doi:10.1093/brain/awl241 *Brain* (2007), **130**, 314–333

REVIEW ARTICLE

Seizure prediction: the long and winding road

Florian Mormann,[1] Ralph G. Andrzejak,[3] Christian E. Elger[1] and Klaus Lehnertz[1,2]

[1]Department of Epileptology, [2]Helmholtz-Institute for Radiation and Nuclear Physics, University of Bonn, Bonn, Germany and [3]Department of Technology, University Pompeu Fabra, Barcelona, Spain

Correspondence to: Florian Mormann, MD, PhD, Division of Biology 216–76, California Institute of Technology, Pasadena, CA 91125, USA
E-mail: fmormann@yahoo.de

The sudden and apparently unpredictable nature of seizures is one of the most disabling aspects of the disease epilepsy. A method capable of predicting the occurrence of seizures from the electroencephalogram (EEG) of epilepsy patients would open new therapeutic possibilities. Since the 1970s investigations on the predictability of seizures have advanced from preliminary descriptions of seizure precursors to controlled studies applying prediction algorithms to continuous multi-day EEG recordings. While most of the studies published in the 1990s and around the turn of the millennium yielded rather promising results, more recent evaluations could not reproduce these optimistic findings, thus raising a debate about the validity and reliability of previous investigations. In this review, we will critically discuss the literature on seizure prediction and address some of the problems and pitfalls involved in the designing and testing of seizure-prediction algorithms. We will give an account of the current state of this research field, point towards possible future developments and propose methodological guidelines for future studies on seizure prediction.

Keywords: seizure anticipation; algorithm; performance; statistical validation; methodology; guidelines

Received March 28, 2006. Revised July 1, 2006. Accepted August 10, 2006, Advance Access publication September 28, 2006.

Introduction

Epilepsy is one of the most common neurological disorders, second only to stroke, with a prevalence of 0.6–0.8% of the world's population (Annegers, 1996). Two-thirds of the patients achieve sufficient seizure control from anti-convulsive medication, and another 8–10% could benefit from resective surgery. For the remaining 25% of patients, no sufficient treatment is currently available.

For epilepsy patients who do not achieve complete seizure control, it is the sudden, unforeseen way in which seizures strike 'like a bolt from the blue' that represents one of the most disabling aspects of the disease. Apart from the risk of serious injury, there is often an intense feeling of helplessness that has a strong impact on the everyday life of a patient. A method capable of predicting the occurrence of seizures could significantly improve the therapeutic possibilities (Elger, 2001) and thereby the quality of life for epilepsy patients.

A question of particular interest is whether apart from clinical prodromi, which are found only in some patients (Hughes *et al.*, 1993; Rajna *et al.*, 1997; Schulze-Bonhage *et al.*, 2006), characteristic features can be extracted from the continuous EEG that are predictive of an impending seizure. If it were possible to reliably predict seizure occurrence from dynamical changes in the EEG of epilepsy patients, fully automated closed-loop seizure-prevention systems could be envisioned. Treatment concepts could move from preventive strategies (e.g. long-term medication with anti-epileptic drugs) towards an EEG-triggered on-demand therapy [e.g. by excretion of fast-acting anticonvulsant substances (e.g. Stein *et al.*, 2000) or by electrical or other stimulation in an attempt to reset brain dynamics to a state that will no longer develop into a seizure (e.g. Theodore and Fisher, 2004; Morrell, 2006)].

In principle, there are two different scenarios of how a seizure could evolve (Lopes da Silva *et al.*, 2003). It could be caused by a sudden and abrupt transition, in which case it would not be preceded by detectable dynamical changes in the EEG. Such a scenario would be conceivable for the initiation of seizures in primary generalized epilepsy. Alternatively, this transition could be a gradual change

(or a cascade of changes) in dynamics, which could in theory be detected. This type of transition could be more likely in focal epilepsies.

Clinical findings in support of the existence of a pre-seizure state include an increase in cerebral blood flow (Weinand *et al.*, 1997; Baumgartner *et al.*, 1998), oxygen availability (Adelson *et al.*, 1999) and blood oxygen-level-dependent signal (Federico *et al.*, 2005) as well as changes in heart rate (Delamont *et al.*, 1999; Novak *et al.*, 1999; Kerem and Geva, 2005) before seizure occurrence.

On the level of neuronal networks, focal seizures are assumed to be initiated by abnormally discharging neurons, so-called bursters (Yaari and Beck, 2002 and references therein) that recruit and entrain neighbouring neurons into a critical mass. This build-up might be mediated by an increasing synchronization of neuronal activity that is accompanied by a loss of inhibition, or by processes that facilitate seizures by lowering the threshold for excitation or synchronization. In this context, the term 'critical mass' might be misleading in the sense that it implies an increasing number of neurons that are entrained into an abnormal firing pattern. This mass phenomenon would be easily accessible to conventional EEG analysis, which, to date, has failed to detect it. Rather, the seizure-initiating process might better be visualized as a process by which an increasing number of critical interactions between neurons in a focal region and connected units in an abnormal functional network unfold over time.

On the basis of these concepts, a number of studies have been carried out aiming to characterize this collective neuronal behaviour from the gross EEG in order to allow definition of a transitional pre-ictal phase. In this review, we will give an overview of the literature on this topic and the current state of this rather young research field as well as critically discuss some of the methodological problems and pitfalls involved in the design and testing of seizure-prediction algorithms.

The history of seizure prediction

For a better understanding of the practical problems in this field, we have categorized previous studies on seizure prediction according to methodological standards. A chronological overview of studies on seizure prediction from the past 10 years along with relevant characteristics (type of epilepsy and EEG, type of EEG analysis, number of patients and seizures, amount of data analysed, etc.) is given in Table 1. Detailed mathematical descriptions of the most common characterizing measures used in EEG analysis can be found in Appendix C (Supplementary online material).

Early approaches

After some early work on the predictability of seizures dating back to the 1970s (Viglione and Walsh, 1975), attempts to extract seizure precursors from surface EEG recordings of absence seizures were carried out by different groups using linear approaches. Using autoregressive modelling, Rogowski *et al.* (1981) and Salant *et al.* (1998) reported pre-ictal changes in the modelled parameters within up to 6 s before seizure onset. Siegel *et al.* (1982) found characteristic changes between the 1-min epochs directly preceding a seizure and control epochs for individual patients. Remarkably, in this early study, the authors already assessed the statistical confidence of their findings and discussed the influence of different vigilance states.

A further group of studies examined the predictive value of spike occurrence rates in the EEG. While an early study reported a decreased focal spiking rate along with an increased rate of bilateral spikes before seizures (Lange *et al.*, 1983), other studies carried out on more extended databases showed no systematic changes in spike rates before seizures (Gotman and Marciani, 1985; Gotman and Koffler, 1989; Katz *et al.*, 1991).

Pre-ictal phenomena

With the advent of the physical–mathematical theory of non-linear systems in the 1980s, novel approaches were introduced that were aimed at a better characterization of dynamical systems exhibiting complex behaviour than hitherto possible with conventional linear approaches. Soon time series analysts became aware of seizure prediction as a potential field of application. In the early 1990s, Iasemidis *et al.* (1990) estimated the largest Lyapunov exponent as an indicator for chaotic behaviour from the intracranial EEG of epilepsy patients by means of a moving window analysis and reported a decrease in chaoticity in the minutes before an epileptic seizure. Some years later, a French group of researchers reported a pre-ictal decrease in spatiotemporal complexity as measured by the correlation density before seizures in a larger group of patients (Martinerie *et al.*, 1998). The same group developed another measure named the dynamical similarity index, which quantified changes in dynamics relative to a constant reference window at the beginning of a pre-ictal recording. They found a decreased dynamical similarity before seizures in both intracranial (Le Van Quyen *et al.*, 1999, 2000) and scalp EEG recordings (Le Van Quyen *et al.*, 2001*a*).

However, what was common for all of these studies is that their focus of interest was entirely limited to the pre-ictal period and that they did not include an evaluation of inter-ictal control recordings (i.e. periods from the seizure-free interval other than the presumed pre-ictal period). By thus neglecting the issue of specificity, these studies rendered an incomplete evaluation of the investigated measures' suitability for seizure prediction.

Proof-of-principle studies

Another group of studies tackled the issue of specificity by comparing pre-ictal changes in dynamics to inter-ictal control recordings, although the findings reported in these studies remained on an exemplary level. Navarro *et al.* (2002)

used selected examples of five of their patients to show that drops in their similarity measure occurred more frequently before seizures than during the inter-ictal EEG. Mormann *et al.* (2000) reported changes in phase synchronization between different brain areas before seizures that were not found in exemplary seizure-free recordings. In two reviews of their own work, Le Van Quyen *et al.* (2001*b*, 2003) referred to a submitted study including eight patients with neocortical epilepsy that seemed to confirm these findings. In 2003, Chavez *et al.* published exemplary results using phase synchronization analysis after band-pass filtering of the EEG and reported pre-ictal changes in synchronization to occur predominantly in the beta band.

Controlled studies on predictability

In the first controlled studies comprising defined groups of patients with pre-ictal and inter-ictal control recordings, measures like the correlation dimension (Elger and Lehnertz, 1998; Lehnertz and Elger, 1998) (as a measure for dynamical complexity), dynamical entrainment (Iasemidis *et al.*, 2001) (defined by the authors as the convergence of largest Lyapunov exponents in certain selected channels), accumulated signal energy (Litt *et al.*, 2001; Gigola *et al.*, 2004), simulated neuronal cell models (Schindler *et al.*, 2002) or phase synchronization (Mormann *et al.*, 2003*a*, *b*) were shown to be suitable for distinguishing inter-ictal from pre-ictal data.

The rise of scepticism

Starting in 2003, a number of studies were published (most of them carried out on extensive databases) that found a substantially poorer predictive performance than presumable from earlier optimistic reports. Studies by De Clercq *et al.* (2003) and Winterhalder *et al.* (2003) challenged the reliability of the optimistic results reported for the similarity index (Le Van Quyen *et al.*, 2001*a*). The ability of the correlation dimension for seizure prediction (Lehnertz and Elger, 1998) was questioned by other studies using this measure (Aschenbrenner-Scheibe *et al.*, 2003; Harrison *et al.*, 2005*b*). Similarly, the promising performance of the accumulated energy (Litt *et al.*, 2001) could not be reproduced in later studies (Maiwald *et al.*, 2004, Harrison *et al.*, 2005*a*). Studies by Lai *et al.* (2003, 2004) raised doubts about the suitability of the Lyapunov exponent (Iasemidis *et al.*, 1990) for seizure prediction.

The advantage of nonlinear measures such as the correlation density (Martinerie *et al.*, 1998) was questioned by McSharry *et al.* (2003). The authors re-evaluated the data from this study and showed that this measure merely reflected the variance of the EEG signals. They pointed out that the presence of non-linearity in a signal does not in itself justify the use of non-linear, complicated measures to characterize dynamical changes and emphasized the importance of showing that these complicated methods indeed outperform simpler linear measures in order to justify their use (see also Andrzejak *et al.*, 2006).

Taken together, these studies indicated that earlier optimistic findings obtained by applying highly optimized algorithms to small, selected data sets could not be reproduced on unselected, more extended EEG recordings that are more closely related to the real-life challenge of predicting seizures prospectively from the continuous EEG.

Continuous multi-day recordings

Around the turn of the millennium, when mass storage capacity became more widely available, epilepsy centres were able to store the complete data acquired during pre-surgical monitoring without the necessity of selecting sample recordings. In 2005, different groups published a series of studies that were carried out on a set of five continuous multi-day recordings provided by different epilepsy centres for the First International Collaborative Workshop on Seizure Prediction (Lehnertz and Litt, 2005) held in Bonn in April 2002. The aim of this workshop was to have different groups test and compare their methods on a joint data set. Results from the different groups for the most part showed a poor performance of univariate measures (D'Alessandro *et al.*, 2005; Esteller *et al.*, 2005; Harrison *et al.*, 2005*a*; Jouny *et al.*, 2005; Mormann *et al.*, 2005). A better performance was reported for bi- and multi-variate measures (Iasemidis *et al.*, 2005; Le Van Quyen *et al.*, 2005; Mormann *et al.*, 2005), although the observed pre-ictal changes were found to be locally restricted to specific channels rather than occurring as a global phenomenon. One of these studies (Mormann *et al.*, 2005) contained an extensive comparison of the predictive performance of a number of univariate and bivariate measures, comprising both linear and non-linear approaches, using the concept of seizure times surrogates (Andrzejak *et al.*, 2003; *see* Appendix B) to assess the statistical significance of the results obtained. For this concept, the seizure-onset times of the original EEG recordings are replaced by artificial seizure-onset times that are generated by randomly shuffling the original onset times. If a measure's predictive performance for the original seizure-onset times is higher than that for a number of realizations of the surrogate onset times, then the performance of this measure can be considered significantly better than a random prediction. In their comparison, the authors found a significant predictive performance for measures of synchronization, whereas univariate measures, including the correlation dimension, the Lyapunov exponent and the signal energy, were not able to discriminate the pre-ictal from the inter-ictal period above chance level. Non-linear measures were not found to exhibit a higher predictive performance than linear measures.

Prospective studies

The first attempts for testing seizure-prediction algorithms in a prospective manner were carried out by Iasemidis *et al.* (2003) and D'Alessandro *et al.* (2005). The sensitivity and specificity rates obtained, however, were unacceptable for

Table 1 Studies on seizure prediction and their relevant characteristics (see text)

Authors	Year	Type of epilepsy	Type of EEG	Characterizing measure	Patients	Seizures	Total EEG (h)	Inter-ictal controls (h)
Lehnertz and Elger	1998	MTLE	iEEG	Correlation dimension	16	16	21	16.9
Martinerie *et al.*	1998	MTLE	iEEG	Correlation density	11	19	13	0
Le Van Quyen *et al.*	1999	MTLE	iEEG	Similarity index	13	23	15	0
Le Van Quyen *et al.*	2000	MTLE	iEEG	Similarity index	9	17	11	0
Mormann *et al.*	2000	MTLE	iEEG	Phase synchronization	2	3	4	1.8
Cerf *et al.*	2000	Focal	iEEG	Lerner density	7	9	n.s.	1.8
Hively *et al.*	2000	Focal	sEEG	Dissimilarity measures	n.s.	20	40	0
Le Van Quyen *et al.*	2001*a*	TLE	sEEG	Similarity index	23	26	26–35	0
Iasemidis *et al.*	2001	TLE	iEEG	Dynamical entrainment	5	58	266	53.9
Litt *et al.*	2001	MTLE	iEEG	Accumulated energy	5	30	>312	50
Le Van Quyen *et al.*	2001*b*	Neocortical	iEEG	Phase synchronization	8	n.s.	n.s.	n.s.
Lehnertz *et al.*	2001	Focal	iEEG	Correlation dimension	59	95	>145	>115
Protopopescu *et al.*	2001	Focal	sEEG	Dissimilarity measure	41	46	261	73.9
Jerger *et al.*	2001	Children	iEEG	Seven different measures	4	12	1	0
Navarro *et al.*	2002	Neocortical	s+iEEG	Similarity index	11	41	53–142	12–60[c]
Schindler *et al.*	2002	Focal	sEEG + FO	Simulated neuronal cells	7	15	144	n.s.
Mormann *et al.*	2003*a*	MTLE	iEEG	Synchronization/correlation	10	14	31	15
Mormann *et al.*	2003*b*	Focal	iEEG	Phase synchronization	18	32	117	49
De Clercq *et al.*	2003	MTLE	sEEG	Similarity index	12	n.s.	n.s.	0
Niederhauser *et al.*	2003	Focal	iEEG	Sign periodogram transf.	5[a]	31	336	335
Chávez *et al.*	2003	Neocortical	iEEG	Phase synchronization	2	6	22	9
Hively and Protopopescu	2003	Focal	sEEG	Dissimilarity measure	41	46	261	73.9
D'Alessandro *et al.*	2003	MTLE	iEEG	Feature selection	4	46	n.s.	160
Iasemidis *et al.*	2003	TLE	iEEG	Dynamical entrainment	5	28[b]	214	n.s.
Winterhalder *et al.*	2003	Focal	iEEG	Similarity index	21	88	588	509
Aschenbrenner *et al.*	2003	Focal	iEEG	Correlation dimension	21	88	588	509
Van Drongelen *et al.*	2003	Children	s+iEEG	Kolmogorov entropy	5	5	5	0
Li *et al.*	2003	MTLE	sEEG	Marginal predictability	8	24	37	13.3
Drury *et al.*	2003	MTLE	sEEG	Marginal predictability	14	44	59	14.7
Maiwald *et al.*	2004	Focal	iEEG	Accumulated energy	21	88	588	509
Gigola *et al.*	2004	Focal	iEEG	Accumulated energy	4	13	26	10.5
D'Alessandro *et al.*	2005	MTLE	iEEG	Feature selection	2	19[b]	177	140
Esteller *et al.*	2005	MTLE	iEEG	Accumulated energy	4	42	294	>168
Harrison *et al.*	2005*a*	MTLE	iEEG	Accumulated energy	5	51	311	<92
Iasemidis *et al.*	2005	MTLE	iEEG	Dynamical entrainment	2	11[b]	41	>8
Jouny *et al.*	2005	MTLE	iEEG	Complexity/synchrony	2	25	177	n.s.
Le Van Quyen *et al.*	2005	MTLE	iEEG	Phase synchronization	5	52	305	25–120
Mormann *et al.*	2005	MTLE	iEEG	30 different measures	5	51	311	>107
Kalitzin *et al.*	2005	TLE	iEEG	Phase clustering	3	20	>75	n.s.
Navarro *et al.*	2005	Focal	iEEG	Similarity index	13	129	227	0
Chaovalitwongse *et al.*	2005	TLE	iEEG	Dynamical entrainment	10	64[b]	597	>404
Harrison *et al.*	2005*b*	Focal	iEEG	Correlation dimension	20	960	2347	n.s.
Schelter *et al.*	2006	MTLE	iEEG	Phase synchronization	4	20	112	96

Authors	Year	Type of analysis	In-sample parameter optimization	Retrospective best channel selection	Prospective	Assumed pre-ictal period (min)	Sensitivity (%)	False-positive rate (FP/h)	Mean prediction time (min)	Statistical validation of performance
Lehnertz and Elger	1998	Statistical	Yes	Yes	No	30	94	0	12	No
Martinerie *et al.*	1998	Algorithmic	No	Yes	No	20	89	n.a.	3	No
Le Van Quyen *et al.*	1999	Algorithmic	No	Yes	No	20	83	n.a.	6	No
Le Van Quyen *et al.*	2000	Algorithmic	No	Yes	No	20	94	n.a.	4	No
Mormann *et al.*	2000	Proof of principle	No	Yes	No	n.s.	100	0	n.s.	No
Cerf *et al.*	2000	Statistical	Yes	Yes	No	60	100	0	n.s.	No
Hively *et al.*	2000	Algorithmic	Yes	No	No	262.5	100	n.a.	52	No
Le Van Quyen *et al.*	2001*a*	Algorithmic	No	Yes	No	60	96	n.a.	7	No
Iasemidis *et al.*	2001	Statistical	Yes	Yes	No	Variable	91	n.s.	49	No
Litt *et al.*	2001	Statistical	Yes	No	No	180	90	0.12	19	No
Le Van Quyen *et al.*	2001*b*	n.s.	n.s.	n.s.	No	n.s.	77	n.s.	Several min	No
Lehnertz *et al.*	2001	Algorithmic	Yes	Yes	No	n.s.	47	0	19	No
Protopopescu *et al.*	2001	Algorithmic	Yes	Yes	No	60	95	0	n.s.	No
Jerger *et al.*	2001	Algorithmic	Yes	Yes	No	3	100	n.a.	2	No
Navarro *et al.*	2002	Algorithmic	No	Yes	No	90	83	0.31[c]	8	No
Schindler *et al.*	2002	Algorithmic	Yes	No	No	Variable	100	n.s.	83	No
Mormann *et al.*	2003*a*	Algorithmic	Yes	No	No	240	86	0	86/102[h]	Yes
Mormann *et al.*	2003*b*	Algorithmic	Yes	No	No	240	81	0	4–221	No
De Clercq *et al.*	2003	Algorithmic	No	Yes	No	60	0	n.a.	–	No
Niederhauser *et al.*	2003	Algorithmic	Yes	Yes	No	2	94	0.08[f]	5–80 s	No
Chávez *et al.*	2003	Proof of principle	Yes	Yes	No	90	n.s.	n.s.	>>30	No
Hively and Protopopescu	2003	Algorithmic	Yes	No	No	60	88	0.02	35	No
D'Alessandro *et al.*	2003	Algorithmic	Yes	Yes	No[d]	10	63	0.28	3	No
Iasemidis *et al.*	2003	Algorithmic	No	No	Yes	180	83	0.17[f]	100	No
Winterhalder *et al.*	2003	Algorithmic	Yes	Yes	No	30[e]	42	0.15	n.s.	No
Aschenbrenner *et al.*	2003	Algorithmic	Yes	Yes	No	50[e]	34	0.10	n.s.	No
Van Drongelen *et al.*	2003	Algorithmic	Yes	No	No	60	60	n.a.	21	No
Li *et al.*	2003	Statistical	No	No	No	60	n.s.	n.s.	n.s.	No
Drury *et al.*	2003	Statistical	No	No	No	60	n.s.	n.s.	30	No
Maiwald *et al.*	2004	Algorithmic	Yes	Yes	No	32[e]	30	0.15	n.s.	No
Gigola *et al.*	2004	Statistical	Yes	n.s.	No	70	92	0	n.s.	No
D'Alessandro *et al.*	2005	Algorithmic	No	No	Yes	10	100/13[g]	1.10/0.71[g]	2/n.s.[g]	No
Esteller *et al.*	2005	Algorithmic	Yes	Yes	No[d]	180[e]	71	0.11[f]	85	No
Harrison *et al.*	2005*a*	Statistical	No	No	No	60	0	–	–	No
Iasemidis *et al.*	2005	Algorithmic	No	No	Yes	120	82	0.15[f]	78	No
Jouny *et al.*	2005	Statistical	No	No	No	60	0	–	–	No
Le Van Quyen *et al.*	2005	Algorithmic	Yes	No	No[d]	Variable	69	n.s	187	No
Mormann *et al.*	2005	Statistical	Yes	Yes	No	5–240[e]	n.s.	n.s.	–	Yes
Kalitzin *et al.*	2005	Statistical	Yes	Yes	No	–	n.s.	n.s.	–	No
Navarro *et al.*	2005	Algorithmic	No	Yes	No	120	64	n.a.	>13	No
Chaovalitwongse *et al.*	2005	Algorithmic	No	No	Yes	180	69	0.15[f]	72	Yes[i]
Harrison *et al.*	2005*b*	Statistical	Yes	No	No	90/15[e]	0	–	–	No
Schelter *et al.*	2006	Statistical	No	Yes	No	40[e]	70	0.15	n.s.	Yes

Adapted from Mormann *et al.* (2006*a*). n.s.: not specified; n.a.: not analysed; MTLE: medial temporal lobe epilepsy; TLE: temporal lobe epilepsy; iEEG: intracranial EEG; sEEG: surface EEG; FO: foramen ovale electrodes.

[a]Selected out of a group of 10 patients.

[b]Results listed are those obtained for out-of-sample testing data after in-sample optimization on training data.

[c]Only from five selected patients.

[d]Algorithm designed to run prospectively, but results are reported for training and testing data together.

[e]Various predefined prediction horizons were analysed.

[f]Uncorrected false prediction rate including pre-ictal periods.

[g]Separate results reported for two different patients.

[h]Separate results reported for two different measures.

[i]Inconclusive validation: surrogate seizure times are not treated in the same way as original seizure times.

clinical implementation. Whether the performance of the algorithms was at all better than chance was not investigated. A recent study by Chaovalitwongse *et al.* (2005) attempted such a validation based on the method of seizure times surrogates proposed 2 years earlier by Andrzejak *et al.* (2003). However, the authors re-used analysis parameters that were optimized for the original seizure-onset times in their analysis of the surrogate onset times, so the results must be regarded as inconclusive (Mormann *et al.*, 2006*b*; Chaovalitwongse *et al.*, 2006; *see* also Winterhalder *et al.*, 2006).

The current state of the field

During the 1990s and around the turn of the millennium, a number of studies were highly optimistic about seizure prediction becoming feasible for clinical application in the near future. However, the focus of these studies was limited to analysing short and selected EEG recordings, and numerous methodological caveats were not addressed properly. In the past 5 years, many studies have been published that questioned both the validity and reliability of these findings by showing that earlier optimistic results could not be reproduced.

While many early studies reported pre-ictal changes in channels within or close to the seizure-onset zone (Elger and Lehnertz, 1998; Lehnertz and Elger, 1998; Martinerie *et al.*, 1998; Le Van Quyen *et al.*, 1999, 2000), more recent ones found channels in more remote and, in some cases, even contralateral areas to carry the relevant information (D'Alessandro *et al.*, 2003, 2005; Mormann *et al.*, 2003*b*, 2005; Esteller *et al.*, 2005; Le Van Quyen *et al.*, 2005). This finding would support the notion of an epileptic network whose interactions extend over large regions of the brain rather than the concept of a localized and well-defined epileptic focus. For the field of seizure prediction to advance towards clinical applications, it is inevitable that future studies on seizure prediction place a strong emphasis on sound methodology and include a rigorous statistical validation. Some of the methodological issues and caveats involved in the designing and testing of seizure prediction algorithms are addressed in Appendix A and B, respectively.

Future perspectives

Prospective out-of-sample algorithms with statistical validation

The next milestone in the field of seizure prediction is to prove that seizure-prediction algorithms can be designed to run prospectively on unselected, out-of-sample data with a performance that is better than that of a random prediction process. If such an algorithm requires a training phase in which some seizures are used to adjust patient–individual parameters and perform a feature or channel selection, the requirements on the individual data sets increase since a larger number of seizures per individual data set will be needed. Performance results should be reported only for the testing data.

Before addressing the question as to whether an obtained performance might be sufficient for clinical application, it needs to be tested whether a performance is at all better than chance. To this aim, methods for statistical validation are inevitable. These methods can be based on Monte Carlo simulations (Andrzejak *et al.*, 2003, Mormann *et al.*, 2003*a*, Kreuz *et al.*, 2004; Jerger *et al.*, 2005; Mormann *et al.*, 2005) or on comparison with analytical results derived from naïve (random or periodic) prediction schemes (Winterhalder *et al.*, 2003; Schelter *et al.*, 2006).

Confounding variables

Another key to the improvement of algorithms could be a better understanding of the inter-ictal period and all of its confounding variables that may influence the characterizing measures used in the algorithms and may thereby decrease the algorithm's sensitivity or specificity. Studies on continuous multi-day recordings have revealed distinct circadian fluctuations of measures characterizing the EEG (Kreuz *et al.*, 2004). In particular, different vigilance states (e.g. slow-wave sleep) seem to have an influence on such measures (Navarro *et al.*, 2005). A further confounding influence on characterizing measures has been described for the blood levels of carbamazepine (Lehnertz and Elger, 1997). To date, little is known about the influence of different cognitive or emotional states (Lehnertz, 1999). Once the influence of confounding variables is better understood, it can be taken into account by an algorithm to increase its predictive performance.

Mechanisms of ictogenesis

While many studies on seizure prediction focused on algorithmic prediction, they paid comparably little interest in the underlying mechanisms of seizure generation. In light of the rather poor performance of the seizure-prediction algorithms designed to date, it is questionable whether any inference from these algorithms to the underlying mechanisms of a presumed pre-ictal transition can be regarded as meaningful. Instead, the mechanisms of ictogenesis remain largely unknown except for certain distinct types of epilepsies such as reflex epilepsies (Kalitzin *et al.*, 2002; Parra *et al.*, 2003). Furthermore, it is conceivable that there may be different mechanisms underlying the initiation of seizures in different brain structures (e.g. hippocampus, neocortex), and thus different seizure-predicting algorithms may be necessary. This may also be true of different pathologies (e.g. dysplasias, malformations, post-traumatic lesions, etc.). The predictive changes in the EEG before a seizure and the best methods for detecting them could thus vary considerably from patient to patient.

A number of recent studies have attempted to increase our understanding of the dynamics of ictogenesis in humans. In both temporal lobe and neocortical epilepsies, high-frequency oscillations were found to play a role in the initiation of epileptiform potentials and seizures (e.g. Bragin

et al., 1999, 2002; Schiff *et al.*, 2000; Worrell *et al.*, 2004). In another recent study on patients with temporal lobe epilepsy, a measure for phase demodulation of intracranial EEG recorded inter-ictally during intermittent electrical stimulation was found to yield important clues for possible dynamical scenarios that lead to seizure onsets (Kalitzin *et al.*, 2005). In extension to previous studies on seizure dynamics (e.g. Pijn *et al.*, 1991; Franaszczuk *et al.*, 1994, 1998; Bartolomei *et al.*, 2004; Stam, 2005 and references therein) Schiff *et al.* (2005) successfully used canonical discrimination analysis to search for dynamically distinct stages of epileptic seizures in humans. A further promising approach is to model EEG signals to gain insight into the dynamical processes involved in seizure generation (Wendling *et al.*, 2002, 2003; Suffczynski *et al.*, 2005, 2006).

A better understanding of the mechanisms of ictogenesis that takes into consideration the complex spatiotemporal interactions between different brain regions for different types of epilepsy may eventually stimulate the design of improved methods and algorithms.

Closed-loop intervention systems

The ultimate goal in designing a reliable seizure-prediction algorithm can be seen in a system capable of not only warning of an impending seizure but actually taking measures to prevent it from occurring. An ideal intervention system would be able to abort the development of a seizure before the onset of clinical symptoms. The tolerance of false alarms leading to unnecessary interventions would depend on the side-effects caused by the intervention.

The principal feasibility of different seizure-intervention strategies such as local application of short-acting, powerful drugs (Stein *et al.*, 2000), electrical stimulation techniques (Morrell, 2006 and references therein), local cooling (Hill *et al.*, 2000) or biofeedback operant conditioning (Sterman, 2000) has been outlined in the literature.

Presently, much research is directed towards the design of a responsive intervention system using deep brain or cortical stimulation (Osorio *et al.*, 2001; Theodore and Fisher, 2004; Morrell, 2006). In its simplest form, such a system could include a single depth recording electrode, a processing unit, and then apply stimulation via the recording electrode at a critical time to alter the local brain state from that of a pre-ictal or pro-convulsive condition to a more stable, non-ictogenic state. Such an EEG-based responsive stimulation system could in principle be based either on prediction algorithms or on algorithms for early seizure detection (*see* Appendix A). At present, prediction algorithms are still too limited in performance to justify clinical trials with responsive stimulation based on these approaches. For early seizure-detection algorithms, the question is whether after the onset of electrographical seizure activity, the seizure can indeed be aborted by stimulation or whether the brain has already passed the 'point of no return' and is in a state that will inevitably progress into a clinical seizure manifestation. First studies using early seizure-detection algorithms for responsive brain stimulation have reported promising results, but these must yet be regarded as preliminary and allow no definite conclusion (Kossoff *et al.*, 2004; Fountas *et al.*, 2005; Osorio *et al.*, 2005). In particular, it remains to be seen whether closed-loop (i.e. responsive) brain stimulation is indeed superior to open-loop (i.e. chronic or scheduled) stimulation in terms of efficacy and tolerability.

For any responsive brain-stimulation system, a crucial issue is where to place both afferent and efferent electrodes, that is, electrodes for detection of a pre-seizure state and stimulation electrodes, respectively. Number and location of electrodes may be critical to achieve a sufficiently early detection of an impending seizure and to apply stimulation locally restricted so that the intervention is not consciously perceived by the patient.

Conclusion

The more rigorous methodological design in many recent seizure-prediction studies has shown that many of the measures previously considered suitable for prediction perform no better than a random predictor. On the other hand, evidence has accumulated that certain measures, particularly measures quantifying relations between recording sites to characterize interaction between different brain regions, show a promising performance that exceeds the chance level as evidenced by statistical validation.

The few studies that have used prediction algorithms in a quasi-prospective manner (i.e. without the use of *a posteriori* information) either did not include a statistical validation or did not apply it correctly.

The design and evaluation of prospective seizure-prediction algorithms involve numerous caveats that need to be considered. The current literature allows no definite conclusion as to whether seizures are predictable by pro-spective algorithms. To answer this question, future studies need to rely on sound and strict methodology and include a rigorous statistical validation.

In order to assure the methodological quality of future studies on seizure prediction, we propose the following guidelines:

- Prediction algorithms should be tested on unselected continuous long-term recordings covering several days of EEG in order to comprise the full spectrum of physiological and pathophysiological states for an individual patient.
- Studies should assess both sensitivity and specificity and should report these quantities with respect to the applied prediction horizon. Rather than false prediction rates, the portion of time under false warning should be reported. If false prediction rates are reported, they should be reported only for the seizure-free interval.
- Results should be tested using statistical validation methods based on Monte Carlo simulations or naïve prediction

schemes to prove that a given prediction algorithm performs indeed above chance level. This is particularly important for studies that contain in-sample optimization such as retrospective adjustment of parameters or selection of EEG channels.

- If prediction algorithms are optimized using training data (in-sample), they should be tested on independent testing data (out-of-sample). If part of the data from an individual patient are used for patient-specific parameter adjustment or EEG channel selection, these data must be excluded when evaluating the performance out-of-sample. Performance of an algorithm should always be reported separately for the testing data.

The next logical step in the field of seizure prediction will be to test on long-term recordings whether any of the prediction algorithms devised to date are able to perform better than a random prediction in a quasi-prospective setting on out-of-sample data. This step is an indispensable prerequisite for justifying prospective clinical trials involving invasive seizure-intervention techniques such as electrical brain stimulation in patients based on seizure prediction.

Supplementary Material

Supplementary data are available at *Brain* online.

Acknowledgements

F.M., C.E.E. and K.L. acknowledge support from the Deutsche Forschungsgemeinschaft and the intramural research fund BONFOR of the University of Bonn. R.G.A. acknowledges support from the Alexander von Humboldt-Foundation. We thank the three anonymous referees for their helpful comments and suggestions.

References

Adelson PD, Nemoto E, Scheuer M, Painter M, Morgan J, Yonas H. Noninvasive continuous monitoring of cerebral oxygenation periictally using near-infrared spectroscopy: a preliminary report. Epilepsia 1999; 40: 1484–9.

Andrzejak RG, Widman G, Lehnertz K, Rieke C, David P, Elger CE. The epileptic process as nonlinear deterministic dynamics in a stochastic environment—an evaluation of mesial temporal lobe epilepsy. Epilepsy Res 2001; 44: 129–40.

Andrzejak RG, Mormann F, Kreuz T, Rieke C, Kraskov A, Elger CE, et al. Testing the null hypothesis of the non-existence of the pre-seizure state. Phys Rev E 2003; 67: 010901.

Andrzejak RG, Mormann F, Widman G, Kreuz T, Elger CE, Lehnertz K. Improved spatial characterization of the epileptic brain by focusing on nonlinearity. Epilepsy Res 2006; 69: 30–44.

Annegers JF. The epidemiology in epilepsy. In: Wyllie E, editor. The treatment of epilepsy: principle and practice. Baltimore: Williams & Wilkins; 1996. p. 165–72.

Arnhold J, Grassberger P, Lehnertz K, Elger CE. A robust method for detecting interdependencies: application to intracranially recorded EEG. Physica D 1999; 134: 419–30.

Aschenbrenner-Scheibe R, Maiwald T, Winterhalder M, Voss HU, Timmer J, Schulze-Bonhage A. How well can epileptic seizures be predicted? An evaluation of a nonlinear method. Brain 2003; 126: 2616–26.

Bartolomei F, Wendling F, Regis J, Gavaret M, Guye M, Chauvel P. Pre-ictal synchronicity in limbic networks of mesial temporal lobe epilepsy. Epilepsy Res 2004; 61: 89–104.

Bauer J, Burr W. Course of chronic focal epilepsy resistant to anticonvulsant treatment. Seizure 2001; 10: 239–46.

Baumgartner C, Serles W, Leutmezer F, Pataraia E, Aull S , Czech T, et al. Preictal SPECT in temporal lobe epilepsy: regional cerebral blood flow is increased prior to electroencephalography-seizure onset. J Nucl Med 1998; 39: 978–82.

Bragin A, Engel J Jr, Wilson CL, Fried I, Buzsaki G. High-frequency oscillations in human brain. Hippocampus 1999; 9: 137–42.

Bragin A, Mody I, Wilson CL, Engel J Jr. Local generation of fast ripples in epileptic brain. J Neurosci 2002; 22: 2012–21.

Broomhead DS, King GP. Extracting qualitative dynamics from experimental data. Physica D 1986; 20: 217–36.

Cerf R, el-Ouasdad EH. Spectral analysis of stereo-electroencephalograms: preictal slowing in partial epilepsies. Biol Cybern 2000; 83: 399–405.

Cerf R, el-Ouasdad EH, Kahane P. Criticality and synchrony of fluctuations in rhythmical brain activity: pretransitional effects in epileptic patients. Biol Cybern 2004; 90: 239–55.

Chaovalitwongse W, Iasemidis LD, Pardalos PM, Carney PR, Shiau DS, Sackellares JC. Performance of a seizure warning algorithm based on the dynamics of intracranial EEG. Epilepsy Res 2005; 64: 93–113.

Chaovalitwongse W, Iasemidis LD, Pardalos PM, Carney PR, Shiau DS, Sackellares JC. Performance of a seizure warning algorithm based on the dynamics of intracranial EEG [Author Reply]. Epilepsy Res 2006, Epub ahead of print doi:10.1016/j.eplepsyres.2006.06.014.

Chávez M, Le Van Quyen M, Navarro V, Baulac M, Martinerie J. Spatio-temporal dynamics prior to neocortical seizures: amplitude versus phase couplings. IEEE Trans Biomed Eng 2003; 50: 571–83.

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

D'Alessandro M, Vachtsevanos G, Esteller R, Echauz J, Cranstoun S, Worrell G, et al. A multi-feature and multi-channel univariate selection process for seizure prediction. Clin Neurophysiol 2005; 116: 506–16.

De Clercq W, Lemmerling P, Van Huffel S, Van Paesschen W. Anticipation of epileptic seizures from standard EEG recordings. Lancet 2003; 361: 971.

Delamont R, Julu P, Jamal G. Changes in a measure of cardiac vagal activity before and after epileptic seizures. Epilepsy Res 1999; 35: 87–94.

Drury I, Smith B, Li D, Savit R. Seizure prediction using scalp electroencephalogram. Exp Neurol 2003; 184: 9–18.

Elger CE, Lehnertz K. Seizure prediction by non-linear time series analysis of brain electrical activity. Eur J Neurosci 1998; 10: 786–9.

Elger CE. Future trends in epileptology. Curr Opin Neurol 2001; 14: 185–6.

Esteller R, Echauz J, D'Alessandro M, Worrell G, Cranstoun S, Vachtsevanos G, et al. Continuous energy variation during the seizure cycle: towards an on-line accumulated energy. Clin Neurophysiol 2005; 116: 517–26.

Federico P, Abbott DF, Briellmann RS, Harvey AS, Jackson GD. Functional MRI of the pre-ictal state. Brain 2005; 128: 1811–7.

Fountas KN, Smith JR, Murro AM, Politsky J, Park YD, Jenkins PD. Implantation of a closed-loop stimulation in the management of medically refractory focal epilepsy: a technical note. Stereotact Funct Neurosurg 2005; 83: 153–8.

Franaszczuk PJ, Bergey GK, Kaminski MJ. Analysis of mesial temporal seizure onset and propagation using the directed transfer function method. Electroencephalogr Clin Neurophysiol 1994; 91: 413–27.

Franaszczuk PJ, Bergey GK, Durka PJ, Eisenberg HM. Time-frequency analysis using the matching pursuit algorithm applied to seizures originating from the mesial temporal lobe. Electroencephalogr Clin Neurophysiol 1998; 106: 513–21.

Gabor D. Theory of communication. Proc IEEE London 1946; 93: 429–57.

Gigola S, Ortiz F, D'Attellis CE, Silva W, Kochen S. Prediction of epileptic seizures using accumulated energy in a multiresolution framework. J Neurosci Methods 2004; 138: 107–11.

Gotman J, Marciani MG. Electroencephalographic spiking activity, drug levels and seizure occurrence in epileptic patients. Ann Neurol 1985; 17: 597–603.

Gotman J, Koffler DJ. Interictal spiking increases after seizures but does not after decrease in medication. Electroencephalogr Clin Neurophysiol 1989; 72: 7–15.

Grassberger P, Procaccia I. Characterization of strange attractors. Phys Rev Lett 1983a; 50: 346–9.

Grassberger P, Procaccia I. Estimation of the Kolmogorov entropy from a chaotic signal. Phys Rev A 1983b; 28: 2591–3.

Grassberger P, Schreiber T, Schaffrath C. Nonlinear time sequence analysis. Int J Bifurcat Chaos 1991; 1: 521–47.

Hao B-L. Elementary symbolic dynamics and chaos in dissipative systems. Singapore: World Scientific; 1989.

Harrison MA, Frei MG, Osorio I. Accumulated energy revisited. Clin Neurophysiol 2005a; 116: 527–31.

Harrison MA, Osorio I, Frei MG, Asuri S, Lai YC. Correlation dimension and integral do not predict epileptic seizures. Chaos 2005b; 15: 33106.

Haut SR, Swick C, Freeman K, Spencer S. Seizure clustering during epilepsy monitoring. Epilepsia 2002; 43: 711–5.

Hill MW, Wong M, Amarakone A, Rothman SM. Rapid cooling aborts seizure-like activity in rodent hippocampal-entorhinal slices. Epilepsia 2000; 41: 1241–8.

Hively LM, Protopopescu VA, Gailey PC. Timely detection of dynamical change in scalp EEG signals. Chaos 2000; 10: 864–75.

Hively LM, Protopopescu VA. Channel-consistent forewarning of epileptic events from scalp EEG. IEEE Trans Biomed Eng 2003; 50: 584–93.

Hjorth B. EEG analysis based on time domain properties. Electroencephalogr Clin Neurophysiol 1970; 29: 306–10.

Hughes J, Devinsky O, Feldmann E, Bromfield E. Premonitory symptoms in epilepsy. Seizure 1993; 2: 201–3.

Huygens C. Horoloquium Oscilatorium. In: Apud F. Muguet, Parisis, France. 1673 (Engl. translation: The pendulum clocks. Ames: Iowa State University Press; 1986).

Iasemidis LD, Sackellares JC, Zaveri HP, Williams WJ. Phase space topography and the Lyapunov exponent of electrocorticograms in partial seizures. Brain Topogr 1990; 2: 187–201.

Iasemidis LD, Pardalos P, Sackellares JC, Shiau DS. Quadratic binary programming and dynamical system approach to determine the predictability of epileptic seizures. J Comb Optim 2001; 5: 9–26.

Iasemidis LD, Shiau DS, Chaovalitwongse W, Sackellares JC, Pardalos PM, Principe JC, et al. Adaptive epileptic seizure prediction system. IEEE Trans Biomed Eng 2003; 50: 616–27.

Iasemidis LD, Shiau DS, Pardalos PM, Chaovalitwongse W, Narayanan K, Prasad A, et al. Long-term prospective on-line real-time seizure prediction. Clin Neurophysiol 2005; 116: 532–44.

Jerger KK, Netoff TI, Francis JT, Sauer T, Pecora L, Weinstein SL, et al. Early seizure detection. J Clin Neurophysiol 2001; 18: 259–68.

Jerger KK, Weinstein SL, Sauer T, Schiff SJ. Multivariate linear discrimination of seizures. Clin Neurophysiol 2005; 116: 545–51.

Jouny CC, Franaszczuk PJ, Bergey GK. Signal complexity and synchrony of epileptic seizures: is there an identifiable preictal period? Clin Neurophysiol 2005; 116: 552–8.

Kalitzin S, Parra J, Velis D, Lopes da Silva FH. Enhancement of phase clustering in the EEG/MEG gamma frequency band anticipates transitions to paroxysmal epileptiform activity in epileptic patients with known visual sensitivity. IEEE Trans Biomed Eng 2002; 49: 1279–86.

Kalitzin S, Velis D, Suffczynski P, Parra J, Lopes da Silva F. Electrical brain-stimulation paradigm for estimating the seizure onset site and the time to ictal transition in temporal lobe epilepsy. Clin Neurophysiol 2005; 116: 718–28.

Kantz H. Quantifying the closeness of fractal measures. Phys Rev E 1994a; 49: 5091–7.

Kantz H. A robust method to estimate the maximal Lyapunov exponent of a time series. Phys Lett A 1994b; 185: 77–87.

Kantz H, Schreiber T. Nonlinear time series analysis. Cambridge: Cambridge University Press; 1997.

Kaplan DT, Glass L. Direct test for determinism in a time series. Phys Rev Lett 1992; 68: 427–30.

Kasper F, Schuster HG. Easily calculable measure for the complexity of spatio-temporal patterns. Phys Rev A 1987; 36: 842–8.

Katz A, Marks DA, McCarthy G, Spencer SS. Does interictal spiking rate change prior to seizures? Electroencephalogr Clin Neurophysiol 1991; 79: 153–6.

Kerem DH, Geva AB. Forecasting epilepsy from the heart rate signal. Med Biol Eng Comput 2005; 43: 230–9.

Kossoff EH, Ritzl EK, Politsky JM, Murro AM, Smith JR, Duckrow RB, et al. Effect of an external responsive neurostimulator on seizures and electrographic discharges during subdural electrode monitoring. Epilepsia 2004; 45: 1560–7.

Kreuz T, Andrzejak RG, Mormann F, Kraskov A, Stogbauer H, Elger CE, et al. Measure profile surrogates: a method to validate the performance of epileptic seizure prediction algorithms. Phys Rev E 2004; 69: 061915.

Lachaux JP, Rodriguez E, Martinerie J, Varela FJ. Measuring phase synchrony in brain signals. Hum Brain Mapp 1999; 8: 194–208.

Lai YC, Harrison MA, Frei MG, Osorio I. Inability of Lyapunov exponents to predict epileptic seizures. Phys Rev Lett 2003; 91: 068102.

Lai YC, Harrison MA, Frei MG, Osorio I. Controlled test for predictive power of Lyapunov exponents: their inability to predict epileptic seizures. Chaos 2004; 14: 630–42.

Lange HH, Lieb JP, Engel J Jr, Crandall PH. Temporo-spatial patterns of pre-ictal spike activity in human temporal lobe epilepsy. Electroencephalogr Clin Neurophysiol 1983; 56: 543–55.

Le Van Quyen M , Martinerie J, Baulac M, Varela F. Anticipating epileptic seizure in real time by a nonlinear analysis of similarity between EEG recordings. Neuroreport 1999; 10: 2149–55.

Le Van Quyen M, Adam C, Martinerie J, Baulac M, Clemenceau S, Varela F. Spatio-temporal characterization of non-linear changes in intracranial activities prior to human temporal lobe seizures. Eur J Neurosci 2000; 12: 2124–34.

Le Van Quyen M, Martinerie J, Navarro V, Boon P, D'Have M, Adam C, et al. Anticipation of epileptic seizures from standard EEG recordings. Lancet 2001a; 357: 183–8.

Le Van Quyen M, Martinerie J, Navarro V, Baulac And M, Varela FJ. Characterizing neurodynamic changes before seizures. J Clin Neurophysiol 2001b; 18: 191–208.

Le Van Quyen M, Navarro V, Martinerie J, Baulac M, Varela FJ. Towards a neurodynamical understanding of ictogenesis. Epilepsia 2003; 44: 30–43.

Le Van Quyen M, Soss J, Navarro V, Robertson R, Chavez M, Baulac M, et al. Preictal state identification by synchronization changes in long-term intracranial EEG recordings. Clin Neurophysiol 2005; 116: 559–68.

Lehnertz K. Non-linear time series analysis of intracranial EEG recordings in patients with epilepsy—an overview. Int J Psychophysiol 1999; 34: 45–52.

Lehnertz K, Elger CE. Neuronal complexity loss in temporal lobe epilepsy: effects of carbamazepine on the dynamics of the epileptogenic focus. Electroencephalogr Clin Neurophysiol 1997; 103: 376–80.

Lehnertz K, Elger CE. Can epileptic seizures be predicted? Evidence from nonlinear time series analysis of brain electrical activity. Phys Rev Lett 1998; 80: 5019–23.

Lehnertz K, Litt B. The first international collaborative workshop on seizure prediction: summary and data description. Clin Neurophysiol 2005; 116: 493–505.

Lehnertz K, Andrzejak RG, Arnhold J, Kreuz T, Mormann F, Rieke C, et al. Nonlinear EEG analysis in epilepsy: its possible use for interictal focus localization, seizure anticipation, and prevention. J Clin Neurophysiol 2001; 18: 209–22.

Lempel A, Ziv J. On the complexity of finite sequences. IEEE Trans Inf Theory 1976; 22: 75–81.

Lerner DE. Monitoring changing dynamics with correlation integrals: case study of an epileptic seizure. Physica D 1996; 97: 563–76.

Li D, Zhou W, Drury I, Savit R. Non-linear, non-invasive method for seizure anticipation in focal epilepsy. Math Biosci 2003; 186: 63–77.

Litt B, Esteller R, Echauz J, D'Alessandro M, Shor R, Henry T, et al. Epileptic seizures may begin hours in advance of clinical onset: a report of five patients. Neuron 2001; 30: 51–64.

Lopes da Silva FH. EEG analysis: theory and practice. In: Niedermayer E, Lopes da Silva FH, editors. Electroencephalography, basic principles, clinical applications and related fields. Baltimore: Urban and Schwarzenberg; 1987. p. 871–97.

Lopes da Silva FH, Mars NJI. Parametric methods in EEG analysis. In: Gevins AS, Remond A, editors. Methods of analysis of brain electrical and magnetic signals. EEG Handbook (Revised Series, Vol. 1). Amsterdam: Elsevier; 1987. p. 243–60.

Lopes da Silva F, Blanes W, Kalitzin S, Parra J, Suffczynski P, Velis DN. Epilepsies as a dynamical disease of brain systems: basic models of the transition between normal and epileptic activity. Epilepsia 2003; 44: 72–83.

Maiwald T, Winterhalder M, Aschenbrenner-Scheibe R, Voss HU, Schulze-Bonhage A, Timmer J. Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic. Physica D 2004; 194: 357–68.

Makhoul J. Linear prediction: a tutorial review. Proc IEEE 1973; 63: 561–80.

Manuca R, Savit R. Stationarity and nonstationarity in time series analysis. Physica D 1996; 99: 134–61.

Martinerie J, Adam C, Le Van Quyen M, Baulac M, Clemenceau S, Renault B, et al. Epileptic seizures can be anticipated by non-linear analysis. Nat Med 1998; 4: 1173–6.

McSharry PE, Smith LA, Tarassenko L. Prediction of epileptic seizures: are nonlinear methods relevant? Nat Med 2003; 9: 241–2.

Mormann F, Lehnertz K, David P, Elger CE. Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients. Physica D 2000; 144: 358–69.

Mormann F, Andrzejak RG, Kreuz T, Rieke C, David P, Elger CE, et al. Automated detection of a pre-seizure state based on a decrease in synchronization in intracranial EEG recordings from epilepsy patients. Phys Rev E 2003a; 67: 021912.

Mormann F, Kreuz T, Andrzejak RG, David P, Lehnertz K, Elger CE. Epileptic seizures are preceded by a decrease in synchronization. Epilepsy Res 2003b; 53: 173–85.

Mormann F, Kreuz T, Rieke C, Andrzejak RG, Kraskov A, David P, et al. On the predictability of epileptic seizures. Clin Neurophysiol 2005; 116: 569–87.

Mormann F, Elger CE, Lehnertz K. Seizure anticipation: from algorithms to clinical practice [Review]. Curr Opin Neurol 2006a; 19: 187–93.

Mormann F, Elger CE, Lehnertz K. Performance of a seizure warning algorithm based on the dynamics of intracranial EEG [Comment Letter]. Epilepsy Res 2006b, Epub ahead of print doi:10.1016/j.eplepsyres.2006.06.011.

Morrell M. Brain stimulation for epilepsy: can scheduled or responsive neurostimulation stop seizures? [Review]. Curr Opin Neurol 2006; 19: 164–8.

Navarro V, Martinerie J, Le Van Quyen M, Clemenceau S, Adam C, et al. Seizure anticipation in human neocortical partial epilepsy. Brain 2002; 125: 640–55.

Navarro V, Martinerie J, Le Van Quyen M, Baulac M, Dubeau F, Gotman J. Seizure anticipation: do mathematical measures correlate with video-EEG evaluation? Epilepsia 2005; 46: 385–96.

Niederhauser JJ, Esteller R, Echauz J, Vachtsevanos G, Litt B. Detection of seizure precursors from depth-EEG using a sign periodogram transform. IEEE Trans Biomed Eng 2003; 50: 449–58.

Novak V, Reeves AL, Novak P, Low PA, Sharbrough FW. Time-frequency mapping of R-R interval during complex partial seizures of temporal lobe origin. J Auton Nerv Syst 1999; 77: 195–202.

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

Osorio I, Frei MG, Manly BF, Sunderam S, Bhavaraju NC, Wilkinson SB. An introduction to contingent (closed-loop) brain electrical stimulation for seizure blockage, to ultra-short-term clinical trials, and to multidimensional statistical analysis of therapeutic efficacy. J Clin Neurophysiol 2001; 18: 533–44.

Osorio I, Frei MG, Sunderam S, Giftakis J, Bhavaraju NC, Schaffner SF, et al. Automated seizure abatement in humans using electrical stimulation. Ann Neurol 2005; 57: 258–68.

Ott E. Chaos in dynamical systems. Cambridge: Cambridge University Press; 1993.

Panter P. Modulation, noise, and spectral analysis. New York: McGraw-Hill, 1965.

Parra J, Kalitzin SN, Iriarte J, Blanes W, Velis DN, Lopes da Silva FH. Gamma-band phase clustering and photosensitivity: is there an underlying mechanism common to photosensitive epilepsy and visual perception? Brain 2003; 126: 1164–72.

Pijn JP, Van Neerven J, Noest A, Lopes da Silva FH. Chaos or noise in EEG signals; dependence on state and brain site. Electroencephalogr Clin Neurophysiol 1991; 79: 371–81.

Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical recipes in C: the art of scientific computing (second edition). Cambridge: Cambridge University Press. 1992.

Protopopescu VA, Hively LM, Gailey PC. Epileptic seizure forewarning from scalp EEG. J Clin Neurophysiol 2001; 18: 223–45.

Quian Quiroga R, Kraskov A, Kreuz T, Grassberger P. Performance of different synchronization measures in real data: a case study on electroencephalographic signals. Phys Rev E 2002; 65: 041903.

Rajna P, Clemens B, Csibri E, Dobos E, Geregely A, Gottschal M, et al. Hungarian multicentre epidemiologic study of the warning and initial symptoms (prodrome, aura) of epileptic seizures. Seizure 1997; 6: 361–8.

Rieke C, Sternickel C, Andrzejak RG, Elger CE, David P, Lehnertz K. Measuring nonstationarity by analyzing the loss of recurrence in dynamical systems. Phys Rev Lett 2002; 88: 244102.

Rieke C, Andrzejak RG, Mormann F, Lehnertz K. Improved statistical test for nonstationarity using recurrence time statistics. Phys Rev E 2004; 69: 046111.

Rogowski Z, Gath I, Bental E. On the prediction of epileptic seizures. Biol Cybern 1981; 42: 9–15.

Rosenblum MG, Pikovsky AS, Kurths J. Phase synchronization of chaotic oscillators. Phys Rev Lett 1996; 76: 1804–7.

Rosenblum MG, Pikovsky AS, Kurths J. From phase to lag synchronization in coupled chaotic oscillators. Phys Rev Lett 1997; 78: 4193–6.

Rosenblum M, Pikovsky A, Kurths J, Schäfer C, Tass PA. Phase synchronization: from theory to data analysis. In: Moss F, Gielen S, editors. Handbook of biological physics. Amsterdam: Elsevier; 2001. p. 279–321.

Rosenstein MT, Collins JC, De Luca CJ. A practical method for calculating the largest Lyapunov exponents from small datasets. Physica D 1993; 65: 117–34.

Rulkov NF, Sushchik MM, Tsimring LS, Ababarnel HDI. Generalized synchronization of chaos in directionally coupled chaotic systems. Phys Rev E 1995; 51: 980–94.

Salant Y, Gath I, Henriksen O. Prediction of epileptic seizures from two-channel EEG. Med Biol Eng Comput 1998; 36: 549–56.

Savit R, Green M. Time series and dependent variables. Physica D 1991; 50: 95–116.

Schelter B, Winterhalder M, Maiwald T, Brandt A, Schad A, Schulze-Bonhage A, et al. Testing statistical significance of multivariate time series analysis techniques for epileptic seizure prediction. Chaos 2006; 16: 013108.

Schiff SJ, Colella D, Jacyna GM, Hughes E, Creekmore JW, Marshall A, et al. Brain chirps: spectrographic signatures of epileptic seizures. Clin Neurophysiol 2000; 111: 953–8.

Schiff SJ, Sauer T, Kumar R, Weinstein SL. Neuronal spatiotemporal pattern discrimination: the dynamical evolution of seizure. Neuroimage 2005; 28: 1043–55.

Schindler K, Wiest R, Kollar M, Donati F. EEG analysis with simulated neuronal cell models helps to detect pre-seizure changes. Clin Neurophysiol 2002; 113: 604–14.

Schreiber T. Efficient neighbor searching in nonlinear time series analysis. Int J Bifurcat Chaos 1995; 5: 349–58.

Schreiber T, Schmitz A. Improved surrogate data for nonlinearity tests. Phys Rev Lett 1996; 77: 635–8.

Schreiber T, Schmitz A. Surrogate time series. Physica D 2000; 142: 346–82.

Schulze-Bonhage A, Kurth C, Carius A, Steinhoff BJ, Mayer T. Seizure anticipation by patients with focal and generalized epilepsy: a multicentre assessment of premonitory symptoms. Epilepsy Res 2006; 70: 83–8.

Schuster HG. Deterministic chaos: an introduction. 2nd revised version. Weinheim: VCH; 1989.

Siegel A, Grady CL, Mirsky AF. Prediction of spike-wave bursts in absence epilepsy by EEG power-spectrum signals. Epilepsia 1982; 23: 47–60.

Stam CJ. Nonlinear dynamical analysis of EEG and MEG: review of an emerging field [Review]. Clin Neurophysiol 2005; 116: 2266–301.

Stanski DR, Hudson RJ, Homer TD, Saidman LJ, Meathe E. Pharmacodynamic modeling of thiopental anesthesia. J Pharmacokinet Biopharm 1984; 12: 223–40.
Stein AG, Eder HG, Blum DE, Drachev A, Fisher RS. An automated drug delivery system for focal epilepsy. Epilepsy Res 2000; 39: 103–14.
Sterman MB. Basic concepts and clinical findings in the treatment of seizure disorders with EEG operant conditioning. Clin Electroencephalogr 2000; 31: 45–55.
Suffczynski P, Lopes da Silva F, Parra J, Velis D, Kalitzin S. Epileptic transitions: model predictions and experimental validation. J Clin Neurophysiol 2005; 22: 288–99.
Suffczynski P, Lopes da Silva FH, Parra J, Velis DN, Bouwman BM, van Rijn CM, et al. Dynamics of epileptic phenomena determined from statistics of ictal transitions. IEEE Trans Biomed Eng 2006; 53: 524–32.
Takens F. Detecting strange attractors in turbulence. In: Rand DA, Young LS, editors. Dynamical systems and turbulence lecture notes on mathematics 898. New York: Springer; 1981. p. 366–81.
Tass P, Rosenblum MG, Weule J, Kurths J, Pikovsky AS, Volkmann J, et al. Detection of n:m phase locking from noisy data: application to magnetoencephalography. Phys Rev Lett 1998; 81: 3291–4.
Theiler J. Spurious dimension from correlation algorithms applied to limited time series data. Phys Rev A 1986; 34: 2427–33.
Theodore WH, Fisher RS. Brain stimulation for epilepsy. [Review]. Lancet Neurol 2004; 3: 111–8.
van Drongelen W, Nayak S, Frim DM, Kohrman MH, Towle VL, Lee HC, et al. Seizure anticipation in pediatric epilepsy: use of Kolmogorov entropy. Pediatr Neurol 2003; 29: 207–13.
Viglione S, Walsh G. Proceedings. Epileptic seizure prediction [Abstract]. Electroencephalogr Clin Neurophysiol 1975; 39: 435–6.
Weinand ME, Carter LP, El-Saadany WF, Sioutos PJ, Labiner DM, Oommen KJ. Cerebral blood flow and temporal lobe epileptogenicity. J Neurosurg 1997; 86: 226–32.
Wendling F, Bartolomei F, Bellanger JJ, Chauvel P. Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition. Eur J Neurosci 2002; 15: 1499–508.
Wendling F, Bartolomei F, Bellanger JJ, Bourien J, Chauvel P. Epileptic fast intracerebral EEG activity: evidence for spatial decorrelation at seizure onset. Brain 2003; 126: 1449–59.
Whitney H. Differentiable manifolds. Ann Math 1936; 37: 645–80.
Winterhalder M, Maiwald T, Voss HU, Aschenbrenner-Scheibe R, Timmer J, Schulze-Bonhage A. The seizure prediction characteristic: a general framework to assess and compare seizure prediction methods. Epilepsy Behav 2003; 4: 318–25.
Winterhalder M, Schelter B, Schulze-Bonhage A, Timmer J. Performance of a seizure warning algorithm based on the dynamics of intracranial EEG [Comment Letter]. Epilepsy Res 2006, Epub ahead of print doi:10.1016/j.eplepsyres.2006.06.012.
Wolf A, Swift JB, Swinney L, Vastano A. Determining Lyapunov exponents from a time series. Physica D 1985; 16: 285–317.
Worrell G, Parish L, Cranstoun S, Jonas R, Baltuch G, Litt B. High frequency oscillations and seizure generation in neocortical epilepsy. Brain 2004; 127: 1496–506.
Yaari Y, Beck H. 'Epileptic neurons' in temporal lobe epilepsy. Brain Pathol 2002; 12: 234–9.

Appendix A: Conceptual issues

This appendix addresses issues that need to be resolved before designing a study on seizure prediction.

Prediction, forecasting or anticipation?

In the strict sense of the words predicting or forecasting an event means the ability to determine in advance the time of its occurrence with a certain precision. The term anticipation implies more of an uncertainty as to when exactly an event will occur. This latter concept better fits the design of seizure-prediction algorithms, which usually assume a seizure to occur within a certain time period after an alarm is issued without knowing its exact onset time. As in the majority of publications in this field, however, we will use the three different terms interchangeably.

The events to be predicted: clinical or electrographical seizures?

An important issue is the selection of the ictal events that are to be anticipated by an algorithm. While the benchmark for clinical application would clearly be the forecasting of clinical seizure events, subclinical seizures today are mostly regarded not as a different entity, but rather as a milder variant of the same dynamical event that constitutes a clinical seizure. It is therefore arguable whether it is reasonable to exclude subclinical ictal events in a prediction algorithm. Nevertheless, most studies so far have restricted themselves to the analysis of clinical seizures.

Similarly, the onset time of a seizure can be determined either from the first clinical signs or from the first visible EEG changes. Since there is often some uncertainty in the assessment of clinical symptoms, particularly in complex partial and absence seizures, it is reasonable to determine the seizure onset electrographically, especially if intracranial recordings from the seizure-onset zone are available.

Seizure prediction versus early seizure detection

Algorithms that aim at an early detection of the electrographical seizure onset, which may occur several seconds before the first clinical symptoms, should not be regarded as seizure-prediction algorithms, but rather as early seizure-detection algorithms (e.g. Osorio *et al.*, 1998). In contrast to seizure prediction, which aims at the identification of a preictal state sufficiently long before the electrographical seizure onset, early seizure detection does not provide an extensive time for intervention if any at all. Early seizure detection alone, without an intervention to immediately abort a seizure, may provide little clinical benefit to a patient, aside from alerting the patient that a seizure is imminent.

If implemented within a closed-loop intervention system (*see* Future perspectives) endowed with sufficient seizure-abatement strategies, however, early detection algorithms may prove useful as a basis for responsive intervention, provided the epileptic brain is not yet beyond a 'point of no return' from which it will inevitably evolve into a clinical seizure.

The type of EEG: intracranial or surface recordings?

While the majority of seizure-prediction studies to date have been carried out on intracranial recordings, there are some

studies that analysed surface recordings (Table 1). Intracranial recordings bear the advantage of a higher signal-to-noise ratio and a better spatial resolution, and the data can be considered mostly artefact-free. They also bear the potential advantage of allowing one to record directly from the seizure-generating region. On the other hand, surface recordings are less invasive and could, in principle, be used in an ambulatory setting to monitor a patient's seizure situation in his/her usual environment. This would, however, require a high degree of compliance on the part of a patient owing to the inconvenience of constantly wearing an EEG cap.

Furthermore, if seizure-anticipation algorithms proved to be successful, they would most likely be implemented in an implantable, closed-loop warning or intervention system. The technical feasibility of intracranial intervention systems has already been proven by responsive brain-stimulation devices that are currently being tested in clinical trials for their ability to reduce seizure frequency (cf. Morrell, 2006). Many groups in the field therefore regard the usefulness of scalp EEG recordings for studies on seizure prediction as rather limited compared with intracranial recordings.

Data requirements

While it is on the one hand desirable to use data sets for analysis that contain a large number of seizures, it is also desirable to have a sufficient time interval between consecutive seizures, so that they can be regarded as independent events. If seizures are too closely spaced (clustered seizures) it becomes difficult to separate the post-ictal period from a presumed pre-ictal state (cf. Jouny *et al.*, 2005) as the exact duration of either of the two is unknown. It may be noted in this context that the average seizure frequency in a monitoring unit of up to three events per day (Haut *et al.*, 2002) is ~30 times higher than the mean seizure frequency of three per month under normal circumstances (Bauer and Burr, 2001). If a certain false prediction rate in the epilepsy monitoring unit corresponds to a situation where every other alarm is a false alarm (positive predictive value of 50%), then the same false prediction rate under normal circumstances would mean that only 1 out of 60 alarms is a correct warning (positive predictive value of 1.7%) (cf. Winterhalder *et al.*, 2003).

EEG recordings used for studies on seizure prediction should ideally comprise EEG data recorded continuously over several days. Recording gaps due to diagnostic procedures during the pre-surgical work-up (e.g. structural MRI to verify electrode placement) are usually unavoidable and are not considered a major drawback. Since during the pre-surgical monitoring, patients are constantly undergoing changes that could have a confounding influence on characterizing measures of the EEG (e.g. tapering of medication), it is advisable to use all inter-ictal control data available since a restriction (e.g. the first 24 h of an EEG) could introduce a confounding bias.

Appendix B: Assessing the performance of a prediction algorithm

In order to compare the relative merit of the different studies on seizure prediction published to date, it is necessary to realize how the performance of a seizure-prediction technique is assessed. In this appendix we will therefore discuss some of the problems and pitfalls involved in the evaluation of an algorithm for seizure prediction.

Moving window analysis

Most of the prediction techniques published up to now use a so-called moving window analysis in which some (linear or non-linear) characterizing measure is calculated from a window of EEG data with a pre-defined length, then the subsequent window of EEG is analysed, and so forth. The duration of these analysis windows usually ranges between 10 and 40 s. Depending on whether the employed measure is used to characterize a single EEG channel or relations between two or more channels, it is referred to as a univariate, bivariate or multivariate measure, respectively. The moving window analysis thus renders *time profiles* of a characterizing measure for different channels or channel combinations, respectively.

Statistical versus algorithmic approaches

The analysis design used to evaluate these time profiles in the following step can be either statistical or algorithmic (cf. Table 1). A *statistical* design is retrospective by nature and compares the amplitude distributions of the characterizing measures from the inter-ictal with those from the *assumed pre-ictal period* in one way or another. The temporal structure of the time profiles is usually not preserved in this type of analysis. Such a design can be useful for investigating and comparing the potential predictive performance of different characterizing measures under different conditions.

On the other hand, an *algorithmic* analysis uses a design that produces a time-resolved output (i.e. an output for every point of a time profile). With respect to practical application, the algorithm should ideally be prospective (i.e. its output for a given time should be a function of the information available at this time). Prediction algorithms usually employ certain thresholds. If the time profile of a characterizing measure crosses the threshold, the algorithm produces an alarm. This alarm can be either true or false, depending on whether it is actually followed by a seizure or not. For this distinction, it is necessary to define a *prediction horizon* (i.e. the period after an alarm within which a seizure is expected). If an alarm is followed by a seizure within the prediction horizon, it is classified as a true alarm (true positive); otherwise it is regarded as a false alarm (false positive). In addition, it may be useful to require a minimum time interval between an alarm and a seizure occurrence in order to count this alarm as a successful prediction if the algorithm is to be used for seizure prevention. This *minimum intervention time* can be introduced as an additional constraint. [It may be noted

that in the literature, different definitions are sometimes used for these quantities, e.g. one group has used the term 'seizure occurrence period' instead of prediction horizon and 'seizure prediction horizon' instead of minimum intervention time (Aschenbrenner-Scheibe *et al.*, 2003; Winterhalder *et al.*, 2003; Maiwald *et al.*, 2004)]. In studies that employ a statistical instead of an algorithmic design, the prediction horizon corresponds to the assumed pre-ictal period.

Sensitivity and specificity

If a seizure is not preceded by an alarm within the prediction horizon, this will be counted as a false negative. A less trivial question is how to quantify true negatives. In principle, every single window of the moving window analysis that is outside the duration of the assumed pre-ictal period (i.e. one prediction horizon before a seizure) and does not produce an alarm could be counted as a true negative. However, since *sensitivity* is usually quantified as the number of seizures with at least one alarm within the preceding prediction horizon divided by the total number of seizures, it is reasonable to define *specificity* based on the prediction horizon, too. If, for instance, the prediction horizon is 3 h, the sensitivity quantifies the fraction of correctly classified pre-ictal 3-h segments, while the specificity measures the fraction of correctly classified (consecutive) inter-ictal 3-h segments.

In order to avoid any ambiguity in statistically quantifying the specificity of a prediction algorithm, most groups have instead reported specificity rates measured as false predictions per hour. Unfortunately, even for false prediction rates, different definitions are found in the literature. Several groups have determined false prediction rates by counting all false positives and dividing this number by the total duration of the analysed recording (Iasemidis *et al.*, 2003; Niederhauser *et al.*, 2003; Chaovalitwongse *et al.*, 2005; Esteller *et al.*, 2005; Iasemidis *et al.*, 2005). This definition ignores the fact that for each seizure contained in the recording, there is a pre-ictal period (i.e. the prediction horizon) during which every alarm is counted as a true prediction, and false predictions cannot occur by definition. Therefore, other groups have used corrected false prediction rates that were calculated only for the inter-ictal period (Aschenbrenner-Scheibe *et al.*, 2003; Mormann *et al.*, 2003*a*, *b*; Winterhalder *et al.*, 2003; Maiwald *et al.*, 2004).

In this context it is important to realize that a reported false prediction rate cannot be judged independent from the prediction horizon, since in a prospective prediction algorithm a false alarm will leave the patient mistakenly awaiting a seizure for the duration of the prediction horizon. It is only after this duration that the patient will know if the alarm was a false warning or not.

As an example from the literature, consider an algorithm with a 2-h prediction horizon that yields a sensitivity of 9/11 = 82% and an uncorrected false prediction rate of 6/41 h = 0.15/h (Iasemidis *et al.*, 2005). If we take into account that the uncorrected false prediction rate includes the pre-ictal periods during which no false prediction can occur by definition, the corrected false prediction rate (assuming that the pre-ictal periods of the different seizures are non-overlapping) is 6/19 h = 0.32/h, thus, more than twice as high. Furthermore, if we consider that after each false prediction, the patient needs to wait for 2 h before knowing if it was a false prediction, the algorithm of our example may (assuming that false predictions are not spaced closer than the prediction horizon) leave a patient spending $6 \times 2/19 = 63\%$ of the inter-ictal period waiting for a seizure that will not occur while still failing to anticipate every fifth seizure. An algorithm yielding the same results for a prediction horizon of 10 min would instead leave the patient in futile expectation of a seizure only in 3% of his/her seizure-free time. This example shows that a prediction rate should be judged in view of the prediction horizon used by the algorithm and that it is the product of these two quantities that should be compared across studies.

A better way to assess the specificity of a prediction algorithm would therefore be to report the portion of time from the inter-ictal period (i.e. the inter-seizure interval without the pre-ictal period) during which the patient is not in the state of falsely awaiting a seizure.

In general, any algorithm can be tuned (e.g. by varying the alarm threshold) to yield a higher sensitivity at the cost of a lower specificity and vice versa. For a closed-loop intervention system, the desired relation between these two quantities will depend on the invasiveness of the intervention technique under consideration. If the intervention does not impair the patient, a higher false prediction rate will be tolerated up to the point where even a constant intervention (such as a chronic or scheduled stimulation from implantable brain-stimulation devices; cf. Theodore and Fisher, 2004) is possible and could be performed without a prediction algorithm.

The problem of in-sample optimization

Another important issue in the evaluation of a prediction algorithm is the use of *a posteriori* information. For a prospective prediction algorithm, this type of information is not available. Two typical cases of using *a posteriori* information are found in the literature: (i) in-sample optimization of parameters of the algorithm; and (ii) *a posteriori* selection of one or more channels with optimum performance.

In-sample optimization or training of parameters is present whenever parameters used for the calculation of the characterizing measure of the EEG or of the prediction algorithm itself are adjusted to produce optimal performance of the algorithm for a given set of data. Such an optimization is likely to result in an over-estimated performance that will not be reproducible when applying the algorithm to other, out-of-sample testing data that were not used in the optimization process. In order to assess the true performance of a prediction algorithm, it is therefore mandatory to test it on out-of-sample data.

Another way of using *a posteriori* information relates to the selection of channels that are able to discriminate an inter-ictal from a pre-ictal state. The great majority of studies have shown that out of the available number of recording channels, only a limited number carry information that can actually be used for the detection of a pre-seizure state, while the remaining channels are likely to increase the number of false detections without contributing to the detection sensitivity of an algorithm. The task at hand is to decide in advance which channels are best suited for the purpose. Several studies have attempted to tackle the problem of channel selection by using the first few seizures to select the appropriate channels and/or parameters for the algorithm before trying to detect precursors of the seizures that follow (D'Alessandro *et al.*, 2003, 2005; Esteller *et al.*, 2005; Le Van Quyen *et al.*, 2005). Such a procedure implies that the spatiotemporal dynamics preceding a seizure do not change from seizure to seizure. Iasemidis *et al.* (2003, 2005) designed an algorithm using a selection of channels that is readjusted after every seizure so that it would have been optimal for the seizure that has just occurred. Such a procedure is based on the implicit assumption that pre-ictal dynamics change to a certain degree from seizure to seizure, but the pre-ictal dynamics of a seizure still depend on the dynamics of the previous one. If these algorithms reliably proved to be better than a random prediction, they could, in addition to being beneficial for patients, provide valuable clues for new theories on the mechanisms involved in ictogenesis.

The need for statistical validation

If an algorithm is designed to run prospectively, its quasi-prospective out-of-sample performance can be tested retrospectively on continuous long-term recordings that were not previously used for parameter optimization or channel selection. Once this quasi-prospective performance (in terms of correct alarms and false alarms with respect to the given prediction horizon) has been assessed, it remains to be tested whether it is indeed superior to naïve prediction schemes such as periodic or random predictors. For this aim, researchers designed a framework to assess the performance of such a random predictor (Winterhalder *et al.*, 2003; Schelter *et al.*, 2006).

In retrospective statistical studies on predictability, however, it may be desirable to investigate and compare the potential predictive performance of different characterizing measures for various thresholds and parameters. In this case, the use of a random predictor for statistical validation would require corrections for multiple testing that can be difficult to perform since the data used for the different tests are usually not independent. Here, the concept of seizure time surrogates as introduced by Andrzejak *et al.* (2003) can provide a means for statistical validation. In this process, artificial seizure-onset times are generated by randomly shuffling the original inter-seizure intervals. Using these surrogate seizure-onset times instead of the original onset times, the EEG data are then subjected to the same algorithm or prediction statistics that was used for the original onset times. Only if the performance of the algorithm for the original seizure times is significantly better than the performance for a number of independent realizations of the surrogate seizure times, can the null hypothesis, namely, that a given algorithm cannot detect a pre-seizure state with a performance above chance level, be rejected. The advantage of this type of statistical validation is that it can be applied to any type of analysis, algorithmic or statistical. A modification of this surrogate test has recently been proposed on the basis of a constrained randomization of the time profile of the characterizing measure (Kreuz *et al.*, 2004).

Appendix C: Characterizing measures of the EEG

This appendix contains an overview of measures described in this review that are commonly used to characterize electroencephalographic time series including detailed mathematical descriptions. The described measures are typically calculated from EEG epochs of ~20 s in a moving window analysis (*see* Appendix B). Some measures may require pre-processing of EEG epochs such as demeaning or filtering.

Univariate linear measures

The information contained in consecutive amplitude values of a signal $x(t)$ that is sampled in the form of a discrete time series $x(t_i) = x(t_0 + i{\cdot}\Delta t) = x_i$ (with $i = 1, \ldots, N$ and Δt denoting the sampling interval) can also be encoded by amplitudes and phases of harmonic oscillations with a range of different frequencies. The map that translates between these representations in the time domain $\{x_i\}$ and the frequency domain $\{s_k\}$ is called Fourier transform *FT*. The periodogram of a real signal provides an estimate of the power spectrum (cf. Press *et al.*, 1992) and is given by the square of the amplitudes of the Fourier transform: $\{p_k\} = \{|s_k|^2\}$ with $k = 1, \ldots, N/2$ for any frequency $f = {}^{k}/_{N} \cdot f_s$, where $f_s = {}^{1}/_{\Delta t}$ is the sampling rate.

The total power of the time series is given by $P = \sum_{k=1}^{N/2} p_k = \sum_{f=0}^{f_s/2} p_f$. In the following, we assume that the time series' mean values were set to zero before analysis.

Statistical moments

Statistical moments characterize the amplitude distribution of a time series $\{x_i\}$. The second moment is the variance $\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N} x_i^2$, the third moment is the skewness $\chi = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i}{\sigma}\right)^3$ and the fourth moment is the kurtosis $\kappa = \left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i}{\sigma}\right)^4\right] - 3$. The skewness is zero for symmetric amplitude distributions and non-zero for asymmetric distributions. The kurtosis measures the relative peakedness or flatness of an amplitude distribution.

Spectral band power

Different physiological and pathological processes are reflected by activity in different frequency ranges of the

power spectrum $\{p_f\}$ of the EEG. According to these ranges, a set of power spectral bands (δ, ϑ, α, β, γ) were defined in classical EEG analysis. The relative power contained in these bands can be defined as

$$\delta_r = \frac{1}{P}\sum_{f=0.5\text{Hz}}^{4\text{Hz}} p_f; \quad \vartheta_r = \frac{1}{P}\sum_{f=4\text{Hz}}^{8\text{Hz}} p_f; \quad \alpha_r = \frac{1}{P}\sum_{f=8\text{Hz}}^{13\text{Hz}} p_f;$$
$$\beta_r = \frac{1}{P}\sum_{f=13\text{Hz}}^{30\text{Hz}} p_f; \quad \gamma_r = \frac{1}{P}\sum_{f=30\text{Hz}}^{100\text{Hz}} p_f,$$

where P is the total power of the signal.

Spectral edge frequency

In a typical EEG signal, most of the power is contained within the frequency band from 0 Hz up to 40 Hz: $P_{40\text{ Hz}} \approx P$. As a characterizing measure for the power distribution, the so-called spectral edge frequency can be used (Stanski *et al.*, 1984), which is defined as the minimum frequency up to which 50% the spectral power up to 40 Hz is contained in the signal:

$$f_{50} = \min\left\{ f^* \;\middle|\; \sum_{f=0\text{Hz}}^{f^*} p_f > P_{40\,\text{Hz}} \cdot 0.50 \right\}.$$

Accumulated energy

The accumulated energy (Litt *et al.*, 2001) is computed from EEG time series by integrating the broadband signal power (or energy) over a sequence of (possibly overlapping) windows in a moving window analysis. According to Parseval's theorem, the average power of a signal is given by the variance, so the accumulated energy for the t-th time window is obtained by cumulatively summing the variance of all past time windows:

$$\text{AE}(t) = \sum_{k=1}^{t} \sigma_k^2,$$

where σ_k^2 is the variance of the k-th time window.

Characteristics of the autocorrelation function

The autocorrelation function of a time series is defined as

$$A(\tau) = \frac{1}{(N-1)\sigma^2}\sum_{i=1}^{N-\tau} x_i x_{i-\tau}$$

for $\tau = 0, \ldots, N-1$ with σ^2 denoting the variance of the signal. By definition, $A(\tau)$ ranges between -1 and 1 with $A(0) = 1$. Provided that the time series is non-periodic, the autocorrelation function decays from $A(0)$ with increasing values of τ, and fluctuates around zero for larger τ-values. The slower $A(\tau)$ decays initially, the stronger are the linear correlations of the time series. Hence, an estimate of the strength of linear correlations can be defined using the first zero crossing

$$\tau_0 = \min\{\tau \mid A(\tau) = 0\}$$

of the autocorrelation function.

Hjorth parameters

Hjorth defined activity, mobility and complexity as 'a set of parameters intended as a clinically useful tool for the quantitative description of an EEG' (Hjorth, 1970). The activity is proportional to the variance of a signal. The mobility is defined as the variance of the slopes of the EEG normalized by the variance of the amplitude distribution of the time series. The complexity quantifies the variance of the rate of slope changes with reference to an ideal sine curve. In the frequency domain, the mobility and complexity can be estimated from the second and fourth statistical moment of the power spectrum:

$$\text{HM} = \sum_{k=1}^{N/2} p_k k^2 j \quad \text{HC} = \sum_{k=1}^{N/2} p_k k^4 j$$

Autoregressive modelling

The most general linear (univariate) model for a time series is the autoregressive moving average (ARMA) model. It is composed of three linear model processes: a purely random process (white noise), an autoregressive (AR) process and a moving average (MA) process. An AR process is defined by

$$x_i = \sum_{l=1}^{p} a_l x_{i-l} + \varepsilon_i$$

and indicates that the value of the time series at time point i is a linear combination of its p past values and a purely random process ε_i. In order to account for possible correlations in the noise, ε_i may be modelled equivalently as an MA process

$$\varepsilon_i = \sum_{l=1}^{q} b_l \varepsilon_{i-l},$$

indicating that the noise at time point i is a linear combination of its q past values. Hence an ARMA model reads

$$x_i = \sum_{l=1}^{p} a_l x_{i-l} + \sum_{k=1}^{q} b_k \varepsilon_{i-k},$$

where the coefficients $\{a_l\}$ and $\{b_k\}$ are to be determined by fitting the data, typically using a least-squares or an information-theoretic criterion. Identification of an appropriate ARMA model allows the design of special filters, forecasting time series or estimation of the power spectrum and derived measures such as the so-called transfer function (Makhoul, 1973; Lopes da Silva, 1987; Lopes da Silva and Mars, 1987). The evolution of different parameters [model order, coefficients, prediction error or characteristics

of the transfer function (Rogowski *et al.*, 1981) or of the power spectrum] over time can be used as measure profiles.

Univariate non-linear measures

While linear measures are calculated directly from the time series or its power spectrum, a number of non-linear measures have been derived from the theory of dynamical systems (Schuster, 1989; Ott, 1993; Kantz and Schreiber, 1997) that are designed to quantify different properties of so-called state space trajectories in a Cartesian space. Calculation of these measures therefore requires reconstruction of the state space trajectory from the scalar time series $\{x_i\}$, where $i = 1, \ldots, N$. This reconstruction (*time-delay embedding*) can be achieved by means of delay coordinates $\vec{x}_i = (x_i, x_{i+\tau}, \ldots, x_{i+(m-1)\tau})$ (Takens, 1981) with $i = 1, \ldots, M = N - (m - 1)\,\tau$, where $\{\vec{x}_i\}$ defines the reconstructed state space trajectory. Here τ is a time delay and m is the embedding dimension, which, according to Whitney's theorem, must be chosen as $m \geqslant 2\,d + 1$ (where d is the dimension of the geometrical object formed by the genuine trajectory in state space) if any exact determinism present in the original (multi-variate) system is to be preserved (Whitney, 1936). In case of simultaneous multi-channel EEG recordings, an alternative embedding scheme would be to use each channel as an axis of the Cartesian space (*spatial embedding*). In this case the embedding dimension m is fixed and equals the number of recording channels. Alternatively, a combination of time-delay and spatial embedding can be used.

Measures based on the correlation sum

The correlation sum (Grassberger and Procaccia, 1983*a*) is an estimate of the local probability density in state space (also referred to as correlation integral). It counts the number of pairs of vectors in state space that are closer than a given hypersphere radius ε:

$$C(\varepsilon) = \frac{2}{(M-W)(M-W-1)} \sum_{i=1}^{M} \sum_{j=i+W}^{M} \Theta(\varepsilon - \|\vec{x}_i - \vec{x}_j\|),$$

where $\|\cdot\|$ indicates some norm (e.g. the maximum norm) in m dimensions and Θ is the Heaviside step function ($\Theta(a) = 0$ for $a \leqslant 0$ and $\Theta(a) = 1$ for $a > 0$). The exclusion of pairs closer in time than the length of the so-called Theiler window W is essential to reduce the unwanted influence of temporal correlations on $C(\varepsilon)$ (Theiler, 1986).

Correlation dimension

For deterministic dynamics the correlation dimension (Grassberger and Procaccia, 1983*a*) allows to estimate the number of active degrees of freedom. From the local slope of the correlation sum

$$d(\varepsilon) = \frac{d \ln C(\varepsilon)}{d \ln \varepsilon},$$

the correlation dimension is defined as

$$D_2 = \lim_{N\to\infty} \lim_{\varepsilon\to 0} d(\varepsilon).$$

From the limits it follows that the calculation of the correlation dimension would require an infinite length N and an unlimited accuracy of the time series. However, an estimate of an *effective* correlation dimension (Grassberger *et al.*, 1991; *see* also Lehnertz and Elger, 1998) can be obtained if an almost constant value of $d(\varepsilon)$ is found at least for a limited range of ε values, the so-called quasi-scaling region.

Correlation density

The correlation density is defined as the correlation sum for some fixed hypersphere radius $\varepsilon = \varepsilon_0$. Martinerie *et al.* (1998) calculated this measure using a combination of time-delay and spatial embedding of EEG time series. In the literature (Cerf *et al.*, 2000, 2004) the correlation density is also referred to as Lerner density (Lerner, 1996).

Correlation entropy

The correlation entropy h_2 (Grassberger and Procaccia, 1983*b*) is a lower bound of the Kolmogorov–Sinai entropy, which describes the level of uncertainty about the future state of the system, and therefore relates to predictability. Provided a scaling region exists, h_2 can be estimated from the correlation sum as

$$h_2 \approx \ln \frac{C_m(\varepsilon)}{C_{m+1}(\varepsilon)},$$

using an extrapolation to large embedding dimensions m. Alternatively, an entropy estimate can be derived from the sum of the positive Lyapunov exponents (Pesin's identity).

Marginal predictability

On the basis of the correlation sum for different embedding dimensions m, Savit and Green (1991) defined predictability as

$$S_m = \frac{C_{m+1}}{C_m}.$$

As a more sensitive discriminator of non-linear structure in time series, Manuca and Savit (1996) proposed the ratio of successive S_m values, defined as

$$R_m = \frac{S_m}{S_{m-1}} = \frac{C_{m+1} C_{m-1}}{C_m^2}.$$

Li *et al.* (2003) and Drury *et al.* (2003) defined marginal predictability as

$$\delta_m = \frac{R_m - 1}{R_m}$$

with the correlation sum C_m estimated for some fixed hypersphere radius $\varepsilon = \varepsilon_0$.

Dynamical similarity index

This measure was designed by Le Van Quyen *et al.* (1999) to measure the dynamical similarity between a running test window and a reference period, usually selected from the beginning of a recording and with a length of l times the length of the test window. Before analysis, the signals from every channel are transformed from the amplitude domain into the domain of inter-event intervals. An event is defined as the crossing from negative to positive amplitude values of the original time series, and the sequence of intervals between the times of subsequent crossings is used as signal representation in the inter-event-interval domain: $I_i = T_{i+1} - T_i$ for $i = 1, \ldots, N^* - 1$, where N^* is the total number of events, which generally varies for different windows. From the inter-event-interval representation the dynamics are reconstructed using conventional delay coordinates: $\vec{a}_i = (I_i, I_{i+\tau}, \ldots, I_{i+(m-1)\tau})$ for $i = 1, \ldots, N^* - (m-1)\tau$ using some high value of the embedding dimension m.

The reconstructed data from the reference period are then transformed and projected into a reduced state space spanned by the first $\tilde{m} < m$ principal components obtained from a singular value decomposition: $\tilde{\vec{a}}_i = A\vec{a}_i$ (Broomhead and King, 1986). The same linear transformation matrix A is subsequently applied to the embedded state space vectors of every running test window yielding $\{\vec{x}^t\}_{k=1,\ldots,N_t}$, where N_t denotes the number of transformed state vectors in the test window t. In order to compare the dynamics of the reference periods with that of the test window, the transformed state vectors from the reference period are down-sampled to a random subset $\{\vec{y}_i\}_{i=1,\ldots,N_r} \subset \{\tilde{\vec{a}}_i\}_{i=1,\ldots,N^*}$ with $N_r = N^* \operatorname{div} l$ such that the resulting number of state vectors N_r corresponds approximately to the different values of N_t.

The dynamical similarity index between the reference period r and a test window t is then defined as

$$\gamma_t = \frac{C_{rt}}{\sqrt{C_{rr}C_{tt}}},$$

where C_{rt} is the cross-correlation sum (Kantz, 1994*a*) given by

$$C_{rt} = \frac{1}{N_r N_t} \sum_{k=1}^{N_t} \sum_{i=1}^{N_r} \Theta(\varepsilon - \|\vec{y}_i - \vec{x}_k^t\|),$$

and C_{rr} and C_{tt} are the (auto-)correlation sums of the reference and test window, respectively:

$$C_{tt} = \frac{1}{N_t^2} \sum_{k=1}^{N_t} \sum_{i=1}^{N_t} \Theta(\varepsilon - \|\vec{x}_i^t - \vec{x}_k^t\|) \quad \text{and}$$

$$C_{rr} = \frac{1}{N_t^2} \sum_{k=1}^{N_r} \sum_{i=1}^{N_r} \Theta(\varepsilon - \|\vec{y}_i - \vec{y}_k\|).$$

State space dissimilarity measures

Since the correlation sum estimates the local probability density in state space, it provides a convenient way to measure dissimilarity between two EEG time series. In a statistical sense, one can quantify dissimilarity from the inconsistency of two samples with the same distribution. This can be based on the well known χ^2-test or on distance measures such as the L_1 norm (Hively *et al.*, 2000; Hively and Protopopescu, 2003):

$$\chi^2 = \sum_i \frac{(Q_i - R_i)^2}{Q_i + R_i},$$

$$L_1 = \sum_i |Q_i - R_i|,$$

where Q_i and R_i denote local probability estimates in state space based on the correlation sum for a running test window and a reference period, respectively.

Largest Lyapunov exponent

The exponential divergence of nearby trajectories in state space is conceptually the most basic indicator of deterministic chaos and can be estimated using the largest Lyapunov exponent $L_{\max}$. The first proposed algorithm to compute $L_{\max}$ from a time series (Wolf *et al.*, 1985) suffers from severe drawbacks that occur particularly with short and noisy time series. Moreover, it strongly depends on parameters used for the state space reconstruction and is computationally highly expensive (Rosenstein *et al.*, 1993). In order to avoid these shortcomings, a combination of improved algorithms can be used (Rosenstein *et al.*, 1993; Kantz, 1994*b*) according to which the $L_{\max}$ can be estimated from

$$d_j(i) \approx C_j e^{L_{\max} i \Delta t},$$

where $d_j(i)$ denotes the average divergence between two trajectory segments at time t_i. C_j with $j = 1, \ldots, M$ is a constant that is given by the initial separation of a reference vector $\vec{z}_j$ in state space and its nearest neighbour. In order to improve statistics Kantz (1994*b*) proposed to search for *all* neighbours starting within a hypersphere of radius ε around $\vec{z}_j$ using a box-assisted algorithm (Schreiber, 1995). On the basis of the relation

$$\ln d_j(i) \approx \ln C_j + L_{\max} \cdot i \cdot \Delta t$$

the largest Lyapunov exponent is then calculated using a least-squares fit to an average line defined by $y(i) = \frac{1}{\Delta t} \langle \ln d_j(i) \rangle$, where $\langle \ldots \rangle$ denotes the average over j.

In order to reduce the unwanted influence of temporal correlations Rosenstein *et al.* (1993) suggested to choose a Theiler window of a length given by the reciprocal of the mean frequency of the power spectrum.

Local flow Λ^*

The local flow, a measure derived from the coarse-grained flow average (Kaplan and Glass, 1992), aims at discriminating deterministic from stochastic dynamics. For this technique the reconstructed m-dimensional state space is divided into b^m non-overlapping hyper-cubes. If the hyper-cube with index j is passed n_j times by the trajectory, a normalized

vector $v_{j,k}$ will be generated for each pass ($k = 1, \ldots, n_j$) whose direction is determined by connecting the points where the trajectory enters and leaves the hyper-cube. Summing up all vectors of passes through hyper-cube j, the resultant vector V_j, normalized by the number of passes n_j, is $V_j = \frac{1}{n_j} \sum_{k=1}^{n_j} v_{j,k}$. The coarse-grained flow average Λ is then defined as

$$\Lambda = \sum_j \frac{V_j^2 - R^2}{1 - R^2}$$

with $R \propto \frac{1}{\sqrt{n}}$ being the expected value for a vector addition of n vectors of unit length yielded by a random walk in m dimensions.

Rather than using a fixed time delay τ, Andrzejak *et al.* (2001) determined the local flow by summing up the coarse-grained flow average for different values of τ:

$$\Lambda^* = \sum_{\tau=\tau_{\min}}^{\tau_{\max}} \Lambda(\tau).$$

Algorithmic complexity

Another approach to characterize time series is based on the theory of symbolic dynamics (Hao, 1989). For this approach, the time series is transformed into a sequence of A symbols by partitioning the range of sampling values and assigning a different symbol S to each interval of this binning. Then each value of the time series is replaced by the symbol of its interval. The thresholds of the partition are chosen separately for each time series to yield a homogeneous distribution of symbols. The resulting symbol sequence $\{S_i\}$ with $i = 1, \ldots, N$ is then investigated for its complexity by estimating the size $c(\{S_i\})$ of its vocabulary. This size is defined as the number of different words in a Lempel–Ziv parsing (Lempel and Ziv, 1976) of the symbol sequence. In this algorithm the symbol sequence is scanned from the beginning to its end, and its complexity $c(\{S_i\})$ is increased by one unit as soon as a new subsequence of consecutive symbols is encountered in the scanning process (Kasper and Schuster, 1987), and the following symbol is regarded as the beginning of the next symbol sequence. This value is normalized by the expected asymptotic value for a random sequence of symbols of length N to yield the algorithmic complexity:

$$\mathrm{AC} = \frac{\log_A N}{N} \cdot c(\{S_i\}).$$

Surrogate time series and surrogate correction

The method of surrogate time series allows to test a specified null hypothesis about the dynamics underlying a given time series (for an overview, *see* Schreiber and Schmitz, 2000). For this purpose, an ensemble of surrogate time series is constructed from the original time series in such a way that the surrogates have all properties included in the null hypothesis in common with the original, but are otherwise random. Then a certain measure, which has to be sensitive to at least one property that is not included in the null hypothesis, for example, non-linearity, is calculated for the original and the surrogates. If the result for the original time series deviates significantly from the distribution of the surrogates, the null hypothesis can be rejected. The probability of false rejections, that is, the nominal size of the test, is adjustable by the number of surrogates. A common type of surrogates are iterative amplitude-adjusted surrogates (Schreiber and Schmitz, 1996). These types of surrogates allow testing of the null hypothesis that the time series were measured from a Gaussian linear stochastic and stationary dynamics by means of a static and invertible but possibly non-linear measurement function. Starting from a random permutation of the original amplitudes of the time series, the surrogates are constructed by an iterative algorithm that alternately adjusts the power spectrum and the amplitude distribution to the original values, resulting in a deviation of the respective other quantity. After a sufficient number of iterations (typically 20–50), deviations of both quantities from values of the original time series will be reduced to negligibly small values.

In Andrzejak *et al.* (2001) and Mormann *et al.* (2005) surrogate-corrected measures are defined as

$${}^{S}NM = |NM_{EEG} - \overline{NM_{SUR}}|$$

where NM is a placeholder for any non-linear univariate measure and the over-bar denotes average over several surrogates. The absolute value is taken to ensure that indications of non-linear structure in a time series are always reflected by an increase in values of the surrogate-corrected measures regardless of whether this structure is reflected by higher (as for Λ^*, AC) or lower (as for D_2, $L_{\max}$) values of the respective non-linear measure. It may be noted that this use of surrogates differs a little from their original purpose since taking these differences can be regarded as an 'offset correction' rather than as a hypothesis test, with the offset given by linear properties of the dynamics.

Loss of recurrence

The loss of recurrence can be used to quantify the degree of non-stationarity within a time series (Rieke *et al.*, 2002, 2004). This measure analyses the distribution of distances in time between reference vectors and their neighbouring vectors in state space. A system is regarded as stationary if the time index of a neighbour is statistically independent from that of the reference. For non-stationary systems, the absence of distant time indices in the neighbourhood of the reference, that is, a loss of recurrence is expected.

Let $U_\varepsilon(\vec{x}_r) = \{\vec{x}_n : \|\vec{x}_r - \vec{x}_n\| \leqslant \varepsilon\}$ define a set of vectors in the ε-neighbourhood of $\vec{x}_r$ in an m-dimensional reconstructed state space. ε_r is defined in dependence of the reference $\vec{x}_r$ using a fixed number of k nearest neighbours and the maximum norm as a metric. The lags

$l_r^i = |n_r^i - r|$ of the i-th nearest neighbour of $\vec{x}_r$ are transformed using the distribution function $\tilde{l} = \Phi_r(l)$, that is, the *a priori* probability under stationary conditions that the observed distance in time is less than or equal to l. The distribution $f(\tilde{l})$ of all transformed time distances $\tilde{l}_n$ reflects the non-stationarity of the system. For a stationary system $f(\tilde{l})$ is uniformly distributed in the interval [0,1], and the median μ equals 0.5, whereas in case of non-stationarity, the recurrence of related state space vectors is reduced, that is, the neighbourhood of $\vec{x}_r$ depends on the time index r and furthermore the indices of the neighbouring vectors n_r^i are clustered around r. For non-stationary signals, the observed distances in time are therefore on average smaller than expected and thus the amount of lower values $\tilde{l}$ is increased, whereas higher values are reduced, and the median μ of this distribution $f(\tilde{l})$ is <0.5. There is no need for a surrogate correction of this measure as surrogate time series are stationary by construction and the median μ of the distribution $f(\tilde{l})$ always matches 0.5 up to statistical fluctuations.

Bivariate linear measures

Maximum linear cross-correlation

In order to quantify the similarity of two signals $\{x_i\}$ and $\{y_i\}$ the maximum of a normalized cross-correlation function can be used as a measure for lag synchronization (Rosenblum *et al.*, 1997):

$$C_{\max} = \max_{\tau}\left\{\left|\frac{C_{xy}(\tau)}{\sqrt{C_{xx}(0)\cdot C_{yy}(0)}}\right|\right\}$$

where

$$C_{xy}(\tau) = \begin{cases} \dfrac{1}{N-\tau}\displaystyle\sum_{i=1}^{N-\tau} x_{i+\tau}y_i & \tau \geqslant 0 \\ C_{yx}(-\tau) & \tau < 0 \end{cases}$$

is the well-known linear cross-correlation function. $C_{\max}$ is confined to the interval [0, 1] with high values indicating that the two signals have a similar course in time (though possibly shifted by a time lag τ) while dissimilar signals will result in values close to zero.

Linear coherence

The coherence function measures the linear synchronization between two signals $\{x_i\}$ and $\{y_i\}$ for a given frequency f and is defined as

$$\Gamma(f) = \left|\frac{G_{xy}(f)}{\sqrt{G_{xx}(f)\cdot G_{yy}(f)}}\right|$$

where

$$G_{xy}(f) = \mathrm{FT}[x](f)\cdot \mathrm{FT}[y]^*(f)$$

is the sample cross-spectrum, with FT denoting the Fourier transform and the asterisk denoting complex conjugation. The coherence function also ranges between 0 (low coherence) and 1 (high coherence) and is useful when synchronization is limited to some particular frequency band, as it is usually the case in EEG signals (*see* Quian Quiroga *et al.*, 2002).

Bivariate non-linear measures

Non-linear interdependence

The non-linear interdependence (Arnhold *et al.*, 1999) as a measure for generalized synchronization (Rulkov *et al.*, 1995) between two EEG signals $\{x_i\}$ and $\{y_i\}$ is calculated after reconstruction of the state space trajectories $\{\vec{x}_i\}$ and $\{\vec{y}_i\}$ for these signals. Let α_{ij} and β_{ij} with $j = 1, \ldots, k$ denote the time indices of the k nearest neighbours in state space of $\vec{x}_i$ and $\vec{y}_i$, respectively. For each $\vec{x}_i$ the squared mean Euclidean distance to its k nearest neighbours is given by

$${}^{x}R_i^{(k)} = \frac{1}{k}\sum_{j=1}^{k}(\vec{x}_i - \vec{x}_{\alpha_{ij}})^2$$

while the *y-conditioned* mean-squared Euclidean distance is constructed by replacing the nearest neighbours by the equal time partners of the closest neighbours of $\vec{y}_i$:

$${}^{x|y}R_i^{(k)} = \frac{1}{k}\sum_{j=1}^{k}(\vec{x}_i - \vec{x}_{\beta_{ij}})^2.$$

${}^{y}R_i^{(k)}$ and ${}^{x|y}R_i^{(k)}$ are defined accordingly.

As measures for non-linear interdependence Arnhold *et al.* (1999) proposed

$${}^{x|y}S = \frac{1}{M}\sum_{i=1}^{M}\frac{{}^{x}R_i^{(k)}}{{}^{x|y}R_i^{(k)}}$$

and

$${}^{x|y}H = \frac{1}{M}\sum_{i=1}^{M}\log\frac{{}^{x}R_i^{(M)}}{{}^{x|y}R_i^{(k)}}$$

with

$${}^{x}R_i^{(M)} = \frac{1}{M-1}\sum_{j=1, j\neq i}^{M}(\vec{x}_i - \vec{x}_j)^2.$$

${}^{y|x}S$ and ${}^{y|x}H$ are defined accordingly. Both measures yield high values for high degrees of non-linear interdependence and low values for independent time series. While ${}^{x|y}S$ is restricted to the interval [0, 1], ${}^{x|y}H$ is not normalized and might also have slightly negative values.

Dynamical entrainment

Iasemidis *et al.* (2001) defined a measure they termed 'entrainment between two brain regions' as the statistical difference between the largest Lyapunov exponents $L_{\max}$ (*see* section Univariate non-linear measures) over a number

of l consecutive time windows for two signals recorded from electrode sites x and y by using the T-index derived from a paired t-test for comparison of means:

$$T_{xy} = \sqrt{l}\frac{|\langle L_{\max,x} - L_{\max,y}\rangle|}{\sigma_{xy}},$$

with $\langle \ldots \rangle$ denoting the mean over l and σ_{xy} the corresponding standard deviation.

According to their interpretation of this measure, a low T-index corresponds to a high entrainment and vice versa.

By minimizing the function

$$f(\vec{c}) = \vec{c}^{t} A \vec{c}$$

$$\text{with } \vec{c} \in \{0,1\}^n \text{ and } \sum_{i=1}^{n} c_i = k \text{ and } A = (T_{xy})_{x,y=1,\ldots,n}$$

where $\vec{c}^{t}$ denotes the transposed vector $\vec{c}$, they selected those k electrode sites out of a total of n that showed the highest mutual entrainment.

Measures for phase synchronization

Phase synchronization (Huygens, 1673) is traditionally defined as *phase locking* [$\phi_x(t) - \phi_y(t) = \text{const}$] or, in the case of noisy and/or chaotic systems (Rosenblum *et al.*, 1996), as *phase entrainment* [$|\phi_x(t) - \phi_y(t)| < \text{const}$], with $\phi_x(t)$ and $\phi_y(t)$ denoting the phase variables of two oscillating signals $x(t)$ and $y(t)$.

Three different measures for phase synchronization have been proposed. The first measure, the *mean phase coherence* (Mormann *et al.*, 2000), is defined as

$$R = \left| \frac{1}{N} \sum_{j=1}^{N} e^{i[\phi_x(t_j) - \phi_y(t_j)]} \right|.$$

The second and third measure are termed *index based on conditional probability* and *index based on Shannon entropy* (Tass *et al.*, 1998). For these measures, an equidistant binning of the interval $[0, 2\pi]$ is required where the number of bins is given by $L = e^{0.626 + 0.4\ln(N-1)}$ as in Rosenblum *et al.* (2001).

The index based on conditional probability is then defined as

$$\lambda_{cp} = \frac{1}{L} \sum_{l=1}^{L} |r_l|$$

where

$$r_l = \frac{1}{M_l} \sum_{\substack{j \\ \phi_x(t_j) \in [\frac{l}{L}2\pi, \frac{l+1}{L}2\pi]}} e^{i\phi_y(t_j)}$$

with

$$M_l = \left| \left\{ \phi_x(t_j) \mid \phi_x(t_j) \in \left[\frac{l}{L}2\pi, \frac{l+1}{L}2\pi \right] \right\} \right|$$

denoting the number of phase values $\phi_x(t)$ contained in the bin l. ($|\{\ldots\}|$ denotes the number of elements contained in the set $\{\ldots\}$.)

The index based on Shannon entropy is given by

$$\rho_{se} = 1 + \frac{1}{\ln L} \sum_{l=1}^{L} p_l \ln p_l$$

with

$$p_l = \frac{|\{\phi_x(t_j) - \phi_y(t_j) \mid \phi_x(t_j) - \phi_y(t_j) \in [\frac{l}{L}2\pi, \frac{l+1}{L}2\pi]\}|}{|\{\phi_x(t_j) - \phi_y(t_j)\}|}$$

denoting the relative frequency of finding a phase difference in a certain bin l.

All three phase-synchronization measures are confined to the interval [0, 1] where high values indicate a high degree of phase synchronization and low values correspond to unsynchronized signals.

In order to measure changes in phase synchronization of two signals $x(t)$ and $y(t)$ over time, it is first of all necessary to determine their phases $\phi_x(t)$ and $\phi_y(t)$. To this aim, two different approaches are frequently used:

One is the *analytic signal* approach (Gabor, 1946; Panter, 1965), which defines an *instantaneous phase*

$$\phi(t) = \arctan \frac{\tilde{s}(t)}{s(t)}$$

for an arbitrary signal $s(t)$ using the *Hilbert transform*

$$\tilde{s}(t) = \frac{1}{\pi} p.v. \int_{-\infty}^{+\infty} \frac{s(t')}{t - t'} dt'$$

(*p.v.* denoting the Cauchy principal value).

This phase definition can be used for broad-band synchronization analysis or after band-pass filtering of the original signals.

The second approach (Lachaux *et al.*, 1999) uses a definition based on the Wavelet transform, yielding a band-specific synchronization index. Here the phase variable is defined as

$$\phi(t) = \arctan \frac{\mathrm{Im} W(t)}{\mathrm{Re} W(t)}$$

using the Wavelet coefficients

$$W(t) = \int_{-\infty}^{+\infty} \Psi(t - t') s(t') dt'$$

of a complex Morlet Wavelet

$$\Psi(t) = (e^{if_0 t} - e^{-f_0^2 \sigma^2/2}) \cdot e^{-t^2/2\sigma^2},$$

where f_0 is the centre frequency and σ the decay rate of the wavelet that governs the width of the frequency band centred around f_0.