Research Repository UCD

Title	Improved patient specific seizure detection during pre-surgical evaluation
Authors(s)	Chua, Eric C.-P., Patel, Kunjan, Fitzsimons, Mary, Bleakley, Chris J.
Publication date	2011-04
Publication information	Chua, Eric C.-P., Kunjan Patel, Mary Fitzsimons, and Chris J. Bleakley. “Improved Patient Specific Seizure Detection during Pre-Surgical Evaluation” 122, no. 4 (April, 2011).
Publisher	Elsevier
Item record/more information	http://hdl.handle.net/10197/7034
Publisher's statement	This is the author’s version of a work that was accepted for publication in Clinical Neurophysiology. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Clinical Neurophysiology (VOL 122, ISSUE 4, (2011)) DOI: 10.1016/j.clinph.2010.10.002
Publisher's version (DOI)	10.1016/j.clinph.2010.10.002

Downloaded 2024-04-19 00:37:44

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)

© Some rights reserved. For more information

*Manuscript

Improved Patient Specific Seizure Detection during Pre-Surgical Evaluation

Eric C-P Chua, PhD[1]

Kunjan Patel, MSc[1]

Mary Fitzsimons, MSc[2]

Chris J. Bleakley, PhD[1]

[1]Complex and Adaptive Systems Laboratory and School of Computer Science and Informatics, University College Dublin, Dublin 4, Ireland

[2]Beaumont Hospital, Dublin 9, Ireland

* Eric C-P Chua is now with Duke-NUS Graduate Medical School, Singapore

Address for Correspondence:

Eric C-P Chua, PhD

Complex and Adaptive Systems Laboratory and School of Computer Science and Informatics

University College Dublin

Dublin 4, Ireland

Tel: (353) 1 716 5353

Email: eric.chua.2@gmail.com

Acknowledgement:

This work was supported by Science Foundation Ireland via the EEDSP Strategic Research Cluster (ref: 07/SRC/I1169). All authors indicate no financial conflict of interest.

Keywords: seizure detection, interactive machine learning, subject-specific, adaptation

Abstract

Objective: There is considerable interest in improved off-line automated seizure detection methods that will decrease the workload of EEG monitoring units. Also, subject-specific approaches have been demonstrated to perform better than subject-independent ones. However, for pre-surgical diagnostics, the traditional method of obtaining *a priori* data to train subject-specific classifiers is not practical. We present an alternative method that works by adapting the threshold of a subject-independent to a specific subject based on feedback from the user.

Methods: A subject-independent quadratic discriminant classifier incorporating modified features based partially on the Gotman algorithm was first built. It was then used to derive subject-specific classifiers by determining subject-specific posterior probability thresholds via user interaction. The two schemes were tested on 529 hours of intracranial EEG containing 63 seizures from 15 subjects undergoing pre-surgical evaluation. To provide comparison, the standard Gotman algorithm was implemented and optimised for this dataset by tuning the detection thresholds.

Results: Compared to the tuned Gotman algorithm, the subject-independent scheme reduced the false positive rate by 51 % (0.23 h^{-1} to 0.11 h^{-1}) while increasing sensitivity from 53 % to 62 %. The subject-specific scheme further improved sensitivity to 78 %, but with a small increase in false positive rate to 0.18 h^{-1}).

Conclusions: The results suggest that a subject-independent classifier scheme with modified features is useful for reducing false positive rate, while subject adaptation further enhances performance by improving sensitivity. The results also suggest that

the proposed subject-adapted classifier scheme approximate the performance of the subject-specific Gotman algorithm.

Significance: The proposed method could potentially help increase productivity of offline EEG analysis. The approach could also be generalised to enhance the performance of other subject independent algorithms.

1 Introduction

Epilepsy is a neurological condition characterised by a recurring and random tendency of the brain to produce sudden bursts of abnormal electrical activity that disrupt other brain functions (Waterhouse 2003). Such episodes are called seizures; clinical manifestations include loss of awareness or consciousness, and disturbances of movement and sensation (WHO 2009). Epilepsy is highly prevalent, affecting at least 8.2 per 1,000 of the general population, and has profound social, physical and psychological consequences (WHO 2009).

The electroencephalogram (EEG) is currently one of the most important diagnostic tools for epilepsy, as seizures usually manifest characteristic signatures on the EEG (Smith 2005). The routine EEG, typically 45 minutes to 1 hour in duration, is the most commonly performed EEG investigation on an individual with a suspected seizure disorder (Cascino 2002). However, in a significant number of patients, such as those with medically refractory epilepsy and who are candidates for surgical resection, long-term, continuous EEG evaluation is subsequently required (Cascino 2002). The large patient numbers and sheer volume of data generated makes analysing long-term EEG records extremely resource-intensive, and puts pressure on many epilepsy units, who already face resource constraints (Waterhouse 2003). Automated seizure detection algorithms have therefore been sought to alleviate the workload and to increase capacity.

The most familiar of these is the Gotman algorithm (Gotman 1999). Several algorithms have been proposed since the Gotman algorithm. For example, Meier *et al* recently introduced a promising subject independent algorithm that takes into account EEG seizure morphology (Meier *et al.* 2008). However, to our knowledge, the new

algorithms have not found widespread adoption, and the Gotman algorithm remains the current mainstay in most epilepsy units.

For automated seizure detection algorithms, subject-specific approaches have been shown to perform better than global, subject independent approaches (Qu and Gotman 1997) (Shoeb *et al.* 2004). The traditional approach to building subject-specific systems has been to first record and score data from individuals, and then use this training data to develop subject-specific classifiers. While this approach could be feasible in other settings, it is not practical for the typical application in epilepsy units, i.e. pre-surgical evaluation of long-term EEG. An alternate approach is therefore desired.

To our knowledge there is limited prior work on methods to adapt subject-independent seizure detection algorithms into subject-specific solutions. Haas *et al* (Haas *et al.* 2007) proposed a method to adapt the Osorio–Frei algorithm (Osorio *et al.* 1998). Although promising, its performance across a large number of subjects is unclear as only results for a few example subjects were presented. Furthermore, the scope for generalisation of the method is somewhat restricted as it is specific to algorithms employing spectral filters exclusively.

From a technical point of view, all previously-proposed seizure detection algorithms that we are aware of were developed using the classical, i.e. automated approach, where the algorithm is implemented in software and the classifier is applied solely by executing the software (Duda *et al.* 2000). Recently, a new approach, termed 'interactive machine learning', has been proposed, where users actually engages in generating the classifier (Ware *et al.* 2001). The key advantage of this approach is that it offers a natural way of integrating *a priori* and expert knowledge into the modelling

stage, and it has demonstrated its potential to be more successful than classifiers built using the traditional approach (Ware *et al.* 2001).

In the clinical and technical contexts discussed, we present a method that adapts a subject-independent quadratic discriminant classifier to each subject. The method involves some user interaction, and delivers enhanced performance over the subject-independent classifier with minimal user overhead and overall, improved productivity. It could also be generalised to enhance the performance of other subject independent algorithms where classification is made by thresholding a generated confidence measure.

This paper is structured as follows: we first describe the data used to develop the method (Section 2.1), and the features used for classification (Section 2.2). Following that, we describe how the subject independent classifier was developed (Section 2.3), and then the proposed method to adapt the subject independent classifier to subject specific classifiers is explained (Section 2.4). Reference configurations built to provide comparison with the proposed schemes are described in Sections 2.5 and 2.6. We then present the results (Section 3), and this is followed by a discussion of the method and results (Section 4).

2 Method

2.1 Subjects and Data

De-identified data was obtained from the Freiburg EEG Database (Freiburg 2009). Briefly, the database contains long-term, intracranial EEG obtained while subjects underwent pre-surgical monitoring at the Epilepsy Centre of the University Hospital of Freiburg, Germany. Six channels of EEG were available for each subject. The ictal and

inter-ictal periods were determined by experienced epileptologists by visual inspection, based on identification of typical seizure patterns preceding clinically manifest seizures (Freiburg 2009). Ictal and inter-ictal annotations were provided as part of the database. EEG was recorded at 256 Hz using a Neurofile NT digital video-EEG system and a 16-bit ADC.

Data from 15 subjects who exhibited generalised tonic-clonic seizures (M/F 5/10, age 33 ± 12 yrs) was analysed. On average, each subject contributed 9.5 (± 3.0) hours of ictal data (which included on average 4 ± 1 seizures), and 24.0 (± 1.0) hours of inter-ictal data. A total of 529 hours of EEG (63 seizures) was analysed.

2.2 Feature Extraction

For each subject, each EEG channel was bandpass filtered between 0.3 and 30 Hz, and then divided into continuous 2-second, non-overlapping epochs (segments). For each epoch, a number of features were computed (details are provided in the subsequent paragraphs). Each EEG channel was processed similarly, and the respective features were then averaged across the channels.

2.2.1 Standard Gotman Feature Set

The standard Gotman features (Gotman 1999) (Chang *et al.* 2005) are a) average half-wave amplitude, b) average half-wave duration and c) coefficient of variation of half-wave duration. For each epoch, the peaks and troughs of the filtered EEG were first identified, and spurious detections were removed. Spurious detections are low-amplitude EEG fluctuations that are falsely detected as peaks or troughs. They occur between actual peaks and troughs. A threshold was used to identify these spurious detections. The amplitudes of all detected peak-trough pairs were first calculated, and the threshold was taken as half of the median of all amplitudes. Pairs

with amplitude smaller than the threshold were identified as spurious detections and removed. Each EEG segment between a trough and the next peak, and between a peak and the next trough was then taken as a half-wave. Half-wave amplitude was computed as the peak to trough amplitude, and half-wave duration as the peak to trough time interval. Half-wave amplitude and duration values in each epoch were then averaged to obtain the average half-wave amplitude and duration features for that epoch respectively. The coefficient of variation feature was obtained as the ratio of the variance of half-wave durations to the square of average half-wave duration.

2.2.2 Line Length

Line length has been proposed as a potential feature for seizure detection in both adults (D'Alessandro *et al.* 2003) and neonates (Greene *et al.* 2008). For each EEG epoch x of length N, line length L was computed using (1)

$$L = \sum_{k=1}^{N} | x[k] - x[k-1] | \qquad (1)$$

2.2.3 Rectified Zero Crossings

This feature was motivated by observing that seizures were often associated with EEG frequency changes from the baseline. However, changes can be either increases or decreases. We therefore use 'rectified zero crossings', which estimates the absolute value of any change in EEG frequency from the baseline.

To compute this feature, the number of EEG zero crossings in each epoch were first determined. This time series of zero crossings was then centred by subtracting its own median, and the absolute value then taken.

Rectified zero crossings is conceptually similar to existing frequency-based features, such as Gotman's average half-wave duration, and can be considered a modified Gotman feature. Its advantage, is that its value always increases regardless whether the EEG frequency increases or decreases relative to the baseline, whereas existing features such as Gotman's average half-wave duration follow the polarity of the change (positive or negative). Figure 1 shows a graphical comparison between rectified zero crossings and Gotman's average half-wave duration for the current dataset. As shown in Figure 1, the feature is more useful for classification as it provides a consistent positive separation between seizure and non-seizure epochs (i.e., the seizure epochs tend to have a higher feature value) regardless of the polarity of change in frequency.

We decided to use rectified zero crossings instead of rectified average halfwave duration as determining zero crossings from the EEG signal is computationally simpler and we found it to be more reliable than determining peaks and troughs when computing halfwave duration.

2.3 Subject Independent Classification

2.3.1 Quadratic Discriminant Classifier

A modified feature set consisting of four features was used together with a Quadratic Discriminant Analysis (QDA) classifier to distinguish between seizure and non-seizure. The four features were: a) relative half-wave amplitude, b) rectified zero crossings, c) coefficient of variation of half-wave duration and d) line length. Each feature was log-transformed to closer approximate Gaussian distributions (McDonald 2009), and normalised to zero mean and unit standard deviation to prevent subsequent numerical difficulties.

Briefly, for each epoch (with associated feature vector **x**), its discriminant values $g_i(\mathbf{x})$ for each class ω_i, i=1,2 (Seizure, Non-seizure), were computed according to (Eq. 2) (Duda *et al.* 2000)

$$g_i(\mathbf{x})=\mathbf{x}^T\mathbf{W}_i\mathbf{x}+\mathbf{w}_i^T\mathbf{x}+\omega_{i0} \tag{2}$$

where

$$\mathbf{W}_i=-\frac{1}{2}\Sigma_i^{-1} \tag{3}$$

$$\mathbf{w}_i=\Sigma_i^{-1}\mu_i \tag{4}$$

and

$$\omega_{i0}=-\frac{1}{2}\mu_i^T\Sigma_i^{-1}\mu_i-\frac{1}{2}\,ln\,|\Sigma_i|+ln\,P(\omega_i) \tag{5}$$

where μ_i are the class means, Σ_i are the class covariance matrices and $P(\omega_i)$ are the class *a priori* probabilities, and these were empirically estimated from the entire dataset. The class posterior probabilities $P(\omega_i|\mathbf{x})$ were then obtained using (Eq. 6)

$$P(\omega_i|\mathbf{x})=\frac{exp(g_i)}{\sum_{i=1}^{2} exp(g_i)} \tag{6}$$

As the *a priori* probability of seizure was very low (~0.004 based on empirical data from the Freiburg database), setting the posterior probability threshold to 0.5 for classification caused a significant number of seizures to be missed. In addition, there is temporal dependency in the posterior probability as seizures occur in episodes that are sustained and continuous. We therefore assumed that very short and isolated detections are physiologically unlikely. Hence, to boost classifier performance, we used a) posterior probability smoothing, and b) setting the classification threshold to less than 0.5. The optimal threshold and amount of smoothing were determined empirically from the data (details described in the next sub-section).

2.3.2 Training and Validation

A leave one subject out cross validation scheme (de Chazal *et al.* 2003) was used. In brief, one subject was omitted each time, and data from all other subjects was used to train a classifier. This involved estimating the class means, covariance matrices and *a priori* probabilities. The trained classifier was then applied on the omitted subject.

To determine the optimal threshold and smoothing window size, the posterior probability of seizure (without smoothing) for each omitted subject was first determined using the classifier trained using data from all the other subjects. A series of thresholds (step size 0.01) and smoothing window sizes (step size 10) was then applied to the computed posterior probability and the corresponding classification performance (in terms of sensitivity and false positive rate) for that omitted subject was computed. After cycling through all of the omitted subjects, the subject-averaged classification performance was then used to identify the optimal threshold/window size combination, i.e. the combination with the shortest Euclidean distance to the ideal point (i.e. 100 % sensitivity, 0 hr^{-1} false positive rate).

A simple moving average was used for smoothing. For a given threshold, epochs whose posterior probabilities exceeded the threshold were classed as seizure episodes. A successful seizure detection was considered made if a detected episode occurred between the (reference) annotated start and end times of the seizure; otherwise the episode was considered a false positive. Sensitivity was calculated as the proportion of seizure episodes detected. For example, if a particular subject had five annotated seizures and four were successfully detected by the algorithm, sensitivity for this subject would be 4/5 = 80 %. The false positive rate was taken as the number of false positive episodes per hour of the whole recording. For example, for a 36-hour recording with 3 false positive episodes, false positive rate for this subject would be 3/36 = 0.08 hr^{-1}.

2.4 Subject-Specific Adaptation

Using annotated reference data, experimentation with subject-specific Gaussian parameters (i.e. class means, covariance matrices and *a priori* probabilities), smoothing window sizes and thresholds led us to the observation that setting subject-specific posterior probability thresholds was the most effective approach (results not presented). The question therefore became one of determining the optimum subject-specific thresholds in such a way that will enhance classification performance with minimal user overhead, and deliver overall improved EEG analysis productivity.

To estimate the optimum threshold for each subject, a probability density function of subject-specific thresholds was first estimated using training data (Figure 2 shows an example; the method is described in detail later). Threshold values that divided the probability density to the left and right side of the global threshold (T0 in Figure 2) into equal areas were then determined (Figure 2), and used as test threshold values.

Figure 3 presents an overview of the adaptation process. Each subject's EEG recording is first analysed using the automated subject independent classifier described in Section 2.3 to determine 'possible' seizure episodes. The EEG technologist's experience in interpreting EEGs then used to determine the optimum subject-specific threshold interactively. The technologist is presented with a 'possible' episode and asked to score it as 'seizure' or 'non-seizure'. Based on the technologist's input, the optimum subject-specific threshold is adjusted using a decision tree algorithm (Figure 4, details described later). This adaptation process terminates whenever a 'possible' episode is positively identified as a 'seizure' episode, or when the technologist has been presented with four 'possible' episodes.

Figure 4 describes the adaptation process in detail. In general, for each threshold value being tested, the 'possible' episode with the posterior probability nearest to the test threshold value is presented to the technologist for scoring. In Step 1, two episodes, one on either side of the global threshold T0, are presented to the technologist. If both episodes are scored as non-seizures (i.e false positives), the threshold is adjusted upwards to T1U (Figure 4, upper path). If both episodes are scored as seizures, or if no more episodes with higher posterior probabilities are available, the threshold is set at T1D (Figure 4, lower path). If only one episode is positive, the global threshold is left unchanged.

In Step 2, the episode with posterior probability closest to T1U is presented. If it is positive, the threshold is set at T2D; otherwise the threshold is adjusted upwards to T2U. In Step 3, the episode nearest to T2U is presented. If it is positive, the threshold is set at T3D; otherwise the threshold is set at T2U.

The probability density function of subject-specific thresholds (Figure 2) was estimated using the same leave one subject out cross validation methodology (Section

2.3.2). For each left out subject, the trained subject-independent classifier was applied, with various threshold levels, to data from each training subject to identify the optimum subject-specific threshold for that training subject. A beta distribution (Panik 2005) was then fitted to the empirical distribution of the obtained subject-specific thresholds. This was achieved by evaluating various beta distribution parameter combinations and selecting the one with lowest least-square error with respect to the empirical distribution.

2.5 Standard Gotman Algorithm

The standard Gotman algorithm was implemented for comparison based on descriptions by Gotman (Gotman 1999) and Chang *et al* (Chang *et al.* 2005). Briefly, for each epoch, the average half-wave amplitude relative to a background value, i.e. the 'relative half-wave amplitude', was computed. The background was taken as the 16-second segment starting from the point 28 seconds before each epoch. A '*channel detection*' was made if a) the relative amplitude exceeded 3, b) the average half-wave duration was between 25 and 150 ms, and c) the coefficient of variation of half-wave duration was less than 0.6.

A '*seizure detection*' was then made if two or more *channel detections* occurred in the same or adjacent epochs, i.e. a) channel detections in adjacent epochs of the same channel, or b) channel detections in the same epoch of different channels. If the two channel detections were in the same epoch but in two different channels, then the average amplitude of the successive epoch in one of the two channels must be at least 80 % that of the detection epoch.

2.6 Reference Configurations

To differentiate the contributions of the various components of the proposed schemes (i.e. feature set, classifier model and subject adaptation), a number of reference configurations were built for comparison purposes.

2.6.1 Subject-Independent Configurations

Tuned Gotman Algorithm. The channel detection thresholds of the standard Gotman algorithm were tuned for the current dataset. This is because, although the proposed schemes were developed using cross-validation techniques to avoid over-fitting, they are based retrospectively on the given dataset. On the other hand, the standard Gotman algorithm is *a priori* non-tuned if applied as-is, and was therefore optimised for fair comparison.

Specifically, the detection thresholds for the relative amplitude and coefficient of variation of half-wave duration were tuned. Sensitivity and false positive rates for each subject were determined for various combinations of the two thresholds. The combination that minimised the mean Euclidean distance to the ideal point (100 % sensitivity, 0 h^{-1} false positive rate) was taken as the operating point.

We interpreted the thresholds for average half-wave duration (between 25 and 150 ms) as broad thresholds for valid EEG frequencies (between 3.3 and 20 Hz), and these thresholds were therefore not tuned.

Standard Gotman Feature Set + QDA. To investigate the contribution of the QDA classifier, we replaced the decision model of the standard Gotman algorithm with a QDA classifier. The classifier was built using the same methodology as described in

Section 2.3 except that here the feature set is the standard Gotman feature set (as described in Section 2.2.1).

2.6.2 Subject-Specific Configurations

Subject-Specific Gotman Algorithm. Here we sought to estimates the 'ground truth' performance of a subject-specific Gotman algorithm by obtaining individual sets of optimal detection thresholds for each subject. For each subject, sensitivity and false positive rates for various combinations of the two thresholds were determined. The combination that minimised the Euclidean distance to the ideal point was taken as the optimal thresholds for that subject.

Subject-Specific Standard Gotman Feature Set + QDA. Subject-specific QDA classifiers using the standard Gotman feature set were built using cross-validation. Briefly, for each subject, the record was divided into *n* folds, *n* being the number of seizures in that record. The proportion of ictal and inter-ictal data in each fold was kept constant. We cycled through the folds, each time leaving one fold out. A classifier was trained on all other folds, and tested on the left-out fold. Classification results (in terms of Euclidean distance from the ideal point) for various combinations of posterior probability threshold and smoothing window size were then averaged across folds, and the combination with minimum distance was selected.

We note that these classifiers are potentially biased, as there wasn't separate training and testing data. Indeed, our intention was not to develop practical subject-specific classifiers using this approach, but instead, what we wanted was to estimate the 'ground truth' performance of such a subject-specific classifier for comparison purposes.

Subject-Specific Modified Feature Set + QDA. Here we sought to estimate the 'ground-truth' performance of the subject-specific QDA classifier that uses the

modified feature set. This serves as comparison to the proposed subject-adaptation scheme, as the goal of the proposed scheme is to approximate this 'ground truth' performance. These 'ground truth' classifiers were built using the same methodology as described in the previous paragraph, except here the feature set was the modified feature set.

2.7 Statistical Analysis

Sensitivity and false positive rate performance were combined into a single metric by calculating the Euclidean distance from each sensitivity and false positive rate pair to the ideal point. Paired t tests were for statistical analysis. Where normality was not satisfied, Wilcoxon signed rank test was used. Statistical analysis was carried out using SigmaPlot (Systat Software Inc, USA).

3 Results

Figure 5 presents the performance of the proposed schemes and various reference configurations in a receiver-operator characteristic (ROC) curve layout, while Figure 6 summarises the corresponding ROC distance, sensitivity and false positive rate results.

Tuned Gotman Algorithm. The standard Gotman algorithm provided sensitivity and false positive rates of 55.8 ± 8.8 % and 0.51 ± 0.21 hr^{-1} respectively (black inverted triangle, ROC distance 0.83 ± 0.18). After optimising the detection thresholds of the Gotman algorithm for the current dataset (i.e. shift along black ROC curve to point with shortest ROC distance), ROC distance improved to 0.62 ± 0.10, although the effect did not reach statistical significance (black circle, Wilcoxon signed rank test: z = -1.852, p

= 0.068). False positive rate decreased to 0.23 ± 0.09 hr^{-1}, while sensitivity remained at similar levels (53.4 ± 10.0 %).

QDA Classifier. Replacing the decision model of the standard Gotman algorithm with a QDA classifier provided outward expansion of the ROC curve starting around the 50 % sensitivity, 0.25 hr^{-1} false positive rate point (red vs. black curves). Compared to the tuned Gotman algorithm (red circle vs. black circle), sensitivity increased to 67.3 ± 8.2 %, but with a concurrent increase in false positive rate to 0.34 ± 0.08 hr^{-1}. ROC distance improved from 0.62 ± 0.10 to 0.55 ± 0.08, but we did not manage to detect a statistically significant effect (paired t test: $t = 0.953$, $p = 0.357$, power < 0.8).

Modified Feature Set. Modifying the feature set further resulted in a general outward expansion of the ROC curve (blue vs. red curves). The optimal global posterior probability threshold and smoothing window size were 0.05 and 125 respectively. Compared to the tuned Gotman algorithm (blue circle vs. black circle), there was a simultaneous increase in sensitivity (to 61.7 ± 9.1 %) and a reduction in false positive rate (to 0.11 ± 0.04 hr^{-1}). ROC distance improved from 0.62 ± 0.10 to 0.46 ± 0.08, although the effect was just below statistical significance (paired t test: $t = 1.796$, $p = 0.094$, power < 0.8). When compared to the standard Gotman feature set + QDA classifier (blue circle vs. red circle), the results suggest that modifying the feature set had the effect of reducing false positive rate (by 67.4 %) while generally maintaining sensitivity.

Subject-Specific Classifiers. There are two subject-independent classifier configurations, one using the standard Gotman feature set (red circle) and the other using the modified feature set (blue circle). Comparing these subject-independent classifiers and their subject-specific counterparts (red and blue circles vs. squares respectively), the results suggest that adopting a subject-specific approach improves

performance by achieving operating points nearer the ideal point than the respective subject-independent ROC curves (blue and red curves), i.e. decreased ROC distance. This corresponds to increased sensitivity, although there was also an increase in false positive rates.

Subject-Adaptation. The proposed subject-adaptation scheme seeks to approximate the 'ground truth' subject-specific classifier during offline EEG analysis, without needing *a priori* subject training data. The adaptation algorithm terminated after 2 and 3 steps in 2 and 11 subjects respectively. In two subjects, the global posterior probability threshold remained unchanged. As shown in Figure 5, the proposed scheme (green triangle) closely approximates the 'ground truth' (blue square) (ROC distance: 0.37 ± 0.06 vs. 0.34 ± 0.07. Paired t-test: $t = 0.532$, $p = 0.603$). At the same time, the proposed scheme also approximates the 'ground truth' subject-specific Gotman algorithm (black square) (ROC distance: 0.37 ± 0.06 vs. 0.35 ± 0.07. Paired t-test: $t = 0.277$, $p = 0.786$). When compared to the tuned Gotman algorithm (green triangle vs. black circle), the subject-adapted scheme achieves increased sensitivity and reduced false positive rates (ROC distance: 0.37 ± 0.06 vs. 0.62 ± 0.10. Paired t-test: $t = 2.709$, $p = 0.017$).

4 Discussion

We presented a method for adapting a subject independent classifier to create subject-specific classifiers during offline EEG analysis. Compared to the commonly-used Gotman algorithm, the method provides significant improvements in sensitivity and false positive rate, and the adaptation only requires minimum user overhead in the form of the EEG technologist scoring up to four likely seizure episodes (as described in Section 2.4). Given that the standard Gotman algorithm is the current

mainstay in most epilepsy units (Chang *et al.* 2005), the proposed algorithm is potentially clinically useful as it would significantly increase the productivity of scoring EEG recordings.

For the subject-independent classifier, the selected values are 0.05 for posterior probability threshold and 125 for smoothing window size. The results suggest that weighing the posterior probability threshold very much in favor of the 'seizure' class and smoothing posterior probability values with a 250-minute window (125 x 2-minute epochs) is conducive for detection and reducing detection noise.

Recently, Meier *et al* introduced a promising subject independent algorithm that takes into account seizure morphology (Meier *et al.* 2008). Compared with the Meier algorithm, our proposed algorithm provides comparable false detection rate, albeit with lower sensitivity. Nonetheless, our proposed method adds to the body of knowledge as we have described a method that provides a way to approximate the performance of a subject-specific classifier without requiring *a priori* training data. In addition, the proposed adaptation technique can potentially be generalised as a technique to adapt subject independent classifiers. It can be applied to any algorithm that produces a confidence measure of a desired class, and where it is desirable to obtain subject specific thresholds for subsequent classification. Therefore it could in theory be applied to another subject independent classifier that performs better than the one described here, to further enhance that classifier's performance.

At the moment, there are some limitations to the proposed algorithm in a practical setting. First, we need to establish with further studies how well the method will generalise to the 20-channel surface electrode configuration commonly used in epilepsy units. Second, the usefulness of the algorithm will be limited in epilepsy units that practise real-time seizure scoring. For example, it is usual to try to capture three

typical events during the monitoring period. So by the time the four suspect events required for adaptation is recorded, the monitoring period could be ending. Third, the number of subjects available in the current dataset is rather small, causing limitations such as limited statistical power in establishing performance differences between algorithms. Further evaluation using a larger dataset would be necessary to confirm the advantages of some aspects of the proposed scheme.

In conclusion, we presented a method for improving the performance of a subject independent quadratic discriminant classifier by adapting it to subject-specific classifiers through limited user interaction. The results suggest that such a method could be clinically useful for increasing productivity of EEG analysis. The method could also be generalised to enhance the performance of other subject independent algorithms.

References

Cascino, G., 2002. Clinical indications and diagnostic yield of video-electroencephalographic monitoring in patients with seizures and spells. Mayo Clin Proc 77 (10), 1111-20.

Chang, B., Schachter, S., Schomer, D., 2005. Atlas of ambulatory EEG. Elsevier Academic Press.

de Chazal, P., Heneghan, C., Sheridan, E., Reilly, R., Nolan, P., O'Malley, M., 2003. Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea. IEEE Trans Biomed Eng 50 (6), 686-96.

Duda, R., Hart, P., Stork, D., 2000. Pattern classification. Wiley-Interscience.

D'Alessandro, M., Esteller, R., Vachtsevanos, G., Hinson, A., Echauz, J., Litt, B, 2003. Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: a report of four patients. IEEE Trans Biomed Eng 50(5), 603-15.

Freiburg, 2009. Freiburg EEG database. https://epilepsy.uni-freiburg.de/freiburg-seizure-prediction-project/eeg-database.

Gotman, J., 1999. Automatic detection of seizures and spikes. J Clin Neurophysiol 16 (2), 130-40.

Greene, B., Faul, S., Marnane, W., Lightbody, G., Korotchikova, I.,Boylan, G., 2008. A comparison of quantitative EEG features for neonatal seizure detection. Clin Neurophysiol 119 (6), 1248-61.

Haas, S., Frei, M., Osorio, I., 2007. Strategies for adapting automated seizure detection algorithms. Med Eng Phys 29 (8), 895-909.

McDonald, J.H. 2009. Handbook of Biological Statistics (2nd ed.). Sparky House Publishing, Baltimore, Maryland. pp. 160-164.

Meier, R., Dittrich, H., Schulze-Bonhage, A., Aertsen, A., 2008. Detecting epileptic seizures in long-term human EEG: A new approach to automatic online and real-time detection and classification of polymorphic seizure patterns. J Clin Neurophysiol 25 (3), 119-31.

Osorio, I., Frei, M., Wilkinson, S., 1998. Real-time automated detection and quantitative analysis of seizures and short-term prediction of clinical onset. Epilepsia 39 (6), 615-27.

Panik, M. J., 2005. Advanced statistics from an elementary point of view. Academic Press.

Qu, H., Gotman, J., 1997. A patient-specific algorithm for the detection of seizure onset in long-term EEG monitoring: possible use as a warning device. IEEE Trans Biomed Eng. 44 (2), 115-22.

Shoeb, A., Edwards, H., Connolly, J., Bourgeois, B., Ted Treves, S., Guttag, J., 2004. Patient-specific seizure onset detection. Epilepsy Behav 5 (4), 483-98.

Smith, S., 2005. EEG in the diagnosis, classification, and management of patients with epilepsy. J Neurol Neurosurg Psychiatry 76 Suppl 2, ii2-7.

Ware, M., Frank, E., Holmes, G., Hall, M., Witten, I., 2001. Interactive machine learning: letting users build classifiers. International Journal of Human Computer Studies 55 (3), 281-92.

Waterhouse, E., 2003. New horizons in ambulatory electroencephalography. IEEE Eng Med Biol Mag 22 (3), 74-80.

WHO, 2009. Epilepsy. http://www.who.int/mediacentre/factsheets/fs999/en/.

Figure Legends

Figure 1: Per-subject feature values for a) Gotman's average halfwave duration and b) rectified zero crossings for the current dataset. Values shown are per-subject means ± SD.

Figure 2: Example of a probability density function of subject-specific thresholds, estimated from training data. The dots mark threshold values that divide the probability density on either side of the global threshold (T0), by the amount stated in the parentheses. For example, the threshold T1U divides the probability density to the right of T0 by half.

Figure 3: Overview of the adaptation process. Compared to current workflow, the additional adaptation step helps to enhance the accuracy of the automated algorithm with minimal user overhead, and increases productivity overall. Please see text for detailed descriptions.

Figure 4: Decision tree algorithm for estimating subject-specific posterior probability thresholds. T*: Threshold values being tested at each step. Y, N: If positive seizure identification has been made. Please see text for detailed descriptions.

Figure 5: Performance of the proposed schemes and reference configurations, presented in a receiver-operator characteristics (ROC) layout. Black curve: ROC for standard Gotman algorithm on current dataset. Black inverted triangle: Standard detection thresholds. Black circle: Tuned detection thresholds for current dataset (point with

shortest ROC distance on the curve). Black square: Operating point for ground truth subject-specific Gotman algorithm. Red curve: ROC curve for standard Gotman feature set + QDA classifier. Red dot: Operating point for subject-independent classifier. Red square: Estimated ground truth operating point for subject-specific classifier. Blue curve: ROC curve for modified feature set + QDA classifier. Blue dot: Operating point for subject-independent classifier. Blue square: Estimated ground truth operating point for subject-specific classifier. Green triangle: Operating point of proposed subject-adapted classifier.

Figure 6: Summary of the performance of the proposed schemes and reference configurations. Panel A: Receiver-operator characteristic (ROC) distance. Panel B: Sensitivity and false positive rates. Algorithm 1: Tuned Gotman algorithm. 2: Standard Gotman feature set + subject-independent QDA classifier. 3: Modified feature set + subject-independent QDA classifier. 4: Modified feature set + subject-adapted QDA classifier. Values shown are means ± SEM. *: $p < 0.05$.

Figure1a

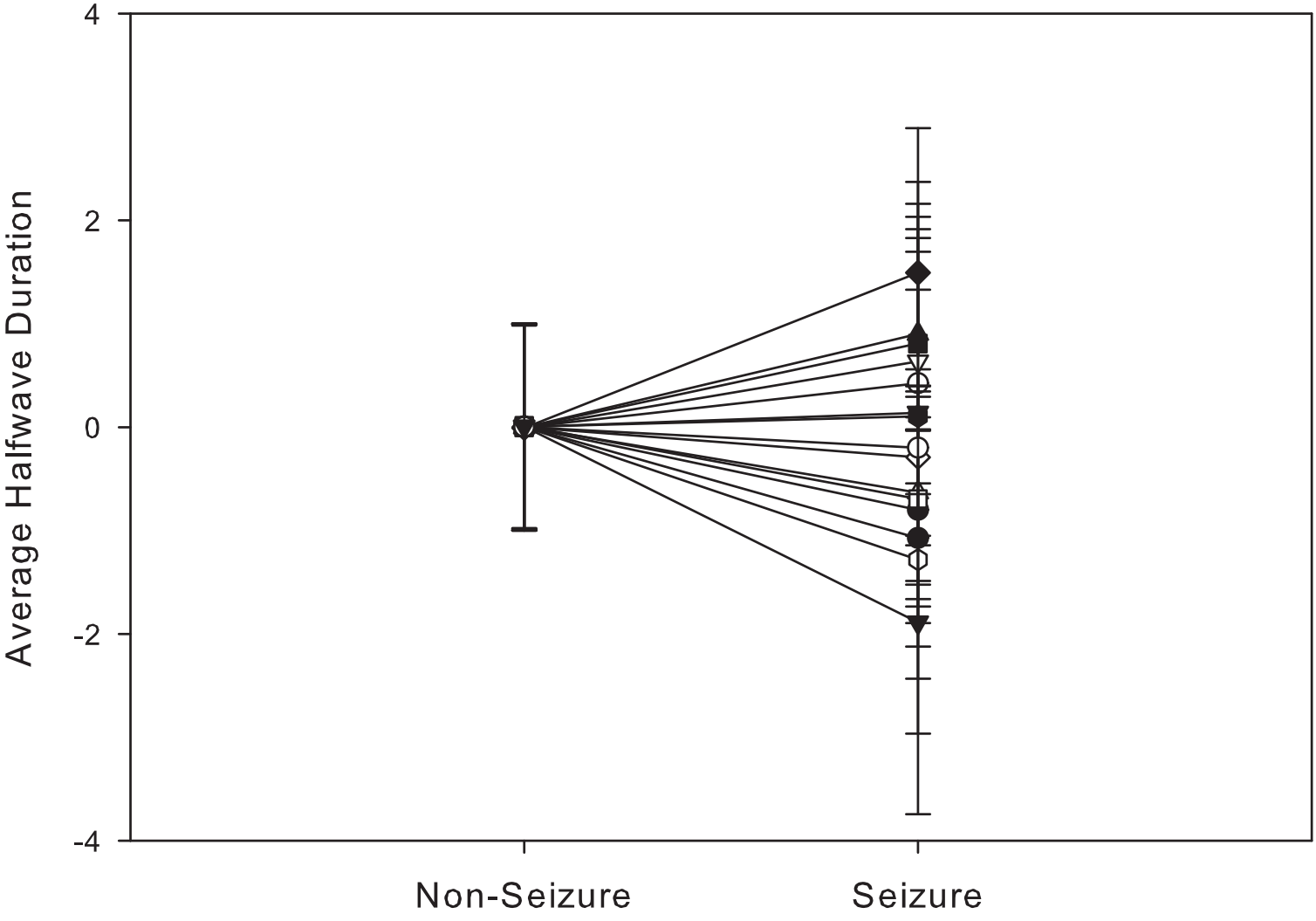

Figure1b

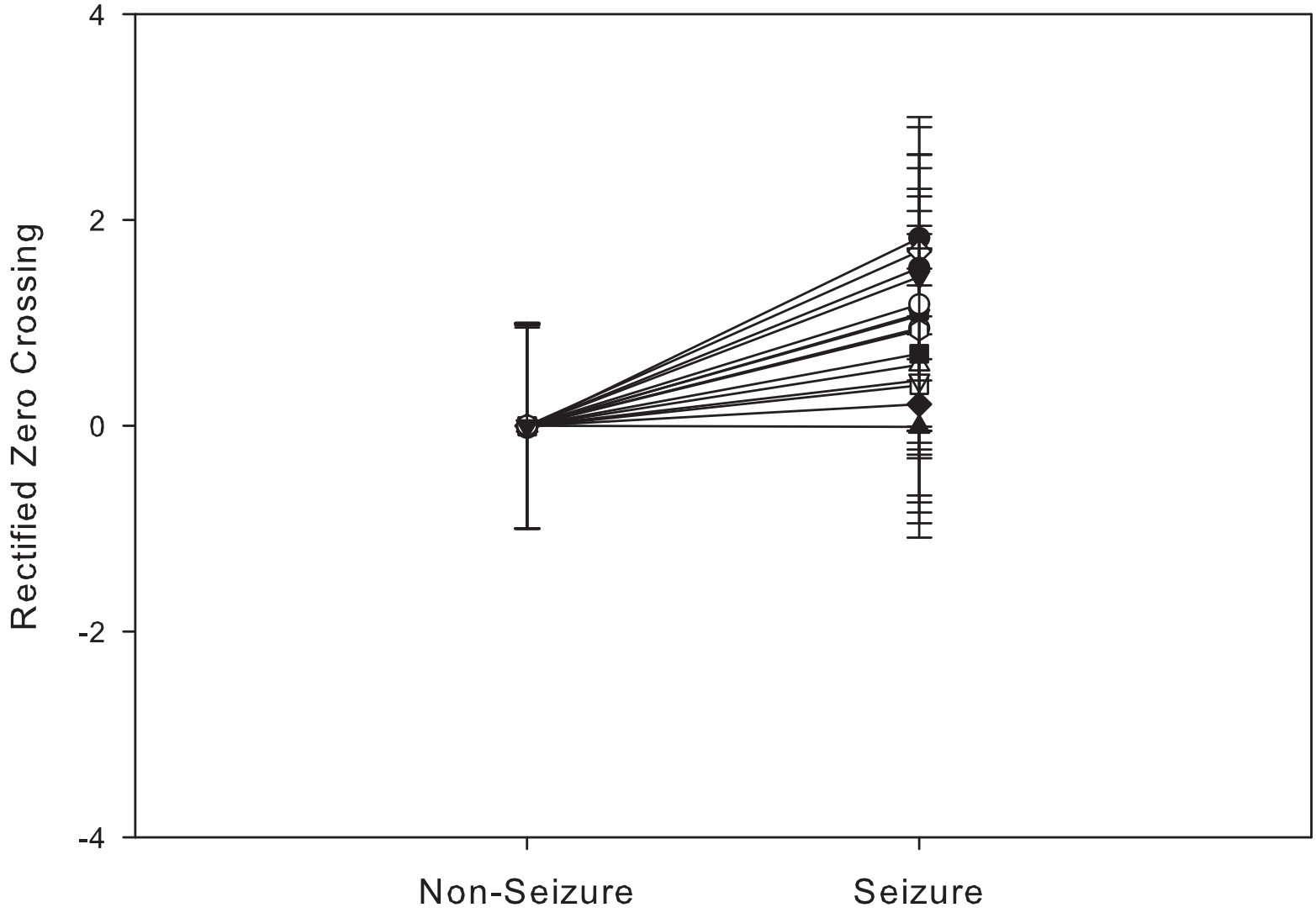

Figure2

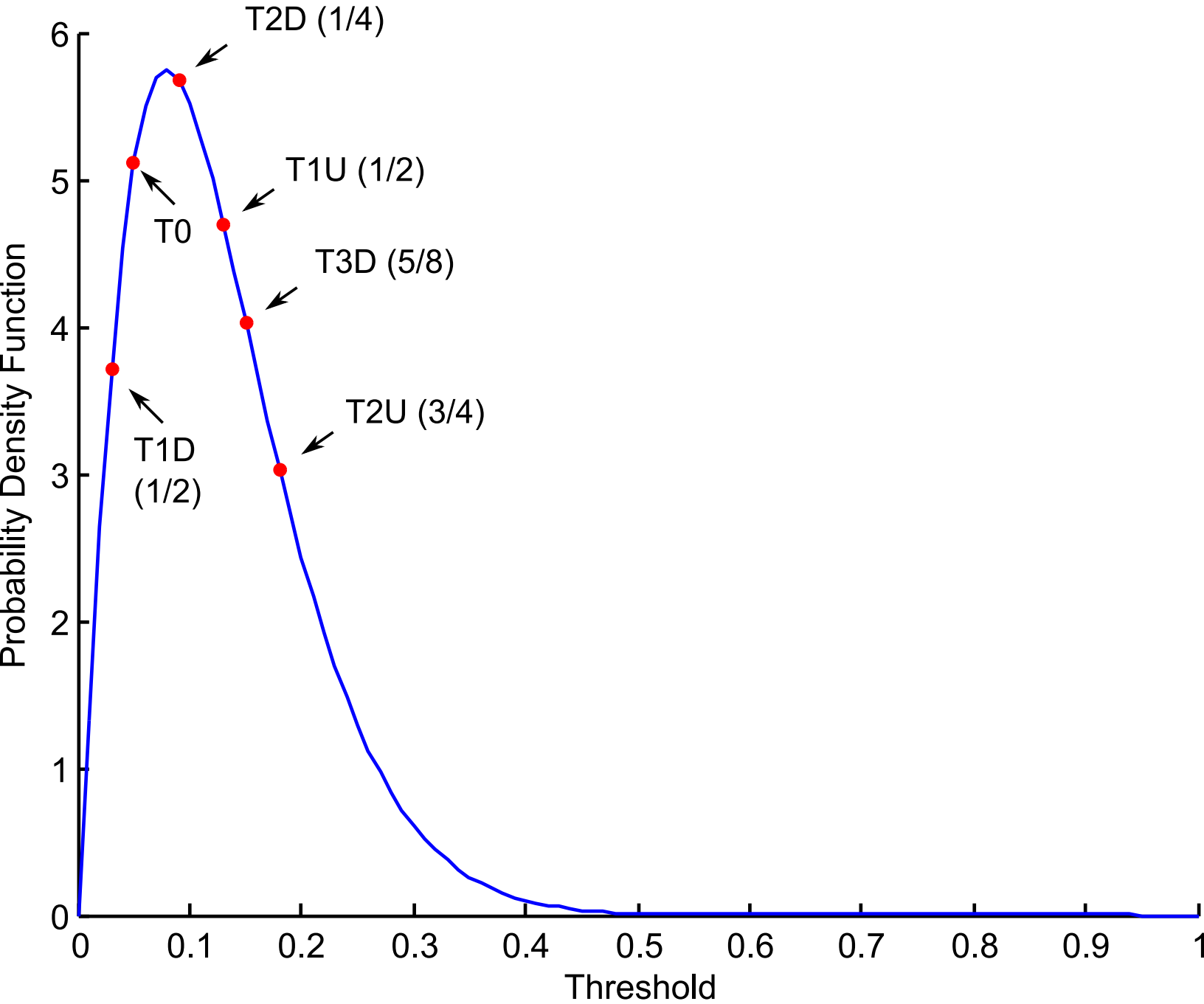
T2D (1/4)
T1U (1/2)
T0
T3D (5/8)
T2U (3/4)
T1D
(1/2)
Probability Density Function
Threshold
0
1
2
3
4
5
6
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure3

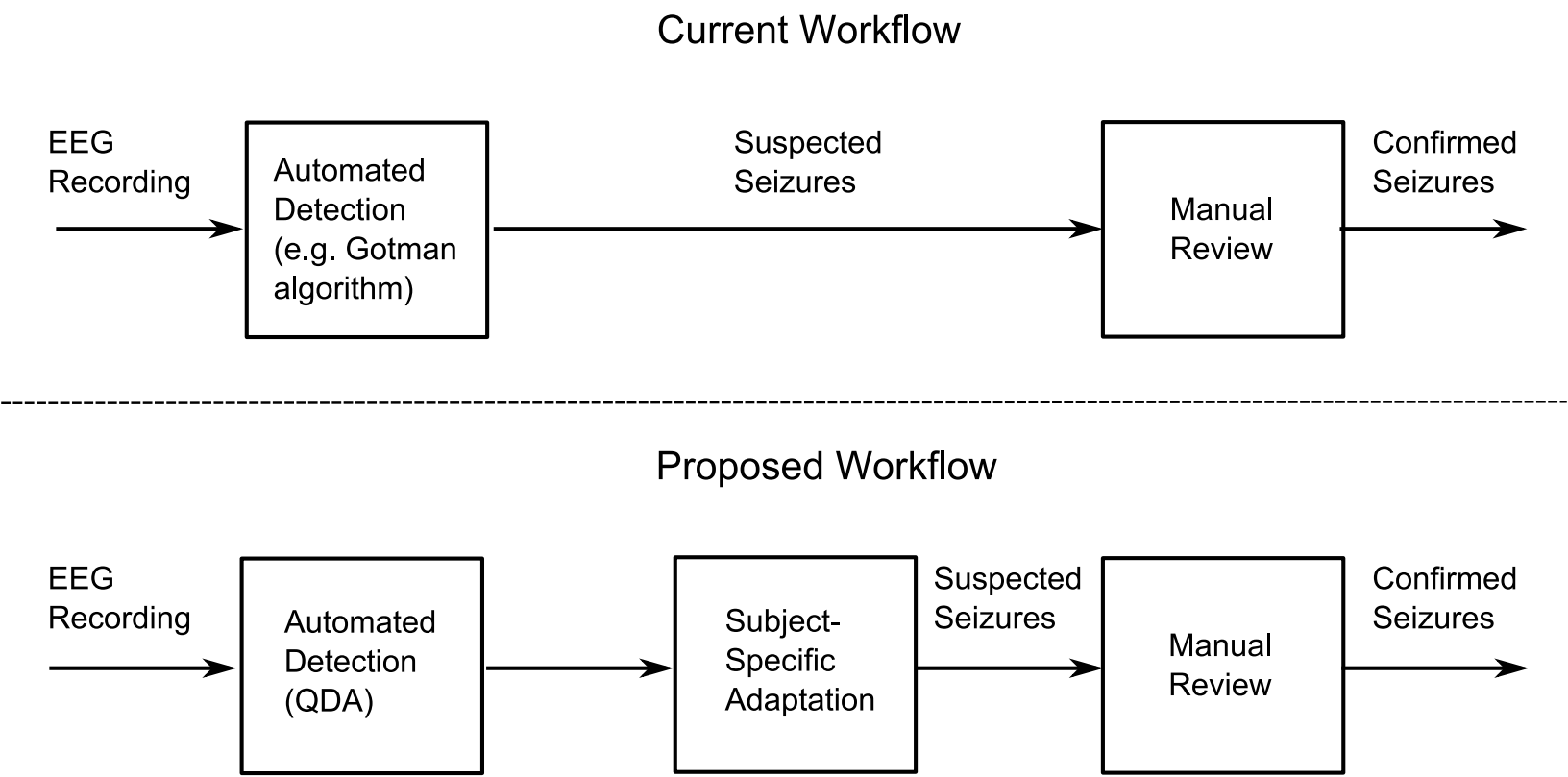
Current Workflow
EEG Recording
Automated Detection (e.g. Gotman algorithm)
Suspected Seizures
Manual Review
Confirmed Seizures
Proposed Workflow
EEG Recording
Automated Detection (QDA)
Subject-Specific Adaptation
Suspected Seizures
Manual Review
Confirmed Seizures

Figure4

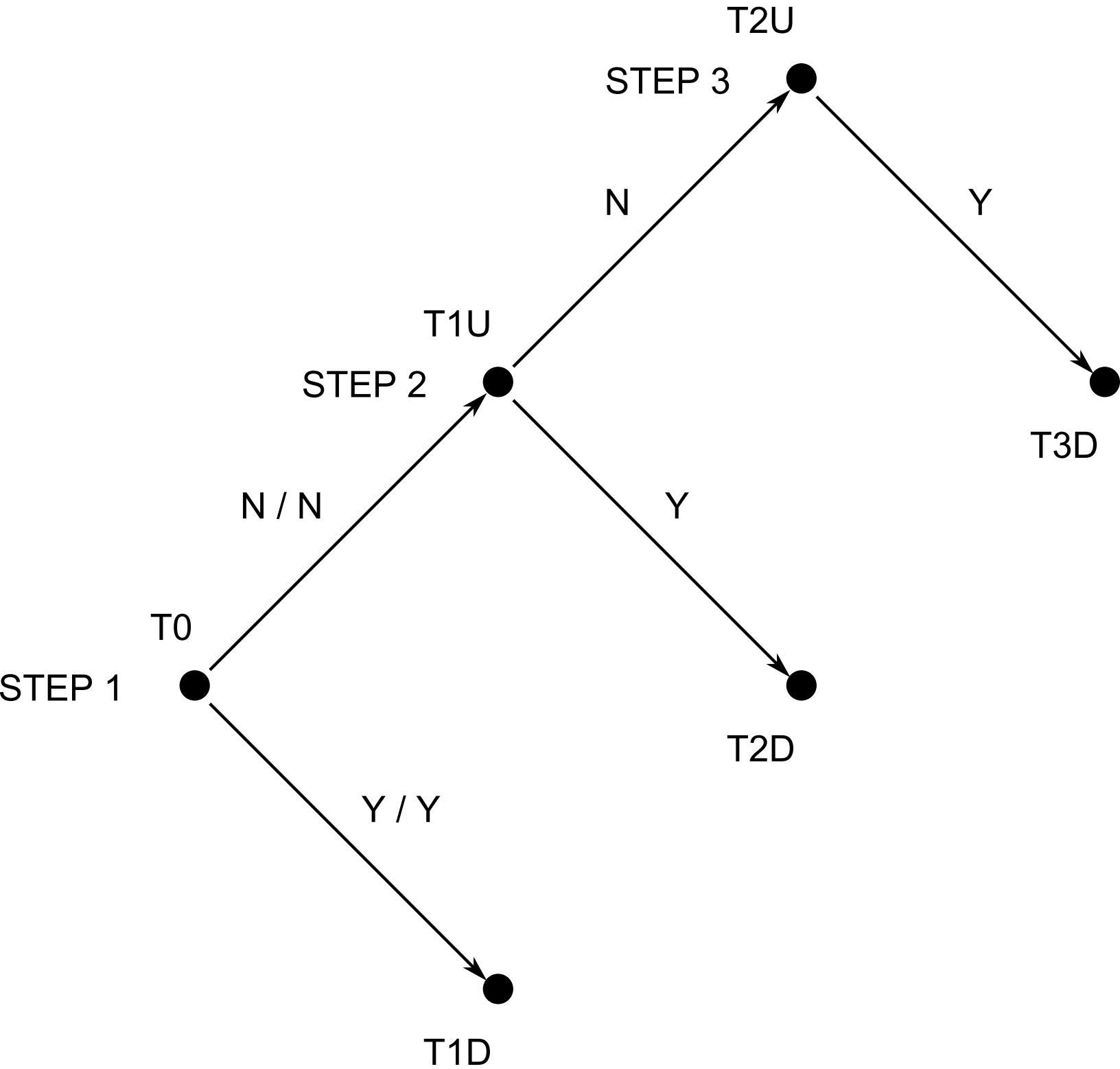

Figure5
Click here to download high resolution image

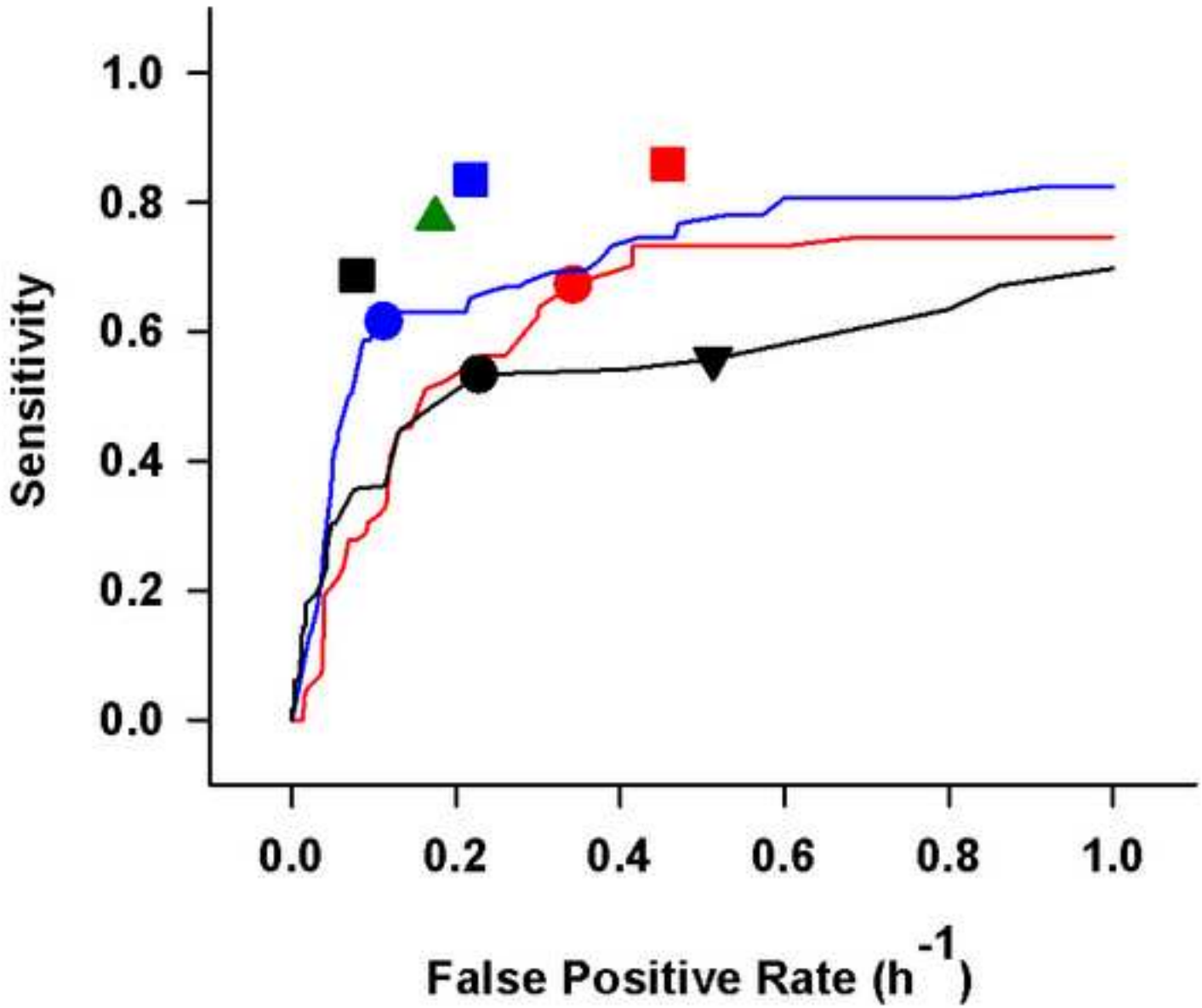

Figure6
Click here to download high resolution image