Real-time Detection of Epileptiform Activity in the EEG: A Blinded Clinical Trial

Michael A. Black, Richard D. Jones, Grant J. Carroll, Alison A. Dingle, Ivan M. Donaldson and Philip J. Parkin

Key Words
Clinical Trial
Electroencephalography
Epileptiform Activity
Gold Standard
Rule-based
Spike Detection

INTRODUCTION

Both rule-based and artificial-neural-network approaches have been applied to the task of automating the detection of epileptiform activity (EA) in the EEG.[1] To date, however, few have been introduced into routine clinical use and, even then, only with limited success.[2] The central problem is false detections arising from various sources (muscle, ECG, electrode, movement, and eye blink artifacts, and sharp background activity), which have proven almost impossible to eliminate automatically. This not only reflects the difficulty in distinguishing spikes from background activities and artifacts but is also due to a lack of attention to spatial and temporal contextual cues present in the EEG.[3-7] In contrast to the electroencephalographer (EEGer) — who utilizes amplitude, sharpness, and polarity of spikes on adjacent channels, as well as the presence and localization of EA elsewhere in the EEG — automated systems have tended to consider waves individually or in a very limited context. Consequently, high false detection rates have largely restricted clinical application of spike detectors to long-term EEG monitoring, where they act as data reduction systems with all detections needing review by an EEGer.[2,8] In addition, these systems are not accurate enough to be of real benefit in standard EEG recording.

We have developed a system for automated detection and topographical mapping of EA in scalp EEG recordings. The primary goal of this system is to assist the EEGer in the reporting of *routine* EEGs by decreasing time spent on and increasing uniformity of interpretation. A central component of this system is a multi-stage rule-based system incorporating mimetic and expert system approaches. In contrast to most other systems, it relies heavily on both spatial and temporal EEG contextual information and is particularly successful at rejecting nonepileptiform activity.[3,6,9-11] This paper presents the results of a clinical trial comparing accuracy and utility of the system to conventional visual EEG reading.

METHODS

I. EEG data

The data comprised EEGs from 521 consecutive patients referred to the Neurology Department at Christchurch Hospital for routine EEG recordings. The patients were aged from 2 weeks to 86 years. Each EEG was of approximately 20-minutes duration (making a total of 173 hours of 16-channel EEG data) from seven bipolar and referential montages. Recordings were usually made while the patient was awake but resting and included periods of eyes open, eyes closed, hyperventilation, and photic stimulation. Amplification was provided by a Siemens Minograph Universal EEG machine. Many EEGs contained substantial artifact, particularly EMG, electrode, movement, and eye movement. The EEGs were bandpass-filtered between 0.5 and 70 Hz (5-pole Butterworth filter) and sampled at 200 Hz. The criterion for stopping the data collection was the recording of 50 EEGs judged by one EEGer (GC) to contain definite EA.

II. System description

The spike detector runs in real-time on a 486/586 PC and comprises six stages (Figure 1): (1) Data acquisition: Low-pass filtering (70 Hz, 5 pole), sampling at 200 Hz and digitization of 16 channels of bipolar or referential EEG. (2) Global amplitude: Calculation of average amplitude over all channels during the first 60 sec of EEG. This eliminates the need to know absolute amplitudes in the detection process. (3) Mimetic (feature extraction): Data reduction, via a single channel process, by calculating parameters of individual waves (e.g., duration, amplitude and peak sharpness) and comparing these with measures of background activity. Waves whose parameters exceed a set of thresholds are put forward as candidate epileptiform transients

From the Department of Medical Physics & Bioengineering, Christchurch Hospital (Black, Jones, Dingle); Department of Medicine, Christchurch School of Medicine, University of Otago (Jones, Donaldson); Department of Electrical & Electronic Engineering, University of Canterbury (Jones, Dingle); and the Department of Neurology, Christchurch Hospital (Carroll, Donaldson, Parkin), Christchurch, New Zealand.

Requests for reprints should be addressed to Dr. Richard D. Jones, Department of Medical Physics & Bioengineering, Christchurch Hospital, Private Bag 4710, Christchurch, New Zealand.

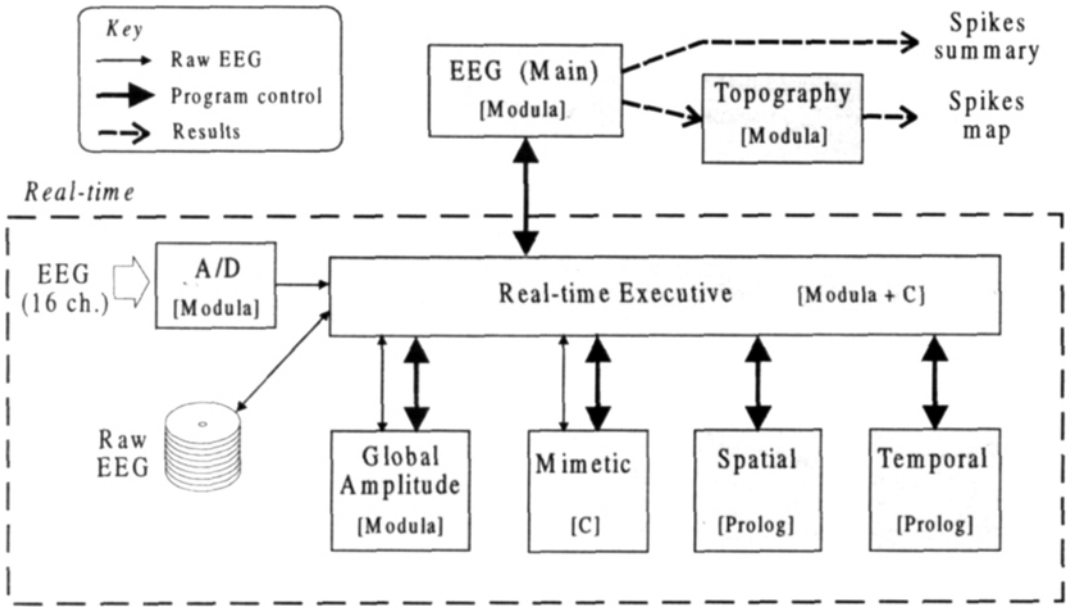

Figure 1.
Schematic of real-time spike detector.

(CETs) at single-channel level. (4) Spatial context (expert system): Use of multi-channel spatial cues (e.g., presence of synchronous waves of sufficient amplitude, sharpness and polarity on adjacent channels) to determine whether CETs are part of definite or possible epileptiform *events* — i.e., spikes or sharp-waves. (5) Temporal context (expert system): Using the presence of definite or possible spikes with a similar distribution elsewhere in the EEG to upgrade possible spikes to definite spikes. (6) Printout - Definite events are reported as numbers of major events (considered definite on spatial grounds alone) and minor events (probable events upgraded to definite due to temporal context). If no definite spikes are detected, 6.0 sec segments of the raw EEG centered around any remaining possible events (usually 1 or 2 at most) can be printed out for review by an EEGer (Figure 2). In addition, if one or more definite spikes are detected a topographic map is produced with discs indicating the electrodes(s) at which the spikes were detected (Figure 3). The size of the disc at the electrode at which most spikes were detected is set to the maximum size. The area of the discs at other electrodes indicates the *proportion* of spikes detected at these electrodes. The final output is classification of the EA as one of definite, questionable or none. For a fuller description of the detection system see Dingle et al.[6]

The first five of the above stages have been integrated to attain spike detection in real-time.[9] This was achieved by pseudo-multitasking these stages via a *real-time executive*. Once the global amplitude has been calculated, and while continuing to sample and store the incoming data, the program goes back to the start of the data and commences the spike detection process. Initially the raw EEG for this must come from disc but, after a minute or so, the system has caught up and is able to process incoming data directly. During spike detection, control alternates between the mimetic and spatial stages to detect spikes on spatial grounds alone. Full temporal processing takes only a few seconds and, hence, can be carried out frequently. In a standard EEG this need only be done at the end of the recording. Conversely, for long-term monitoring temporal updates about every 5 minutes or so are adequate.

III. Review of raw EEG data

The conventional EEG chart recordings were read by three experienced EEGers (EEGers-I): GC, EW, and SD. GC read all 521 EEGs, whereas EW and SD read 312 EEGs and 315 EEGs, respectively. All three EEGers read 106 EEGs.

The EEGers reported on two levels about EA. Firstly, the *presence* of EA, with options being none, questionable or definite. Secondly, if EA was present (definite or questionable) the *distribution* of activity was listed as generalized, lateralized, focal or multifocal.

IV. Review of detection system output

Two further experienced EEGers (EEGers-II), ID and PP, each reported the detection system's output for all 521 EEGs. The system's assessment of the presence of EA was used directly by these EEGers unless it detected only questionable spikes, in which case 6.0 sec segments of raw EEG centered around each questionable event were examined and the presence of EA was determined directly

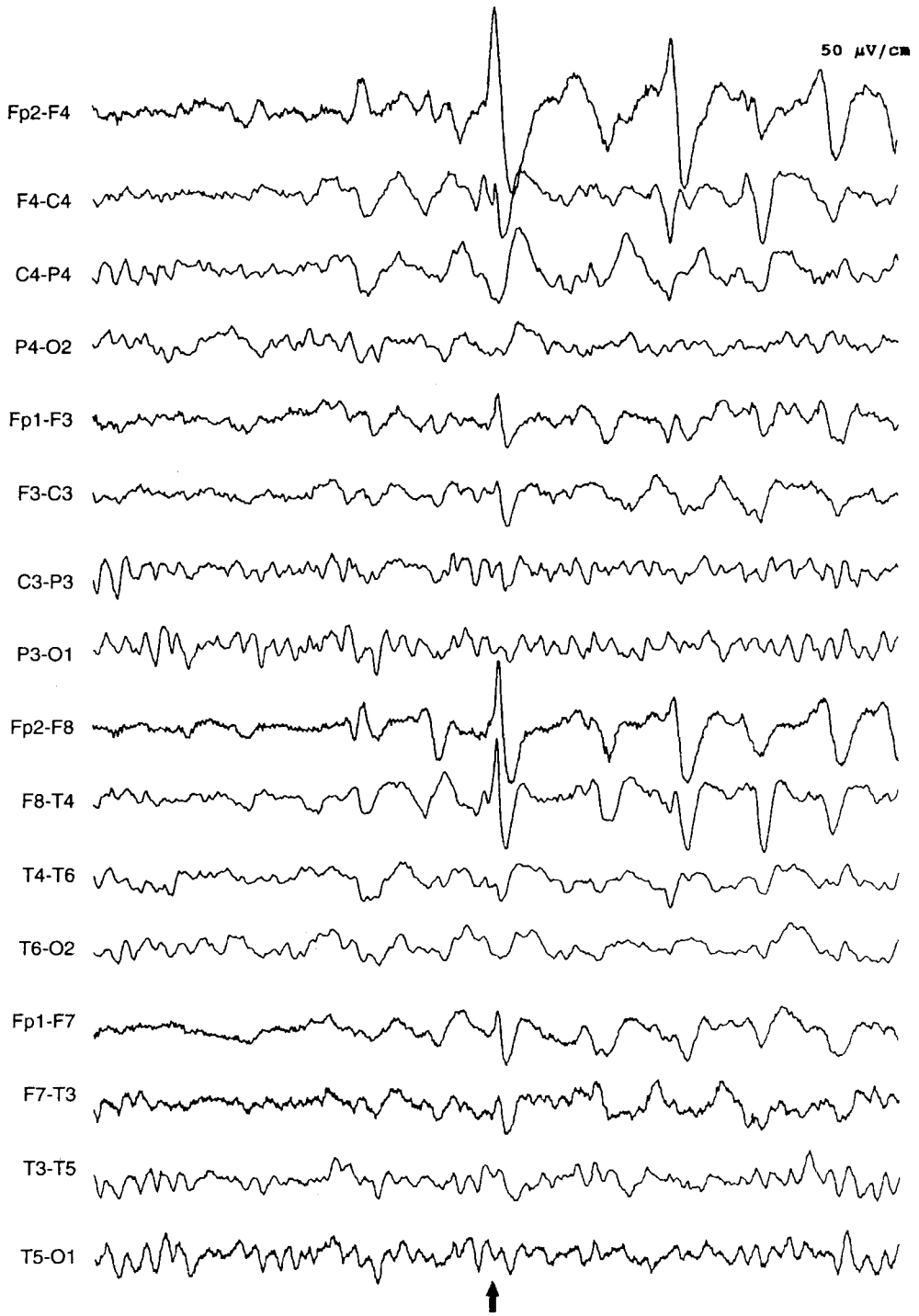

Figure 2.
A 6.0 sec segment of EEG centered on a questionable epileptiform event reported by the detection system. The event was reported as definitely epileptiform by the 2 EEGers in EEGers-I who read this EEG and had the benefit of temporal context of the entire 20 minute EEG. In contrast, on being able to view only the EEG segment, both EEGers in EEGers-II were only able to report the event as questionable. Although reasonably high amplitude, the epileptiform event has been confounded by the surrounding background showing generalized slowing, due to the patient being asleep.

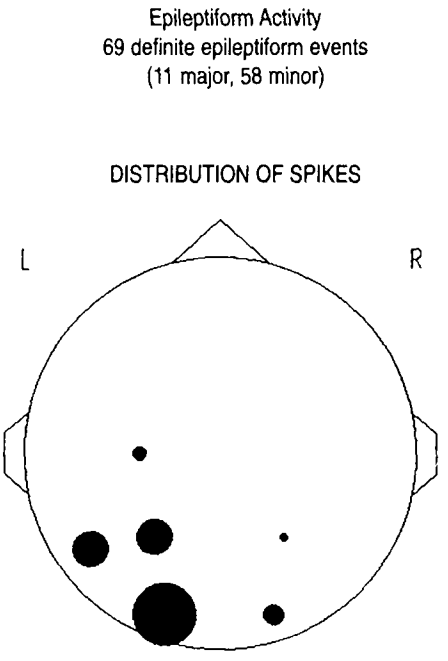

Figure 3.
System output for an EEG in which definite EA was detected by the PC. The definite events are subdivided into major and minor according to whether they were considered definite on spatial grounds alone or whether both spatial and temporal context was needed. A topographic map displays the distribution of spikes over the head. The area of a particular disc indicates the proportion of spikes arising at that electrode. In this case, the discs indicate a spike focus under or close to the left occipital electrode.

from these. The topographic map of spikes was used to determine the distribution of the EA.

V. Data organization

To create a gold standard against which performance of the detection system could be assessed, the data from EEGers-I was combined for each patient. This involved the use of a subjective scheme that was felt to best represent the certainty of there being EA in each patient's EEG based on how each EEGer had rated it. There were 16 possible combinations of rating by EEGers, assuming that ratings could be regarded exchangeably. Each combination of ratings was then placed in one of three combined categories — none, questionable, or definite — that was felt to best represent the combined EEGers assessment of the presence of EA in the EEG. Table 1 shows the combining scheme used for EEGers-I. In essence, the combining scheme was designed to place the EEGs in the questionable category unless there was good agreement for or against the presence of EA. A similar scheme was used to combine the reports from EEGers-II, with EEGs placed in the questionable category unless there was exact agreement.

Table 1

Combining scheme for detection of epileptiform activity by EEGers-I and number of each combination observed

EEGers' classifications	Combined classification	Number
N, N, N	N	80
N, N, —	N	348
Q, N, N	N	9
D, N, N	Q	1
Q, N, —	Q	21
D, N, —	Q	7
Q, Q, N	Q	1
Q, Q, —	Q	3
Q, Q, Q	Q	1
D, Q, —	Q	10
D, Q, N	Q	2
D, Q, Q	Q	0
D, D, N	D	3
D, D, —	D	26
D, D, Q	D	0
D, D, D	D	9
	Total	521

D = Definite; Q = Questionable; N = None; — = Not read

VI. Data analysis

The performance of the system was measured in terms of sensitivity and selectivity in a similar manner to that described by Webber et al[1] but at a global rather than individual event level. Thus the results reported parallel more closely the global approach taken by the EEGer when reading EEGs in clinical practice. The quantities used to calculate percentage sensitivity and selectivity were *EA count* (number of EEGs in which EEGers-I reported EA), *detector count* (number of EEGs in which the PC reported EA), and *match count* (number of EEGs for which both EEGers-I and the system reported the presence of EA). Using these quantities the percentage sensitivity and selectivity were defined as:

$$Sensitivity = \frac{100 \times match\ count}{EA\ count}$$

$$Selectivity = \frac{100 \times match\ count}{detector\ count}$$

The sensitivity and selectivity of the system relative to EEGers-I were calculated for EEGs containing definite EA as well as for EEGs containing either definite or questionable EA. The rate of false detections at the individual spike level was also calculated to facilitate comparison of performance with other spike detection systems.

RESULTS

I. Detection of epileptiform activity

Of the 521 EEGs, GC reported 70 (13.4%) as containing definite or questionable EA, with 50 (9.6%) considered definite. EW read 312 EEGs and reported the presence of

Table 2

Detection of epileptiform activity — EEGers-I versus detection system

EEGers-I	System			
	None	Questionable	Definite	Total
None	303	105	29	437
Questionable	20	13	13	46
Definite	2	7	29	38
Total	325	125	71	521

Table 3

Detection of epileptiform activity — EEGers-I versus EEGers-II

EEGers-I	EEGers-II (including system)			
	None	Questionable	Definite	Total
None	388	20	29	437
Questionable	29	4	13	46
Definite	2	7	29	38
Total	419	31	71	521

EA in 50 (16%), with 26 (8.3%) EEGs considered definite. SD read 315 EEGs and reported EA in 36 (11.4%), with 29 (9.2%) considered definite.

For the 106 EEGs read by all three of EEGers-I there was agreement in 85% of cases, while for the 415 EEGs read by only two of EEGers-I the mean level of agreement on the classification of EA was 89%. A substantial proportion of this inter-EEGer agreement was due to the large number of EEGs (428 or 82%) in which no EEGer reported the presence of EA. If these cases are disregarded, the level of agreement drops markedly. For cases in which two EEGers read the EEG, their average agreement on the presence of definite or questionable EA was only 55% for EEGs in which at least one EEGer reported EA. For the cases where three EEGers read the EEG their agreement on the presence of EA was only 39% for EEGs in which at least one EEGer reported EA. Table 1 contains the frequency data for each combination of EEGer interpretations observed in this study.

The average percentages of missed and false definite epileptiform EEGs for EEGers-I were calculated on a pairwise basis and were both found to be 12%. The reports of the three EEGers were combined to obtain a gold standard for the 521 EEGs — 437 (83.9%) EEGs containing no EA, 46 (8.8%) EEGs containing questionable EA and 38 (7.3%) EEGs containing definite EA.

The detection system detected the presence of EA in 196 (37.6%) of the 521 EEGs, with 71 (13.6%) classified as definite. Table 2 shows the contingency table of combined EEGers-I *versus* the detection system. By examining the cells on the main diagonal of the Table it can be seen that EEGers-I and the system agreed on 345 (66%) cases.

The system's sensitivity and selectivity were calculated, with the combined data from EEGers-I being considered as the gold standard. The system had a sensitivity of 76% for correctly classifying EEGs reported as containing definite EA by the EEGers. The system's selectivity for EEGs containing definite EA was 41%. The low selectivity of the system reflects the relatively large number of false detections.

Complete disagreement between the system and EEGers-I occurred in 31 (6.0%) cases. Of these, 2 were missed definite detections by the system and 29 were false definite detections. The 2 missed definite detections out of the 38 definite detections reported by EEGers-I gives a missed detection level of 5%, while the 29 false definite detections out of the 71 definite detections reported by the system gives a false detection level of 41%.

The system's false detection rate was determined from the 59 false definite events reported by the system within the 29 false definite EEGs. As there were 437 nonepileptiform EEGs (146 hours) this indicates a false detection rate of 0.41 false definite detections per hour.

EEG segments containing the events in the 125 EEGs that the system reported as questionable were reviewed by EEGers-II. One EEGer classified 114 of these as containing no EA and 11 as containing questionable EA, with none considered to contain definite EA. The other EEGer reported definite EA in 3 EEGs, questionable EA in 22 EEGs and absence of EA in 100 EEGs. Combined results for these two EEGers gave 94 EEGs that contained no EA and 31 EEGs that contained questionable EA. Table 3 shows the contingency Table for combined EEGers-I *versus* combined EEGers-II. Of the 125 EEGs reported as containing questionable EA by the system, EEGers-I reported 105 as containing no EA. After reviewing the printouts of the questionable events, EEGers-II were able to correctly classify 85 of these 105 EEGs as nonepileptiform. EEGers-I reported 13 EEGs as containing questionable EA, of which, on reviewing the printouts, EEGers-II classified 9 as containing no EA and 4 as containing questionable EA. The remaining 7 EEGs were classified as containing definite EA by EEGers-I but classified as questionable by EEGers-II.

The number of EEGs on which there was exact agreement increased to 405 (78%) after the review by EEGers-II. Conversely, the cases of complete disagreement were not affected by the review of the questionable spikes. The sensitivity and selectivity of the combination of the automated system and EEGers-II to definite EEGs were also unaltered by the review. The sensitivity for definite and questionable epileptiform EEGs decreased to 63% while the selectivity rose to 52%.

II. Distribution of epileptiform activity

Of the 29 EEGs in which both EEGers-I and the system reported the presence of definite EA, EEGers-I agreed on 17 as being generalized, 4 as focal, 4 as multifocal and disagreed on the remaining 4 EEGs. The combined EEGers-

Table 4

Categorization of epileptiform activity — EEGers-I versus EEGers-II

EEGers-I	EEGers-II			
	Generalized	Focal	Multifocal	Total
Generalized	13	2	0	15
Focal	0	2	0	2
Multifocal	0	2	0	2
Total	13	6	0	19

II reported 15 as generalized, 8 as focal, none as multifocal and disagreed on 6 EEGs. No lateralized EA was reported by either group.

Table 4 gives the contingency table of combined EEGers-I *versus* combined EEGers-II for categorization of EA with the 10 cases of within-group disagreement removed. The two groups agreed on the distribution of the EA in 15 (79%) of the 19 EEGs. Two cases of incorrect distribution were due to EEGers-II misinterpreting generalized EA on the spike map as being focal and the other 2 were due to inability to distinguish multiple foci due to lack of temporal separation information on maps.

DISCUSSION

I. Conventional Interpretation of EEGs and obtaining a gold standard

EEGers-I had a high level of agreement on the classification of EEGs in the overall data-set, although in part this could be attributed to the large number of EEGs containing no EA. Conversely, for EEGs in which at least one EEGer reported EA there was surprising disagreement. Such disagreement made it difficult to obtain a gold standard against which to judge the system.

The combination scheme was devised to help overcome the difficulties created by this disagreement and was our best estimate of whether EA was actually present. It is important to note that comparing the automated system with this standard is not strictly fair. This is because disagreement between EEGers-I (which occurred for 11% of the EEGs) still led to the EEG being classified as containing no, questionable or definite EA which was different from at least one of the EEGers' interpretations. With this in mind it seems unreasonable to expect the system to agree exactly with the combined EEGers when the EEGers themselves do not agree.

Other studies have used various approaches to try to overcome the problem of not having a gold standard for automated EEG analysis. Hostetler et al[8] weighted EEGers' scores for detecting spikes by experience and obtained an average for the EEGers. This score was used to assess the ability of their detection system to detect spikes. Webber et al[12] incorporated an experience score into their analysis to assess agreement between pairs of EEGers detecting spikes. Pietilä et al[13] allowed their two EEGers to re-evaluate differences in their scoring so as to create a consensus file for use as a reference in determining the performance of two detection systems. Wilson et al[14] used a probability-based scheme to establish continuous-valued definitions of sensitivity, selectivity and specificity, although the validity of this approach has been questioned.[15] A Bayesian approach has been used by Black et al[16] to create a gold standard from which estimates of the performance of EEG readers (human or computer-based) can be obtained. In all of these studies there was considerable variation present between the expert EEGers.

II. Automated detection

The system succeeded in detecting (i.e., reporting as containing definite or questionable EA) 36 of the 38 EEGs containing definite EA, a missed detection level of only 5%, indicating that it misses very few EEGs containing definite EA. Interestingly, EEGers-I had an average within-group missed detection level of 12%.

The biggest area of disagreement between EEGers-I and the system was over the 437 EEGs classified by EEGers-I as containing no EA. These EEGs were reported by the system as containing definite EA in 29 (6.6%) cases and questionable EA in 105 (24%) cases. This indicates that the system was unable to distinguish between EA and artifacts or sharp background activity as accurately as an EEGer. This is in marked contrast to preliminary results obtained with 148 routine EEGs in which the system was found to have a sensitivity and selectivity of 100% at the global EEG level,[11] a reflection of the test data having been previously used for fine-tuning the system. The system has a false detection rate of only 0.41 per hour. Importantly, this was achieved on data containing substantial artifact.

A major problem when comparing studies in the literature is that performance of a detection system is highly dependent on the data used. For example, it is possible for a system with a high false detection rate to also have a high sensitivity and selectivity if the data contains a large amount of EA, as the proportion of correct detections can still outnumber the proportion of false detections. However, as is evidenced by this study, the majority of routine clinical EEGs are nonepileptiform, so a system with high selectivity on data containing a large amount of EA would be expected to have a much lower selectivity for routine clinical EEGs. Thus, we consider the false detection rate to be the most appropriate measure of the selectivity of the system for both standard EEG and long-term EEG monitoring.

So, how does our system compare with others reported in the literature with respect to false detection rate? Glover et al[17] used a rule-based expert system to detect spikes in approximately 90 minutes of EEG from three subjects. The false detection rate was reported to be 23 per hour for this data. Their system has been further improved by the addition of new rules and an emphasis on localization of epilep-

tic foci as opposed to detection of all epileptiform events.[7] Its performance was evaluated on a total of 6.2 hours of EEG from 18 subjects (13 with EA and 5 controls) and indicated an average false detection rate of 17 per hour.

Hostetler et al[8] used the spike detection system developed by Gotman and his colleagues.[18,19] A total of 1160 events were reported in 5 EEGs of 20 minutes duration. An average of 16% of events reported were false detections, giving a false detection rate of 111 per hour. This corresponds to the highest sensitivity setting for the system, which would naturally have the lowest selectivity. Although Hostetler et al have not provided sufficient data to calculate an exact figure, we estimate that the setting giving the highest selectivity would result in a false detection rate of approximately 37 per hour. The 111 false detections per hour corresponds to a rate of 117 false detections per hour (based on an average of 195 false detections per 100 minute EEG from 20 patients) reported in a study by Gotman and Wang.[5] Similarly, Gotman and Wang[20] assessed an improved version of their mimetic system which made use of wide temporal context by classifying EEG into one of five states (degrees of wakefulness and sleep) and applying state-dependent rules to help reject artifacts. This system reduced the false detection rate to 47 per hour.

An artificial neural network (3-layer feed-forward network trained by back-propagation) was used by Gabor and Seyal[21] to detect spikes in 5 epileptiform EEGs. However, as the system had to be trained on predetermined spikes for each patient it would seem to have limited clinical utility. The system had an average false detection rate of 64 per hour.

Webber et al[1] used a similar artificial neural network to detect spikes in 10 EEGs of 1.5 to 2.5 minutes duration. At the crossover, a sensitivity and selectivity of 73.3% was reported and a false detection rate of 592 per hour could be deduced.

Pietilä et al[13] used adaptive segmentation to perform automated EA detection in 12 EEGs from six patients with epilepsy. The false detection rate was not reported but the system was said to have a lower selectivity (although higher sensitivity) than the Gotman system.[18,19]

The spike detection systems reviewed above have false detection rates that are, at best, 41 times higher than our system. This indicates that if they were tested on our data they would have a sensitivity at the global EEG level of 100% for definite epileptiform EEGs but, conversely, a very low selectivity as a result of most nonepileptiform EEGs being reported as containing EA. This highlights the effect that different data — particularly the ratio of normal to epileptiform EEGs — can have on the apparent performance of an automated spike detection system.

It is notable that the two expert systems — the one used by Ramabhadran[7] and that used in this study — achieved the lowest false detection rates of any of the systems reviewed here. Also, contrary to Webber et al's[1] conclusion, this study has demonstrated that it is possible for a rules-based system to detect EA in real-time using software run on a PC. Furthermore, this can be done with reasonable accuracy, although not as yet at a level comparable with expert EEGers.

III. Interpretation of computer-selected EEG segments

In their review of EEGs reported as containing questionable EA by the system, EEGers-II interpreted the 6 sec segments of raw EEG with a good level of agreement. In most cases, the assessments of these EEGers matched that of combined EEGers-I, meaning that the review process was successful at eliminating false questionable detections by the system.

Although EEGers-II were successful in rejecting non-EA reported as questionable by the system, the review process made it more likely that questionable spikes (according to EEGers-I) would also be rejected. This happened in a number of cases and probably reflects the limited temporal information available to EEGers-II during review of EEG segments. Other than questionable events, EEGers-II had by definition to report on the presence of definite EA simply on the basis of the computer outputs.

IV. Distribution of epileptiform activity

EEGers-I agreed well on distribution of EA in the 29 EEGs reported as containing definite EA by both groups of EEGers. EEGers-II also agreed well with each other's classification, indicating a high level of within-group uniformity in the two groups. These high levels of agreement partly reflect the decision to examine only those EEGs in which both combined groups detected definite EA. There was also a high level of inter-group agreement between EEGers-I and EEGers-II, indicating that the spike topographical maps could be interpreted with good accuracy with respect to the distribution of EA.

V. Clinical Utility

Even though our spike detection system has a much lower false detection rate than that reported for other systems, it is still too high to be considered for routine clinical use without review of reported events. Thus, its most appropriate application in its current state is to screen for the presence of EA. This would involve using it to detect possible EA in routine EEGs or in long-term EEG monitoring, and having these EEGs or segments reviewed by an EEGer. The low missed detection rate of the current system would mean that the workload of the EEGer could be reduced by 62% (325/521) without an unacceptably high number of EEGs containing definite EA (approximately 5%) being missed.

There remains considerable scope in our detection system for further improvements in sensitivity and selectivity. These will, in turn, enhance the value of our system as a diagnostic tool. Our system would also be improved by the addition of algorithms specifically tuned for detection of

electrographic seizures, although many are identified already by detection of spikes within bursts.

In addition, we are investigating a number of alternative techniques to incorporate into the spike detection system. Particularly promising are (a) pre-detection spike enhancement by both neural-network-based multireference adaptive noise cancelling[22] or 3-dimensional adaptive spatial filtering of deep EA,[23] and (b) epileptiform waveform recognition using a hybrid approach comprising mimetic, self-organizing neural network, and rule-based fuzzy logic,[24-26] and wavelet analysis.[27]

This study has further strengthened our contention that multi-channel spatial context and wide-temporal context (particularly the presence of EA elsewhere in the EEG) must be utilized in a spike detection system for it to be able to match and, ultimately, better the expert EEGer.

SUMMARY

The aim of this study was to determine the performance of a PC-based system for real-time detection and topographical mapping of epileptiform activity (EA) in the EEG during routine clinical recordings. The system incorporates a mimetic stage to locate candidate spikes (including sharp-waves) followed by two expert-system-based stages, which utilize spatial and wide-temporal contextual information in deciding whether candidate events are epileptiform or not. The data comprised 521 consecutive routine clinical EEG recordings (173 hours). Performance was evaluated by comparison with three independent electroencephalographers (EEGers-I). A second group of two EEGers (EEGers-II) separately interpreted the spike topographical maps and, for EEGs categorized as containing only questionable EA by the detection system, reviewed 6 sec segments of raw EEG centered on each questionable event. Thirty-eight of the EEGs were considered to contain definite EA by at least two of EEGers-I. The false detection rate of the system was 0.41 per hour. The system was found to have a sensitivity of 76% and a selectivity of 41% for EEGs containing definite EA. However, it only missed detection of EA in 5% of the recordings. EEGers-II agreed with EEGers-I on the distribution (generalized, lateralized, focal, multifocal) of EA in 79% of cases. This is by far the largest clinical evaluation of computerized spike detection reported in the literature and the only one to apply this in routine clinical recordings. The false detection rate is the lowest ever reported, suggesting that this multi-stage rule-based system is a powerful and practical tool in clinical electroencephalography and long-term EEG monitoring.

ACKNOWLEDGMENTS

We wish to thank Dr. Elizabeth Walker and Dr. Suzanne Davis for their help in reading EEGs and Karen Burgess for her help in the organization of the raw data and data entry. We also wish to thank New Zealand Lotteries and the Canterbury Medical Research Foundation for their financial support of this project.

REFERENCES

1. Webber WRS, Litt B, Wilson K, Lesser RP. Practical detection of epileptiform discharges (EDs) in the EEG using artificial neural network: a comparison of raw and parameterized EEG data. Electroencephalogr Clin Neurophysiol 1994; 91: 194-204.
2. Dümpelmann M, Elger CE. Visual and automatic investigation of epileptiform spikes in intracranial EEG recordings. Epilepsia 1999; 40: 275-285.
3. Davey BLK, Fright WR, Carroll GJ, Jones RD. Expert system approach to detection of epileptiform activity in the EEG. Med Biol Eng Comput 1989; 27: 365-370.
4. Glover JR, Raghavan N, Ktonas PY, Frost JD. Context-based automated detection of epileptogenic sharp transients in the EEG: elimination of false positives. IEEE Trans Biomed Eng 1989; 36: 519-527.
5. Gotman J, Wang LY. State-dependent spike detection: concepts and preliminary results. Electroencephalogr Clin Neurophysiol 1991; 79: 11-19.
6. Dingle AA, Jones RD, Carroll GJ, Fright WR. A multi-stage system to detect epileptiform activity in the EEG. IEEE Trans Biomed Eng 1993; 40: 1260-1268.
7. Ramabhadran B, Frost JD, Glover JR, Ktonas PY. An automated system for epileptogenic focus localization in the electroencephalogram. J Clin Neurophysiol 1999; 16: 59-68.
8. Hostetler WE, Doller JH, Homan RW. Assessment of a computer program to detect epileptiform spikes. Electroencephalogr Clin Neurophysiol 1992; 83: 1-11.
9. Jones RD, Dingle AA, Carroll GJ, et al. A system for detecting epileptiform activity in the EEG: real-time operation and clinical trial. Proc Ann Int Conf IEEE Eng Med Biol Soc, Amsterdam, The Netherlands 1996; 18: 948-949.
10. Carroll GJ, Dingle AA, Jones RD, Donaldson IM, Parkin PJ, Burgess KL. Automated detection of epileptiform activity in the EEG. Electroencephalogr Clin Neurophysiol 1995; 97: S242.
11. Jones RD, Dingle AA, Carroll GJ, et al. A PC-based system for automated analysis of the EEG. Proc Med Eng Week World Conf, Taipei, Taiwan 1994; 1: 153-157.
12. Webber WRS, Litt B, Lesser RP, Fisher RS, Bankman I. Automatic EEG spike detection: what should the computer imitate? Electroencephalogr Clin Neurophysiol 1993; 87: 364-373.

13. Pietilä T, Vapaakoski S, Nousiainen U, et al. Evaluation of a computerized system for recognition of epileptic activity during long-term EEG recording. Electroencephalogr Clin Neurophysiol 1994; 90: 438-443.
14. Wilson SB, Harner RN, Duffy FH, Tharp BR, Nuwer MR, Sperling MR. Spike detection. I. Correlation and reliability of human experts. Electroencephalogr Clin Neurophysiol 1996; 98: 186-198.
15. Black MA, Jones RD. Sensitivity and selectivity for continuous perception values: a comment. Electroencephalogr Clin Neurophysiol 1998; 106: 457-459.
16. Black MA, Jones RD, Carroll GJ, Smith MH. Bayesian evaluation of automated EEG analysis system in the absence of a gold standard. NZ Med J 1997; 110: 444.
17. Glover JR, Varmazis DN, Ktonas PY. Continued development of a knowledge-based system to detect epileptogenic sharp transients in the EEG. Proc Ann Int Conf IEEE Eng Med Biol Soc, Philadelphia, USA 1990; 12: 1374-1375.
18. Gotman J, Gloor P. Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG. Electroencephalogr Clin Neurophysiol 1976; 41: 513-529.
19. Gotman J, Ives JR, Gloor P. Automatic recognition of interictal epileptic activity in prolonged EEG recordings. Electroencephalogr Clin Neurophysiol 1979; 46: 510-520.
20. Gotman J, Wang LY. State-dependent spike detection: validation. Electroencephalogr Clin Neurophysiol 1992; 83: 12-18.
21. Gabor AJ, Seyal M. Automated interictal EEG spike detection using artificial neural networks. Electroencephalogr Clin Neurophysiol 1992; 83: 271-280.
22. James CJ, Hagan MT, Jones RD, Bones PJ, Carroll GJ. Multireference adaptive noise cancelling applied to the EEG. IEEE Trans Biomed Eng 1997; 44: 775-779.
23. Ward D, Jones R, Bones P, Carroll G. Enhancement of deep epileptiform activity in the EEG via 3-D adaptive spatial filtering. IEEE Trans Biomed Eng 1999; 46: 707-716.
24. James CJ, Jones RD, Bones PJ, Carroll GJ. The self-organizing feature map in the detection of epileptiform transients in the EEG. Proc Ann Int Conf IEEE Eng Med Biol Soc, Amsterdam, The Netherlands 1996; 18: 913-914.
25. James CJ, Jones RD, Bones PJ, Carroll GJ. Spatial analysis of multi-channel EEG recordings through a fuzzy-rule based system in the detection of epileptiform events. Proc Ann Int Conf IEEE Eng Med Biol Soc, Hong Kong 1998; 20: 2175-2178.
26. James CJ, Jones RD, Bones PJ, Carroll GJ. Detection of epileptiform discharges in the EEG by a hybrid system comprising mimetic, self-organized artificial neural network, and fuzzy logic stages. Clin Neurophysiol 1999; 110: 2049-2063.
27. Goelz H, Jones RD, Bones PJ. Wavelets and their application to biomedical signal analysis. Proc Ann Int Conf IEEE Eng Med Biol Soc, Atlanta, USA 1999; 21: (CD-ROM).