00
R-5 vs. GS	0.58	0.33–0.83	0.81	0.40–1.00	0.90	0.49–1.00	0.15	−0.26–0.56
R-1 vs. GS	0.49	0.24–0.74	0.82	0.41–1.00	0.81	0.40–1.00	−0.10	−0.51–0.31
R-2 vs. GS	0.35	0.08–0.62	0.63	0.22–1.00	0.45	0.04–0.86	−0.10	−0.51–0.31
R-3 vs. GS	0.30	0.05–0.55	0.39	−0.02–0.80	0.40	−0.01–0.81	0.23	−0.18–0.64
R-4 vs. GS	0.56	0.31–0.81	0.63	0.22–1.00	0.81	0.40–1.00	0.32	−0.09–0.73

PNES, psychogenic nonepileptic seizures; ES, epileptic seizures.
[a]No reviewers give the response "Other."

the convergence of observation and reasoning. Raters' attention was focused on the presence of elementary signs typically associated with ES or PNES, and at the same time, as discussed below, on how these signs are linked and develop, either in a progressive evolution that follows the central nervous system (CNS) organization or in a disorganized manner, incongruous with neurologic pathophysiology. This conforms to the prelearned model of ES and PNES that includes a number of positive signs associated with, or exclusive of, the diagnosis of either ES or PNES. It also explains why in the raters' comments the term "semiology" was the most frequently mentioned as key to the diagnosis. The term was used directly as a positive statement ("semiology is consistent with... ."), as a negative statement ("not consistent with... ."), or indirectly such as "the sequence of motor manifestation (i.e., the semiology) is not consistent with ES but with PNES." This approach was most likely successful in cases with ES or PNES but almost always failed when the diagnosis was Other NES. There are two possible explanations for this phenomenon: (1) motor manifestations, that we found directly correlated with the ability to match the GS diagnosis, are more likely to be represented in ES and PNES than in Other NES; (2) for the diagnosis of Other NES, in addition to video and EEG data, information about the prodromal symptoms and the results of monitoring other physiologic parameters (ECG telemetry, blood pressure, and so on) is necessary. Raters were aware of these limitations and chose in many such cases to withhold a diagnosis (Cannot Say) stating: "diagnosis impossible" or "cannot be reached with confidence based on video alone."

Our data indicate that an important factor affecting raters' accuracy was the number of diagnostic clues detected. The presence of multiple signs and objective symptoms increased accuracy because objective signs are more reliable than subjective reports and the joint occurrence of various signs increases specificity, as reported by others.[2]

Our study, investigating the reliability of video data alone, duplicates the results of a previous study testing the reliability of combined video-EEG data.[3] Interrater variability in interpreting videotaped events was present in about equal measure, whether the EEG was included or excluded. K values in the two studies for the overall group and for each diagnostic category were very close. Although the content of the video vignettes could not be compared and the number of raters was different from ours, interrater agreement in predicting the GS diagnosis was comparable. This finding was corroborated by the additional observation that in our study, among the three "best" raters, the two (R-1 and R-4), both blinded to patient history and EEG findings, performed just as well as R-5 who was not blinded. This suggests that, in some cases, the video provides key information (multiple signs, certain types of motor manifestations, or interaction with bystanders) so typical that knowledge of the simultaneous EEG has little impact on diagnostic accuracy.

Our findings underscore the position of the ILAE Task Force[4] that in the majority of cases an adequate diagnosis is based on the convergence of (1) history, (2) witnessed semiology, (3) monitoring of multiple physiologic functions (EEG, ECG, and other physiologic functions). At the same time, our data indicate that the three elements are not essential in all cases to reach the diagnosis, confirming the clinical observation that not infrequently epileptologists are quick in differentiating ES from PNES on video before viewing the EEG.

Because it is proven that discriminating one type of seizure from another is a learned skill and requires neurologic training,[6,7,12] one note of caution is that reviewers must be properly trained for the task and experienced. One major cause of concern is the variable degree of competence among reviewers and how it affects interrater agreement.

Our group of five raters, interpreting video data alone, was not exempt from such liability. Two raters performed consistently at a lower level of accuracy compared to the other three. That may be explained, at least for one who was a child neurologist, by lower exposure during clinical practice to the adult population that was the focus of this investigation. However, irrespective of variable degree of expertise, the diagnosis based on video alone showed concordance among all five raters and was correct in almost one third of the videos reviewed. In this study, motor seizure semiology diagnoses fared clearly better than non-motor seizures, confirming an observation reported in a previous study.[13] As reported by others,[14] some seizure types were easier than others to diagnose based on videos. In our sample, these seizures were represented by partial seizures with secondary generalization and complex and rhythmic motor PNES. The reasons for that are clearly stated in the raters' comments, indicating that video alone was not useful, was inadequate, and was at times misleading in the case of non-motor, sensory, and dyscognitive seizures.

A correct prediction of the putative diagnosis also depends on video quality, particularly when the concomitant EEG is not available. Videos recorded in hospital rooms during VEM, like those utilized in this investigation, are not constantly supervised, and capture spontaneous events that often go unnoticed or are only partially witnessed by staff. Indeed, missing relevant features such as the proper display of the subject and subject's behavior, a dysfunctional audio, or lack of intervention by bystanders, were the legitimate and most common reasons for which more than half of the video segments submitted for review were rated below the minimum desirable standard. However, after removing poor quality videos from analysis, interrater agreement improved only slightly. This finding can be explained by the different reasons given by the raters to qualify a video as inadequate, some of which were unlikely to affect the diagnosis, and by the nonunanimous quality assessment among raters. Nonetheless, efforts to improve video quality would be valuable.

The main strength of this investigation is the prospective approach, with inclusion of all eligible candidates consecutively admitted for monitoring, and a collection of data prior to the diagnosis. This provided a cohort of patients with mixed seizure disorders that reflects the current referral pattern to epilepsy monitoring units and allowed investigatation of these patients without bias, contrary to many previous PNES studies based on retrospective data. Our main contribution is a closer insight of how epileptologists utilize the information displayed on video of events recorded during hospital monitoring. We also indicated that such approach has inherent constraints. For instance, we demonstrate that although some events such as rhythmic motor PNES are possible to diagnose on video alone, others such as sensory, non-motor, or autonomic events, are not.

We are aware of this study's limitations. First, the number of cases investigated in this feasibility study is fairly small. Second, our restricted cohort did not include "hypermotor PNES," an uncommon type that is notoriously difficult to differentiate from hypermotor frontal lobe seizures. Therefore, our data will have to be confirmed in larger cohorts.

On the other hand, we felt that our results were worth reporting for a number of reasons. First, they may provide useful information for the power analysis of larger studies. Second, they may generate new hypotheses and stimulate new research. Third, our experience proves that international collaboration is possible in compliance with privacy regulations, and that exchange of personal information such as video images through the Internet is acceptable to participants. This in itself represents an opportunity for exploring patient populations where manifestations and frequency of PNES may vary in relation to cultural differences.[15] Most importantly, it may spur new interest in the video format, once defined as "the closest proxy to witnessed events."[12] A greater use of this modality may be particularly beneficial when VEM is not available. This is relevant not only for future research but also for clinical purposes. For instance, the utilization of homemade videos as a method of screening patients before admission to a full monitoring unit and its diagnostic value could be further explored. Finally, because clinical assessment based on video alone can easily lead to misdiagnosis, further investigations are needed to identify which additional clinical parameters would be necessary to corroborate a video diagnosis in settings where full VEM investigation is not accessible.

Acknowledgments

Drs Massimiliano Beghi, Paola Canevini, and Cesare Maria Cornaggia are gratefully acknowledged for their useful comments and suggestions in developing the project and Ms. Susanna Franceschi for handling the manuscript.

Disclosure of Conflicts of Interest

E. Beghi serves on the editorial boards of *Amyotrophic Lateral Sclerosis, Clinical Neurology & Neurosurgery,* and *Neuroepidemiology*; has been an associate editor of *Epilepsia*; has received money for board membership from VIROPHARMA and EISAI; has received funding for travel and speaker honoraria from UCB-Pharma, Sanofi-Aventis, GlaxoSmithKline; has received funding for educational presentations from GlaxoSmithKline; and reports grants from the Italian Drug Agency and from the Italian Ministry of Health. The remaining authors have nothing to disclose. We confirm that we have read the Journal position on issue involved in ethical publication and affirm that this report is consistent with these guidelines.

References

1. Gedzelman ER, LaRoche SM. Long-term video EEG monitoring for diagnosis of psychogenic nonepileptic seizures. *Neuropsychiatr Dis Treat* 2014;10:1979–1986.
2. Syed TU, LaFrance WC Jr, Kahriman ES, et al. Can semiology predict psychogenic nonepileptic seizures? A prospective study *Ann Neurol* 2011;69:997–1004.
3. Benbadis SR, LaFrance WC Jr, Papandonatos GD, et al. Interrater reliability of EEG-video monitoring. *Neurology* 2009;73:843–846.
4. LaFrance WC Jr, Baker GA, Duncan R, et al. Minimum requirements for the diagnosis of psychogenic nonepileptic seizures: a staged approach: a report from the International League Against Epilepsy Nonepileptic Seizures Task Force. *Epilepsia* 2013;54:2005–2018.
5. Samuel M, Duncan JS. Use of hand-held video camcorder in the evaluation of seizures. *J Neurol Neurosurg Psychiatry* 1994;57:1417–1418.
6. Ristić AJ, Drašković M, Bukumirić Z, et al. Reliability of the witness descriptions of epileptic seizures and psychogenic non-epileptic attacks: a comparative analysis. *Neurol Res* 2015;37:560–562.
7. Seneviratne U, Ding C, Bower S, et al. Video-based training improves the accuracy of seizure diagnosis. *J Neurol Neurosurg Psychiatry* 2014;85:466–470.
8. Commission on Classification and Terminology of the International. League Against Epilepsy. Proposal for revised clinical and electroencephalographic classification of epileptic seizures. *Epilepsia* 1981;22:489–501.
9. Seneviratne U, Reutens D, D'Souza W. Stereotypy of psychogenic nonepileptic seizures: insights from video-EEG monitoring. *Epilepsia* 2010;51:1159–1168.
10. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76:378–382.
11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–174.
12. MacDonald T, Hill A, Phan T, et al. Psychiatry versus general physicians: who is better at differentiating epileptic from psychogenic nonepileptic seizures? *Australas Psychiatry* 2012;20:379–383.
13. King DW, Gallagher BB, Murvin AJ, et al. Pseudoseizures: diagnostic evaluation. *Neurology* 1982;32:18–23.
14. Seneviratne U, Rajendran D, Brusco M, et al. How good are we at diagnosing seizures based on semiology? *Epilepsia* 2012;53:e63–e66.
15. Asadi-Pooya AA, Sperling MR. Epidemiology of psychogenic nonepileptic seizures. *Epilepsy Behav* 2015;46:60–65.

doi: 10.1111/epi.13351