AUTOMATED MARKERS FOR ENHANCED EPILEPSY DIAGNOSIS, TREATMENT, AND SAFETY MONITORING

Evelien Geertsema

AUTOMATED MARKERS FOR ENHANCED EPILEPSY DIAGNOSIS, TREATMENT, AND SAFETY MONITORING

ISBN	978-94-028-1276-3
PRINTED BY	Ipskamp Printing, Enschede
COVER ILLUSTRATION	Victorus
LAYOUT & DESIGN	Evelien Geertsema

The research in this thesis was financially supported by Margaret Knip Fonds.

The author gratefully acknowledges financial support for the printing of this thesis by Stichting Epilepsie Instellingen Nederland (SEIN), UMC Utrecht Brain Center Rudolf Magnus, UCB Pharma B.V., and Chipsoft B.V.

AUTOMATED MARKERS FOR ENHANCED EPILEPSY DIAGNOSIS, TREATMENT, AND SAFETY MONITORING

AUTOMATISCHE MARKERS VOOR VERBETERDE EPILEPSIE DIAGNOSE, BEHANDELING EN VEILIGHEIDSMONITORING

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
donderdag 20 december 2018 des middags te 4.15 uur

door

Evelien Elisabeth Geertsema

geboren op 4 juli 1988 te Hengelo

Promotoren:	Prof.dr.ir. M.A. Viergever Prof.dr. J.W. Sander
Copromotoren:	Dr. G.H. Visser Dr. S.N. Kalitzin

CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

INTRODUCTION

EPILEPSY

Epilepsy is the most common serious neurological disorder, affecting up to 1.0% of the population at any given time.[1] It is a chronic brain disorder characterized by recurrent epileptic seizures, defined as a transient occurrence of symptoms due to excessive or synchronous neuronal activity.[2] Seizure manifestations can range from short involuntary muscle twitches to generalized convulsions with loss of consciousness, depending on the brain regions involved.[3] Seizures can be very disabling and are often unpredictable, resulting in stress, social isolation, and decreased quality of life.[4]

The diagnosis of epilepsy relies solely on medical history, obtained from the patient and from witnesses. Home videos of the events are routinely requested when diagnosis is uncertain. Laboratory procedures such as EEG, MRI, blood tests and genetic testing can support a diagnosis, determine seizure type and syndrome, and establish etiology.[5]

Approximately two thirds of people with epilepsy can be treated effectively with (a combination of) anti-epileptic drugs.[6] The remaining have refractory epilepsy, with seizures which are not controlled by anti-epileptic medication. In cases of refractory epilepsy, other treatment modalities such as surgical treatment, neurostimulation or diets maybe be considered.[7]

DIAGNOSTICS IN THE EPILEPSY MONITORING UNIT

If a diagnosis or disease etiology cannot be ascertained via (hetero-) anamnesis, complemented with a short EEG recording and/or an MRI, long-term EEG-video recording may be performed. Such recordings are made in an epilepsy monitoring unit (EMU), where a patient stays for 24 hours up to 1-2 weeks, depending on the objective. Recordings are used to determine seizure type, to examine therapeutic options (e.g., surgery), or to distinguish epilepsy from non-epileptic seizures.[8]

Patient safety and quality of diagnostics in the EMU depends on the staff response to seizures. In most EMUs, staff are continuously present to monitor patients. When seizures are detected, EMU staff attend the person to reduce the risk of or treat injuries resulting from seizures.[9] Additionally, staff performs standardized tests to assess consciousness and neurological functioning during seizures, aimed at determining seizure semiology and

type.[10] Fast detection of seizures upon occurrence may improve safety and aid diagnosis in the EMU.

EPILEPSY SURGERY AND BIOMARKERS TO GUIDE RESECTION

Epilepsy surgery may be considered if seizure control could not be attained with multiple trials of antiepileptic drugs, and the person has a focal epileptic syndrome which is remediable by surgery. Surgery is currently the only curative treatment option for refractory focal epilepsy. Early surgical intervention may also prevent harmful effects of continuing seizures, notably changes in brain network topology,[11] as well as of prolonged use of antiepileptic drugs.[12] The goal of epilepsy surgery is to remove the epileptogenic zone, or to disconnect this zone from the brain network.[13] While epilepsy surgery is often successful, about a third of patients having had a resection do not achieve seizure freedom.[14]

There is need for biomarkers to delineate the epileptogenic tissue to be resected to assure seizure freedom post-surgically. Even in the presence of an epileptogenic lesion on MRI the resection boundaries are not always clear, e.g. in cases of neuronal tumors, in which the epileptogenic zone may extend beyond the lesion, or when the hippocampus is secondarily affected.[15] Ictal (i.e. seizures) and interictal events (e.g., epileptic spikes or high-frequency oscillations) found in noninvasive and sometimes invasive electrophysiological recordings can be used to delineate epileptogenic tissue before surgery. An intra-operative electrocorticogram (ECoG), in which epileptogenic tissue is delineated by showing areas with epileptiform activity, may help increase surgery success rate by enabling tailored resection. Recordings can be made quickly and can provide direct feedback on the adequateness of resection during surgery. Currently resection tailoring is done by identifying spikes in the ECoG, but this method is controversial.[16,17] High-frequency oscillations (HFOs), especially fast ripples (250-500 Hz), are likely more specific than spikes and may thus be more suited to the tailoring of epilepsy surgery.[17] Surgical removal of cortex showing fast ripples has been associated with post-operative seizure freedom.[18]

SAFETY MONITORING OF HIGH-RISK EVENTS

Up to a third of people with refractory epilepsy keep having seizures regularly because of unsatisfactory seizure control by medication, neurostimulation or surgical treatment,[19] which puts them at risk. People with epilepsy have a

higher probability to die prematurely, compared with the general population (more than double in chronic epilepsy).[20,21] Sudden unexpected death in epilepsy (SUDEP) is usually a seizure-related event and convulsive seizures are considered an important risk factor.[22–24] After major seizures the person is often incapacitated or in need of first aid owing to (non-life-threatening) injury, but not able to alert anyone.

Monitoring people who are at risk because of their seizures could increase their safety.[25,26] When a caregiver is alerted to the occurrence of high-risk events that can result from epileptic seizures, care can be provided. Convulsive seizures for example often leave subjects incapacitated or injured and SUDEP is often preceded by this type of seizure. Interventions such as repositioning the subject, stimulation, or clearing of the airway may have a protective effect in preventing SUDEP.[27] Some types of seizures can also cause high-impact falls which may lead to serious injury, requiring first aid.[28] And while apneas are self-limiting when occurring during seizures, they may, when occurring after a seizure, lead to asystole and SUDEP.[29]

Continuous real-time video monitoring is sometimes performed manually to detect high-risk events, but this is very time-consuming and privacy-sensitive, whilst dangerous seizures or after-seizure effects may be missed. Automated seizure detection devices can help alert caregivers to seizures. The currently available devices are, however, not suitable for all patient groups: Devices attached to the bed are only usable when patients are in bed, and wearable detection devices are not tolerated by some patient groups, like children or people with intellectual disability. An alternative solution is remote detection, using sensors that can cover a living area. Video cameras can be used as sensitive, versatile and relatively cheap remote sensors to quantify movement, and automated online analysis may enable remote detection of high-risk events.

DATA STREAM MONITORING AND AUTOMATED MARKERS

In specialized epilepsy care, data streams are often monitored and analyzed real-time by a human observer to detect events. Such events can be ictal or interictal transient signals produced by the brain and specific for epilepsy. The events usually occur unexpectedly. Knowledge about these events can aid diagnosis or direct treatment, and can also indicate the need of immediate assistance. It might be of vital importance that these events are noted directly upon occurrence.

Event monitoring may take place in epilepsy monitoring units, in operating rooms, and at homes of people with epilepsy. In the EMU, seizures need to be detected by staff so they can go to the patient to reduce risks arising from seizures (e.g., treat injuries) and to perform testing during the seizure to aid diagnosis. In the operating room, the electrocorticogram is monitored for spikes or high-frequency oscillations when a tailored surgical epilepsy resection is performed. The locations and frequencies of these markers help delineate the area of brain tissue that needs to be resected for the patient to become seizure free. In homes of people with epilepsy, monitoring for (results of) dangerous seizures can improve safety by indicating whether someone is in need of assistance and at risk of sudden death.

For events to be detected, expert human observers need to inspect online – often multiple – data streams (e.g., EEG channels or video streams). During monitoring, the observers have to be continuously vigilant for occurring events of interest. This visual observation is time-consuming, subjective, and sensitive to distractions. Human brains are incapable of processing all incoming information in parallel.[30] The attention we use to select the important information to detect events is limited in capacity,[31] over time as well as spatially.[32] Consequentially, important events could be missed, which might impair safety and the quality of diagnosis and treatment.

Automated markers have the potential to solve these issues. Automated markers might detect occurrences and characteristics (e.g., timing and location) of events in the data streams, and in addition might be able to identify data streams (e.g., epochs on specific channels) that are likely to contain events.

OUTLINE OF THIS THESIS

The aim of this thesis is to improve situations of real-time data monitoring for event detection in epilepsy, by constructing and validating automated algorithms to detect markers of epilepsy. The algorithms should detect seizures in the EMU, delineate epileptogenic tissue during epilepsy surgery, and detect high-risk (consequences of) seizures in patient's homes. For each of these markers the starting point for our work is different according to the state of the art in the specific field. Our additions to the respective fields range from algorithm construction to validation and investigation of the added value of algorithms when applying them in real life.

PART I: AUTOMATED MARKERS TO ENHANCE DIAGNOSIS

Online seizure detection algorithms might help staff on the EMU to detect seizures that are otherwise missed or recognized late. Detection performance of such algorithms has been studied thoroughly, showing good sensitivity and low false alarm rates.[33] Knowing about the sensitivity and specificity of algorithms alone is, however, not sufficient when considering implementation. In most EMUs such algorithms would be used in addition to staff, who may also detect seizures. The goal of automated detection is not substitution of EMU staff by automated detection, but to aid the detection of seizures. In **Chapter 2** we investigate the added value of applying algorithms for online seizure detection in the EEG, in terms of extra detected seizures or faster detection time compared to staff.

PART II: AUTOMATED MARKERS TO ENHANCE TREATMENT

During surgery, an automated algorithm is needed that can independently, reliably, and reproducibly delineate the epileptogenic zone to be resected. The algorithm needs to be able to function in real time, and should be insensitive to artefacts. In **Chapter 3** we present a novel *autoregressive model residual variation (ARR)* algorithm, aiming to automatically delineate the epileptogenic zone. In **Chapter 4** we adjust the ARR algorithm to reduce the influence of typical intra-operative artefacts. We also test the potential of the new algorithm to identify epileptogenic tissue during surgery.

PART III: AUTOMATED MARKERS TO ENHANCE SAFETY MONITORING

We aim to monitor people at risk because of their seizures by designing a system with three modules that can detect convulsive seizures, falls, and apneas. We use video cameras as sensitive, versatile and relatively cheap remote sensors that can quantify movement patterns of interest throughout the home.

Previously, our research group presented an algorithm aiming to discern convulsive seizures from normal behavior in video recordings.[34] A detection threshold had, however, not yet been established and the detection performance had not been determined in a real-life setting. This is required

to show the algorithm's usefulness in daily practice and to provide guidelines to enable its use. In **Chapter 5** we establish performance of a non-contact convulsive seizure detection algorithm, by determining a detection threshold and by investigating detection performance as a function of several variables.

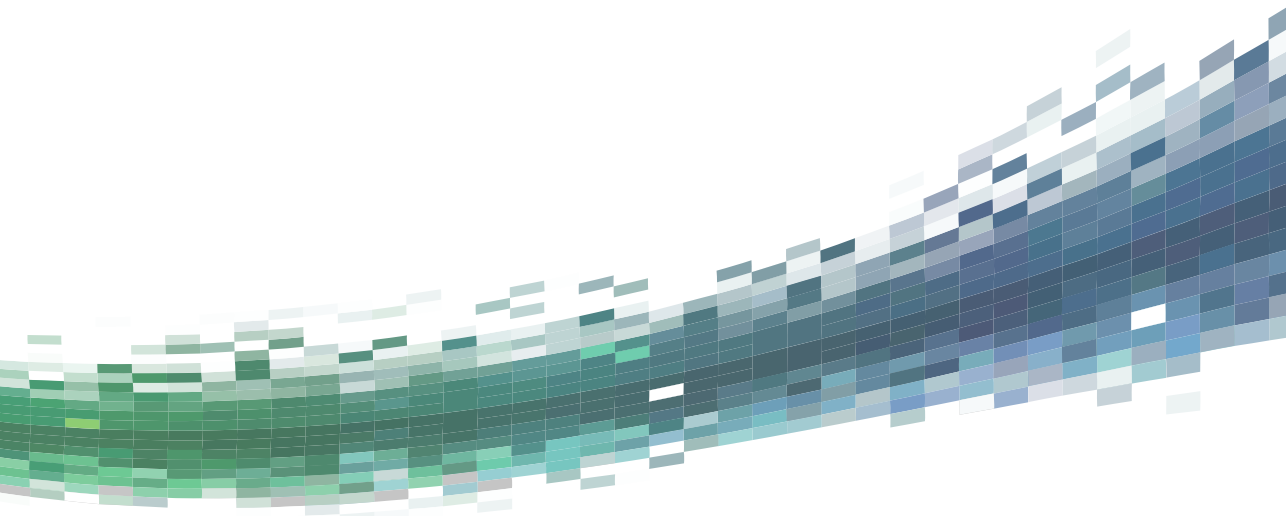
The few papers written on fall and apnea detection in video sequences do not show sufficient detection performance of current algorithms for our clinical application, in which fast detection, high detection sensitivity, and low false positive rates are needed. Furthermore, to keep computing time of the 3-module detection system short, it would be beneficial to have the same preprocessing step for each module. In **Chapter 6** we present an automated algorithm for remote detection of falls, based on a physical model of a fall, aiming at universality and robustness. In **Chapter 7** we present a novel algorithm for real-time detection of apnea events in video, aiming at fast detection when the subject is immobile.

In **Chapter 8** the findings of this thesis are summarized and discussed.

REFERENCES

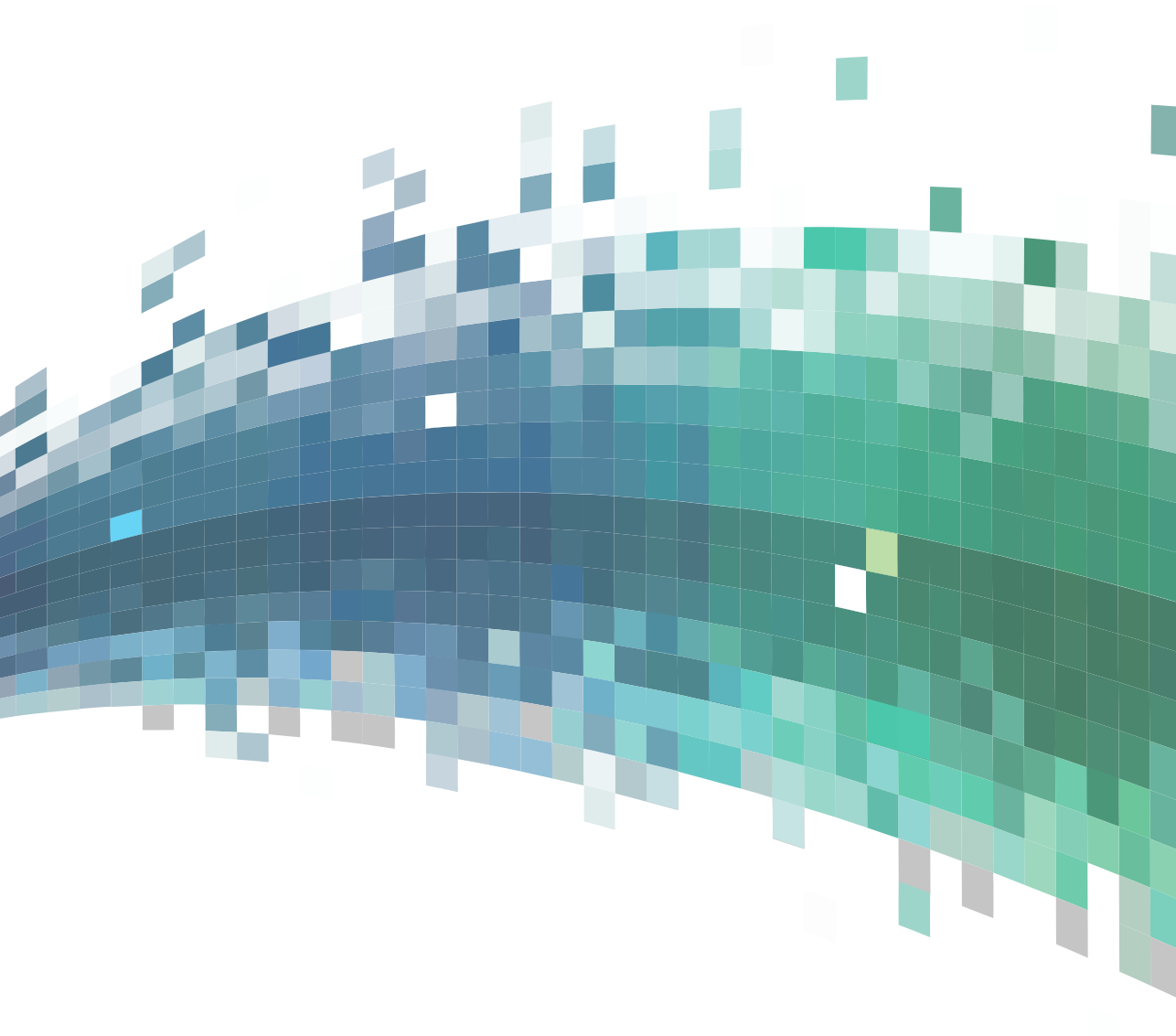
1. Sander JW. The epidemiology of epilepsy revisited. *Curr Opin Neurol*. 2003;16:165–70.
2. Fisher RS, van Emde Boas W, Blume W, et al. Epileptic Seizures and Epilepsy: Definitions Proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia*. 2005;46:470–2.
3. Tufenkjian K, Lüders HO. Seizure semiology: its value and limitations in localizing the epileptogenic zone. *J Clin Neurol*. 2012;8:243–50.
4. de Boer HM, Mula M, Sander JW. The global burden and stigma of epilepsy. *Epilepsy Behav*. 2008;12:540–6.
5. Panayiotopoulos C. General aspects of epilepsies. In: A Clinical Guide to Epileptic Syndromes and their Treatment. 2nd ed. London: Springer; 2010. p. 1–20.
6. Schmidt D, Schachter SC. Drug treatment of epilepsy in adults. *Bmj*. 2014;348:g254.
7. Panayiotopoulos C. Principles of Therapy in Epilepsies. In: A Clinical Guide to Epileptic Syndromes and their Treatment. London: Springer; 2010.
8. Ghougassian DF, D'Souza W, Cook MJ, et al. Evaluating the utility of inpatient video-EEG monitoring. *Epilepsia*. 2004;45:928–32.
9. Fahoum F, Omer N, Kipervasser S, et al. Safety in the epilepsy monitoring unit: A retrospective study of 524 consecutive admissions. *Epilepsy Behav*. 2016;61:162–7.
10. Beniczky S, Neufeld M, Diehl B, et al. Testing patients during seizures: A European consensus procedure developed by a joint taskforce of the ILAE – Commission on European Affairs and the European Epilepsy Monitoring Unit Association. *Epilepsia*. 2016;57:1363–8.
11. Braun KPJ. Preventing cognitive impairment in children with epilepsy. *Curr Opin Neurol*. 2017;30:140–7.
12. Mula M, Trimble MR. Antiepileptic drug-induced cognitive adverse effects: potential mechanisms and contributing factors. *CNS Drugs*. 2009;23:121–37.
13. Lüders HO, Najm I, Nair D, et al. The epileptogenic zone: general principles. *Epileptic Disord*. 2006;8:S1–9.
14. West S, Nolan S, Cotton J, et al. Surgery for epilepsy. *Cochrane Database Syst Rev*. 2015;CD010541.
15. Gallentine WB, Mikati MA. Intraoperative electrocorticography and cortical stimulation in children. *J Clin Neurophysiol*. 2009;26:95–108.
16. Kuruvilla A, Flink R. Intraoperative electrocorticography in epilepsy surgery: Useful or not? *Seizure*. 2003;12:577–84.
17. van 't Klooster MA, Van Klink NEC, Leijten FSS, et al. Residual fast ripples in the intraoperative corticogram predict epilepsy surgery outcome. *Neurology*. 2015;85:120–8.
18. Wu JY, Sankar R, Lerner JT, et al. Removing interictal fast ripples on electrocorticography linked with seizure freedom in children. *Neurology*. 2010;75:1686–94.
19. Schuele SU, Lüders HO. Intractable epilepsy: management and therapeutic alternatives. *Lancet Neurol*. 2008;7:514–24.
20. Cockerell O, Hart Y, Sander J. Mortality from epilepsy: results from a prospective population-based study. *Lancet*. 1994;84:71–2.
21. Mohanraj R, Norrie J, Stephen LJ, et al. Mortality in adults with newly diagnosed and chronic epilepsy: a retrospective comparative study. *Lancet Neurol*. 2006;5:481–7.

22. Nilsson L, Farahmand BY, Persson PG, et al. Risk factors for sudden unexpected death in epilepsy: a case-control study. *Lancet*. 1999;353:888–93.
23. Langan Y, Nashef L, Sander JW. Case-control study of SUDEP. *Neurology*. 2005;64:1131–3.
24. Hesdorffer DC, Tomson T, Benn E, et al. Do antiepileptic drugs or generalized tonic-clonic seizure frequency increase SUDEP risk? A combined analysis. *Epilepsia*. 2012;53:249–52.
25. Lamberts RJ, Thijs RD, Laffan A, et al. Sudden unexpected death in epilepsy: people with nocturnal seizures may be at highest risk. *Epilepsia*. 2012;53:253–7.
26. Massey CA, Sowers LP, Dlouhy BJ, et al. Mechanisms of sudden unexpected death in epilepsy: the pathway to prevention. *Nat Rev Neurol*. 2014;10:271–82.
27. Surges R, Thijs RD, Tan HL, et al. Sudden unexpected death in epilepsy: risk factors and potential pathomechanisms. *Nat Rev Neurol*. 2009;5:492–504.
28. Russell-Jones DL, Shorvon SD. The frequency and consequences of head injury in epileptic seizures. *J Neurol Neurosurg Psychiatry*. 1989;52:659–62.
29. Ryvlin P, Nashef L, Lhatoo SD, et al. Incidence and mechanisms of cardiorespiratory arrests in epilepsy monitoring units (MORTEMUS): A retrospective study. *Lancet Neurol*. 2013;12:966–77.
30. Lennie P. The Cost of Cortical Computation. *Curr Biol*. 2003;13:493–7.
31. Kahneman D. Basic issues in the study of attention. In: Attention and Effort. Englewood Cliffs, New Jersey: Prentice-Hall; 1973. p. 1–12.
32. Eriksen CW, Yeh YY. Allocation of attention in the visual field. *J Exp Psychol*. 1985;11:583–97.
33. Tzallas AT, Tsipouras MG, Tsalikakis DG, et al. Automated Epileptic Seizure Detection Methods: A Review Study. In: Stevanovic D, editor. Epilepsy - Histological, Electroencephalographic and Psychological Aspects. InTech; 2012.
34. Kalitzin S, Petkov G, Velis D, et al. Automatic segmentation of episodes containing epileptic clonic seizures in video sequences. *IEEE Trans Biomed Eng*. 2012;59:3379–85.



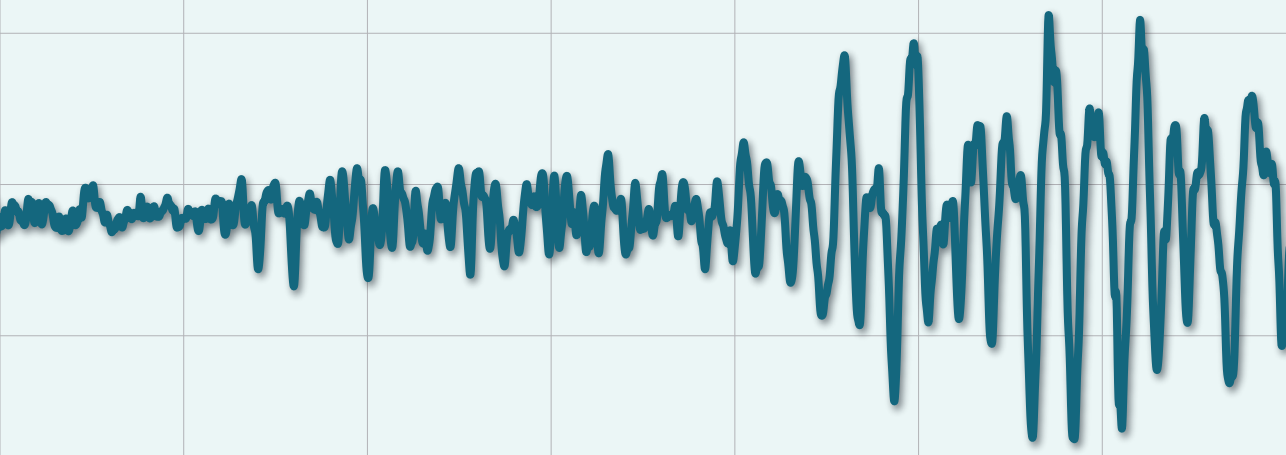
PART I

AUTOMATED MARKERS TO ENHANCE DIAGNOSIS



CHAPTER 2

IMPROVING STAFF RESPONSE TO SEIZURES ON THE EPILEPSY MONITORING UNIT WITH ONLINE EEG SEIZURE DETECTION ALGORITHMS



Based on:
Rommens N, Geertsema E, Jansen Holleboom L, Cox F, Visser G. Improving staff response to seizures on the epilepsy monitoring unit with online EEG seizure detection algorithms. Epilepsy & Behavior. 2018; 84: 99-104.

ABSTRACT

OBJECTIVE

User safety and the quality of diagnostics on the epilepsy monitoring unit (EMU) depend on reaction to seizures. Online seizure detection might improve this. While good sensitivity and specificity is reported, the added value above staff response is unclear. We ascertained the added value of two electroencephalograph (EEG) seizure detection algorithms in terms of additional detected seizures or faster detection time.

METHODS

EEG-video seizure recordings of people admitted to an EMU over one year were included, with a maximum of two seizures per subject. All recordings were retrospectively analyzed using Encevis EpiScan and BESA Epilepsy. Detection sensitivity and latency of the algorithms were compared to staff responses. False positive rates were estimated on 30 uninterrupted recordings (~24 h per subject) of consecutive subjects admitted to the EMU.

RESULTS

EEG-video recordings used included 188 seizures. The response rate of staff was 67%, of Encevis 67%, and of BESA Epilepsy 65%. Of the 62 seizures missed by staff, 66% were recognized by Encevis and 39% by BESA Epilepsy. Median latencies were 31 s (staff), 10 s (Encevis), and 14 s (BESA Epilepsy). After correcting for walking time from the observation room to the subject, both algorithms detected faster than staff in 65% of detected seizures. The full recordings included 617 h of EEG. Encevis had a median false positive rate of 4.9 per 24 h and BESA Epilepsy of 2.1 per 24 h.

CONCLUSIONS

EEG-video seizure detection algorithms may improve reaction to seizures by improving the total number of seizures detected and the speed of detection. The false positive rate is feasible for use in a clinical situation. Implementation of these algorithms might result in faster diagnostic testing and better observation during seizures.

INTRODUCTION

Video-EEG monitoring in an epilepsy monitoring unit (EMU) is widely used as a diagnostic tool in people suspected of having a seizure disorder. It can be used to determine seizure type and classification, to distinguish epilepsy from nonepileptic seizures, or to examine or evaluate therapeutic options.[1,2] People are continuously monitored by staff in a separate observation room, using real-time video, audio, and EEG recordings. When seizures are detected, nursing staff enter the subject's room to reduce the risk of adverse events such as falls, respiratory compromise, and injuries.[3] Standardized tests are also performed to assess consciousness and cognition during seizures, which helps to determine seizure semiology and type.[4,5]

Staff supervision demands skills and uninterrupted attentive observation for any sign of a seizure, as otherwise, they may be missed. One study showed a response rate of 41% to seizures with a mean latency over 2 min.[6] While response rate and time may vary between EMUs, response rates are limited by human abilities. Seizures are often recognized by clinical manifestations, so seizures showing subtle or no clinical semiology are more often missed.

Online seizure detection algorithms might help to detect seizures that could have otherwise been missed or recognized too late. Seizures can be detected with a variety of signals, such as movement, electrodermal activity, heart rate, and EEG. We focused on EEG seizure detection as it is closest to the source of epilepsy, specific to epilepsy, and measured as standard on every EMU. EEG seizure detection has been ascertained since 1982, and much research has since been performed on various approaches to seizure detection.[7–10]

Recently, EEG seizure detection software, such as Encevis EpiScan and BESA Epilepsy, has become commercially available. Encevis EpiScan uses two modules that detect epileptiform activity.[11,12] To detect seizures, the extracted features are continuously compared with past information from the EEG. The BESA Epilepsy software estimates normalized energy and integrated power for different frequency bands.[13,14] This algorithm is based on the hypothesis that seizure activity manifests itself by a change in frequency and amplitude that is distinct from non-seizure or background activity. When extracted features are above a threshold for longer than 10 s, a seizure is detected. BESA Epilepsy has been developed and tested for adults.

Seizure detection algorithms have been studied thoroughly and show good detection sensitivity and low false response rates.[8,9] For Encevis EpiScan

a sensitivity of 81% with a false-detection rate of 0.30 per hour has been reported;[15] BESA Epilepsy was reported to have sensitivity of 87% with a false-detection rate of 0.22 per hour.[14] It is, however, unclear what added value the seizure detection algorithms provide to EMU seizure monitoring, as these algorithms are not widely implemented.[16] It is important to know the added value, as seizure detection systems are not standalone but aids for staff already present. We investigated this added value by assessing: 1) the current response rate and latency of staff to seizures; 2) the sensitivity, latency, and false positive rates of Encevis EpiScan and BESA Epilepsy; 3) the value added to the current response in terms of additional detected seizures and shorter latency; and 4) which monitoring cases could benefit from these algorithms.

METHODS

EMU SETTING

The added value of a detection algorithm depends on the work setting and the staffing; to allow comparisons, we describe here our setting: It is an 8-bed unit, where each individual stays in a separate room, for up to 5 days. Three to four remote control cameras are installed in each room to capture the whole room. Individuals have call buttons to alert staff.

Subjects are monitored continuously by staff (specialized nurses) in an observation room, where a real-time EEG, electrocardiogram (ECG), video, and audio stream is shown for each room. An intercom system can be used for communication. When a seizure is noticed, the subject is attended to ensure safety and execute standardized diagnostic tests. Three nurses are present during the daytime and two during the night. No automated seizure detection techniques are used.

EEG RECORDINGS

A Micromed EEG system (Micromed, Mogliano Veneto, Italy) was used to record EEGs with a sampling frequency of 256 Hz. The international 10–20 electrode placement system was used. Some individuals had additional electrodes to provide higher spatial sampling. After recording and reporting, it is standard practice to cut EEG and video files to decrease storage space. Only diagnostically relevant parts of the registration, for example diagnostic tests and seizures, are stored.

DATA SELECTION

Seizures between May 2014 and April 2015 were included retrospectively in a seizure database. Only seizures confirmed as epileptic in the corresponding EEG report and longer than 5 s were included. To prevent overrepresentation only two seizures per subject were included. If more than two seizures were present, two of the first five were randomly selected. It's important to perform diagnostic tests in these initial seizures, so staff response is required; this might not be the case for later seizures. The seizure database encompasses a representative sample of all seizure types occurring in the EMU. Seizures where staff was already present at the start of the seizure were excluded, as response could not be evaluated. Seizures where the patient alerted the staff were not excluded. Call buttons will not be removed from clinics when using EEG-based seizure detection and are therefore an important addition to visual recognition of seizures by staff. The EEG file duration of the seizures could vary depending on how files were cut.

An additional database (the 24-hour database) was collected, consisting of nonstop EEG recordings. These recordings represent all registrations occurring on an EMU and can therefore be used to calculate false positives. The 24-hour database included 30 consecutive recordings from September 2016. For every subject, a section of 16 to 24 consecutive hours of the recording was randomly included.

This study was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans.

SCORING OF REGISTRATIONS

Seizures' start and end in both databases were identified by trained reviewers. Four different time points were scored: clinical seizure onset (CSO), clinical seizure end (CSE), electrographic seizure onset (ESO), and electrographic seizure end (ESE), as illustrated in Figure 2.1. The ESO was defined as the moment where the first EEG seizure pattern could be seen and the ESE where it ends. The CSO was defined as the start of the first clinical symptom. The CSE was defined as the time when subjects were able to resume normal activities, as up to that point, it is of value to respond to seizures. The CSO–CSE period may therefore include postictal symptoms.

For the seizure database, electrographic and clinical seizure characteristics were also scored to evaluate how easily changes could be detected by an

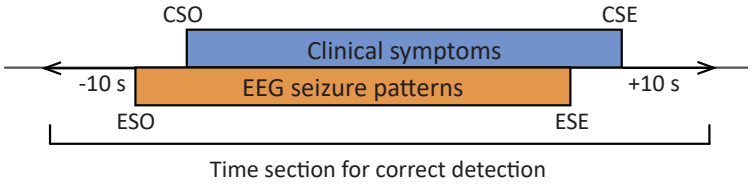


FIGURE 2.1.
Seizure marker timing, which might vary between subjects. CSO (clinical seizure onset), CSE (clinical seizure end), ESO (electrographic seizure onset), and ESE (electrographic seizure end) were scored for every seizure.

observer. Both sets of characteristics were scored using values between 1 and 4, representing no visible manifestations (1) to very clear manifestations (4), from the perspective of the nurses who monitor the subjects. The characteristics were scored every 5 s until staff responded, up to the first 60 s of the seizure. From these scores, a mean value was calculated. Seizure classification was also collected from EEG reports.

The interictal EEG in the 24-hour database was evaluated to investigate whether epileptiform activity would influence the false positive rate. Four categories were used: 'Normal interictal EEG', 'Abnormal interictal EEG with nonspecific nonepileptiform abnormalities', 'EEG with some epileptiform abnormalities', and 'EEG with frequent epileptiform abnormalities', based on the EEG report.

Staff response was evaluated by retrospectively reviewing the videos from the seizure database. A response was defined as staff entering the room of the subject or using the intercom any time from the seizure onset until 10 s after the end of the seizure (when EEG and clinical manifestations have both stopped).

All recordings were analyzed retrospectively using Encevis EpiScan and BESA Epilepsy. The detection algorithms should operate the same in an online situation, but due to unavailability of online functioning this could not be tested.

SENSITIVITY

We calculated the detection sensitivity of staff, Encevis EpiScan, and BESA Epilepsy. A correct detection was defined as occurring within the period from 10 s before the start of a seizure (CSO or ESO) to 10 s after the end of a seizure (ESE or CSE) (Figure 2.1).

LATENCY

Latency of staff, Encevis EpiScan, and BESA Epilepsy detections were calculated from electrographic seizure onset (ESO). For BESA Epilepsy, 10 s were added to account for the delay in the algorithm's online functioning; the algorithm places detection markers at seizure onset after having registered 10 s of the seizure. The median and p5–p95 percentile ranges of the latencies were calculated.

FALSE POSITIVES

The false positive rate of Encevis EpiScan and BESA Epilepsy was calculated on the 24-hour database. A false positive is defined as a detection that does not take place during a seizure, i.e., beyond 10 s before the start of a seizure (CSO or ESO) and 10 s after the last end (ESE or CSE).When a false positive occurred, a black-out period of 10 s was defined, in which no new false positives could occur. The median false positive rate and p5–p95 percentile ranges were calculated.

STATISTICAL ANALYSIS

The difference in sensitivity between seizure characteristics (adulthood, seizure classification, clinical characteristics, and electrographic characteristics) (separately for staff, Encevis EpiScan, and BESA Epilepsy) were tested for statistical significance with a Chi-square test. If a characteristic could not be determined, the recording would be removed from this analysis.

We also assessed the effect of subject age and the amount of interictal abnormalities in the EEG on the false positive rate with a Kruskal–Wallis test. The significance level was set at $p \leq 0.05$. All analyses were performed using MATLAB (R2017a, The MathWorks Inc.).

RESULTS

In total, 188 seizures in 115 subjects were included in the seizure database and 617 h of 30 subjects in the 24-hour database. The mean age in the seizure database was 28.7 years (SD 17 years) and was 24.2 years (SD 15.5 years) in the 24-hour database. Included seizures were generalized onset seizures (9.6%), focal onset seizures with temporal lobe semiology (42.6%), focal onset seizures with extratemporal lobe semiology (45.7%), and seizures that could not be classified (2.1%).

TABLE 2.1.
Performance of staff, Encevis EpiScan and BESA Epilepsy.

	Staff	Encevis EpiScan	BESA Epilepsy
Sensitivity	67.0%	77.6%	65.4%
Median detection latency in seconds (p5 to p95)	31 (-5 to 98)	10 (-4 to 50)	14 (6 to 68)
Median false positive rate per 24 hours (p5 to p95)	-	4.9 (1.2 to 13.8)	2.1 (0 to 222.7)

SENSITIVITY

The sensitivity of staff, Encevis EpiScan, and BESA Epilepsy are shown in Table 2.1. Of the 62 seizures missed by staff, 41 were recognized by Encevis EpiScan and 24 by BESA Epilepsy. Sixteen seizures were recognized only by staff. The comparison of sensitivity of Encevis EpiScan and BESA Epilepsy for all seizures and all seizures undetected by staff are shown in Figure 2.2. The influence of different seizure characteristics on the sensitivity are shown in Table 2.2.

LATENCY

Latency results are shown in Table 2.1. Figure 2.3 shows the time of detection of Encevis EpiScan and BESA Epilepsy compared to staff response time. In 83.5% of the 103 seizures detected by staff and Encevis EpiScan, the algorithm detected the seizure faster than the staff response. In 81.6% of the 98 seizures detected by staff and BESA Epilepsy, the algorithm detected the seizure faster than the staff response. This would lead to a median improvement of 18.1 s for Encevis EpiScan and a median improvement of 15.6 s for BESA Epilepsy. When correcting for walking time of 10 s from the observation room to the subject, Encevis EpiScan was still faster in 65.0% of detected seizures and BESA Epilepsy in 65.3% of detected seizures.

FALSE POSITIVES

Median false positive rates can be found in Table 2.1, and a histogram of the false positive rates per subject is shown in Figure 2.4. Encevis EpiScan had low false positive rates for almost every subject. BESA Epilepsy had zero false positives for most subjects but also some outliers with many false positives. Most false positives with Encevis EpiScan occurred during the first analysis hours. This is probably due to a learning period, in which the algorithm needs to establish a baseline. If the alarm of Encevis EpiScan was turned off

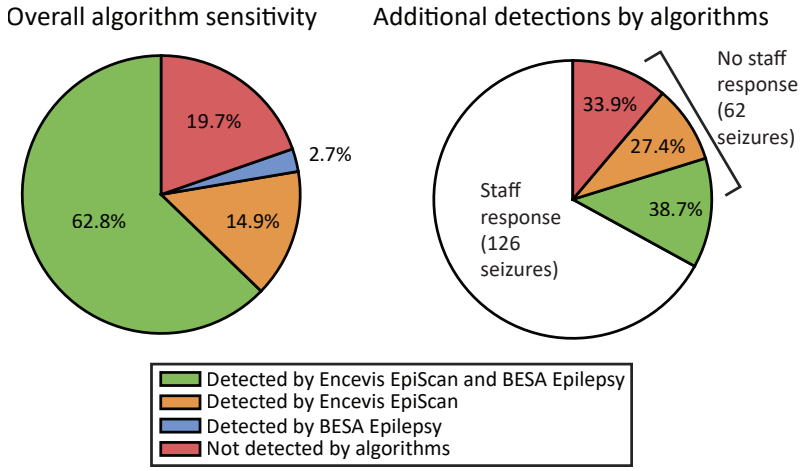


FIGURE 2.2.
Overall sensitivity of seizure detection algorithms Encevis EpiScan & BESA Epilepsy on all 188 seizures (left) and sensitivity for seizures missed by staff (right).

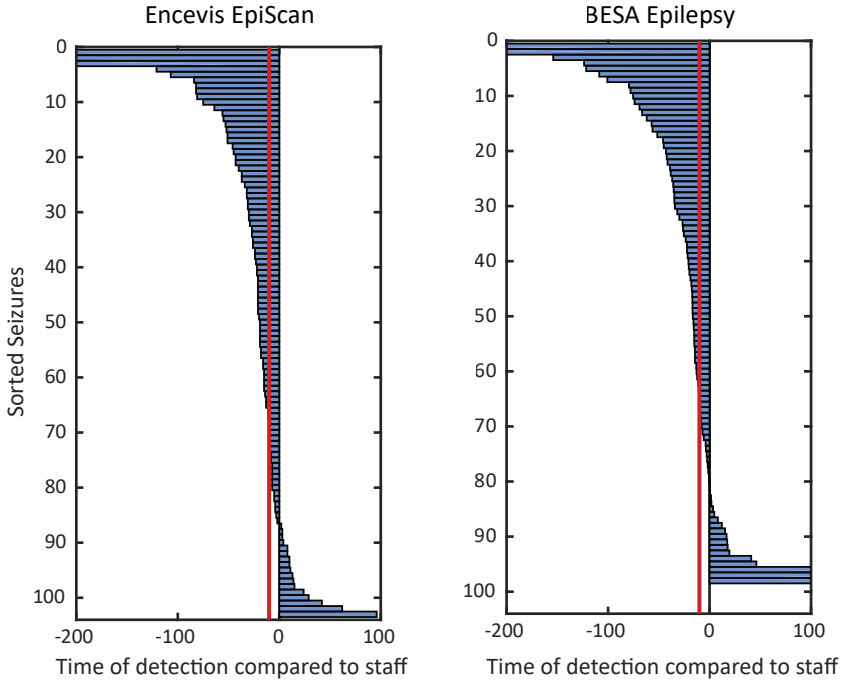


FIGURE 2.3.
Time of algorithm detection compared to staff response time (at 0 s) for the seizures detected by the algorithm and staff. For BESA Epilepsy, 10 s was added to account for the delay in the algorithm's online functioning. A red line is drawn to take walking time of staff into account (10 s). Some outliers are out of axes bounds; –1579, –611 and – 215 s for Encevis Episcan and –552, –253, 309, 392 and 3173 s for BESA Epilepsy.

TABLE 2.2.
Sensitivity of staff, Encevis EpiScan, and BESA Epilepsy for seizures with different characteristics. The difference in sensitivity for specific characteristics was tested for statistical difference per detection method with a Chi-square test. Values under the significance level ($p \leq 0.05$) are presented in bold.

			Staff		Encevis EpiScan		BESA Epilepsy	
			Sensitivity	p-Value	Sensitivity	p-Value	Sensitivity	p-Value
All subjects			67.0%	-	77.6%	-	65.4%	-
Age	Children under 18 years (n=54)		59.3%	**0.03**	79.6%	0.68	72.2%	0.21
	Adults (n=134)		70.1%		76.9%		62.7%	
Seizure classification	Generalized onset (n=18)		88.9%	**0.004**	100%	**<0.001**	100%	**<0.001**
	Focal onset seizures	Temporal (n=80)	75.0%		86.3%		71.3%	
		Extra-temporal (n=86)	55.8%		64.0%		53.5%	
	Unclear classification (n=4)		-		-		-	
Clinical characteristics	No visible changes (n=27)		37.0%	**<0.001**	85.2%	0.51	66.7%	0.31
	Subtle clinical symptoms (n=119)		67.2%		77.3%		61.3%	
	Clear clinical symptoms (n=34)		82.4%		70.6%		73.5%	
	Very clear clinical symptoms (n=8)		100%		87.5%		87.5%	
Electrographic characteristics	No visible changes (n=6)		66.7%	**0.03**	66.7%	**<0.001**	0%	**<0.001**
	Subtle changes (n=73)		54.8%		58.9%		39.7%	
	Clear focal changes (n=73)		72.6%		89.0%		83.6%	
	Clear diffuse changes (n=36)		80.6%		94.4%		91.7%	

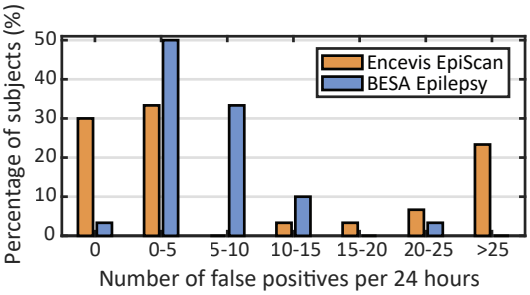


FIGURE 2.4.
Distribution of false positive rates for all subjects.

during the first hour, the false positive rate would decrease by 36.4% and by 46.1% if it were turned off for the first 2 h.

Children had a higher false positive rate than adults for Encevis EpiScan ($p = 0.0430$), with a median false positive rate of 6.99 per 24 h for children and 4.47 per 24 h for adults. The difference in false positive rate between registrations with different interictal abnormality categories was not statistically significant. BESA Epilepsy also had a higher false positive rate in children than in adults ($p = 0.0057$), with a median false positive rate of 39.1 per 24 h for children and 1.28 per 24 h for adults. Additionally, EEGs with frequent interictal epileptiform abnormalities had a significantly higher false positive rate ($p = 0.0308$), with a median false positive rate of 33.1 per 24 h (compared to 1.16–9.86 for EEGs with fewer abnormalities).

DISCUSSION

Reaction to seizures can be improved by online EEG seizure detection algorithms by improving the number of detected seizures and the response latency after seizure onset. We were able to show that more than half of the undetected seizures could be recognized by EEG seizure detection algorithms. For most seizures the detections by both algorithms preceded detection by staff. The algorithms had acceptable median false positive rates.

The staff response and latency that we found is better than the previously described response rate of 41% with an average latency of 142.3 s.[6] This may be due to differences between EMU settings, e.g., staff experience or EMU layout. The sensitivity and specificity of the algorithms that we found are comparable to those from previous reports.[14,15] Latency has not been previously described. The added value above staff already present has not

previously been reported. This information is key, since a seizure detection algorithm will not be a stand-alone system but an addition to current staff.

Staff response and sensitivity of detection algorithms are influenced by characteristics of seizures and the individuals. When there is a low staff response rate but algorithms are able to detect seizures, people could particularly benefit from their use. Staff response was highly dependent on clinical characteristics of seizures, as response is based on symptoms seen on the video stream. Conversely, algorithms were mostly dependent on the presence of electrographic changes. This influence is also reflected in the sensitivity for different seizure classifications. For example, generalized seizures are electrographically and clinically very clear and therefore have a high response rate by staff and algorithms. Focal onset seizures with extra-temporal lobe semiology, on the other hand, were short with few clinical and electrographic changes, which explains the lower response rate by staff and algorithms. Thus, people with less clear seizures showing electrographic changes could benefit from these algorithms, for example in case of seizures with temporal lobe semiology. Children might benefit from seizure detection algorithms, as staff sensitivity was lower in this group. A higher false positive rate was, however, found for children. This might be due to variations in EEG patterns in children, making it more challenging to differentiate normal EEG from ictal patterns.[17] Children in our dataset more often had EEG abnormalities, which also influenced the specificity of the algorithms. Encevis EpiScan still had an acceptable false positive rate for children, so its use could be preferred over BESA Epilepsy in children.

The algorithms could not be tested online, as at the time of the study they were not ready for online implementation. Therefore, the true effect of these algorithms could not be assessed. Sensitivity and false positive rates were tested in two different databases, because a dataset of full recordings including all type of seizures was not available at the time of the study. Testing in the same full recordings allows estimation of the false positives relative to the true positives and calculation of positive predictive values. This study did not include any nonepileptic events, as these events cannot be detected by EEG-based algorithms. Lastly, depending on how staff are trained and subjects are monitored, the staff response rate may differ between different EMU settings. Since lower response rates were found in another center, the added value in other centers might be higher than we described.[6]

Future research should focus on testing these algorithms online on continuous unselected data. Additionally, performance of EEG seizure detection algorithms might increase when using multisensor seizure

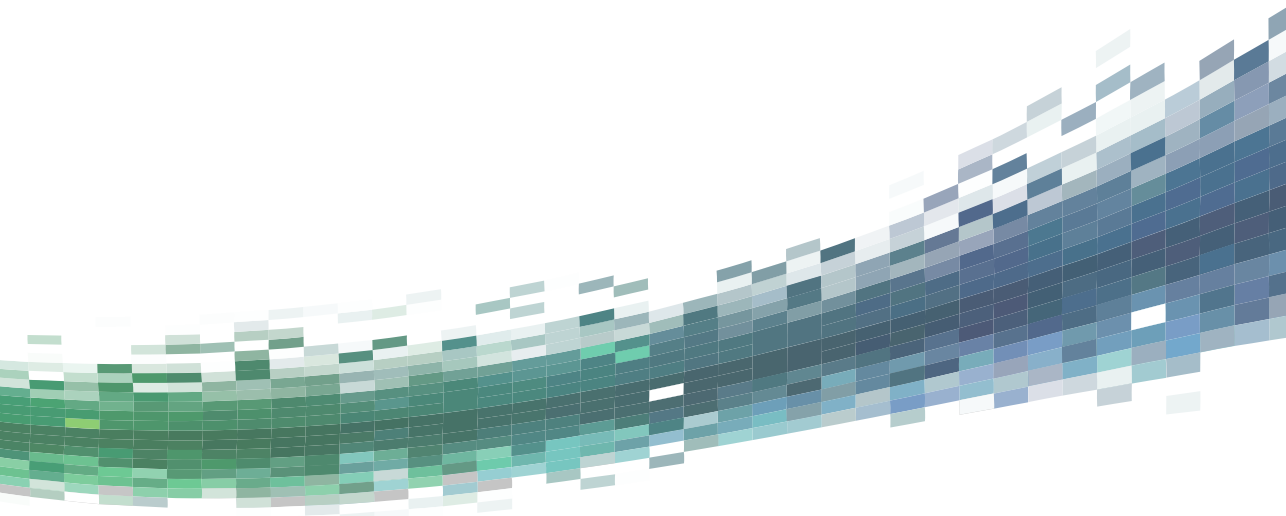
detection. We do not see substantial benefit from adding movement-based sensors or electromyographic sensors. These detectors perform best on tonic-clonic seizures or hypermotor seizures, which are already recognized by staff and algorithms. There might be an improvement when adding ECG seizure detection. Heart rate changes occur in all type of seizures and mostly in the beginning or even before the electrographic start of a seizure.[18,19] Adding this modality might increase sensitivity and latency, but more research on this topic is necessary.

CONCLUSION

Online EEG seizure detection algorithms can improve the staff response to seizures by detecting additional seizures and improving latency. The false positive rate is reasonable for use in a clinical setting. Implementation of these algorithms may help to ensure patient safety and improve the quality of diagnostics by assessing consciousness and cognition in a timely manner.

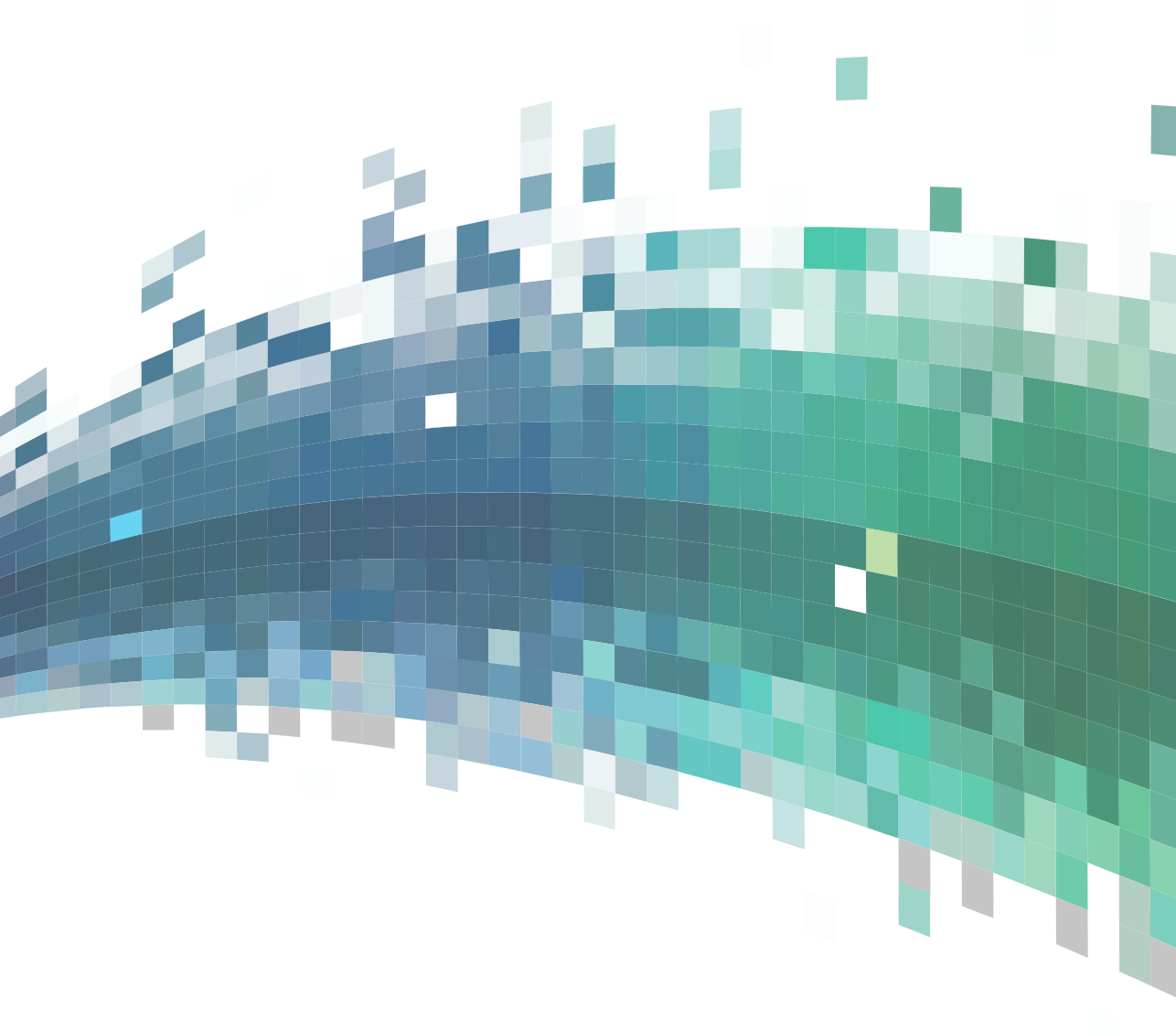
REFERENCES

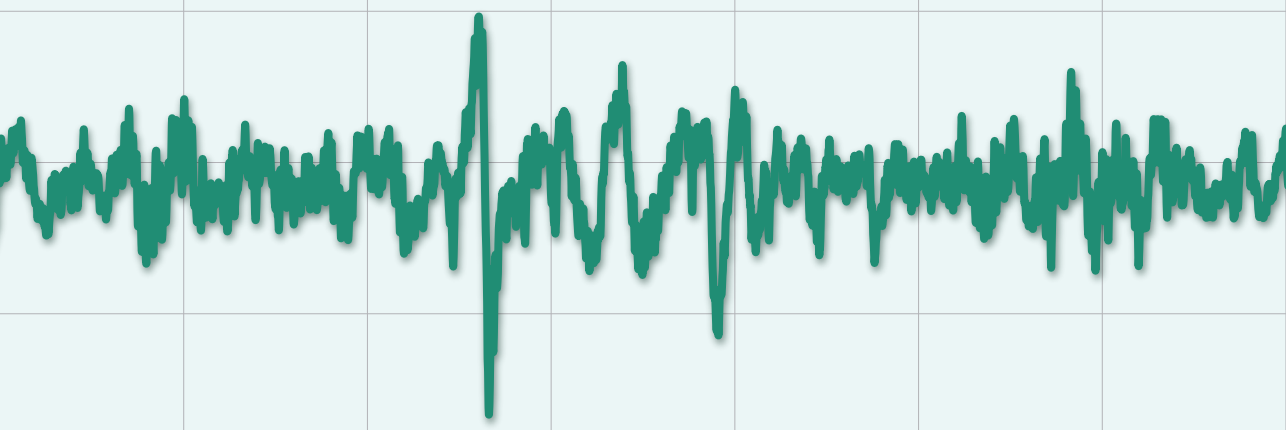
1. Cascino GD. Video-EEG monitoring in adults. *Epilepsia*. 2002;43:80–93.
2. Ghougassian DF, D'Souza W, Cook MJ, et al. Evaluating the utility of inpatient video-EEG monitoring. *Epilepsia*. 2004;45:928–32.
3. Fahoum F, Omer N, Kipervasser S, et al. Safety in the epilepsy monitoring unit: A retrospective study of 524 consecutive admissions. *Epilepsy Behav*. 2016;61:162–7.
4. Johanson M, Valli K, Revonsuo A. How to assess ictal consciousness? 2011;24:11–20.
5. Bauerschmidt A, Koshkelashvili N, Ezeani CC, et al. Prospective assessment of ictal behavior using the revised Responsiveness in Epilepsy Scale (RES-II). *Epilepsy Behav*. 2013;26:25–8.
6. Atkinson M, Hari K, Schaefer K, et al. Improving safety outcomes in the epilepsy monitoring unit. *Seizure*. 2012;21:124–7.
7. Gotman J. Automatic recognition of epileptic seizures in the EEG. *Electroencephalogr Clin Neurophysiol*. 1982;54:530–40.
8. Acharya UR, Vinitha Sree S, Swapna G, et al. Automated EEG analysis of epilepsy: A review. *Knowledge-Based Syst*. 2013;45:147–65.
9. Alotaiby TN, Alshebeili SA, Alshawi T, et al. EEG seizure detection and prediction algorithms : a survey. *EURASIP J Adv Signal Process*. 2014;2014:1–21.
10. Ramgopal S, Thome-Souza S, Jackson M, et al. Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy. *Epilepsy Behav*. 2014;37:291–307.
11. Hartmann MM, Fürbaß F, Perko H, et al. EpiScan: Online seizure detection for epilepsy monitoring units. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2011. p. 6096–9.
12. Fürbass F, Hartmann M, Perko H, et al. Combining time series and frequency domain analysis for a automatic seizure detection. In: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2012. p. 1020–3.
13. Hopfengärtner R, Kerling F, Bauer V, et al. An efficient, robust and fast method for the offline detection of epileptic seizures in long-term scalp EEG recordings. *Clin Neurophysiol*. 2007;118:2332–43.
14. Hopfengärtner R, Kasper BS, Graf W, et al. Automatic Seizure Detection in Long-term Scalp EEG Using an Adaptive Thresholding Technique: A Validation Study for Clinical Routine. *Clin Neurophysiol*. 2014;125:1346–52.
15. Fürbass F, Ossenblok P, Hartmann M, et al. Prospective multi-center study of an automatic online seizure detection system for epilepsy monitoring units. *Clin Neurophysiol*. 2015;126:1124–31.
16. Rubboli G, Beniczky S, Claus S, et al. A European survey on current practices in epilepsy monitoring units and implications for patients' safety. *Epilepsy Behav*. 2015;44:179–84.
17. Eisermann M, Kaminska A, Moutard ML, et al. Normal EEG in childhood: From neonates to adolescents. *Neurophysiol Clin*. 2013;43:35–65.
18. Leutmezer F, Schernthaner C, Lurger S, et al. Electrocardiographic changes at the onset of epileptic seizures. *Epilepsia*. 2003;44:348–54.
19. Zijlmans M, Flanagan D, Gotman J. Heart rate changes and ECG abnormalities during epileptic seizures: Prevalence and definition of an objective clinical sign. *Epilepsia*. 2002;43:847–54.



PART II

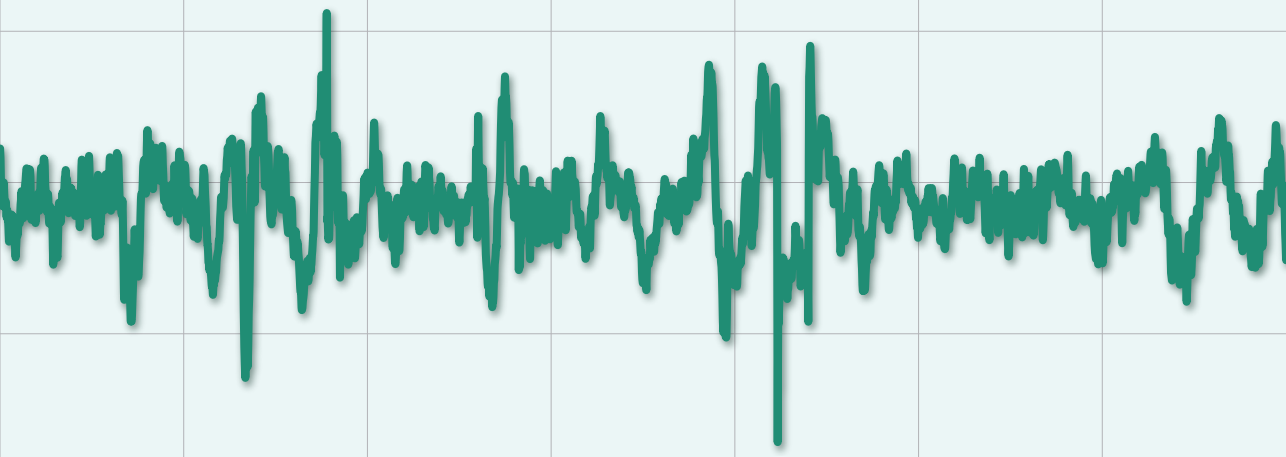
AUTOMATED MARKERS TO ENHANCE TREATMENT





CHAPTER 3

AUTOMATED SEIZURE ONSET ZONE APPROXIMATION BASED ON NONHARMONIC HIGH-FREQUENCY OSCILLATIONS IN INTERICTAL INTRACRANIAL EEGS



Based on:
Geertsema EE, Visser GH, Velis DN, Claus SP, Zijlmans M, Kalitzin SN. Automated seizure onset zone approximation based on nonharmonic high-frequency oscillations in human interictal intracranial EEGs. International Journal of Neural Systems. 2015; 25 (5): 1550015.

ABSTRACT

OBJECTIVE

The objective of epilepsy surgery is to resect completely or to disconnect the epileptogenic zone (EZ). We propose a novel automated algorithm to approximate the EZ in presurgical intracranial electroencephalograms (iEEG), while providing reproducible output.

METHODS

The seizure onset zone (SOZ), a surrogate marker for the EZ, was approximated from iEEGs of nine people with temporal lobe epilepsy (TLE), using three methods: 1) total ripple length (TRL): manually segmented high-frequency oscillations, 2) rippleness (R): area under the curve of the autocorrelation functions envelope, and 3) autoregressive model residual variation (ARR, novel algorithm): time-variation of residuals from autoregressive models of iEEG windows. TRL, R, and ARR results were compared in terms of separability, using Kolmogorov-Smirnov tests, and detection performance, using receiver operating characteristic (ROC) curves, to the gold standard for SOZ delineation: visual observation of ictal video-iEEGs.

RESULTS

TRL, R, and ARR can distinguish signals from iEEG channels located within the SOZ from those outside it ($p < 0.01$). The ROC area under the curve was 0.82 for ARR, while it was 0.79 for TRL, and 0.64 for R.

CONCLUSIONS

ARR outperforms TRL and R, and may be applied to identify channels in the SOZ automatically in interictal iEEGs of people with TLE. ARR, interpreted as evidence for nonharmonicity of high-frequency EEG components, could provide a new way to delineate the EZ, thus contributing to presurgical workup.

INTRODUCTION

Epilepsy is a chronic brain disorder characterized by recurrent epileptic seizures.[1] When a patient has an epileptic syndrome that is possibly remediable by surgery, and seizure control could not be attained with multiple trials of antiepileptic drugs, the patient qualifies to enter presurgical evaluation. The objective of epilepsy surgery is to resect completely or to disconnect the epileptogenic zone (EZ), which is defined as the minimum amount of cortical tissue that must be resected to produce seizure-freedom.[2] The EZ is a theoretical concept. In practice, surrogate markers for the EZ are used, such as the seizure onset zone (SOZ), defined as the area of the cortex from which clinical seizures are generated.[3]

The gold standard method of determining the SOZ is visual observation of ictal recordings from intracranial electroencephalograms (iEEG) and synchronized video.[3] During such observations a representative sample of habitual seizures needs to be recorded with electrodes positioned in the SOZ. Therefore, several days of video-EEG recording may be required (at least one week). As a consequence, review of the attained long-term video-EEG is both laborious and time consuming.

When validated, using interictal high-frequency oscillations (HFOs) to delineate the SOZ may result in a reduced need for seizure recordings, thus shortening EEG registration time.[4] Interictal HFOs have been associated with epileptic properties of neuronal tissue in the brain.[5–8] HFOs have been shown to be reliable markers of the SOZ,[6,9–12] and to provide better SOZ localization than interictal spikes.[9,11,13] Visual detection of HFOs is subjective with poor inter-rater reliability and reproducibility,[14] and time consuming, due to the small time frame needed to observe the oscillations with their short duration and small amplitude.[15] Automatic detection of HFOs could possibly solve these problems, and several possibilities to implement such a detector have been proposed in the literature.[11,14,16–20]

HFOs, and especially ripples (80–250 Hz oscillations), are not necessarily a sign of epileptogenicity alone; they can also have a physiological origin. Physiological ripples (~200 Hz) have been measured in the CA1 region of the hippocampus, the entorhinal cortex,[21] the mesiotemporal lobe, and the extratemporal neocortex.[22] Therefore, when using HFOs to identify the SOZ, it is important to realize that HFOs may also be found in areas outside the SOZ; e.g. in the contralateral hippocampus. Furthermore, "false ripples" can result from filtering sharp epileptic transients or nonsinusoidal oscillations

in the EEG.[23,24] Consequently, separation of epileptic HFOs and other HFO-like signals remains a challenge.

We propose a novel methodology, for which the main underlying assumption is that epileptic HFOs are of the nonharmonic type. In other words, they cannot be exactly described as solutions of a linear differential equation of a certain order. There can be a variety of reasons for deviation from harmonic behavior, such as noise input, nonstationary parameters, and nonlinear dynamics. Various quantifiers that characterize for example nonlinearity related to a single recording site have proved effective in aiding the lateralization and localization of the EZ.[25–33] The capacity for SOZ localization (not limited to lateralization) of the corresponding algorithms is, however, reported only in a sub-selection of those studies.[27,32,33]

We aimed to design an automated algorithm that can independently, reliably, and reproducibly approximate the SOZ, as a surrogate marker for the EZ. The proposed algorithm quantifies certain nonharmonic features of high-frequency EEG components to select regions that may be closest to the SOZ. As a secondary goal, we investigated if 'rippleness', obtained with an algorithm previously proposed by our group,[33] could possibly be used for fast automated estimation of the amount of HFOs. We investigated if the previously found correspondence between rippleness and the total length of manually scored HFOs (80-250 Hz), can be reproduced in our data set. Subsequently, the new algorithm is compared with manually scored HFOs and rippleness in terms of their ability to approximate the SOZ.

METHODS

SUBJECTS

Nine individuals with intractable temporal lobe epilepsy were included. The selected individuals were candidates for epilepsy surgery between 2008 and 2012, and the registered iEEGs were part of the presurgical work up. The iEEG registrations were performed at Stichting Epilepsie Instellingen Nederland (SEIN), in Heemstede, The Netherlands. No further exclusion criteria were applied beyond the sampling frequency of the stored data (>1000 Hz). Data were analyzed retrospectively; individuals had epilepsy surgery at least one year prior to conducting this study. Relevant clinical information on the subjects in this study is shown in Table 3.1.

TABLE 3.1.
Subject information. Age is defined at time of registration; registration day is defined as the day, during registration, after medication lowering. Surgical outcome is defined in terms of the UCLA classification, and follow up period is given in years. UCLA 1A: Free of disabling seizures and auras, UCLA 1B: Free of disabling seizures, UCLA 2: Rare disabling seizures, UCLA 3: Worthwhile improvement, UCLA 4: No worthwhile improvement.[34] For electrode locations and corresponding abbreviations, see Supplementary Table 3.S1.

Subject	Age	Sex	Day	Seizure onset zone	Resected area (year)	Surgical outcome (follow up)
1	22	F	3	ATHCL1, MTHCL1, IAIL2-4, AML1-2, IAPL2-4, TPL1-3	L temporal + insula (2011)	UCLA IA (1.5)
2	27	M	8	LesAntR1-2, LesSupR2-3, LesAnIR1-3, InsAntR1-5	R temporo-insular frontal (2012)	UCLA IA (1)
3	63	M	2	1) HAR1-3, HPR1-3, AR1-4; 2) HAR1-3 IAIR1	R temporal + amygdalo-hippocampectomy (2011)	UCLA III (2)
4	28	M	7	1) HCR1-5; 2) HCL1-5; (1>2)	R temporal + amygdalo-hippocampectomy (2008)	UCLA II (4.5)
5	13	F	3	FODL4-7 *	Superior to earlier resection area (L frontal, 2009) + L SMA (2010)	UCLA IV (3.5)
6	35	F	6	HCL1-2, AML1-2	L fronto-temporal + amygdalohippocampectomy +fronto-opercular + frontobasal (2008)	UCLA IB (5)
7	26	F	4	1) ATR1-2, AHR1-2, MHR1-2, IAR2-4, PHR1-2; 2) MHL1-2; (1>2)	R hippocampectomy (2011)	UCLA II (2)
8	26	F	3	HCR1-3	R temporal + leasionectomy + amygdalohippocampectomy (2010)	UCLA IB (3)
9	39	M	10	TBL1-3, HC1L1-3, HC2L1-3, AML1-2	L temporal (2011)	UCLA IA (2)

* There is possibly a second, less frequent seizure type (involving PSMAL, PSMAR, TOL, and PFL bundles in its SOZ), but due to many artefacts and unclear seizure semiology, this hypothesis could not be substantiated.

DATA

INTRACRANIAL EEGS

All patients underwent iEEG registration for at least one week. Depth electrodes were implanted at the Academic Center for Neurosurgery of the Free University Medical Center (VUmc) in Amsterdam, The Netherlands. Ad-Tech (Racine, WI, USA) electrodes with a 2.4 mm^2 contact size were used. Electrode placement was guided by instructions from clinical neurophysiologists and authors SC and DV, according to clinical hypotheses regarding the SOZ and ictal spreading patterns. These hypotheses were formed on the basis of scalp video EEG and MRI, supplemented by MEG, PET, and SPECT results. iEEG data was recorded using a 65-channel Schwarzer amplifier (Schwarzer GmBH, Germany) implying 0.016 Hz hardware high-pass filtering. The amplifier was coupled to a workstation with Harmonie 6.2 acquisition software (Stellate Systems, Montreal, PQ Canada). During registration of the iEEGs, a silent electrode situated in the white matter was chosen as a reference. An external ground electrode was placed on the forehead of the subject. A sampling frequency of 1000 Hz was used. Antiepileptic drugs were tapered to increase the chance of a seizure occurring during registration; this was part of the standard operating procedure. Based on visual inspection, EEG channels with severe artefact contamination were excluded from further processing. All recordings were performed using referential montages. Bipolar montages were secondarily constructed by subtracting the signals from the neighboring contacts on each depth electrode.

CLINICAL DATA

The SOZ was determined by visual observation of ictal recordings from the video-iEEG by DV and SC, based on ictal changes in time-related context of the semiology. These findings were used as the gold standard. The outcome of surgery was assessed according to the UCLA classification proposed by Engel et al., simplified to discriminate classes IA, IB, II, III, and IV.[34] For part of the analysis in this study, subjects are subdivided into two outcome groups: subjects with good postoperative seizure control (UCLA class IA, IB, II) and those with poor postoperative seizure control (UCLA class III, IV). The SOZ and surgical outcome are shown in Table 3.1. A total of 77 channels of EEG data were considered to be inside the SOZ, with 364 channels outside the SOZ.

EEG DATA SELECTION

Five consecutive minutes of slow-wave-sleep (SWS) were selected from the iEEG data of each subject. HFOs and spikes occur more frequently during SWS,[9,10] and the signals of the iEEG channels closest to the skull are less influenced by muscle activity in this sleep stage. During the iEEG registrations no surface EEG, electro-oculograms, or electro-myograms were recorded, so we relied on information on sleep staging that could be gathered from the iEEG. SWS epochs were found with a method similar to the one used by Dümpelmann et al.[20] The relative delta power was tracked automatically in the iEEG, with relative delta power being the energy in the delta band (0.5-4 Hz) divided by the energy in a large frequency band (0.5-100 Hz). The module to calculate spectral features in Stellate Harmony (Stellate Systems, Montreal, PQ Canada) was used to obtain the relative delta power trend. The relative delta power trend was obtained throughout a nightly recording in 30-s windows in three superficial iEEG channels from different bundles. A 5-min epoch, in which the relative delta power exceeded 70%, was selected for each subject and used in the analysis for this study.

DATA ANALYSIS

Data were analyzed retrospectively using three methods independently: Visual detection of HFOs, the rippleness algorithm, and the novel autoregressive model residual variation (ARR) algorithm. Results from all three methods were compared with the gold standard (visual SOZ determination). An overview of the analysis steps is shown in Figure 3.1. Data analysis was performed using Matlab (Mathworks, Natick, MA, USA) version 7.5.0.

TOTAL RIPPLE LENGTH (TRL)

Each 5-min iEEG epoch was assessed visually for HFOs using the currently standard method, similar to the one described by Jacobs et al.[13] An 80 Hz high-pass filter (Finite Impulse Response filter, order 63) was applied to discern the HFOs from the relatively high-voltage background activity. HFOs were visualized in eight channels simultaneously, while displaying the traces at maximum time-resolution. Events were regarded as ripples if their amplitude was clearly higher than the high-frequency baseline activity of the channel, and the oscillation consisted of at least four periods. HFOs were distinguished from presumed artefacts based on their frequency content and sharpness and co-occurrence over several channels.[4,35] HFOs were marked by a technical physician (E. Geertsema) and revised by an

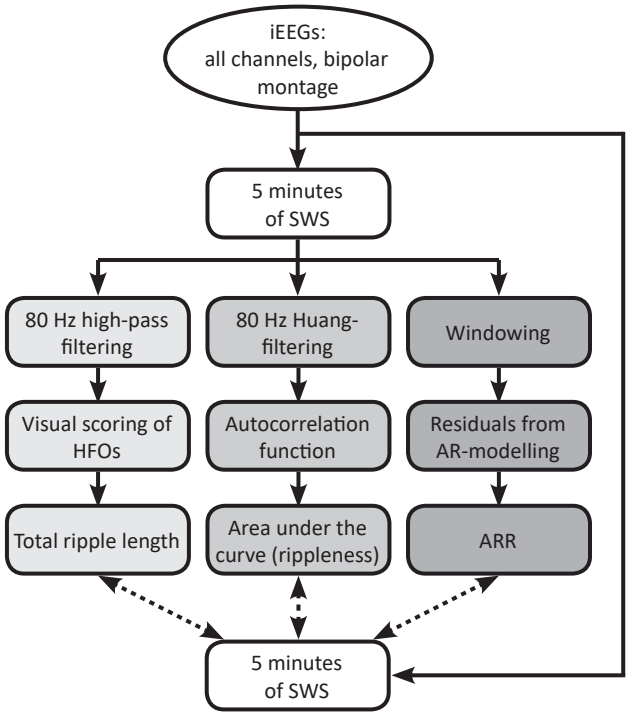


FIGURE 3.1.
Informational flow diagram of how the data in this study was used. Dashed arrows indicate comparison of data. (iEEG: Intracranial electroencephalogram, SWS: slow-wave sleep, HFOs: high-frequency oscillations, ACF: autocorrelation function, AUC: area under the curve, SOZ: seizure onset zone, AR: autoregressive, ARR: autoregressive model residual variation.)

experienced observer (M. Zijlmans). After HFO segmentation the TRL (sum of the lengths of all ripples) was calculated per channel.

By filtering above 80 Hz, manually segmented HFOs lay in the ripple (80-250 Hz) range. With the sampling frequency (1000 Hz) used, it was unfeasible to take fast ripples (250-500 Hz) into account. Digital sampling at five times the oscillation frequency of interest is advisable to sample adequately the temporal dynamics of an HFO.[36]

RIPPLENESS (R)

Channel rippleness was obtained by applying a previously described algorithm that uses the autocorrelation function (*ACF*) and its area under the curve (AUC) to quantify the extent of oscillation in a signal.[33] A brief description is provided here.

First, the selected 5-min epoch of iEEG data was 80 Hz high-pass filtered, using a morphological decomposition of the signal, the Huang-Hilbert transform. This filtering method is applied because unlike commonly used FIR filters, it does not cause spurious oscillations which can be mistaken for ripples.[23,33] Subsequently, the ACF is calculated using

$$ACF(\tau) = \langle S(\tau)S(t+\tau)\rangle_{\tau}, \quad (3.1)$$

where S(t) is the filtered signal through time, S(t+τ) the shifted signal with lag τ (t and τ in samples). The lowest frequency of interest was selected as 80 Hz and a maximal number of 10 cycles to constitute an HFO was postulated. Therefore, a maximum lag of 10/80 of a second, or 125 samples (with a 1000 Hz sampling frequency), was chosen. The envelope of the ACF is obtained by the Hilbert transform, after which the area under the envelope, or curve (AUC), can be calculated. This quantity represents the total power of oscillations above 80 Hz, or rippleness (R).

AUTOREGRESSIVE MODEL RESIDUAL VARIATION (ARR)

The autoregressive model residual variation (ARR) algorithm uses the presence of high residual signal variation after autoregressive (AR) model fit as a biomarker to identify contacts in or close to the SOZ. High ARR values are interpreted as evidence of the presence of nonharmonicity in iEEG signals.

AR modeling was selected because the frequencies and damping coefficients resulting from AR modeling can quantify oscillatory properties of an EEG.[37] A preliminary study in part of the data (subjects 1, 2, 3, 4, 5, and 7, results not shown here) showed that iEEG windows containing ripples measured in areas close to the SOZ exhibited higher residual signal variation after AR model fit than windows with ripples from areas outside the SOZ (e.g., from the contralateral hippocampus). Moreover, when the residual signal variation from all iEEG windows in a 5-min epoch was obtained, channels in the SOZ showed frequent windows with an excessively high residual. By quantifying these frequently occurring high residuals using the coefficient of variation (i.e. a normalized standard deviation), a classifier was found for the approximation of the SOZ.

The processing steps executed by the ARR algorithm are shown in Figure 3.2. For each channel, the 5-min iEEG epochs were divided into windows of 40 samples, with 50% overlap. Three-pole AR models were estimated, thereby obtaining the residual signal variance (r) for each window. Preliminary

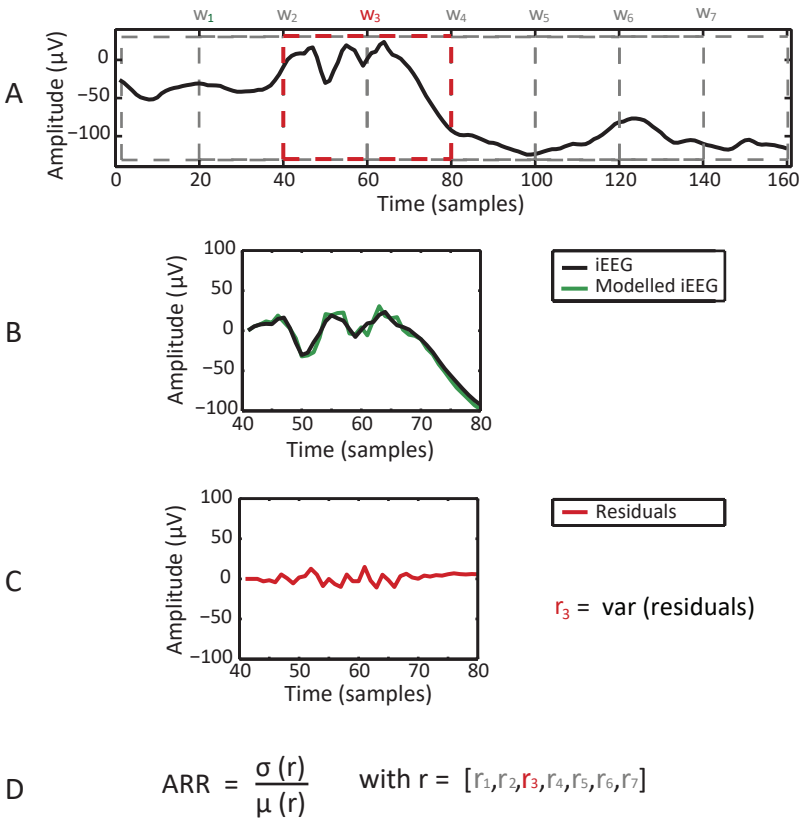


FIGURE 3.2.
The processing steps executed by the ARR algorithm. A) The iEEG epoch, which has a length of 160 samples in this example, is divided into windows of 40 samples, with 50% overlap. B) Of each window, an order 3 AR model is estimated. C) The residual signal variation (r) in the window is obtained. D) ARR is calculated for the iEEG epoch, defined as the standard deviation of the r values from the epoch, divided by the mean of those r values.

investigation in a part of the data (subjects 1, 2, 3, 4, 5, and 7) showed that a 40 sample window length and model order 3 provided the best distinction between channels inside and channels outside the SOZ. High-frequency components of the EEG are presumably quantified with these parameters. Because of the 40-sample window length used, the lowest oscillatory frequency that can be modeled in this window, using a sampling frequency of 1000 Hz, is (1/40) ∗ 1000 = 25 Hz. Nonharmonicities in low frequencies therefore do not influence *r*.

The variation of *r* in time was quantified by calculating the coefficient of variation, providing an ARR value for each channel:

$$ARR = \frac{\sigma(r)}{\mu(r)}. \quad (3.2)$$

Here σ is the standard deviation and μ is the mean for the series of time windows. Using the coefficient of variation instead of the standard deviation provides normalization of values, thus allowing comparison of results between subjects.

STATISTICAL ANALYSIS

COMPARISON BETWEEN R AND TRL RESULTS

The TRL method and R technique are directly comparable; the association between them is a measure of the quality of the automated ripple quantification given the manual segmentation of ripples as gold standard. The R results from the iEEG of all channels were compared with TRL results, using the h^2 index. This index provides a measure for the association between pairs of signals in general (linear or nonlinear) and information on the directionality of the association.[38] We obtained the h^2 index in two directions, to test the fraction of variation of R that can be explained by TRL, and vice versa; to test the fraction of variation of TRL that can be explained by R. This results in a value between 0 and 1, where $h^2(TRL,R) = 1$ implies that all variation of TRL is explained by R. The findings are significant if h^2 exceeds the critical value, or significance margin, in a bootstrapped test with $\alpha = 0.05$.[38]

COMPARISON OF TRL, R, AND ARR RESULTS WITH THE SOZ

TRL, R and ARR results were compared with the SOZ, for each individual subject, and on group level, by calculating separability and detection performance. Separability, or the ability of each method to distinguish between signals from channels in the SOZ as opposed to those outside it, was assessed using a Kolmogorov-Smirnov test (K-S test). The K-S test was used in two ways: 1) to determine whether the distributions of the values from channels inside versus from outside the SOZ were distinctly different, and 2) to test the hypothesis that the values obtained from channels in the SOZ are higher than those obtained outside the SOZ (both calculations with $\alpha = 0.01$).

The performance of TRL, R, and ARR at approximating the SOZ was determined using a receiver operating characteristic (ROC) curve, drawn

with sensitivity and specificity for various thresholds. The ROC's area under the curve (AUC) represents algorithm performance. Sensitivity is defined as the proportion of SOZ channels which are correctly classified as such. Specificity is defined as the proportion of channels not in the SOZ which are correctly classified as such. A channel was classified as positive (i.e. being potentially in the SOZ) by one of the three methods, when its value exceeds a varying threshold. Otherwise, the channel is classified as negative according to that method. Performances were determined from the TRL, R, and ARR results per subject and at group level. In the latter case, all channels from all subjects are taken together.

Algorithm performance was also calculated separately for only the cases with good postoperative seizure control. Assuming that in those cases the SOZ as delineated by the gold standard contains the whole epileptogenic area, this analysis shows possible causal relationships between HFO generation and nonharmonic dynamics, and the processes of transitions to seizures.

ARR ALGORITHM PERFORMANCE AND SURGICAL OUTCOME

We also investigated whether surgical outcome can possibly explain the performance results of the ARR algorithm. For this purpose, ARR performance results are viewed in respect to surgical outcome of all subjects. Poorer performance of the ARR algorithm to approximate the SOZ can be caused by either false positives (FPs) or false negatives (FNs). ARR FPs occurring more in data from subjects with poor postoperative seizure control, could suggest that some of the electrodes corresponding to those FPs were actually situated in an epileptogenic area. This area could have been missed in the (gold standard) delineation of the SOZ, or the area could be outside the SOZ, while situated in the EZ. ARR FNs occurring in data from subjects with poor postoperative seizure control suggest that electrodes corresponding to those FNs were actually situated outside the epileptogenic area. It is therefore also informative to investigate whether diminished performance of the ARR algorithm is caused by FPs rather than FNs, or vice versa. This was done by comparing ARR values from subjects with good postoperative seizure control with ARR values from subjects with poor postoperative seizure control. This comparison is made separately for channels recorded from inside and outside the SOZ.

ARR PERFORMANCE AND ANTIEPILEPTIC DRUGS

Antiepileptic drug (AED) tapering is standard operating procedure to increase the chance of seizures during iEEG registration. The day of

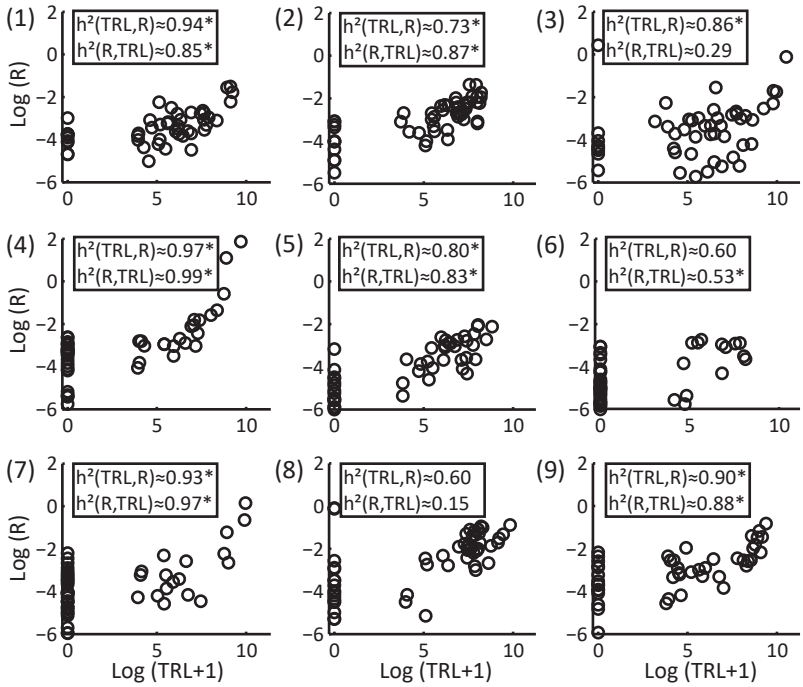


FIGURE 3.3.
Scatterplots for each subject showing total ripple length (TRL) results versus rippleness algorithm (R) results, both logarithmically scaled. Each data point represents an individual channel. The h2 indexes are shown in the boxes, followed by an asterisk if the found association was significant ($p < 0.05$). On the x-axis is log(TRL + 1), so that when no ripples were found visually (TRL = 0), there is a datapoint at log(TRL+1) = 0.

registration after drug tapering differs between subjects (see Table 3.1), because the choice of stored interictal data was limited. We investigated whether the registration day of the used iEEG recording, reflecting the level of AED tapering, could explain performance results of the ARR algorithm.

RESULTS

TRL AND R RESULTS

TRL and R results are shown in scatterplots per subject in Figure 3.3. Results from h^2 signal association analysis are shown in the boxes in each boxplot. The rising trends in the scatterplots, especially where TRL > 0, suggest that R and TRL values are correlated. This correspondence is not strictly monotonic, and some scatter is observed. The h^2 index findings show significant signal association in both directions (R to TRL and TRL to R)

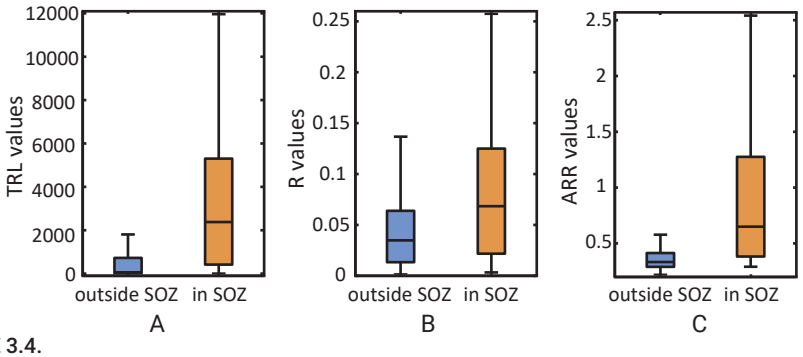


FIGURE 3.4.

Distributions of A) TRL, B) R, and C) ARR values per channel inside versus outside the seizure onset zone. Only iEEG channels of the subjects with a good postoperative seizure control (subjects 1, 2, 4, 6–9) were used to obtain the distributions. Outliers are not shown.

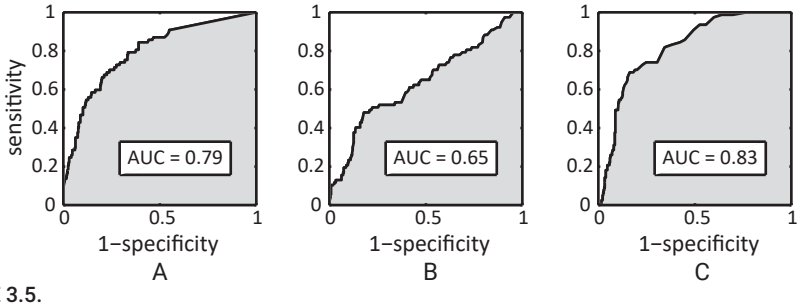


FIGURE 3.5.

ROC curves for the performance of A) TRL, B) R, and C) ARR for predicting the seizure onset zone. ROC area under the (AUC) is indicated in the textboxes. Only results from subjects with good postoperative seizure control (subjects 1, 2, 4, 6–9) were used to obtain ROC curves.

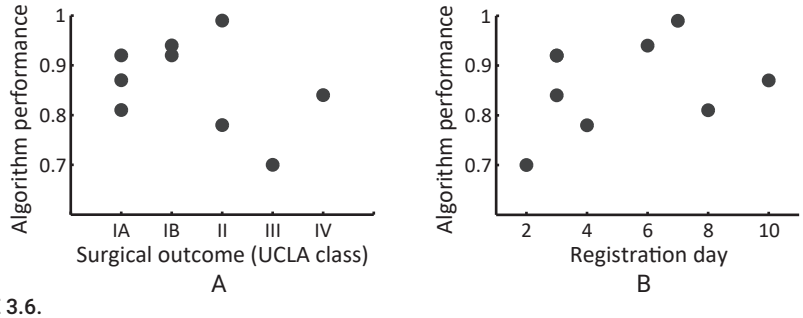


FIGURE 3.6.

Scatter plots showing the influence of A) surgical outcome and B) iEEG registration day after medication lowering on algorithm performance. Performance (in ROC AUC) of the ARR algorithm is shown on the y-axis. Each data point represents the results found for one subject. In panel B), subjects 1 and 8 share the same data point [3, 0.92].

in the iEEG data from subjects 1, 2, 4, 5, 7, and 9. In data from subjects 3 and 6, signal association was only significant in one direction (TRL to R and R to TRL, respectively). Signal association is thus not favored in one particular direction. In subject 8, signal association was not significant in either direction. When combining all R and TRL data points from the total dataset, $h^2(TRL,R) \approx 0.99$, and $h^2(R,TRL) \approx 0.81$ were found, both significant with $p<0.01$.

TRL, R, AND ARR RESULTS IN RELATION TO THE SOZ

The distributions of TRL, R, and ARR values inside and outside the SOZ from individuals with good postoperative seizure control are shown in Figure 3.4. The applied K–S tests showed that TRL, R, and ARR values from inside and outside the SOZ have independent and distinctly different distributions ($p < 0.01$). Furthermore, the values for signals from channels in the SOZ were higher than values for the signals obtained from channels outside the SOZ for all three methods ($p < 0.01$).

The performance of TRL, R, and ARR at approximating the SOZ is provided per subject in Table 3.2. For all subjects except subject 3, ARR provides the highest ROC AUC. Performance of TRL, R, and ARR on group level with only data from subjects with good postoperative seizure control is shown in the ROC curves in Figure 3.5.

ARR PERFORMANCE AND SURGICAL OUTCOME

In Figure 3.6A, ARR performances are plotted against surgical outcome per subject. A declining trend can be observed for ARR performance with poorer surgical outcome. An increasing trend was seen for later registration days. Linear regression performed on both associations was not statistically significant.

In Figure 3.7, distributions of ARR values for channels outside the SOZ (according to the gold standard for SOZ delineation) are shown for good postoperative seizure control and poor postoperative seizure control. The shown distributions hold only alleged true negative and false positive findings. Distributions of ARR values for channels inside the SOZ hold only the alleged true positive and false negative findings (results not shown). Applying K–S tests on the results from channels *outside* the SOZ showed: separate distributions for good versus poor postoperative seizure control ($p < 0.05$), and significantly higher ARR values in subjects with poor surgical outcome than in subjects with good surgical outcome ($p < 0.01$). Applying

TABLE 3.2.
Performance expressed in ROC AUC values per subject for the three classifiers. Group performance for all subjects and for the good postoperative seizure control group (UCLA IA, IB, II) are shown separately.

Subject	Surgical outcome (good/poor)	TRL	R	ARR
1	UCLA IA (good)	0.88	0.60	0.92
2	UCLA IA (good)	0.78	0.66	0.81
3	UCLA III (poor)	0.87	0.48	0.70
4	UCLA II (good)	0.92	0.81	0.99
5	UCLA IV (poor)	0.71	0.53	0.84
6	UCLA IB (good)	0.79	0.46	0.94
7	UCLA II (good)	0.75	0.77	0.78
8	UCLA IB (good)	0.83	0.67	0.92
9	UCLA IA (good)	0.82	0.63	0.87
All		0.79	0.64	0.82
Only UCLA IA, IB, II		0.79	0.65	0.83

K-S tests on the results from channels *inside* the SOZ showed nonsignificant separation of distributions for good versus poor outcome, and ARR values were not significantly higher in subjects with good outcome compared to those in subjects with poor outcome. Thus, there are significantly more false positives, rather than false negatives, in the poor postoperative seizure control group than in the good postoperative seizure control group.

ARR PERFORMANCE AND ANTIEPILEPTIC DRUGS

In Figure 3.6B ARR algorithm performance is plotted against the registration day after medication lowering for each subject. A rising trend can be observed in the ARR results with increasing registration day. Linear regression performed on this data was not statistically significant.

DISCUSSION

The results from this study suggest that ARR, interpreted as evidence for nonharmonicity measured in short interictal epochs of iEEGs, may be used to automatically identify channels in the SOZ, a surrogate marker for the EZ, in people with temporal lobe epilepsy. The ARR algorithm more reliably identified the SOZ than the R algorithm and the TRL, although the difference in performance is smaller in the latter case. The findings in this

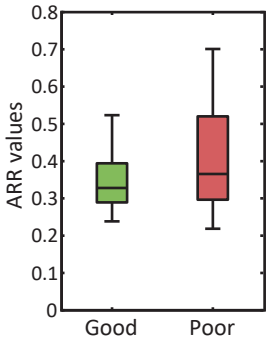


FIGURE 3.7.
Distributions of ARR values for channels outside the SOZ, for good versus poor postoperative seizure control. Outliers are not shown.

study also suggest correspondence between TRL and R values per channel, thus confirming results from our previous study.[33] Therefore, the R measure can possibly be used for fast automated estimation of the TRL.

Three methods for reliable approximation of the SOZ - TRL, R, and ARR - were compared to the current gold standard for determining the SOZ. All three methods possess the ability to distinguish between signals obtained from channels in the SOZ and those from outside it. The ARR algorithm outperforms TRL and R in the extent to which the algorithm achieves correct classification of channels in the SOZ. Moreover, the ARR and R algorithms process a 5-min 1000 Hz sampled epoch of iEEG data rapidly (within a minute). They might therefore be used as an alternative to the current gold standard of SOZ determination (visual observation of ictal recordings from invasive video-EEG) and manual segmentation of HFOs, which are both very time consuming and can take several hours.

SURGICAL OUTCOME AND ARR RESULTS

ARR performance shows a declining trend with poorer surgical outcome. Linear regression was however not significant, possibly due to the small number of subjects. There are more false positives (rather than false negatives) in the poor outcome group than in the good outcome group. These results imply that diminished ARR algorithm performance with poorer surgical outcome is most likely due to false positives. Some of the channels corresponding to these false positives could actually be located in the EZ and could thus potentially be used for improvement of EZ localization.

AED LEVELS AND ARR RESULTS

In the current study it was not possible to control for AED tapering, which could influence TRL, R, ARR, and even gold standard SOZ results.[39,40] A rising trend can be observed in the ARR performance at later recording days, but it was not found to be statistically significant (possibly due of the limited number of subjects in this study). A similar rising trend was observed in the TRL results (results not shown). These findings could imply better agreement between the TRL and ARR algorithms and the SOZ when fewer AEDs or a lower concentration of them is used. HFOs have been found to be attenuated by AEDs,[39] and evidence for the attenuation of the values measured with certain nonlinear quantifiers has also been found.[40] It could be that due to this attenuation, the difference between normal and epileptic brain areas, regarding HFOs and nonharmonicity, is smaller. If this hypothesis is correct, it would be advisable to quantify HFOs and nonharmonicity in an interictal iEEG epoch from a later registration day, when AED levels have decreased. The effect of AEDs on ARR results should therefore be further investigated.

EXAMINATION OF STUDY RESULTS

Because the AR model describes harmonic oscillators of given order that are stationary and linear, the ARR measures the nonharmonic component in the system dynamics. Typically the residual variance of the signal after AR fit is interpreted as noise, but in our application it can be reminiscent of the nonlinear or nonstationary dynamic patterns of the underlying system generating the signal.[41] Assuming that the noisy component is always present, the intermittent generation of nonharmonic events such as ripples will cause variations in the r quantity, or the residual signal variation after AR modeling. Because to obtain an ARR value the coefficient of variation of r rather than the mean value per channel is calculated, we can assume that the noisy source of ARR has been largely suppressed.

Various quantifiers, which characterize nonlinearity related to a single recording site, have been used for lateralization and localization of the EZ. Used quantifiers include neuronal complexity loss,[25,42] correlation dimension,[27,28] Lyapunov exponents,[29] ξ,[30] mean phase coherence,[31] and relative phase clustering index (rPCI).[32] The ARR algorithm provides global evidence of nonharmonicity of system dynamics, meaning that the specific type of nonlinearity or nonstationarity causing the high ARR values, is disregarded.

One might argue that by focusing on the frequently appearing high residual values in EEG data, ARR actually detects interictal spikes, as previous research showed is possible.[43,44] Because spikes were not quantified in this study, it cannot be confirmed whether this is the case or not. However, spikes are also often observed in channels outside the SOZ.[13] The zone with HFOs often has a smaller spatial extent than the area of the cortex generating spikes.[4] Results from this study suggest similar or better performance of ARR compared with manually segmented HFOs. Furthermore, it is noted by Lehnertz that certain nonlinear quantifiers of the EEG do not necessarily coincide with interictal discharges, but seem to coincide quite well with the EZ.[45] It therefore seems unlikely that ARR locates spike-generating cortex rather than the SOZ.

It was observed in the study data that in the signals from channels in the SOZ, the iEEG windows with high r values often coincide with HFOs. In signals from channels contralateral to a SOZ that showed evidence of HFOs, r values were not high in iEEG windows during HFOs. This is consistent with findings from preliminary investigation, where we found that HFOs from the SOZ had higher nonharmonicity (results not shown). We hypothesize therefore, that at least two properties are needed in order for a neuronal network to be prone to producing seizures, i.e. recurrent connections and evidence of nonharmonicity in its behavior. A combination of recurrent connections and nonharmonicity promotes a pathological oscillatory state, and thus increases the chance of seizure occurrence.

VALUE OF THE ARR METHOD

An advantage of the ARR algorithm is that it yields results rapidly, as opposed to visual observation of video-EEG, manual segmentation of HFOs, and some nonlinear time series analysis algorithms[46] that are characterized by high computational complexity. Also, the algorithm's findings are not sensitive to spurious, filter-induced oscillations, which can be mistaken for HFOs when manually segmenting those.

ARR could potentially be used in combination with the current gold standard for delineation of the SOZ as an extra control during presurgical workup. Sub-optimal postoperative seizure control could, among other reasons, be caused by sub-optimal electrode placement or even a SOZ which is unidentifiable with the implantation method used.[47] When the ARR findings and the SOZ as found by the current gold standard are not in agreement, this could suggest that poorer postoperative seizure control

might be expected. Combining ARR with the current gold standard for delineation of the SOZ could therefore be beneficial for presurgical workup.

STUDY LIMITATIONS

ARR results were compared with the current gold standard for delineating the SOZ, which is a surrogate marker for the EZ. Sub-optimal agreement of a new algorithm with an imperfect gold standard is possible, even if the algorithm performs its task well. False positives produced by any algorithm in light of the SOZ could in fact be true positives in light of the EZ.

There were some false positive detections by ARR in data from subjects where the gold standard SOZ was in all probability correct, with good postoperative seizure control. When visually inspecting the EEG data from these channels, two explanations were found. Firstly, in some channels activity resembling continuous rippling can be observed. This continuous high-frequency activity has recently been found to be a physiological rather than pathological finding.[48] Secondly, FPs could be caused by recurrent artefacts, the morphology of which could give rise to spuriously high residual values. Small numbers of artefacts are, however, unlikely to influence the CV used to obtain ARR; the artefacts would have to be recurrent to give rise to high ARR values. Additionally, channels with many artefacts are often excluded from EEG analysis, if the channels are unlikely to be located in the SOZ.

The choice of 5-min slow-wave sleep epochs of iEEG for analysis was based on findings regarding the stability of the HFO rate.[15] Often only one short nightly interictal iEEG recording was stored, thus investigating ARR stationarity during the total registration, which often lasted a week, was not possible.

FUTURE WORK

Further investigations are justified by the promising results from the pilot test performed in this study. Validation of the ARR algorithm for its added value in the identification of the SOZ as well as in the delineation of the EZ is required. In such a validation study ARR findings should be compared with long term surgical outcome, as this is currently the only way to derive the EZ. Before performing such a validation study, a number of properties should be further investigated.

Firstly, optimization of the chosen length of the AR modeling window should be investigated. As mentioned in the methods section, the ARR algorithm possesses filtering capacities in accordance with the chosen window length. One could decrease the window length to focus on even higher frequencies (e.g. in the ripple, or even fast ripple range). In doing this, however, one is limited by the sampling frequency; the modeling window (40 samples in this study) must have sufficient samples to model the signal adequately.

Secondly, general stationarity of the measured EEG characteristic should be further investigated by comparing results from various registration nights and from various epoch lengths. That way the appropriate recording moment and length of an iEEG sample epoch can be determined. It is possible for example, that ARR values in SOZ channels increase when there is an increased chance of a seizure. AED concentration fluctuations during tapering and reintroduction may influence ARR results as well. Furthermore, the choice of a sample of slow-wave sleep as opposed to other epoch sampling methodologies in wakefulness as well as in sleep for ARR determination should be further studied.

REFERENCES

1. Blume WT, Lüders HO, Mizrahi E, et al. Glossary of descriptive terminology for ictal semiology: report of the ILAE task force on classification and terminology. *Epilepsia*. 2001;42:1212–8.
2. Lüders HO, Najm I, Nair D, et al. The epileptogenic zone: general principles. *Epileptic Disord*. 2006;8:S1–9.
3. Rosenow F, Lüders H. Presurgical evaluation of epilepsy. *Brain*. 2001;124:1683–700.
4. Zijlmans M, Jiruska P, Zelmann R, et al. High-frequency oscillations as a new biomarker in epilepsy. *Ann Neurol*. 2012;71:169–78.
5. Bragin A, Engel J, Wilson CL, et al. Hippocampal and entorhinal cortex high-frequency oscillations (100--500 Hz) in human epileptic brain and in kainic acid--treated rats with chronic seizures. *Epilepsia*. 1999;40:127–37.
6. Worrell GA, Parish L, Cranstoun SD, et al. High-frequency oscillations and seizure generation in neocortical epilepsy. *Brain*. 2004;127:1496–506.
7. Urrestarazu E, Jirsch JD, LeVan P, et al. High-frequency intracerebral EEG activity (100-500 Hz) following interictal spikes. *Epilepsia*. 2006;47:1465–76.
8. Bragin A, Engel J, Staba RJ. High-frequency oscillations in epileptic brain. *Curr Opin Neurol*. 2010;23:151–6.
9. Staba RJ, Wilson CL, Bragin A, et al. High-frequency oscillations recorded in human medial temporal lobe during sleep. *Ann Neurol*. 2004;56:108–15.
10. Bagshaw AP, Jacobs J, LeVan P, et al. Effect of sleep stage on interictal high-frequency oscillations recorded from depth macroelectrodes in patients with focal epilepsy. *Epilepsia*. 2009;50:617–28.
11. Jacobs J, Zijlmans M, Zelmann R, et al. High-frequency electroencephalographic oscillations correlate with outcome of epilepsy surgery. *Ann Neurol*. 2010;67:209–20.
12. Wu JY, Sankar R, Lerner JT, et al. Removing interictal fast ripples on electrocorticography linked with seizure freedom in children. *Neurology*. 2010;75:1686–94.
13. Jacobs J, LeVan P, Chander R, et al. Interictal high-frequency oscillations (80-500 Hz) are an indicator of seizure onset areas independent of spikes in the human epileptic brain. *Epilepsia*. 2008;49:1893–907.
14. Gardner AB, Worrell GA, Marsh E, et al. Human and automated detection of high-frequency oscillations in clinical intracranial EEG recordings. *Clin Neurophysiol*. 2007;118:1134–43.
15. Zelmann R, Zijlmans M, Jacobs J, et al. Improving the identification of High Frequency Oscillations. *Clin Neurophysiol*. 2009;120:1457–64.
16. Staba R, Wilson C, Bragin A, et al. Quantitative analysis of high-frequency oscillations (80–500 Hz) recorded in human epileptic hippocampus and entorhinal cortex. *J Neurophysiol*. 2002;88:1743–52.
17. Firpi H, Smart O, Worrell G, et al. High-frequency oscillations detected in epileptic networks using swarmed neural-network features. *Ann Biomed Eng*. 2007;35:1573–84.
18. Crépon B, Navarro V, Hasboun D, et al. Mapping interictal oscillations greater than 200 Hz recorded with intracranial macroelectrodes in human epilepsy. *Brain*. 2010;133:33–45.
19. Birot G, Kachenoura A, Albera L, et al. Automatic detection of fast ripples. *J Neurosci Methods*. 2012;1–14.

20. Dümpelmann M, Jacobs J, Kerber K, et al. Automatic 80-250Hz "ripple" high frequency oscillation detection in invasive subdural grid and strip recordings in epilepsy by a radial basis function neural network. *Clin Neurophysiol*. 2012;123:1721–31.
21. Csicsvari J, Hirase H, Czurkó A, et al. Fast network oscillations in the hippocampal CA1 region of the behaving rat. *J Neurosci*. 1999;19:RC20.
22. Curio G, Mackert B-MM, Burghoff M, et al. Localization of evoked neuromagnetic 600 Hz activity in the cerebral somatosensory system. *Electroencephalogr Clin Neurophysiol*. 1994;91:483–7.
23. Bénar CG, Chauvière L, Bartolomei F, et al. Pitfalls of high-pass filtering for detecting epileptic oscillations: a technical note on "false" ripples. *Clin Neurophysiol*. 2010;121:301–10.
24. Jmail N, Gavaret M, Wendling F, et al. A comparison of methods for separation of transient and oscillatory signals in EEG. *J Neurosci Methods*. 2011;199:273–89.
25. Lehnertz K, Elger CE. Spatio-temporal dynamics of the primary epileptogenic area in temporal lobe epilepsy characterized by neuronal complexity loss. *Electroencephalogr Clin Neurophysiol*. 1995;95:108–17.
26. Elger CE, Widman G, Andrzejak R, et al. Nonlinear EEG analysis and its potential role in epileptology. *Epilepsia*. 2000;41 Suppl 3:S34–8.
27. Casdagli MC, Iasemidis LD, Savit RS, et al. Non-linearity in invasive EEG recordings from patients with temporal lobe epilepsy. *Electroencephalogr Clin Neurophysiol*. 1997;102:98–105.
28. Jing H, Takigawa M, Benasich AA. Relationship of Nonlinear Analysis, MRI and SPECT in the Lateralization of Temporal Lobe Epilepsy. *Eur Neurol*. 2002;48:11–9.
29. Kowalik ZJ, Schnitzler A, Freund H, et al. Local Lyapunov exponents detect epileptic zones in spike-less interictal MEG recordings. *Clin Neurophysiol*. 2001;112:60–7.
30. Andrzejak RG, Widman G, Lehnertz K, et al. The epileptic process as nonlinear deterministic dynamics in a stochastic environment: an evaluation on mesial temporal lobe epilepsy. *Epilepsy Res*. 2001;44:129–40.
31. Mormann F, Lehnertz K, David P, et al. Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients. *Phys D*. 2000;144:358–69.
32. Kalitzin S, Velis D, Suffczynski P, et al. 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.
33. Kalitzin S, Zijlmans M, Petkov G, et al. Quantification of spontaneous and evoked HFO's in SEEG recording and prospective for pre-surgical diagnostics. Case study. In: *34th Annu Int Conf IEEE EMBS*. 2012. p. 1024–7.
34. Engel Jr J, van Ness P, Rasmussen T, et al. Outcome with respect to epileptic seizures. In: Surgical treatment of the epilepsies. 2nd ed. New York: Raven Press; 1993. p. 609–21.
35. Otsubo H, Ochi A, Imai K, et al. High-frequency oscillations of ictal muscle activity and epileptogenic discharges on intracranial EEG in a temporal lobe epilepsy patient. *Clin Neurophysiol*. 2008;119:862–8.
36. Niedermeyer E, Lopes da Silva F. Electroencephalography: Basic Principals, Clinical Applications and Related Fields. Philadelphia: Lipponcott and Wilkins; 2011.
37. Franaszczuk P, Blinowska K. Linear model of brain electrical activity - EEG as a superposition of damped oscillatory modes. *Biol Cybern*. 1985;53:19–25.
38. Kalitzin SN, Parra J, Velis DN, et al. Quantification of unidirectional nonlinear associations between multidimensional signals. *IEEE Trans Biomed Eng*. 2007;54:454–61.

39. Zijlmans M, Jacobs J, Zelmann R, et al. High-frequency oscillations mirror disease activity in patients with epilepsy. *Neurology*. 2009;72:979–86.

40. 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.

41. Rieke C, Mormann F, Andrzejak RG, et al. Discerning nonstationarity from nonlinearity in seizure-free and preseizure EEG recordings from epilepsy patients. *IEEE Trans Biomed Eng*. 2003;50:634–9.

42. Widman G, Lehnertz K, Urbach H, et al. Spatial distribution of neuronal complexity loss in neocortical lesional epilepsies. *Epilepsia*. 2000;41:811–7.

43. Lopes da Silva FH, Dijk A, Smits H. Detection of nonstationarities in EEGs using the autoregressive model: an application to EEGs of epileptics. In: Kunkel H, Dolce G, editors. CEAN: computerized EEG analysis. Stuttgart: Gustav Fischer Verlag; 1975. p. 180–99.

44. Tzallas AT, Oikonomou VP, Fotiadis DI. Epileptic spike detection using a Kalman filter based approach. *Conf Proc IEEE Eng Med Biol Soc*. 2006;1:501–4.

45. Lehnertz K. Epilepsy and nonlinear dynamics. *J Biol Phys*. 2008;34:253–66.

46. Andrzejak RG, Mormann F, Widman G, et al. Improved spatial characterization of the epileptic brain by focusing on nonlinearity. *Epilepsy Res*. 2006;69:30–44.

47. Harroud A, Bouthillier A, Weil AG, et al. Temporal lobe epilepsy surgery failures: a review. *Epilepsy Res Treat*. 2012;2012:1–10.

48. Melani F, Zelmann R, Mari F, et al. Continuous High Frequency Activity : A peculiar SEEG pattern related to specific brain regions. *Clin Neurophysiol*. 2013;124:1507–16.

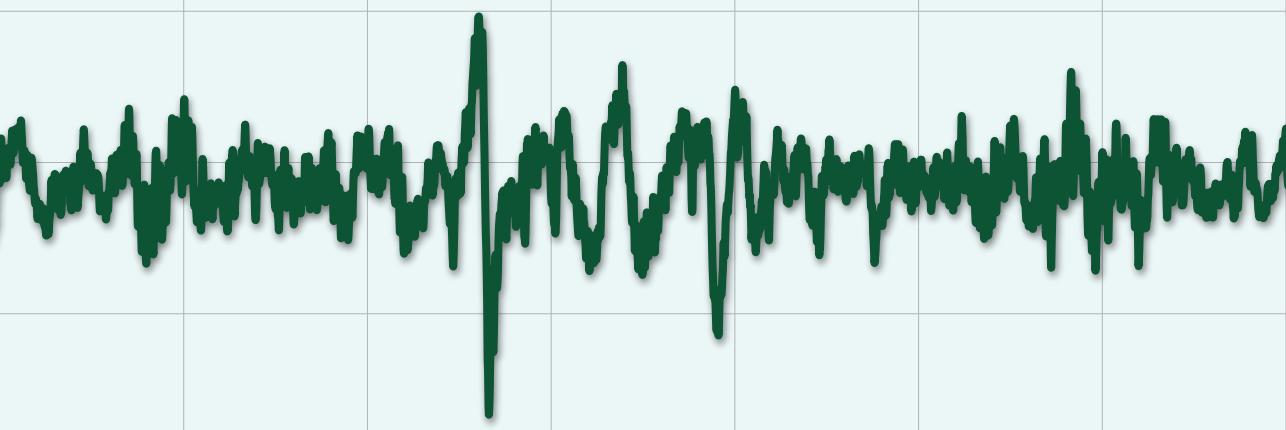
SUPPLEMENTARY MATERIALS

TABLE 3.S1.
Implanted intracranial electrodes per subject, and used abbreviations.

Subj	Abbre-viation	Electrode location
1	AML	Left amygdala
	ATHCL	Left anterior-temporal hippocampus
	IAIL	Left anterior insula
	MTHCL	Left mid-temporal hippocampus
	IAPL	Left insula anterior posterior
	PTlL	Left parieto-temporal
	TPL	Left temporal pole
	FBR	Right fronto-basal
	AMR	Right amygdala
	ATHCR	Right anterior-temporal hippocampus
	IPAR	Right insula posterior anterior
	PTlR	Right parieto-temporal
2	LesAntR	Right lesion anterior
	LesSupR	Right lesion superior
	LesAnIR	Right lesion anterior insula
	TPamR	Right temporal pole amygdala
	HCR	Right hippocampus
	FOrbR	Right fronto-orbital
	FPoIR	Right fronto-polar insula
	DLSMAR	Right dorso-lateral supplementary motor area
	LGCAnR	Right lateral gyrus cinguli anterior
	LGCMdR	Right lateral gyrus cinguli medial
	InsAntR	Right insula anterior
3	FBL	Left fronto-basal
	HAL	Left hippocampus anterior
	FBR	Right fronto-basal
	AR	Right amygdala
	HAR	Right hippocampus anterior
	HPR	Left hippocampus posterior
	TBR	Right temporo-basal
	TPOR	Right temporo-parietal-occipital

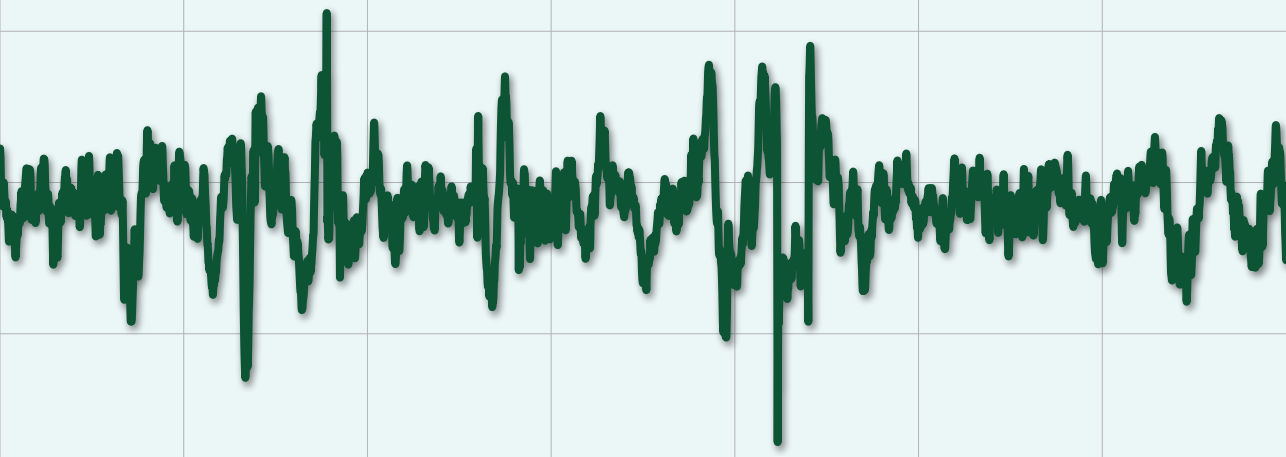
	IAIR	Right insula anterior
	IPIR	Right insula posterior inferior
	TOPAR	Right temporal-occipital posterior amygdala
4	LFL	Left lateral frontal
	INL	Left insula
	HCL	Left hippocampus
	TOR	Right temporo-occipital
	FOR	Right frontal operculum
	LFR	Right lateral frontal
	INR	Right insula
	MFR	Right medial frontal
	HCR	Right hippocampus
5	PSMAL	Left posterior supplementary motor area
	FODL	Left frontal operculum
	PFL	Left parietal frontal
	FORBL	Left fronto-orbital
	TOL	Left temporal operculum
	HCL	Left hippocampus
	HCR	Right hippocampus
	PSMAR	Right posterior supplementary motor area
6	FOL	Left fronto-orbital
	OFL	Left frontal operculum
	GCL	Left gyrus cinguli
	MFL	Left medial frontal
	OTL	Left temporal operculum
	HCL	Left hippocampus
	AML	Left amygdala
	OFR	Right frontal operculum
	MFR	Left medial frontal
	HCR	Right hippocampus
7	LFL	Left lateral frontal
	MHL	Left medial hippocampus
	IAR	Right insula anterior
	IPR	Right insula posterior
	IIR	Right insula inferior
	FCR	Right frontal cingular
	ATR	Right anterior temporal

	AHR	Right anterior hippocampus
	MHR	Right medial hippocampus
	PHR	Right posterior hippocampus
	FBR	Right fronto-basal
	TPOR	Right temporo-parieto-occipital
8	HippL	Left hippocampus
	FrInsR	Right frontal insula
	Peril	Peri lesional
	FrOrbR	Right fronto-orbital
	Antgr	Anterior grid
	Postgr	Posterior grid
	HippR	Right hippocampus
9	FBL	Left fronto-basal
	TPL	Left temporal pole
	HC1L	Left hippocampus
	HC2L	Left hippocampus
	AML	Left amygdala
	IAL	Left insula anterior
	TBL	Left temporo-basal
	TPOL	Left temporo-parieto-occipital
	S2PIL	Left posterior insula
	MIL	Left mid-insula
	FBR	Right fronto-basal
	TPR	Right temporal pole
	HC1R	Right hippocampus
	TBR	Right temporo-basal
	TPOR	Right temporo-parieto-occipital



CHAPTER 4

NON-HARMONICITY IN HIGH-FREQUENCY COMPONENTS OF THE INTRA-OPERATIVE CORTICOGRAM TO DELINEATE EPILEPTOGENIC TISSUE DURING SURGERY



Based on:
Geertsema EE*, van ’t Klooster MA*, van Klink NEC, Leijten FSS, van Rijen PC, Visser GH, Kalitzin SN, Zijlmans M. Non-harmonicity in high-frequency components of the intra-operative corticogram to delineate epileptogenic tissue during surgery. Clinical Neurophysiology. 2017; 128: 153-164.
* Shared first authors

ABSTRACT

OBJECTIVE

We aimed to test the potential of auto-regressive model residual modulation (ARRm), an artefact-insensitive method based on non-harmonicity of the high-frequency signal, to identify epileptogenic tissue during surgery.

METHODS

Intra-operative electrocorticography (ECoG) of 54 patients with refractory focal epilepsy were recorded pre- and post-resection, sampled at 2048 Hz. The ARRm was calculated in one-minute epochs in which high-frequency oscillations (HFOs; fast ripples, 250–500 Hz; ripples, 80–250 Hz) and spikes were marked. We investigated the pre-resection fraction of HFOs and spikes explained by the ARRm (h^2-index). A general ARRm threshold was set and used to compare the ARRm to surgical outcome in post-resection ECoG (Pearson X^2).

RESULTS

ARRm was associated strongest with the number of fast ripples in pre-resection ECoG ($h^2 = 0.80$, $P<0.01$), but also with ripples and spikes. An ARRm threshold of 0.47 yielded high specificity (95%) with 52% sensitivity for channels with fast ripples. ARRm values >0.47 were associated with poor outcome at channel and patient level (both $P<0.01$) in post-resection ECoG. High ARRm results post-resection had a positive predictive value of 88.9% for poor outcome, including in patients without fast ripples.

CONCLUSIONS

The ARRm algorithm might enable intra-operative delineation of epileptogenic tissue. ARRm is the first unsupervised real-time analysis that could provide an intra-operative, 'on demand' interpretation per electrode about the need to remove underlying tissue to optimize the chance of seizure freedom.

INTRODUCTION

Epilepsy is a dynamic condition as seizures occur intermittently and in most cases unexpectedly. This suggests that the underlying processes responsible for the generation and cessation of seizures are non-harmonic, i.e. non-linear and non-stationary. A high level of non-linearity in the system is needed to initiate seizures.[1] The epileptiform EEG contains several biomarkers which all have a strong non-linear component.[2,3] This holds true for epileptiform inter-ictal spikes, for the recently discovered high-frequency oscillations (HFOs; >80 Hz) and for seizures themselves. All these events appear without notice.[4–6]

During epilepsy surgery there is need for biomarkers to delineate the epileptogenic tissue that should be resected to assure seizure freedom after surgery. Even in the presence of an epileptogenic lesion on MRI the resection boundaries are not always clear, e.g. in cases of neuronal tumors, in which the epileptogenic zone may extend beyond the lesion, or when the hippocampus is secondarily affected.[7] Epilepsy surgery can be tailored by identifying spikes in the intra-operative electrocorticogram (ECoG), but this method is controversial.[8,9] Surgical removal of cortex showing fast ripples, between 250 and 500 Hz, has been associated with post-operative seizure freedom.[9,10] HFOs probably result from hyper synchronous, out-of-phase firing of groups of principal neurons,[11] and might be more local to the epileptogenic tissue than spikes.[11,12] It may thus be better for the tailoring of epilepsy surgery with ECoG to rely on HFOs, especially fast ripples, rather than on the currently used spikes.[9]

Visual analysis of HFOs is time consuming and requires the presence of expert reviewers during surgery.[13] The time available to record and analyze HFOs is limited, and HFO occurrence is infrequent, so they can easily be missed.[9,14] Automatic detectors have been developed, but most still require extensive computing time,[12,15–21] making these techniques unsuitable for use in the time frame of surgery. Additionally, it can be difficult to distinguish pathological HFOs from physiological high-frequency activity. High-frequency physiological activity probably shows a harmonic pattern.[22] It can also be difficult to distinguish HFOs from artefacts, while artefact sources, such as the diathermy and the surgical microscope, are ubiquitous in the surgical theatre.[23] Also, detection software for clinical use in the operating theatre needs to function near real-time, since prolonged surgery is associated with increased complication risks.[24] Analysis of non-harmonicity in high-frequency components of the EEG signal, recorded at a high sampling rate

>2000 Hz, may provide an automated and objective solution that solves these issues.

The auto-regressive model residual modulation (ARRm) can be used to predict which intracranial EEG channels are within the presumed epileptogenic area.[25] We developed the ARRm algorithm to identify 'bad' channels. The ARRm reflects the amount of non-harmonicity in the high-frequency components of the signal, in terms of high residual signal variation after autoregressive modelling. The ARRm can be computed rapidly and only requires short epochs of EEG and thereby seems to lend itself to online analysis during surgery.[25] We adjusted the ARRm algorithm to reduce the influence of typical intra-operative ECoG artefacts; artefacts may produce spuriously large residual signal variations, resulting in false positive high ARRm values unrelated to the epileptic tissue. We focused on the rejection of subtle short-lasting electrode and movement artefacts, as these are the most difficult to identify.

We aimed to test the potential of the new ARRm algorithm to identify epileptogenic tissue during surgery. Therefore, we compared the ARRm to the occurrence of fast ripples, ripples and spikes in pre-resection intra-operative ECoG, and studied the relationship between ARRm results in post-resection ECoG and post- surgical outcome.

METHODS

ECOG DATABASE

PATIENTS

The database consisted of intra-operative ECoG recordings, before and after resection, of patients with refractory focal epilepsy who underwent tailored epilepsy surgery between 2008 and 2012 at the UMC Utrecht. General anesthesia was induced using a combination of propofol and a synthetic opioid and maintained using a target-controlled propofol infusion pump. The surgical strategy was based on the results of presurgical diagnostics (e.g. MRI, PET, MEG, video-EEG) together with intra-operative tailoring based on inter-ictal spikes and spike patterns, using conventional EEG settings (70 Hz low pass filter, 10 s/page) in a common average reference montage. HFOs were never analyzed during surgery. The database was collected according to the guidelines of the institutional ethical committee of the UMC Utrecht, The

Netherlands. The institutional ethical committee approved the study and waived the need for written informed consent because of the retrospective character, provided that data were coded and handled anonymously.

The database was constructed based on the inclusion and exclusion criteria as described in a previous study.[9] We included 54 patients (median age 15.5 years, 29 male, 30 right-sided surgeries) with a median follow-up after surgery of 25 months (range 17.0–35.8 months). The resection area was temporal in thirty-three patients, frontal in twelve, peri-central in seven, and parietal and occipital each in one patient. Pathology was classified in four groups; mesial temporal sclerosis (MTS, N = 8), tumors with glial components (including gangliomas and dysembryoplastic neuroepithelial tumors (DNET); N = 23), malformations of cortical development (e.g. focal cortical dysplasia (FCD) and tuberous sclerosis; N = 15) and other pathologies (e.g. cavernomas and gliosis; N = 8). Thirty patients (56%) became seizure free after surgery (Engel 1A).[26] This postsurgical outcome was determined for the most recent follow-up with a minimum of one year.

ECOG RECORDINGS

Intra-operative ECoG recordings were made with a 64-channel EEG system (MicroMed, Veneto, Italy) at 2048 Hz sampling rate using an anti-aliasing filter at 538 Hz. We used 4x5 or 4x8 electrode grids and 1x6 or 1x8 electrode strips (Ad-Tech, Racine, WI). The platinum electrodes are embedded in silicone with a 4.2 mm^2 contact surface and 1 cm inter-electrode distance. Grids and strips were placed in multiple configurations before (preresection ECoG), during extension of the resection, and after resection (post-resection ECoG). Propofol anesthesia was stopped during ECoG registration until a continuous ECoG background pattern was achieved. The anesthesiologist monitors the patient's heart rate and blood pressure closely, restarting the propofol before the patient starts to wake up. To minimize propofol effects we selected the last minute of each pre- and post-resection recording for analysis; on average this was 11 min after propofol was stopped. Analysis was performed on bipolar electrode pairs lengthwise on the grids. Bipolar channels with continuous artefacts that were visible in the raw ECoG signal were excluded.

MARKING OF HFOS AND SPIKES

We used an automated detector previously developed, adapted for our intra-operative ECoG, to detect HFOs.[16] The detector uses a high pass finite impulse response (FIR) filter >80 Hz for ripples and >250 Hz for fast

ripples.[14,16] Subsequently, visual post-processing of the data and the HFO detections in Stellate Harmonie Reviewer (v7.0, Montreal, QC, Canada) was done in consensus by two of three reviewers (MvtK/MZ/NvK) to correct for artefacts, and falsely identified or missed HFOs. A split screen was used to visualize ripples (gain 5 μV/mm) and fast ripples (gain 1 μV/mm) simultaneously at an elongated time interval of 0.4 s/page and the same filter settings as mentioned above.[4] Spikes were visually marked (MvtK) in the same bipolar montage, using conventional infinite impulse response (IIR) filter settings of 0.5–70 Hz at a gain of 75–200 μV/mm and 10 s/page. Marked spikes were checked by a clinical neurophysiologist (FL). Spikes were defined as paroxysmal sharp transients, with a maximum duration of 80 ms[27] and minimum amplitude of twice the baseline. Sharp waves co-occurring with spikes on other bipolar channels were also marked. HFOs were marked independently and blinded for spikes and vice versa. All events were marked blinded for outcome.

AUTOREGRESSIVE RESIDUAL MODULATION

ORIGINAL AUTOREGRESSIVE RESIDUAL MODULATION

We used the original autoregressive model residual modulation ($ARRm_{orig}$) algorithm, as previously described.[25] The $ARRm_{orig}$ algorithm quantifies the variance of residual signal variation (r) after modelling short windows of ECoG data using autoregressive models with model order 3. High ARRm values represent the intermittent occurrence of non-harmonicity in the ECoG signal. The window length was 40 samples, approximately 20 ms of ECoG data, and consecutive windows with 50% overlap were chosen. The algorithm quantifies the high-frequency components of the EEG signal, as frequencies below 50 Hz cannot be modelled within a 40 sample window. These signal components therefore do not influence the residual signal variation.

IMPROVED AUTOREGRESSIVE RESIDUAL MODULATION

Artefacts may produce spuriously large residual signal variations, which result in false positive high ARRm values unrelated to the epileptic tissue. We modified the ARRm algorithm to be less sensitive to subtle artefacts than the original $ARRm_{orig}$ algorithm. Artefacts can be roughly classified into two groups; (1) gross artefacts, which can be observed in the unfiltered ECoG, e.g. from electrode movement, or lack of contact with tissue. Such artefacts are easy to recognize and may be prevented by careful recording.

(2) Subtle short-lasting artefacts may result from electrical interference, touching the electrode or short electrode problems and have a very small amplitude. These artefacts can be short and sharp, resulting in a strong non-linear component, and thus greatly influence the ARRm.

A preliminary study on a learning set showed that these artefacts differ from HFOs when looking at residuals from different autoregressive model orders. The learning set consisted of twelve 5-s ECoG epochs from seven patients chosen from the preresection ECoG dataset. We included examples of artefacts, spikes and HFOs, and data without such events, but cases were otherwise chosen randomly. We observed that r-values stayed high for increasing model orders in windows containing artefacts. Windows containing true events, however, showed r-values declining much faster for increasing model orders. The difference between artefacts and HFOs was largest when looking at the steepness of r decline from the first to the second order (r_1 to r_2). The decline of r over an increasing model order was used to reject high r_3-values resulting from artefacts. Using the same learning set, an appropriate decline threshold for the ARRm algorithm was chosen. Figure 4.1 illustrates the artefact rejection method used in the ARRm algorithm.

ECoG epochs are evaluated per bipolar channel to provide an ARRm value. First, second and third order AR models were estimated, thereby obtaining residual signal variance time series (r_1, r_2 and r_3) with values for each 40-sample window. Residual decline (D) over order one and two was calculated for each window (w), as follows:

$$D(w) = \frac{2(r_1(w) - r_2(w))}{r_1(w) + r_2(w)}. \quad (4.1)$$

Outliers in r_3, i.e. exceeding the 95th percentile, and its two circumjacent r_3 samples, were removed from the r_3 time series if the outlier window meets the condition D<0.9. This provides us with a collection of 'cleaned' windows (w_c).

Subsequently, the improved autoregressive residual modulation ARRm can be calculated, using only the clean windows:

$$ARRm = \frac{\sigma(r_3(w_c))}{\mu(r_3(w_c))}, \quad (4.2)$$

where σ is the standard deviation and μ is the mean.

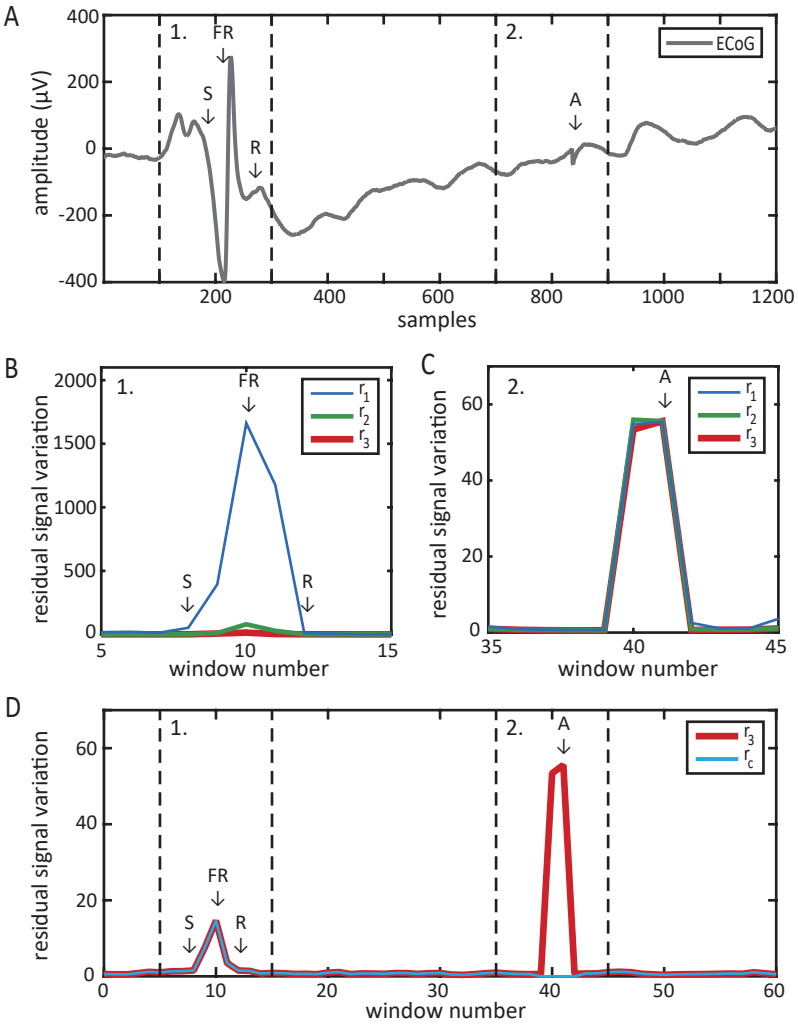


FIGURE 4.1.

Illustration of the artefact rejection method in the ARRm algorithm. A) ECoG raw signal (1200 samples = 0.6 s). Box 1 holds a spike (S), a ripple (R) and a fast ripple (FR), whereas box 2 holds an artefact (A). Arrows indicate the timing of these events. B) Residual signal variations (r) calculated using autoregressive model order 1, 2 and 3, from the ECoG segment in box 1, containing S, R and FR events. The r-values are high for events for all three orders, but the r-value declines steeply when the model order increases. C) r-Values for model order 1, 2 and 3 from the ECoG segment including the artefact in box 2. During this artefact, the r-values are high for all three orders; there is no steep decline of r-values when the model order increases. D) r_3-Values that exceed the 95th percentile of the epoch, and have a small r_1-to-r_2 decline, or modulation, are rejected. Shown are the r-values from model order 3 (r_3) and the resulting clean r_3 time series (r_c). For visualization purposes, the rejected values are shown here as zeros; in reality they are removed from the r_3 time series.

STATISTICAL ANALYSIS

$ARRM_{ORIG}$ AND ARRM VALIDATION IN PRE-RESECTION DATA

The $ARRm_{orig}$ and ARRm results from the ECoG data recorded pre-resection were compared on a channel level with spike, ripple, and fast ripple counts, using the h^2 index. The h^2 index is a measure of the association between pairs of data in general (linear or non-linear).[28] We obtained the h^2 index to test the fraction of variation of $ARRm_{orig}$ and ARRm that can be explained by the number of spikes, ripples and fast ripples. This results in a value between 0 and 1, where $h^2(A,B) = 1$ implies that all variation of A is explained by B. The h^2 findings are significant if h^2 exceeds the critical value, or significance margin, in a bootstrapped test with a = 0.05. We evaluated the ability of $ARRm_{orig}$ and ARRm to correctly classify channels with (+) and without (-) spikes, ripples or fast ripples. ROC curves for all event types were obtained using the sensitivity and specificity for various $ARRm_{orig}$ or ARRm thresholds at channel level. The ROC curve for fast ripples – the most specific biomarker for epileptogenic tissue – provided a clinical threshold. This threshold was used for comparison of ARRm results in postresection ECoG and surgical outcome. For this analysis we dichotomized the ARRm results into channels with values above (ARRm+) and below (ARRm-) the determined threshold.

RELATIONSHIP OF ARRM TO OUTCOME IN POST-RESECTION DATA

In contrast to the rejected subtle artefacts, gross artefacts may be prevented during registration and can be easily observed in the unfiltered ECoG. We, therefore, visually checked the data for an explanation of the ARRm+ value in that specific channel. We removed channels with gross artefacts causing ARRm+ results from the post-resection data analysis.

We investigated the spatial distribution of fast ripples and ARRm values in relation to resection and outcome in three patient groups. Group classification was based on the presence of fast ripples and ARRm+ in any of the ECoG channels. Group a) contains patients who had fast ripples and ARRm+. This group was studied to confirm overlap of the channels that identified the epileptogenic area. Group b) contains patients with fast ripples but ARRm-. This group is used to see whether ARRm is less susceptible to false positives due to physiological fast ripples. Group c) contains patients without fast ripples but ARRm+. In this group we studied if the ARRm identifies new patients with poor prognosis who were missed by identification with fast ripples.

OUTCOME PREDICTION BASED ON ARRM

We performed statistical analysis of the ARRm results at channel level and maximal ARRm value at patient level in comparison to outcome (Mann–Whitney U test). We did this for the dichotomized ARRm results based on the clinical threshold as determined on pre-resection ECoG data (Pearson X^2-test or Fisher exact-test). We stratified for the presence and absence of fast ripples, ripples and spikes both at patient and channel level. We calculated the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of the maximal ARRm value for poor outcome.

Computation of the ARRm results and h^2 association index was performed in Matlab (version 8.5.0; The MathWorks Inc., Natick, MA, USA). Statistical analysis was performed in IBM SPSS Statistics 21 (IBM Corp, Armonk, NY, USA). We considered a P-value <0.05 significant. Illustrations were constructed using Matlab and Adobe Illustrator (CS6).

RESULTS

DATA

All 54 patients in our database of intra-operative ECoG recordings were included in the analysis. We included all 2744 bipolar channels of the pre-resection recordings, resulting in a median (inter quartile range, IQR) of 41.0 (36.0–63.5) channels per patient for pre-ECoG. For the post-resection ECoG we included 1291 of 1325 bipolar channels, resulting in a median (IQR) of 18.5 (range: 15.0–29.0) channels per patient after resection.

ARTEFACT REJECTION

Figure 4.2 shows the effect of several artefact types on $ARRm_{orig}$ and ARRm. Glitches, which are commonly appearing subtle artefacts that have a substantial effect on $ARRm_{orig}$, are successfully rejected (Figure 4.2A). Several other gross artefact types are also shown, which are not rejected by ARRm. Figure 4.3 illustrates the effect of artefact rejection on channel level, considering the presence of fast ripples. The majority of FR- channels get lower ARRm values than their corresponding $ARRm_{orig}$ values, while FR+ channels keep similar values. This indicates successful rejection of high-impact artefacts.

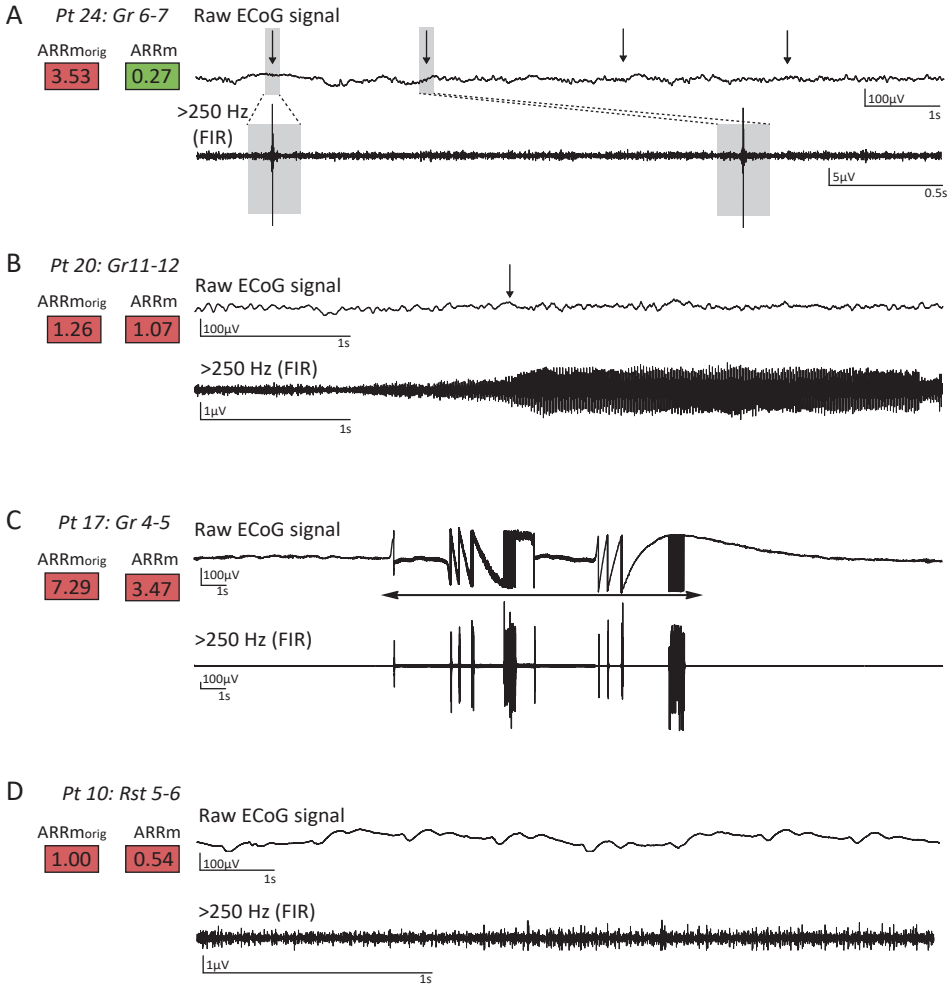


FIGURE 4.2.

Example of common subtle and gross artefacts in the intra-operative ECoG and accompanying $ARRm_{orig}$ and ARRm value. A) Subtle intermittent glitches are one of the most common artefacts in intra-operative ECoG data and are barely visible in the raw data signal. They do not hamper visual analysis of spikes, but may cause spurious artefacts in the higher frequency ranges, and can be easily mistaken for HFOs. The ARRm algorithm, unlike the ARRmorig, was successfully designed to reject this type of artefact. B) Sudden onset or waxing and waning high frequency noise on individual channels can cause undesirably high ARRmorig as well as ARRm values. This subtle artefact is also difficult to spot in the raw signal. C) Typical gross artefact on one electrode due to interference of the surgeon touching the grid. D) Example of a floating channel that does not touch the brain surface and picks up noise, especially visible in the higher frequency ranges. This often occurs for electrodes on corner of grids or the last electrode of a strip. Note that in B), C) and D) the ARRm value was substantially lower compared to the $ARRm_{orig}$ value, as some artefacts in these signals might have fulfilled the rejection criteria. None of the here depicted channels showed fast ripples.

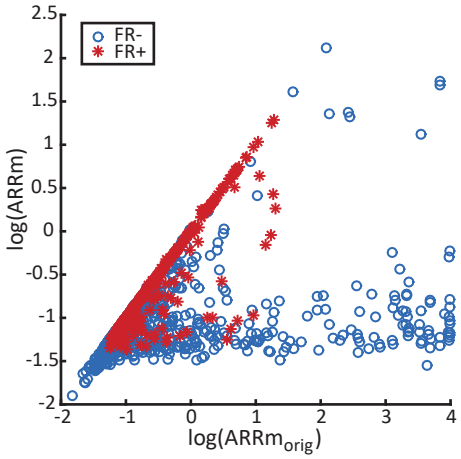


FIGURE 4.3.
Scatterplot of ARRm$_{orig}$ values vs. ARRm values, where each data point corresponds to one channel. Circles correspond to fast ripple negative channels (FR-), where no fast ripples were scored visually. Asterisks correspond to fast ripple positive channels (FR+), where one or more fast ripples were scored visually. Both axes are scaled logarithmically.

ARRM$_{ORIG}$ AND ARRM VALIDATION IN PRE-RESECTION DATA

COMPARISON OF ARRM$_{ORIG}$ AND ARRM TO HFOS AND SPIKES IN PRERESECTION DATA

The ARRm$_{orig}$ and ARRm results from the ECoG data recorded pre-resection were compared on a channel level with spike, ripple, and fast ripple counts. In Figure 4.4, the results of the six comparisons are shown, together with the resulting h^2 values reflecting the amount of association. ARRm showed particularly low values for channels which had high ARRm$_{orig}$ values but no fast ripples, showing that the artefact rejection was successful (see Figure 4.4C and F and Figure 4.3). All h^2 associations between ARRm$_{orig}$ and spikes, ripples, or fast ripples were significant ($P<0.01$), as well as for ARRm ($P<0.01$). Thus ARRm$_{orig}$ and ARRm variation may be explained by the quantity of events, where associations were strongest for fast ripples. ARRm was better associated with fast ripples than ARRm$_{orig}$.

ARRM$_{ORIG}$ AND ARRM AS EVENT CHANNEL CLASSIFIERS

We investigated the ability of ARRm$_{orig}$ and ARRm to correctly classify channels with and without spikes, ripples or fast ripples. The boxplot and ROC curve illustrating the FR+/- channel classification ability of ARRm$_{orig}$ and ARRm are shown in Figure 4.5. Best performance is found

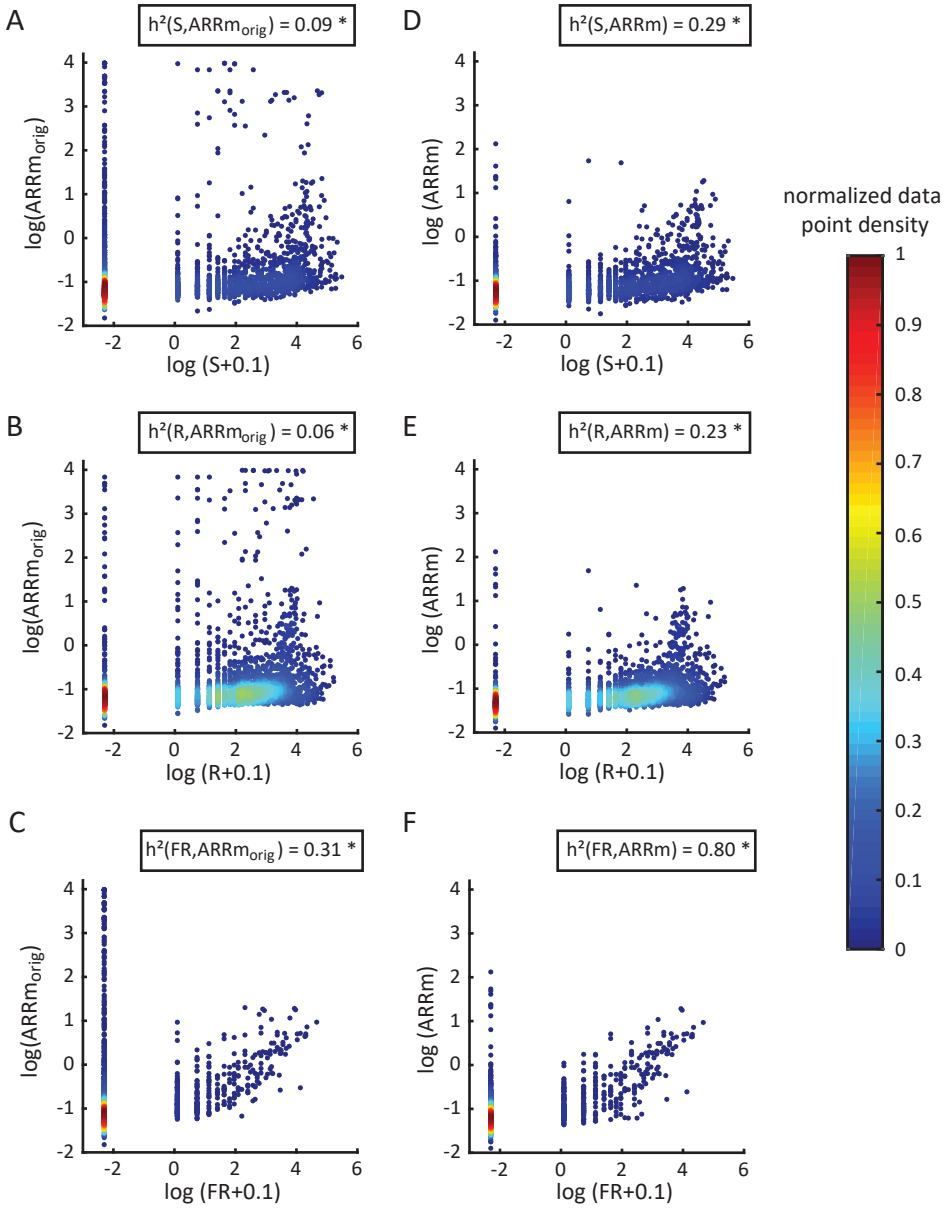


FIGURE 4.4.

Scatter density plots of $ARRm_{orig}$ (panels A, B and C) and ARRm (panels D, E and F) results from the ECoG data recorded pre-resection, compared on a channel level with spike (S), ripple (R), and fast ripple (FR) counts. Both axes are scaled logarithmically. Note that 0.1 is added to all event counts (x-axis) for visualization purposes, for log(0) = 1. Thus, a data point at -2.3 on the x-axis corresponds to zero events to have been scored on a channel. h2 association between the compared results is shown in the boxes next to the figures (*P<0.05 was considered significant). ARRm showed particularly low values for channels which had very high $ARRm_{orig}$ values but no fast ripples, showing that the artefact rejection was successful (see Figure 4.4C and F and Figure 4.3).

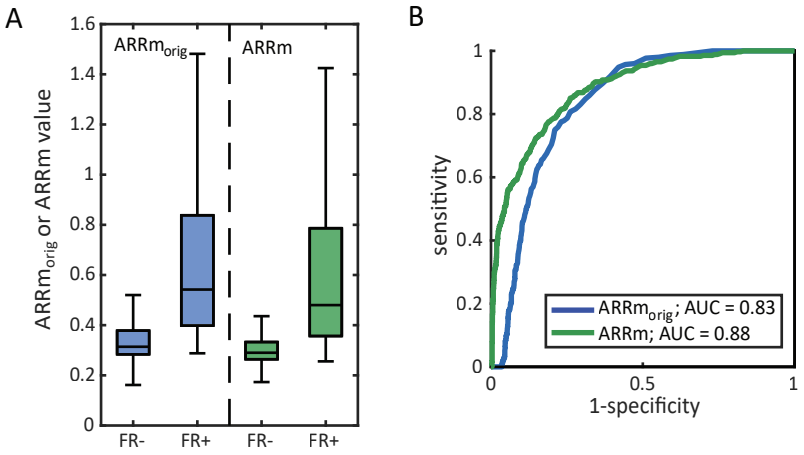


FIGURE 4.5.

Performance of $ARRm_{orig}$ and ARRm in pre-resection ECoG. Boxplot (A) and receiver operating characteristic (ROC) curve (B) illustrating the ability of $ARRm_{orig}$ (in blue) and ARRm (in green) to classify channels negative or positive for fast ripples. The FR- group contains channels without marked fast ripples and the FR+ group contains channels in which one or more fast ripples were marked. Outliers are not shown in the boxplots. The calculated ROC area under the curve (AUC) values for $ARRm_{orig}$ and ARRm are indicated in textboxes in the figure.

for classification of FR+ and FR- channels with a ROC area under the curve (AUC) of 0.83 for $ARRm_{orig}$ and 0.88 for ARRm. For spikes the AUC values were 0.70 and 0.76, and for ripples 0.70 and 0.78 for $ARRm_{orig}$ and ARRm, respectively. ARRm performance exceeds $ARRm_{orig}$ performance for all event types, which implies successful artefact reduction. 100% specificity for FR- channels could not be obtained with any threshold for $ARRm_{orig}$ - there were many FR- channels with high $ARRm_{orig}$ values - but was possible for ARRm (see Figure 4.4C and F). Since the ARRm outperforms the $ARRm_{orig}$, we used the ARRm algorithm throughout the rest of the study.

ARRM CLINICAL THRESHOLD

A threshold is needed to use ARRm as a classifier to find channels located in epileptogenic tissue. We chose this threshold favoring specificity; if a low ARRm value is measured, nonepileptogenicity of the tissue at that location should be concluded with as much certainty as possible. Therefore the 95% specificity point, ARRm >0.47 (corresponding to 52% sensitivity), was chosen as the ARRm threshold to define a channel as suspect, i.e. likely to be located on epileptogenic tissue.

TABLE 4.1.
Performance of ARRm in predicting surgical outcome, based on results in post-resection ECoG.

	Channel level (N = 1291) ARRm			Patient level (N = 54) Max ARRm		
	>0.47	<0.47	P-value	>0.47	<0.47	P-value
# Counts	27	1264		9	45	
# Outcome (Poor/Good)	26/1	518/746	<0.01* (a)	8/1	16/29	<0.01* (a)
ECoG events						
FRs+ (Poor/Good)	20 (20/0)	32 (18/14)	<0.01* (b)	6 (6/0)	6 (3/3)	0.18 (b)
FRs- (Poor/Good)	7 (6/1)	1232 (500/732)	0.02* (b)	3 (2/1)	39 (13/26)	0.29 (b)
R+ (Poor/Good)	27 (26/1)	751 (324/427)	<0.01* (a)	9 (8/1)	42 (16/26)	<0.01* (b)
R- (Poor/Good)	0 (0/0)	513 (194/319)	n.a.	0 (0/0)	3 (0/3)	n.a.
S+ (Poor/Good)	26 (25/1)	468 (225/243)	<0.01* (a)	8 (7/1)	30 (19/11)	<0.01* (b)
S- (Poor/Good)	1 (1/0)	796 (293/503)	0.37 (b)	1 (1/0)	15 (5/10)	0.38 (b)

*Abbreviations: Poor - recurrent seizures (Engel Ic-IV), Good - seizure freedom (Engel Ia,b), a - Pearson X^2-test, b - Fisher exact-test, n.a. - not available, * P-value <0.05 considered significant. Note that the FR- group contains channels or patients without marked fast ripples, while the FR+ group contains channels or patients in which one or more fast ripples were marked, and similar for ripples (R+/R-) and spikes (S+/S-) .*

ARRM COMPARED TO POST-SURGICAL OUTCOME

ARRM IN RELATION TO FAST RIPPLES IN POST-RESECTION ECOG AND CLINICAL DECISIONS

We investigated the spatial extent of residual fast ripples and the ARRm results in post-resection ECoG in relation to outcome (Table 4.1). An illustrative patient example with overlapping fast ripple and ARRm results is shown in Figure 4.6. We dichotomized the ARRm results at channel level based on the above mentioned threshold. We found six patients who had both fast ripples and ARRm+ channels (group A) in the same locations, all had recurrent seizures. Five patients had fast ripples but didn't show ARRm+ channels (group B). This group contains three seizure-free patients with fast ripples after resection. Patient 3, for example, showed fast ripples in Broca's

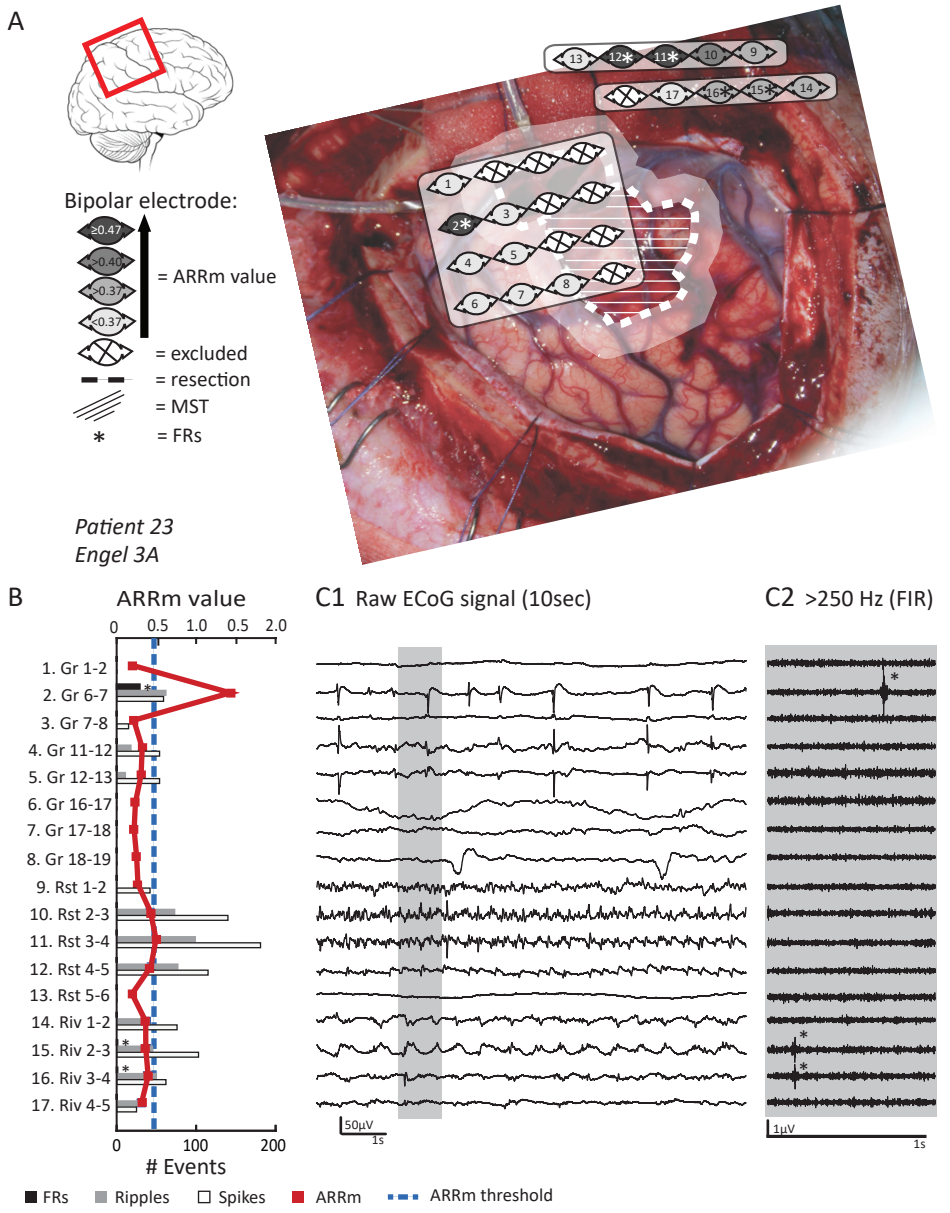


FIGURE 4.6.

Patient example with ARRm results. A) A patient with a neocortical epileptogenic lesion in the right central cortex that was resected. Additionally multi subpial transections were performed (poor outcome). Pathology revealed a mild organisational disorder of the cortex. B) Post-resection recordings showed multiple channels with spikes, but since this was an eloquent area, it was decided not to resect more tissue. However, in only few channels ARRm was elevated, similar to channels with fast ripples (*). C1) Ten seconds of raw unfiltered intra-operative ECoG as recorded. Note that the ARRm results are calculated on this raw signal, while for marking of HFOs the signal has to be filtered. C2) A close-up of the ECoG high-pass filtered at 250 Hz high-pass, showing fast ripples (*).

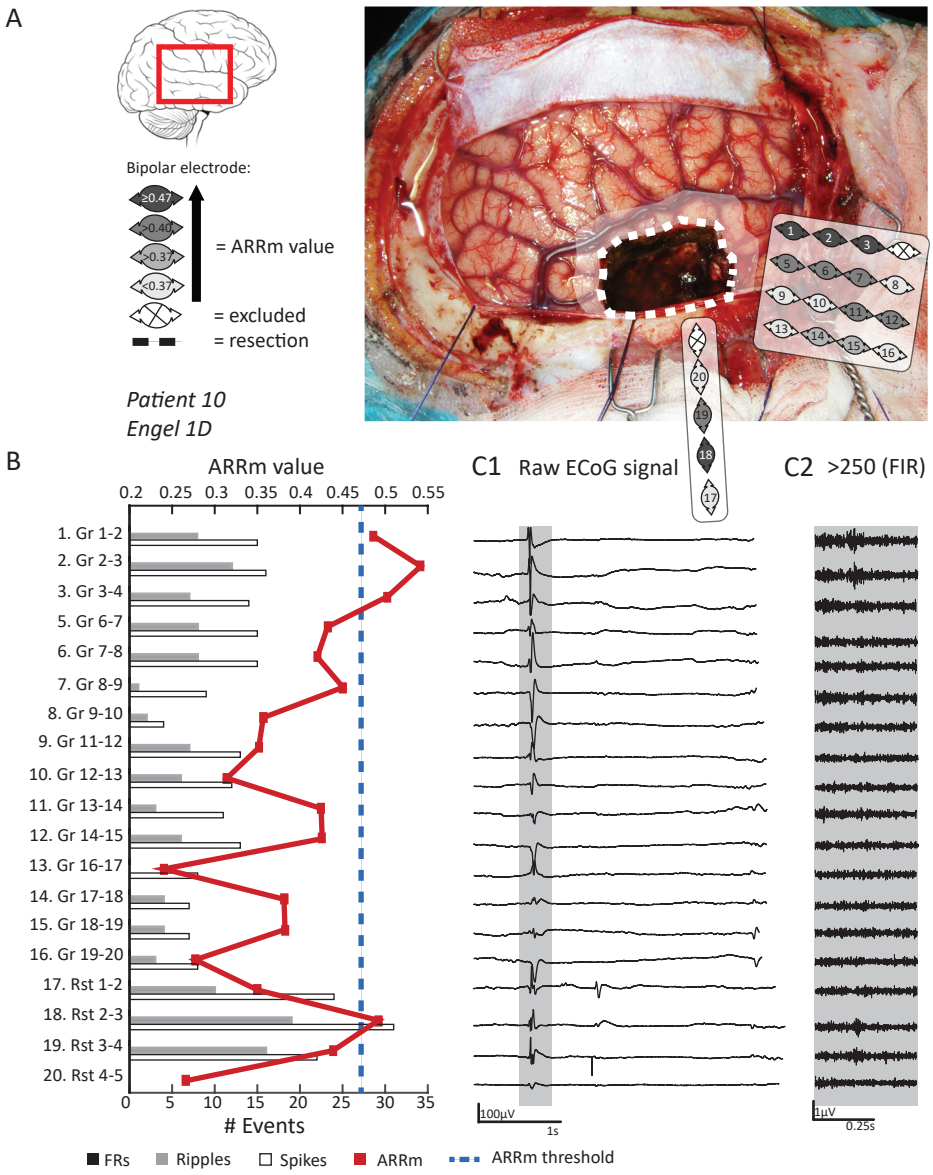


FIGURE 4.7.

Patient example with ARRm results. A) A patient with a right neocortical temporal epileptogenic lesion that was resected (poor outcome). Pathology revealed a ganglioglioma WHO grade II. B) Post-resection recordings showed spikes and ripples, in the anterior temporal area, but no fast ripples (FR). The ARRm was elevated in 4 channels. C1) Segment of raw unfiltered intra-operative ECoG as recorded. In the grey box, an event can be observed. This event was initially marked as an artefact. C2) 250 Hz high-pass filtered ECoG data from the grey box in panel C1, showing fast activity that could be classified as fast ripples. If this is the case, the four channels with high ARRm values should be associated with paroxysmal events. These events were then wrongly discarded during visual HFO marking, and the channels falsely classified as fast ripple negative.

area and the central motor area, which we believe represent physiological fast ripples. This suggests the ARRm might be insensitive to physiological fast ripples. On the other hand, there were three patients with fast ripples and poor outcome but without ARRm+ results. Although the channels with fast ripples exhibited the highest ARRm values when ranked, these values did not exceed the predefined threshold. Three patients had ARRm+ channels but did not show residual fast ripples (group C). Two of these patients had poor outcome. In one patient with recurrent seizures the ARRm+ channels overlapped with an event that was incorrectly discarded as artefact during visual analysis (see Figure 4.7).

OUTCOME PREDICTION WITH ARRM

The ARRm at channel level and the maximal ARRm value at patient level are both indicators of poor outcome if the value is above 0.47 (both $P<0.01$, Pearson X^2 test; Table 4.1). A maximal ARRm >0.47 predicted poor outcome with a sensitivity of 33.3%, a specificity of 96.7% and had a PPV of 88.9% and a corresponding NPV = 64.4%. When stratified for the presence or absence of fast ripples, ripples and spikes, the ARRm performs better at channel level than at patient level. Remarkably, even in the absence of fast ripples (FR-) at channel level, the ARRm >0.47 is an indicator of poor post-surgical outcome ($P = 0.02$, Fisher exact; Table 4.1).

DISCUSSION

The autoregressive model residual modulation (ARRm) algorithm provides a meaningful tool for intra-operative delineation of the epileptogenic zone. The algorithm is inspired by the most characteristic aspect of epilepsy; its unpredictable nature. In the ECoG ARRm is a 'bad' channel identifier rather than an event detector such as HFO detectors. We adjusted the ARRm to become less sensitive to subtle but influential artefacts in the intra-operative ECoG. We showed that brain tissue producing signals with high non-harmonicity corresponds to areas producing fast ripples, and to a lesser extent to ripples and spikes. High ARRm results post-resection had a positive predictive value of 88.9% for poor outcome, including in patients without fast ripples. The ARRm calculated from interictal post-resection ECoG could be an equally good or even better indicator of post-surgical outcome than residual fast ripples, provided that gross peri-surgical artefacts are prevented during recording. We believe that the ARRm algorithm has the potential to replace visual HFO analysis for tailoring of epilepsy surgery, as it can be quickly calculated and thus used in real-time during surgery.

AUTOMATED EPILEPTOGENIC ZONE LOCALIZATION IN LITERATURE

For decades, researchers have focused on the challenge of identifying automatically the precise location of the epileptogenic zone, by quantifying various signal aspects in the intracranial EEG.[29,30] The recent discovery of HFOs has inspired researchers to develop new algorithms for localisation.[15,21,31,32] The overall reported sensitivity of HFO detection algorithms varies between 50% and 100%, and the specificity ranges between 36% and 90%.[32] A drawback of the majority of HFO detectors is false positive detections, due to artefacts that are mistaken for HFOs or physiological high-frequency activity, with false discovery rates reported as low as 13% but as high as 75%.[17,32] The ability of non-linearity quantifiers to localize the epileptogenic zone in brain signals has been extensively studied.[33–37] Despite promising correlations between resection areas identified by such quantifiers and postsurgical seizure freedom,[37,38] many of these algorithms are unsuitable for automated online analysis during surgery, as measurements with an extensive duration are needed for reliable calculations. One of the strengths of the ARRm algorithm is that calculations are very fast with virtually no delay, as the relatively simple calculations are performed on short unfiltered raw ECoG segments, instead of detecting separate events. Therefore, the algorithm could be used real-time during surgery.

NON-HARMONICITY

Traditionally, residual signal variance after an autoregressive fit is regarded as noise, but in our application it indicates nonharmonic characteristics, i.e. the nonlinear or non-stationary dynamics of the underlying system generating the signal. The main assumption underlying ARRm methodology is that epileptic HFOs are, in contrast to physiological high-frequency activity, of the non-harmonic type. In other words, they cannot be well described as solutions of a linear differential equation of a certain order.[25] Quantification of the variance of residual signal variation in the higher frequencies is thus used to delineate epileptogenic tissue. Previously autoregressive fitting was used to quantify interictal spiking.[29,39] One might argue that focusing on high values of the residuals in the ECoG data, which is essentially what the ARR algorithm does, is just another way of detecting interictal spikes, with or without HFOs on top. Our results suggest that the ARRm is more a reflection of liability towards, rather than an indicator of, epileptiform transients measured on the electrodes. We see that not every electrode with spikes has a high ARRm value. Spikes are not necessarily associated with

nonlinear quantifiers, but nonlinear quantifiers coincide quite well with the epileptogenic zone.[40]

Our results give rise to the question if the ARRm algorithm would give the same results on epochs without interictal events. This was suggested by a study in which the authors found tissue epileptogenicity by studying signal properties of very short (1-2 s) epochs that did not exhibit interictal events.[41] This should be studied further.

PATHOPHYSIOLOGY

A common difficulty in working with HFOs is the inability to distinguish physiological from pathological HFOs. Physiological HFOs may arise as a consequence of synchronous interneuron-mediated inhibitory postsynaptic potentials (IPSPs) that coordinate principal cells to fire sparsely.[42] In contrast, pathological ripples are thought to arise due to reduced IPSPs that give rise to brief bursts of population spikes.[43] Pathological fast ripples may result from slightly out-of-phase firing of independent clusters of these pathologically interconnected neurons. Proposed mechanisms for generating these fast pathological oscillations are ephaptic interactions, electrotonic coupling via gap junctions, or fast synaptic transmission.[11,23] It is known that physiological HFOs can occur in functionally eloquent regions such as the sensorimotor[14] and visual cortex[44] and in the hippocampus, where ripples may be involved in memory processing.[45] In a previous study we described a patient who became completely seizure free (and stopped using medication) but showed residual spikes and fast ripples in Broca's area and the central motor area after surgery.[9] We found normal ARRm values in these regions in the same patient, suggesting that the ARRm might be insensitive for what is probably physiological oscillatory brain activity. This finding also confirms that pathological HFOs may have a more non-harmonic signature than physiological HFOs. This potential insensitivity of ARRm to physiological oscillations is also reflected in a better poor outcome prediction compared to fast ripples, as previously described;[9] the PPV was 88.9% for ARRm and 75.0% for fast ripples, while they have an equal NPV; 64.4% and 64.3% respectively.

Analogously, while the exact mechanism behind the nonharmonic behavior quantified by the ARRm remains unclear, we might be able to find some explanatory links between our method and its clinical results. The observation that modulation of the autoregressive residual variation localizes epileptogenic tissue, suggests that the processes causing high ARRm values are related to relatively short bursts of highly non-harmonic

signal features interrupting the background activity. This is supported by the findings by Dümpelmann et al. (2012) who distinguished epileptogenic HFO activity on the basis of its non-stationary appearance.[18] The neuronal processes accountable for such bursts of exceptional activity might be avalanche dischargers as shown in computational models.[46,47] An even more challenging question is how to relate those processes to the epileptic properties of neural circuits. It was recently demonstrated that both epileptic transitions and action potential avalanches could be also be generated by excessive amounts of gap junctions.[1] These discharges also generated high values of residual variance after auto-regressive model fitting. In resemblance to pathological HFOs, we speculate that a possible cause of high ARRm values is the presence of high electric conductance between the axonal compartments in the brain tissue picked up by the electrode.

STUDY LIMITATIONS

We implemented a threshold for the ARRm, based on results from pre-surgical ECoG recordings and the ability of ARRm to predict channels with fast ripples, with a specificity of 95%. Such a general threshold may not fit all individual patients, and could induce false negative results as shown by the results in group B (fast ripples but no ARRm+ channels). With respect to surgical outcome, only 8 of 24 patients with seizure recurrence showed an elevated ARRm in post-resection ECoG. The absence of epileptic biomarkers in the post-resection ECoG may be due to a too strict threshold definition, inadequate sampling of the cortex or ongoing epileptogenesis after surgery. A strength from a clinical perspective is that our unsupervised ARRm algorithm is a channel identifier rather than an event identifier. Surgeons require information about the channels, representing cm^2 of cortical tissue, that need to be resected or not. However, a more gradual scale could be preferred over this black-and-white decision offered by a thresholded quantifier. This way, ARRm values could indicate tissue epileptogenicity risk.

A potential bias in our study could be that high ARRm values are found in eloquent areas close to the resection area, if the eloquent areas produced more physiological high frequency activity.[48] These areas may not have been part of the epileptogenic zone, but patients with a resection close to functionally eloquent areas do have a higher risk of incomplete resection and thus poor surgery outcome. Our study population included only a few patients who underwent a resection close to eloquent areas.

We incorporated artefact rejection designed to reject subtle but influential artefacts. This method did not work for all types of artefacts (Figure 4.2). Intra-operative ECoG recordings are sensitive to artefacts and noise due to the electrical equipment and surgical manipulation. These types of artefacts could all be actively prevented by a 'time-out' procedure, in which all interfering devices are switched off if possible, and all hands are clear from the patient. It then takes only one minute of recording to calculate accurately the ARRm results.

Due to the retrospective nature of this study and the limited amount of data, it was not possible to investigate the stationarity of our results. It was recently suggested that the performance of automatic HFO detection depends on the type of resting state, and benefits from sleep rather than awake data.[49] This fits the findings that slow wave (non-REM) sleep gives the best HFOs signals.[50,51] Intra-operative recordings are a hybrid form, where sedation induced by propofol is subsequently tapered until a continuous EEG pattern is reached. We know that propofol, also used as an anti-epileptic agent in status epilepticus, suppresses the amount of interictal HFOs in the signal, but not their location.[52] If pathological HFOs and non-harmonicity are indeed linked, anesthetics and anti-epileptic drugs would influence nonharmonicity in the signal as well. This would be in line with evidence found for attenuation of non-linear EEG quantifiers by such drugs.[34,53,54]

CONCLUSIONS

We showed that the ARRm predicts poor post-surgical outcome when measured after resection. It performs as well as, or maybe even better than, residual fast ripples. There is an increasing need for implementation of automated algorithms to find the epileptogenic zone in clinical EEG software. The ARRm is the first unsupervised real-time analysis that could be used intraoperatively to provide an 'on demand' interpretation per electrode about whether or not the underlying tissue needs to be removed. Future research should focus on further optimizing the ARRm algorithm and on finding the optimal threshold. Meanwhile, we should aim to enable artefact-free recordings in the operating theatre and know the influence of propofol on the non-harmonicity of the ECoG signal.

REFERENCES

1. Helling RM, Koppert MMJ, Visser GH, et al. Gap Junctions as Common Cause of High-Frequency Oscillations and Epileptic Seizures in a Computational Cascade of Neuronal Mass and Compartmental Modeling. *Int J Neural Syst*. 2015;25:1550021.
2. Elger CE, Widman G, Andrzejak R, et al. Nonlinear EEG analysis and its potential role in epileptology. *Epilepsia*. 2000;41 Suppl 3:S34–8.
3. Andrzejak RG, Schindler K, Rummel C. Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients. *Phys Rev E - Stat Nonlinear, Soft Matter Phys*. 2012;86.
4. Jacobs J, LeVan P, Chander R, et al. Interictal high-frequency oscillations (80-500 Hz) are an indicator of seizure onset areas independent of spikes in the human epileptic brain. *Epilepsia*. 2008;49:1893–907.
5. Staba RJ, Stead M, Worrell GA. Electrophysiological Biomarkers of Epilepsy. *Neurotherapeutics*. 2014;11:334–46.
6. Staba RJ, Worrell GA. What Is the Importance of Abnormal "Background" Activity in Seizure Generation? In: Sharfman H, Buckmaster P, editors. Issues in Clinical Epileptology: A View from the Bench. Dordrecht: Springer; 2014. p. 43–54.
7. Gallentine WB, Mikati MA. Intraoperative electrocorticography and cortical stimulation in children. *J Clin Neurophysiol*. 2009;26:95–108.
8. Kuruvilla A, Flink R. Intraoperative electrocorticography in epilepsy surgery: Useful or not? *Seizure*. 2003;12:577–84.
9. van 't Klooster MA, Van Klink NEC, Leijten FSS, et al. Residual fast ripples in the intraoperative corticogram predict epilepsy surgery outcome. *Neurology*. 2015;85:120–8.
10. Wu JY, Sankar R, Lerner JT, et al. Removing interictal fast ripples on electrocorticography linked with seizure freedom in children. *Neurology*. 2010;75:1686–94.
11. Jefferys JGR, Menendez de la Prida L, Wendling F, et al. Mechanisms of physiological and epileptic HFO generation. *Prog Neurobiol*. 2012;98:250–64.
12. Crépon B, Navarro V, Hasboun D, et al. Mapping interictal oscillations greater than 200 Hz recorded with intracranial macroelectrodes in human epilepsy. *Brain*. 2010;133:33–45.
13. Zelmann R, Zijlmans M, Jacobs J, et al. Improving the identification of High Frequency Oscillations. *Clin Neurophysiol*. 2009;120:1457–64.
14. van Klink NEC, Van't Klooster M a, Zelmann R, et al. High frequency oscillations in intra-operative electrocorticography before and after epilepsy surgery. *Clin Neurophysiol*. 2014;125:2212–9.
15. Blanco JA, Stead M, Krieger A, et al. Unsupervised classification of high-frequency oscillations in human neocortical epilepsy and control patients. *J Neurophysiol*. 2010;104:2900–12.
16. Zelmann R, Mari F, Jacobs J, et al. Automatic detector of High Frequency Oscillations for human recordings with macroelectrodes. In: *32nd Annu Int Conf IEEE EMBS*. 2010. p. 2329–33.
17. Zelmann R, Mari F, Jacobs J, et al. A comparison between detectors of high frequency oscillations. *Clin Neurophysiol*. 2012;123:106–16.
18. Dümpelmann M, Jacobs J, Kerber K, et al. Automatic 80-250Hz "ripple" high frequency oscillation detection in invasive subdural grid and strip recordings in epilepsy by a radial basis function neural network. *Clin Neurophysiol*. 2012;123:1721–31.

19. von Ellenrieder N, Andrade-Valença LP, Dubeau F, et al. Automatic detection of fast oscillations (40-200 Hz) in scalp EEG recordings. *Clin Neurophysiol*. 2012;123:670–80.
20. Worrell GA, Jerbi K, Kobayashi K, et al. Recording and analysis techniques for high-frequency oscillations. *Prog Neurobiol*. 2012;98:265–78.
21. Burnos S, Hilfiker P, Sürücü O, et al. Human Intracranial High Frequency Oscillations (HFOs) Detected by Automatic Time-Frequency Analysis. *PLoS One*. 2014;9:e94381.
22. Kerber K, Dümpelmann M, Schelter B, et al. Differentiation of specific ripple patterns helps to identify epileptogenic areas for surgical procedures. *Clin Neurophysiol*. 2014;125:1339–45.
23. Zijlmans M, Jiruska P, Zelmann R, et al. High-frequency oscillations as a new biomarker in epilepsy. *Ann Neurol*. 2012;71:169–78.
24. Golebiowski A, Drewes C, Gulati S, et al. Is duration of surgery a risk factor for extracranial complications and surgical site infections after intracranial tumor operations? *Acta Neurochir (Wien)*. 2015;157:235–40.
25. Geertsema EE, Visser GH, Velis DN, et al. Automated Seizure Onset Zone Approximation Based on Nonharmonic High-Frequency Oscillations in Human Interictal Intracranial EEGs. *Int J Neural Syst*. 2015;25:1550015.
26. Engel Jr J, van Ness P, Rasmussen T, et al. Outcome with respect to epileptic seizures. In: Surgical treatment of the epilepsies. 2nd ed. New York: Raven Press; 1993. p. 609–21.
27. Alarcon G, Garcia Seoane JJ, Binnie CD, et al. Origin and propagation of interictal discharges in the acute electrocorticogram. Implications for pathophysiology and surgical treatment of temporal lobe epilepsy. *Brain*. 1997;120:2259–82.
28. Kalitzin SN, Parra J, Velis DN, et al. Quantification of unidirectional nonlinear associations between multidimensional signals. *IEEE Trans Biomed Eng*. 2007;54:454–61.
29. Lopes da Silva FH, Dijk A, Smits H. Detection of nonstationarities in EEGs using the autoregressive model: an application to EEGs of epileptics. In: Kunkel H, Dolce G, editors. CEAN: computerized EEG analysis. Stuttgart: Gustav Fischer Verlag; 1975. p. 180–99.
30. Andrzejak RG, David O, Gnatkovsky V, et al. Localization of Epileptogenic Zone on Pre-surgical Intracranial EEG Recordings: Toward a Validation of Quantitative Signal Analysis Approaches. *Brain Topogr*. 2015;28:832–7.
31. López-Cuevas A, Castillo-Toledo B, Medina-Ceja L, et al. An algorithm for on-line detection of high frequency oscillations related to epilepsy. *Comput Methods Programs Biomed*. 2013;110:354–60.
32. Chaibi S, Lajnef T, Sakka Z, et al. A reliable approach to distinguish between transient with and without HFOs using TQWT and MCA. *J Neurosci Methods*. 2014;1–11.
33. Casdagli MC, Iasemidis LD, Savit RS, et al. Non-linearity in invasive EEG recordings from patients with temporal lobe epilepsy. *Electroencephalogr Clin Neurophysiol*. 1997;102:98–105.
34. Widman G, Lehnertz K, Urbach H, et al. Spatial distribution of neuronal complexity loss in neocortical lesional epilepsies. *Epilepsia*. 2000;41:811–7.
35. Kalitzin S, Velis D, Suffczynski P, et al. 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.
36. Kalitzin S, Zijlmans M, Petkov G, et al. Quantification of spontaneous and evoked HFO's in SEEG recording and prospective for pre-surgical diagnostics. Case study. In: *34th Annu Int Conf IEEE EMBS*. 2012 [cited 2013 Nov 4]. p. 1024–7.

37. Rummel C, Abela E, Andrzejak RG, et al. Resected brain tissue, seizure onset zone and quantitative EEG measures: Towards prediction of post-surgical seizure control. *PLoS One*. 2015;10:e0141023.

38. Kim J-Y, Kang H-C, Cho J-H, et al. Combined use of multiple computational intracranial EEG analysis techniques for the localization of epileptogenic zones in Lennox-Gastaut syndrome. *Clin EEG Neurosci*. 2014;45:169–78.

39. Tzallas AT, Oikonomou VP, Fotiadis DI. Epileptic spike detection using a Kalman filter based approach. *Conf Proc IEEE Eng Med Biol Soc*. 2006;1:501–4.

40. Lehnertz K. Epilepsy and nonlinear dynamics. *J Biol Phys*. 2008;34:253–66.

41. Kowalik ZJ, Schnitzler A, Freund H, et al. Local Lyapunov exponents detect epileptic zones in spike-less interictal MEG recordings. *Clin Neurophysiol*. 2001;112:60–7.

42. Schomburg EW, Anastassiou CA, Buzsaki G, et al. The Spiking Component of Oscillatory Extracellular Potentials in the Rat Hippocampus. *J Neurosci*. 2012;32:11798–811.

43. Bragin A, Benassi SK, Kheiri F, et al. Further evidence that pathologic high-frequency oscillations are bursts of population spikes derived from recordings of identified cells in dentate gyrus. *Epilepsia*. 2011;52:45–52.

44. Nagasawa T, Juhász C, Rothermel R, et al. Spontaneous and visually driven high-frequency oscillations in the occipital cortex: Intracranial recording in epileptic patients. *Hum Brain Mapp*. 2012;33:569–83.

45. Kucewicz MT, Cimbalnik J, Matsumoto JY, et al. High frequency oscillations are associated with cognitive processing in human recognition memory. *Brain*. 2014;137:2231–44.

46. Beggs JM, Plenz D. Neuronal Avalanches in Neocortical Circuits. *J Neurosci*. 2003;23:11167–77.

47. Poil S-S, Hardstone R, Mansvelder HD, et al. Critical-state dynamics of avalanches and oscillations jointly emerge from balanced excitation/inhibition in neuronal networks. *J Neurosci*. 2012;32:9817–23.

48. Nonoda Y, Miyakoshi M, Ojeda A, et al. Interictal high-frequency oscillations generated by seizure onset and eloquent areas may be differentially coupled with different slow waves. *Clin Neurophysiol*. 2016;127:2489–99.

49. Liu S, Sha Z, Sencer A, et al. Exploring the time-frequency content of high frequency oscillations for automated identification of seizure onset zone in epilepsy. *J Neural Eng*. 2016;13:026026.

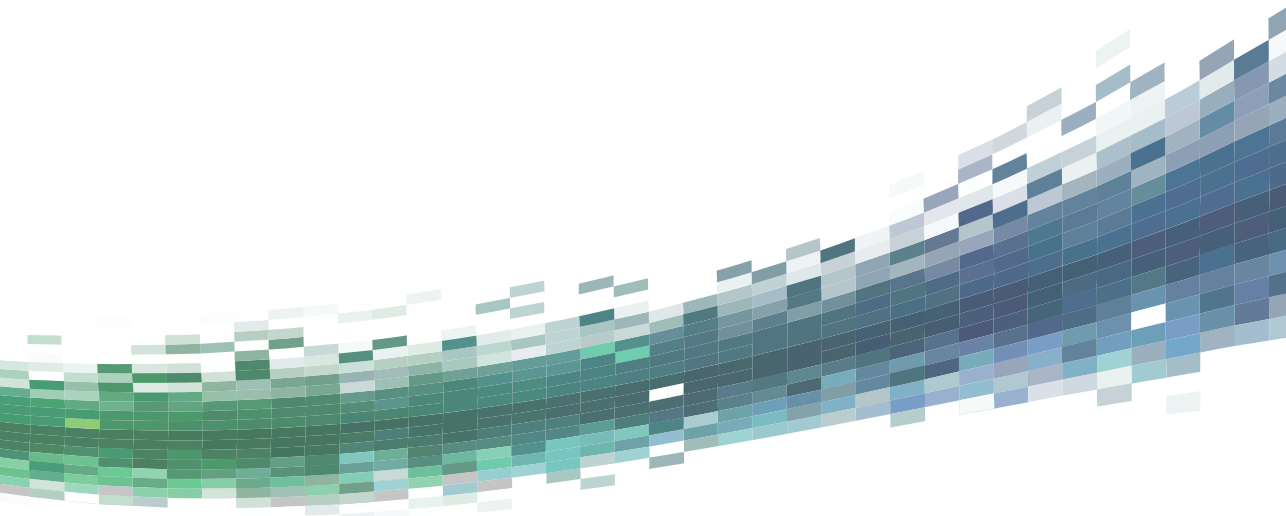
50. Bagshaw AP, Jacobs J, LeVan P, et al. Effect of sleep stage on interictal high-frequency oscillations recorded from depth macroelectrodes in patients with focal epilepsy. *Epilepsia*. 2009;50:617–28.

51. Frauscher B, Von Ellenrieder N, Ferrari-Marinho T, et al. Facilitation of epileptic activity during sleep is mediated by high amplitude slow waves. *Brain*. 2015;138:1629–41.

52. Zijlmans M, Huiskamp GM, Cremer OL, et al. Epileptic high-frequency oscillations in intraoperative electrocorticography: the effect of propofol. *Epilepsia*. 2012;53:1799–809.

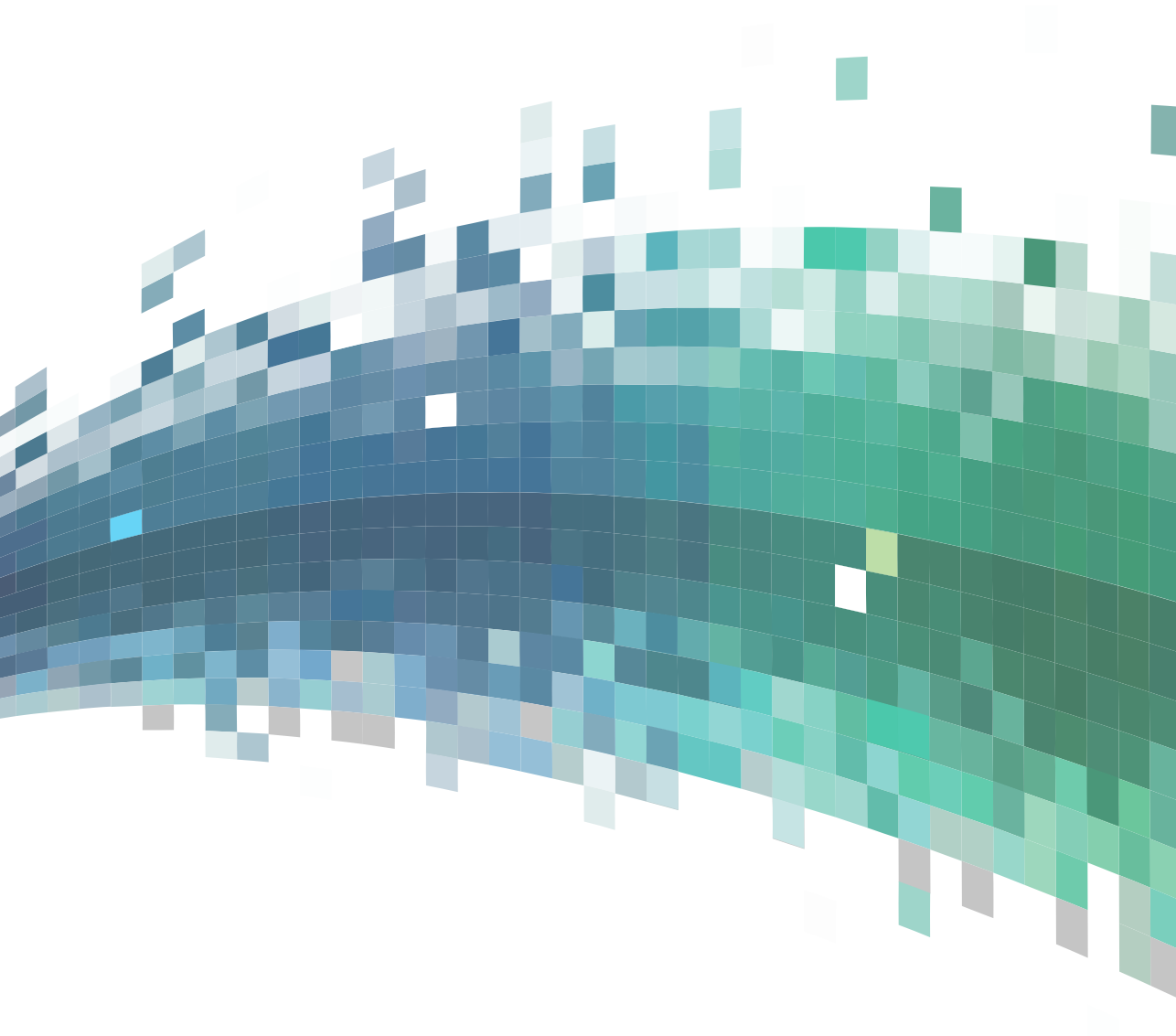
53. Lehnertz K, Elger CE. Neuronal complexity loss in temporal lobe epilepsy: effects of carbamazepine on the dynamics of the epileptogenic focus. *Electroencephalogr Clin Neurophysiol*. 1997;103:376–80.

54. Kim JM, Jung KY, Choi CM. Changes in brain complexity during valproate treatment in patients with partial epilepsy. *Neuropsychobiology*. 2002;45:106–12.



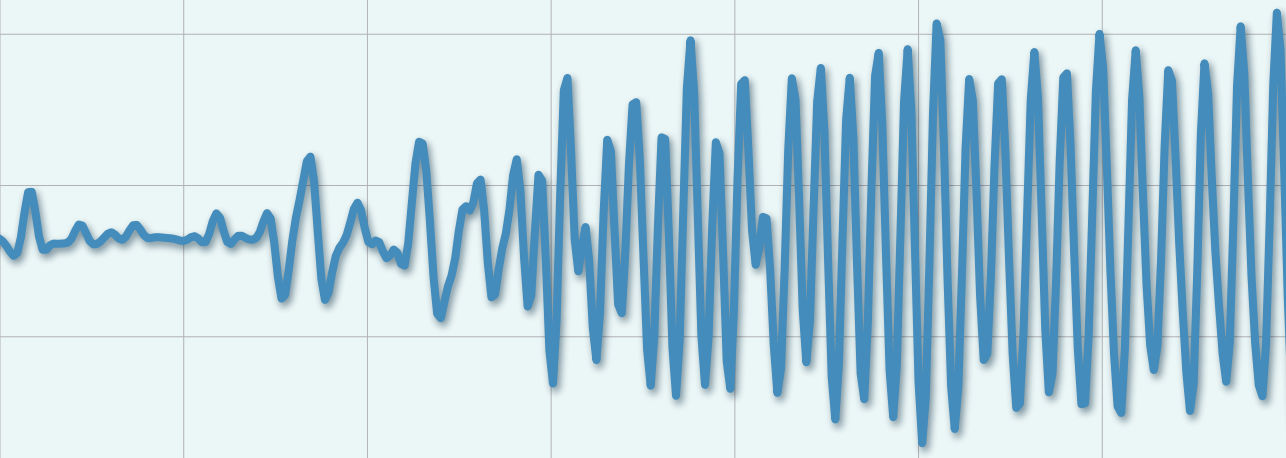
PART III

AUTOMATED MARKERS TO ENHANCE SAFETY MONITORING



CHAPTER 5

AUTOMATED VIDEO-BASED DETECTION OF NOCTURNAL CONVULSIVE SEIZURES IN A RESIDENTIAL CARE SETTING



Based on:
Geertsema EE, Thijs RD, Gutter T, Vledder B, Arends JB, Leijten FS, Visser GH, Kalitzin SN. Automated video-based detection of nocturnal convulsive seizures in a residential care setting. Epilepsia. 2018; 59 (S1): 53-60.

ABSTRACT

OBJECTIVE

People with epilepsy need assistance and are at risk of sudden death when having convulsive seizures (CS). Automated real-time seizure detection systems can help alert caregivers, but wearable sensors are not always tolerated. We determined algorithm settings and investigated detection performance of a video algorithm to detect CS in a residential care setting.

METHODS

The algorithm calculates power in the 2-6 Hz range relative to the total power (0.5-12.5 Hz) in group velocity signals derived from video-sequence optical flow. A detection threshold was found using a training set consisting of video-electroencephalogaphy (EEG) recordings of 72 CS. A test set consisting of 24 full nights of 12 new subjects in residential care and additional recordings of 50 CS selected randomly was used to estimate performance. All data were analyzed retrospectively. The start and end of CS (generalized clonic and tonic-clonic seizures) and other seizures considered desirable to detect (long generalized tonic, hyperkinetic, and other major seizures) were annotated. The detection threshold was set to the value that obtained 97% sensitivity in the training set. Sensitivity, latency, and false detection rate (FDR) per night were calculated in the test set. A seizure was detected when the algorithm output exceeded the threshold continuously for 2 seconds.

RESULTS

With the detection threshold determined in the training set, all CS were detected in the test set (100% sensitivity). Latency was ≤10 seconds in 78% of detections. Three/five hyperkinetic and six/nine other major seizures were detected. Median FDR was 0.78 per night and no false detections occurred in 9/24 nights.

CONCLUSIONS

Our algorithm could improve safety unobtrusively by automated realtime detection of CS in video registrations, with an acceptable latency and FDR. The algorithm can also detect some other motor seizures requiring assistance.

INTRODUCTION

The importance of monitoring people who are at risk because of their seizures has often been stressed.[1–3] After convulsive seizures (CS), defined here as generalized clonic and tonic–clonic seizures, interventions such as repositioning, stimulation, or clearing of the airway may have a protective effect in preventing sudden unexpected death in epilepsy (SUDEP).[4] The person with seizures is often in need of assistance or first aid due to (non–life-threatening) injury, but is not able to alert anyone. Alternative ways of alerting the caregiver are needed.

Several devices for automated seizure detection are on the market. Many seizure detection systems require sensors or complete devices to be attached to the individual.[5] Some patient groups such as children or people with intellectual disability may not tolerate wearable devices and may try to dislodge them. Unless properly concealed, such devices may also contribute to the social stigma associated with epilepsy. Alternatively, unobtrusive wireless sensors could be used, but these require regular charging and a reliable wireless connection to an alerting unit. Movement sensors that can be attached to the bed are widely used in nocturnal seizure monitoring and show fair detection performance for CS.[6,7] Such detectors, however, are only effective if the person is in bed. An alternative solution is remote detection.

Automated online analysis of video recordings can enable remote detection of the rhythmic vibratory or jerklike body movements in CS. Such a system would be privacy-friendly as there is no need for video storage or for someone to monitor output. A number of studies have been performed on detecting CS using video recordings.[8–14] These studies were, however, proofs of principle, showing detection feasibility in small datasets recorded in controlled clinical settings. There is currently no working system available that has been shown to have good performance in real life settings.

Previously, we presented an algorithm aiming to discern CS from normal behavior in video recordings.[14] The algorithm quantifies the oscillatory movements seen as vibrations during the tonic phase, and clonic movements in the clonic phase.[15,16] The algorithm showed promising CS detection performance in a video-electroencephalography (EEG) training set and is suitable for real-time use. There is, however, currently no information on the behavior of our seizure detection algorithm in daily practice. A detection threshold is not yet established, and algorithm performance has not been validated in new test data. This is required to make the algorithm functional and to provide practical guidelines to enable its use.

We aimed to determine a detection threshold and to investigate the detection performance, and variables influencing performance, of our noncontact seizure detection algorithm. We pursued a realistic performance estimate by analyzing long-term nightly video recordings in a residential care setting.

METHODS

VIDEO DATA

We retrospectively analyzed two separate video databases; a training set to find a suitable detection threshold and a test set to study detection performance. The training set is an existing video database that was described previously.[14,17] The detection algorithm was developed in 2012 using this database,[14] and in the present study we reused the database to find suitable detection settings. The test set is a novel video database consisting entirely of new subjects. Test set data were collected under the LICSENSE trial (NTR4115), by the Dutch TeleEpilepsy consortium, a collaboration between University Medical Center Utrecht, Stichting Epilepsie Instellingen Nederland, and Kempenhaeghe.[18] The study protocol was approved by a regional ethics committee and written informed consent was given by all participants or their guardians. All data were handled anonymously.

TRAINING SET

The training set consisted of 50 video-EEG recordings selected randomly from an epilepsy monitoring unit (EMU) database, recorded between 2003 and 2011. The training set contained 72 CS from 50 individuals. Videos were recorded with Bosch (Bosch Security Systems, B.V.) Dinion-LTC 0610, and Ikegami (Tsushinki Co., Ltd., Ohtaku, Tokyo, Japan) B/W CCD ICD-47 E-type cameras. All digitized recorded images were in mpeg2 format with a resolution of 352(H)x288(V) pixels and a fixed frame rate of 25 frames per second.

TEST SET

The test set comprises a selection from the video data collected in the LICSENSE trial, conducted in 2015 and 2016. In this observational study, performance of the Nightwatch (LivAssured BV), a wearable seizure detection system, was tested in residential care settings, where most residents have mild to severe intellectual disability. Those residents with at least one

monthly nocturnal CS were included and monitored at night for a period of 3 months. Caregivers kept a seizure diary. In a random 10% sample of all nights, the full video recording was reviewed by an experienced epilepsy nurse (off-line analysis), and seizures were annotated. Seizure annotations are limited to symptoms visible on the video.

In this study, 2 nights were selected from the 10% screening samples of 12 individuals - one night with one or more annotated CS, and another night without. If no CS were present, a night with CS was selected based on diary entries, or a second night without CS was selected. These nights were screened in the same way and by the same observers as the 10% screening sample. Additional CS recordings were included in the test set to enable accurate estimation of the sensitivity and latency of the algorithm, accounting for the low number of CS in the selected nights. In total, 50 recordings with CS were selected randomly from LICSENSE trial subjects. Overrepresentation of data from a particular subject in the performance estimates was prevented by incrementing the number of seizures selected per subject, until 50 seizures were included.

Videos were recorded with FOSCAM (Shenzhen Foscam Intelligent Technology Co., Ltd., Shenzhen, China) FI9805E Outdoor 960P PoE IP cameras, with infrared illuminator for night-time recordings. Recordings were in mp4 format, with a resolution of 640(H)x480(V) pixels. Frame rate was variable, but had a minimum of 25, and a stable mean of ~30 frames per second over the 4-second windows used in algorithm calculations. Each half hour the recording system recorded a new video, resulting in video epochs of up to 30 minutes.

SEIZURE ANNOTATIONS

Two neurologists (RT, GV, or JA), blinded to the results, independently determined seizure category to establish the detection desirability of the seizures found. Any incongruence was solved by consensus. Detection desirability of specific seizure types was determined as (see Table 5.1): *essential* - for CS (category I), as they are an important SUDEP risk factor[1,19,20]; *desirable* - for long generalized tonic, hyperkinetic, and other major seizures such as series of short myoclonic/tonic seizures (categories IIa, b, and c, respectively), as these seizures may be harmful or require assistance; and *nonclinically vital* - for minor seizures (category III). In the training set, only annotations of CS are used, which were based on the video-EEG report.

TABLE 5.1.
Seizure categories and the need for detection of these seizures.

Category	Description	Detection need
I	Convulsive seizures (CS): Tonic-clonic seizures (may start with a clonic phase) or generalized clonic seizures.	Essential
IIa	Tonic seizures that last longer than 30 seconds.	Desirable
IIb	Hyperkinetic seizures.	Desirable
IIc	Other major seizures: These seizures cannot be classified as tonic-clonic, tonic or hyperkinetic seizures and may include a cluster of short tonic or myoclonic seizures.	Desirable
III	Minor seizures: All other seizures.	Non-clinically vital

The start and end times of category I and II seizures in the test set were annotated by a trained technical physician (EG) who was blinded to detection algorithm results. The moment when the first behavioral change is observed, signifying clinical seizure onset, is annotated as the seizure start. Timing was based on the video and audio recording only. It is therefore possible that seizure activity started before the onset of a seizure is observable in the video. The end-of-seizure annotation is placed where the last observable seizure symptom ends and signifies the start of the postictal period. The start of the oscillatory period of CS was also annotated. The first visible oscillatory movements may either be vibratory movements during a tonic phase or the first clonic jerks of the clonic phase. Five-second annotation margins were applied before the seizure and oscillatory period start points to allow for slightly earlier detections.

CONVULSIVE SEIZURE DETECTION ALGORITHM

The detection algorithm used in this study was described previously.[14] The algorithm consists of 4 steps: 1) optical flow calculation,[21] reconstructing the vector field of velocities from luminance changes, presumably resulting from movements recorded by the camera; 2) reconstruction of group velocity parameters, obtaining 6 time series representing the rates of spatial transformations; the translation (horizontal and vertical), rotation, dilatation, and shear rates (horizontal and vertical); 3) extraction of the "seizureness spectrum," representing the dominant component of the time-frequency spectra of the 6 spatial group velocities. Spectra are calculated using Gabor aperture functions with central frequencies ranging from 0.5-12.5 Hz, in 1-s windows; 4) calculation of the spectral contrast quantity, defined as the power in the 2-6 Hz band relative to the total Gabor power (0.5-

12.5 Hz), in 4-s moving windows with 75% overlap. The 2-6 Hz frequency range was identified as the "spectral footprint" of CS. In some seizures the motion oscillations extend beyond this frequency range, but the range was considered optimal to minimize overlap with normal behavior. After the first 4 s of a registration, each second the algorithm generates a dimensionless output value between 0 and 1. These values correspond to very low (close to 0) to very high (close to 1) proportion of oscillatory movement in the 2-6 Hz frequency range, and with that the likelihood of registering a CS.

DETERMINATION OF ALGORITHM SETTINGS

To construct a functioning detector from the algorithm output, we implemented a detection threshold using the training set. A suitable threshold promotes detection sensitivity for CS as high as possible while keeping the number of false positives low. The detection threshold was set at the third percentile of detection output maxima during all oscillatory phases, obtaining 97% sensitivity in the training set. This was done to account for the possibility that not all CS had good quality recording; for example, caregivers may obstruct the view. A threshold resulting in 100% training set sensitivity would produce significantly more false positives.

After finding the detection threshold, a delay parameter was set to diminish the number of false positives caused by short oscillatory movements in the video. Suprathreshold algorithm output is ignored when the output does not stay above threshold for a duration equal to or longer than the delay parameter. The delay parameter was incrementally increased and set to the highest value where detection sensitivity was maintained, and latency did not increase more than the delay itself.

PERFORMANCE ANALYSIS

Detection performance was measured in terms of sensitivity and latency for CS (category I) and false detection rate (FDR) per night (8 hours). Detection performance for category II seizures is not a goal but considered a helpful sideline and is measured secondarily. Detection of category III seizures implies a false detection.

The algorithm detects a seizure when its output exceeds the threshold equal to or longer than the delay parameter during the seizure event. The seizure event is defined as the period between start and end annotations of the seizure. Detection latency is defined as the time between the start of the seizure and the detection. For CS, latency is also calculated from the start of

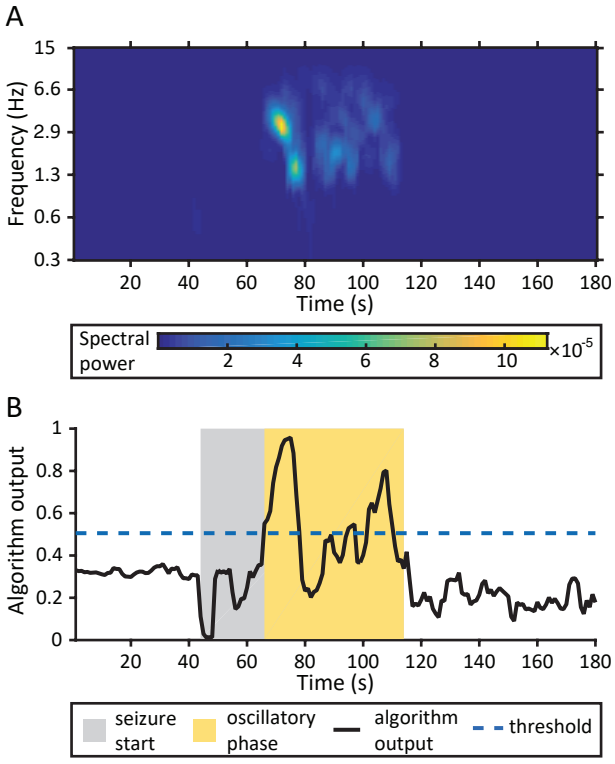


FIGURE 5.1.

Example of the algorithm output around the time of a convulsive seizure. A) Gabor time frequency spectrum, representing the dominant component of the time-frequency spectra of the 6 spatial transformations of the optical flow output. B) Convulsive seizure detection algorithm output (unitless, black solid line), defined as the power in the 2-6 Hz spectrum relative to the total spectral power. The colored boxes together represent the timeframe of the seizure, where the gray box shows the start of the seizure before the oscillatory phase, which is indicated with the yellow box. In this case, the seizure detection threshold (dashed blue line, at 0.51) is exceeded by the algorithm output just when the oscillatory phase of the seizure starts.

the oscillatory phase, as detection of the seizure before the start of this period is not expected due to the algorithm's sensitivity to rhythmic movements. Factors possibly influencing detection latency were investigated visually.

When the algorithm output exceeds the threshold equal to or longer than the delay parameter at times other than during seizure events that were considered essential or desirable for detection, a false detection is generated. If a second false detection occurs within a 10-s blackout period after the first detection, this second false detection is disregarded. After the blackout period, new false detections are again taken into account. False detections were categorized as the following: detections of category III seizures,

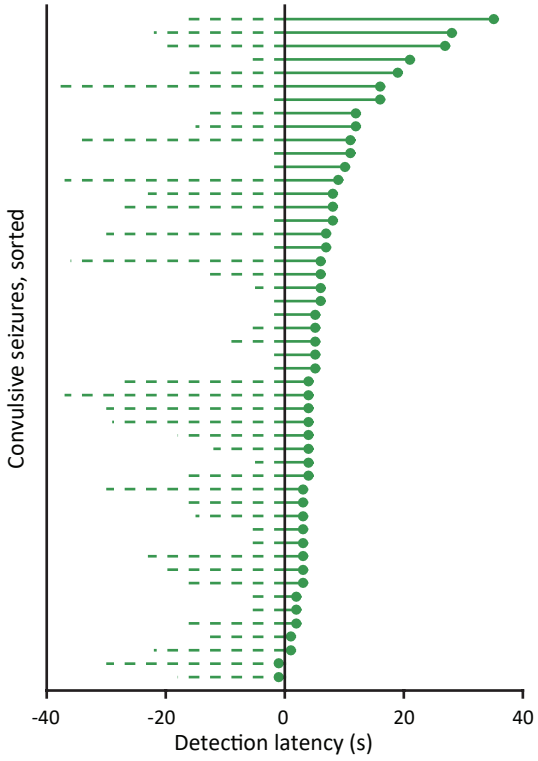


FIGURE 5.2.
Detection latencies for the 50 convulsive seizures in the test set, sorted according to latency. The circles indicate the moment in time when the detection was made, calculated from the start of the oscillatory phase (at time = 0 seconds). The dotted lines indicate the duration of the convulsive seizures before the oscillatory phase. This duration may include symptoms of a focal onset of the seizure.

detections during the postictal phase after a seizure, detections with caregivers present (not during a postictal phase), and other false alarms.

RESULTS

An example of the detection algorithm output is shown in Figure 5.1. The algorithm output threshold that detected 97% of seizures in the training set (threshold: 0.51) and a 2-s detection delay was applied to the algorithm for use in the test set. The test set consisted of recordings from 24 full nights (total duration ~253 hours) of 12 subjects, with 5 hyperkinetic and 9 major seizures recorded in 7 different subjects. Fifty CS were included from 9 different subjects (mean 5.5 seizures per subject, range 1-8). Six of the included seizures occurred within the 24 full-night recordings.

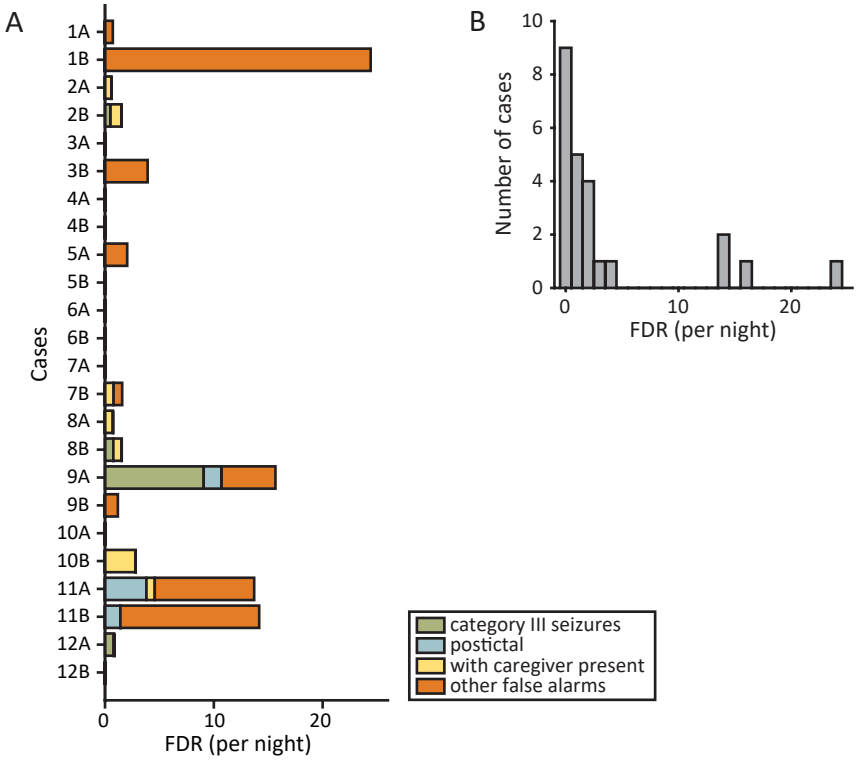


FIGURE 5.3.

A) Test set false detection rates (FDRs) per night (8 hours) for each case. The colors in the stacked bars indicate the situation in which the false detection occurred. Green bars indicate detections of nonclinically vital seizures (category III), blue bars indicate detections during the postictal phase after a seizure, yellow bars indicate detections when a caregiver was present (not during a seizure or postictal phase), and orange bars indicate other false alarms. B cases 2-6, 8, and 11 included a category I seizure. B) Histogram of FDR results per night, where each case is an observation. The first bar holds cases without false alarms.

All CS were detected in the test set (100% sensitivity), with latencies shown in Figure 5.2. Seventy-eight percent of CS were detected within 10 seconds from the start of the oscillatory period. Detection latencies for seizures in subjects covered completely by a blanket (n = 19) were not significantly different from latencies in those who were uncovered (n = 7) or partly covered (n = 24) (2-sample Kolmogorov-Smirnov test, $P>0.05$). In cases where detection latency was longer than 20 seconds, either a fluctuating oscillation amplitude was seen in the tonic phase, or caregivers were present, creating low-frequency "noise" in the video with their movements. Category II seizures were detected with a sensitivity of 57%. Detection latencies varied between 7 and 35 seconds. No tonic seizures longer than 30 seconds (category IIa) were registered in the test set. Three of five hyperkinetic seizures were detected (category IIb), and 5 of 9 other major seizures (category IIc).

False detection rates (FDRs) for all nights are shown in Figure 5.3. Median FDR was 0.78 per night (95% confidence interval [CI] 0-2.0 per night). No false detections occurred in 9/24 nights, which applied to both nights for 2 subjects. In more than half of the cases, an FDR of one or less per night occurred. FDR was high (>5 per night) in 4 cases, corresponding to 3 subjects. Detection of category III seizures caused 12% of false detections, which were all myoclonic jerks (frequent in case 9A). Eight percent of false detections occurred during the postictal period, due to physical restlessness (all after category I or II seizures). In 8% of false detections, the detection occurred when a caregiver was in the room (not during a postictal period). Examples of false detection causes in these cases were movement of a flash light beam or patient manipulation. Other false alarms (72%) were caused by patient behavior (45%), such as scratching and fidgeting (frequently in case 1B), and video disturbances (27%), such as objects (e.g., cobwebs, curtain cord) moving due to airflow fluctuations.

DISCUSSION

This study shows that our noncontact seizure detection algorithm can perform well in a test set of new cases, when applying detection settings that were optimized using a training set. The algorithm detected all CS with an acceptable latency in a test set of nightly video recordings in a residential care setting. Seizures with a 2-6 Hz oscillatory movement pattern, observed for 2 seconds or longer, are detected, even when the subject is covered by a blanket. Detection latency is minimally 2 seconds, plus the time it takes for the seizure activity to manifest in a clear oscillatory movement pattern. Hyperkinetic seizures and other major seizures with a 2-6 Hz movement pattern can also be detected. The algorithm's calculations are computationally light and use only the last registered 4 s of data, making it suitable for real-time use.

In more than half of cases, an FDR of one or less per night was observed. In a small number of cases, however, a high FDR was observed. False detections were most frequently caused by active behavior of an awake or postictal patient, that is, scratching or fidgeting. Myoclonic jerks were in some cases also detected despite their short duration, as oscillatory movements occurred when limbs bounced on the bed after an event. In some cases, video disturbances by objects moving in airflow caused false alarms. Combining detection output from video with other noncontact inputs (eg, sound) might diminish the chance of false positives, but possibly at the cost of sensitivity and latency. In light of the favorable sensitivity and latency findings with the

current algorithm, such a tradeoff to improve the FDR could be considered. Alternatively, awake users could be empowered to disable false detections manually, although giving a user time to disable a false alarm will inevitably lengthen detection latency. Which FDR is acceptable may depend on the individuals involved and their living conditions. If the subject can be observed from a distance via video connection in case of an alarm, like in a residential care setting, a higher FDR could be acceptable compared to a home setting where caregivers are awakened by alarms.

In this phase 2 study (according to the recently proposed standards for seizure detection studies[22]) an extensive dataset with randomly selected long-term recordings was used, where data was not clipped, edited, or filtered before automated processing. Bias was prevented by concealing the algorithm output from the experts annotating seizures. Prior to this, only feasibility of detection has been demonstrated, and the generalizability of detection results, particularly to real-life situations, remained unclear.[8,10,12–14] Performance was often calculated on the same data that was used to select appropriate detector settings, with short video fragments recorded in a controlled clinical environment and only the subject in view.

Most remote CS detection methods reported in literature are, like ours, targeted on movement periodicity (the exception being methods targeted on seizure sounds[23,24]). CS have been detected in video recordings by calculating periodicity in the luminance signal[10–12] and with neural networks trained on optical flow motion tracking output.[8,9] In another study, colored pyjamas were used to facilitate movement quantification for CS detection.[13] Compared to other algorithms targeted on periodicity, the detection delay we used (2 seconds) is shorter than generally applied (10 seconds),[6,10–12] while maintaining a low false detection rate. This can be attributed to the application of spectral contrast (opposed to power) and the output-smoothening effect of the 4-second calculation window.

We used an extensive test set, but all recordings were made during the night and all events of interest were derived from a small number of subjects. Detector settings were, however, based on a much larger sample in the training set, which contains 72 CS from 50 individuals in day- and nighttime recordings. It is likely, therefore, that detector sensitivity and latency in practice will be close to the sensitivity and latency in the training set (97% and <10 seconds in 81% of detected seizures, respectively). Realistic FDR results during daytime could not be derived from the fragmented video registrations in the training set. We expect false detections to be more frequent during the day in individuals with a tendency for false detection–

causing behavior. False detection rates might be different in other target groups that were not in the test set, such as children or adults outside of residential care.

The detection settings of our algorithm are presumably generic, as they were chosen using a data set from different subjects and video data from different hardware than the data used for validation. Although personalized detection settings might improve detection performance, this requires long (supervised) training sessions, making it less practical for direct deployment. Lengthening the detection delay could, for example, prevent false detections caused by short rhythmic movement patterns, while retaining sensitivity for CS if they have rhythmic movement patterns with a longer duration. If needed, personalization should be attempted only by professionals in a controlled setting, where video and EEG recordings enable checking and analysis of missed seizures and false detections.

CONCLUSIONS

Our detection algorithm could be used in a real-time automated noncontact monitoring system to increase the safety of people with epilepsy at home, without intruding on privacy, as no video storage or monitoring is necessary. The algorithm is highly sensitive to CS and false detection rates are low in most cases. For some subjects, application of our algorithm could be unsuitable in practice; that is, subjects with many false detections, who are unable to disable false alarms themselves. Future work should focus on prospectively evaluating real time detection performance of the algorithm in a broad target group of users.

REFERENCES

1. Langan Y, Nashef L, Sander JW. Case-control study of SUDEP. *Neurology*. 2005;64:1131–3.
2. Lamberts RJ, Thijs RD, Laffan A, et al. Sudden unexpected death in epilepsy: people with nocturnal seizures may be at highest risk. *Epilepsia*. 2012;53:253–7.
3. Massey CA, Sowers LP, Dlouhy BJ, et al. Mechanisms of sudden unexpected death in epilepsy: the pathway to prevention. *Nat Rev Neurol*. 2014;10:271–82.
4. Surges R, Thijs RD, Tan HL, et al. Sudden unexpected death in epilepsy: risk factors and potential pathomechanisms. *Nat Rev Neurol*. 2009;5:492–504.
5. van Andel J, Thijs RD, de Weerd A, et al. Non-EEG based ambulatory seizure detection designed for home use: What is available and how will it influence epilepsy care? *Epilepsy Behav*. 2016;57:82–9.
6. Narechania AP, Garić II, Sen-Gupta I, et al. Assessment of a quasi-piezoelectric mattress monitor as a detection system for generalized convulsions. *Epilepsy Behav*. 2013;28:172–6.
7. van Poppel K, Fulton SP, McGregor A, et al. Prospective Study of the Emfit Movement Monitor. *J Child Neurol*. 2013;28:1434–6.
8. Karayiannis NB, Tao G, Xiong Y, et al. Computerized motion analysis of videotaped neonatal seizures of epileptic origin. *Epilepsia*. 2005;46:901–17.
9. Karayiannis NB, Tao G, Frost JD, et al. Automated detection of videotaped neonatal seizures based on motion segmentation methods. *Clin Neurophysiol*. 2006;117:1585–94.
10. Kouamou Ntonfo GM, Ferrari G, Raheli R, et al. Low-complexity image processing for real-time detection of neonatal clonic seizures. *IEEE Trans Inf Technol Biomed*. 2012;16:375–82.
11. Pisani F, Spagnoli C, Pavlidis E, et al. Real-time automated detection of clonic seizures in newborns. *Clin Neurophysiol*. 2014;125:1533–40.
12. Cattani L, Alinovi D, Ferrari G, et al. Monitoring Infants by Automatic Video Processing: A Unified Approach to Motion Analysis. *Comput Biol Med*. 2017;80:158–65.
13. Lu H, Pan Y, Mandal B, et al. Quantifying limb movements in epileptic seizures through color-based video analysis. *IEEE Trans Biomed Eng*. 2013;60:461–9.
14. Kalitzin S, Petkov G, Velis D, et al. Automatic segmentation of episodes containing epileptic clonic seizures in video sequences. *IEEE Trans Biomed Eng*. 2012;59:3379–85.
15. Lhatoo S, Lüders H. Secondary generalised tonic-clonic seizures. In: Lüders HO, editor. Textbook of Epilepsy Surgery. London: CRC Press; 2008. p. 492–500.
16. Panayiotopoulos C. Epileptic seizures and their classification. In: A Clinical Guide to Epileptic Syndromes and their Treatment. Rev. 2nd E. London: Springer; 2010. p. 21–63.
17. Lamberts RJ, Laranjo S, Kalitzin SN, et al. Postictal generalized EEG suppression is not associated with periictal cardiac autonomic instability in people with convulsive seizures. *Epilepsia*. 2013;54:523–9.
18. van Andel J, Ungureanu C, Arends J, et al. Multimodal, automated detection of nocturnal motor seizures at home: is a reliable seizure detector feasible? *Epilepsia Open*. 2017;2:424–32.
19. Nilsson L, Farahmand BY, Persson PG, et al. Risk factors for sudden unexpected death in epilepsy: a case-control study. *Lancet*. 1999;353:888–93.
20. Hesdorffer DC, Tomson T, Benn E, et al. Do antiepileptic drugs or generalized tonic-clonic seizure frequency increase SUDEP risk? A combined analysis. *Epilepsia*. 2012;53:249–52.
21. Horn BKP, Schunck BG. Determining optical flow. *Artif Intell*. 1981;17:185–203.

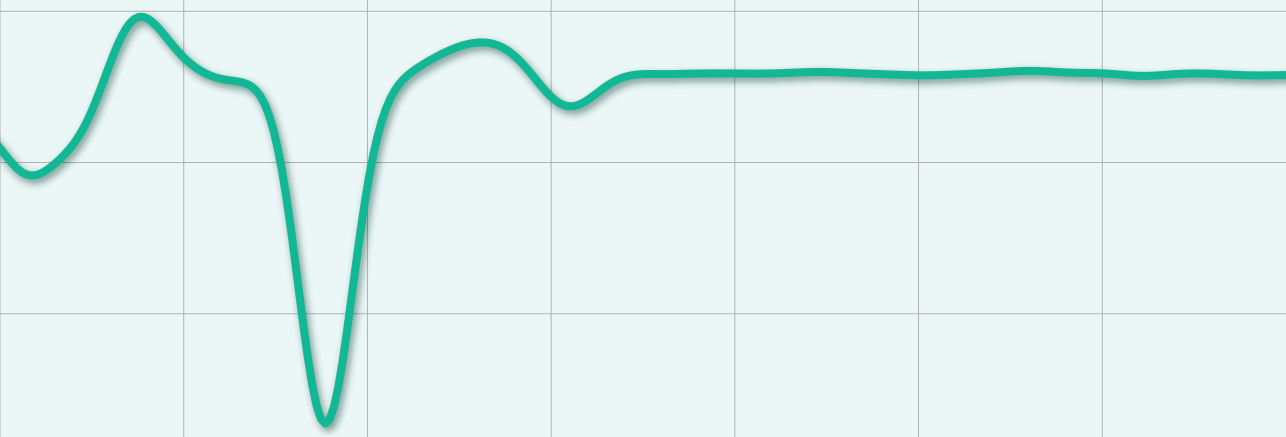
22. Beniczky S, Ryvlin P. Standards for testing and clinical validation of seizure detection devices. *Epilepsia*. 2018;59:9–13.

23. de Bruijne GR, Sommen PCW, Aarts RM. Detection of Epileptic Seizures Through Audio Classification. In: Vander Sloten J, Verdonck P, Nyssen M, Haueisen J, editors. *4th Eur Conf Int Fed Med Biol Eng IFMBE Proc*. Heidelberg: Springer Berlin; 2009. p. 1450–4.

24. Arends JB, van Dorp J, van Hoek D, et al. Diagnostic accuracy of audio-based seizure detection in patients with severe epilepsy and an intellectual disability. *Epilepsy Behav*. 2016;62:180–5.

CHAPTER 6

AUTOMATED REMOTE FALL DETECTION USING IMPACT FEATURES FROM VIDEO AND AUDIO



Based on:
Geertsema EE, Visser GH, Viergever MA, Kalitzin SN. Automated remote fall detection using impact features from video and audio. 2018. [under review]

ABSTRACT

OBJECTIVE

Elderly people and people with epilepsy may need assistance after falling, but may be unable to summon help due to injuries or impairment of consciousness. Several wearable fall detection devices have been developed, but these are not used by all people at risk. We present an automated analysis algorithm for remote detection of high impact falls, based on a physical model of a fall, aiming at universality and robustness.

METHODS

Candidate events are automatically detected and event features are used as classifier input. The algorithm uses vertical velocity and acceleration features from optical flow outputs, corrected for distance from the camera using moving object size estimation. A sound amplitude feature is used to increase detector specificity. We tested the performance and robustness of our trained algorithm using acted data from a public database and real life data with falls resulting from epilepsy and with daily life activities.

RESULTS

Applying the trained algorithm to the acted dataset resulted in 90% sensitivity for detection of falls, with 92% specificity. In the real life data, six/nine falls were detected with a specificity of 99.7%; there is a plausible explanation for not detecting each of the falls missed.

CONCLUSIONS

These results reflect the algorithm's robustness and confirms the feasibility of detecting falls using this algorithm.

INTRODUCTION

Like age-related health problems in elderly people, some seizure types in people with epilepsy cause loss of balance which may result in a fall.[1,2] Falls can cause serious injuries, especially when there is high impact with the floor or some other hard surface.[3] People who have fallen may be unable to summon help due to injuries or impairment of consciousness. Automated detection may help alert caregivers to seizure-related falls.

Several solutions have been proposed for automated detection of falls, often using accelerometers to measure movement patterns. Such devices need to be worn at all times, however, and this may be forgotten or not well tolerated by the subject. An alternative solution is remote fall detection.

Monocular video cameras can be used as sensitive, versatile and relatively cheap sensors to quantify movement, and automated online analysis of the video stream may enable remote detection of falls. Common features for detection include shape-related features such as changes in the person's shape,[4–9] bounding box dimension ratios,[10–14] and ellipsoids modelling posture changes.[15–17] Motion pattern features such as velocity and motion energy of the individual segmented in the frame have also been used to detects falls.[18–24] In many recent studies a combination of motion and shape related features has been used.[15,16,33,25–32] Other methods to detect falls in video recordings include inactivity detection[34] or anomaly/unusual event detection.[35,36]

Remote fall detection has also been performed using audio analysis,[37–39] and combining video and audio features could potentially increase detection performance. Video and audio classifiers have previously been combined, detecting a fall only if both classifiers provided a high fall probability.[40]

There is currently no working system for remote fall detection that has shown good detection performance on real-life data. Promising performance results have been reported in benchmark video datasets with acted falls, with sensitivity and specificity values >95%.[15,26–28,33,41,42] The generalizability of these results is questionable, as results on new data not used for training and, particularly, on real-life data are often lacking. When performed, application of trained algorithms on new data shows more modest performance results.[16,30,41] It has been suggested that real-life falls differ from acted falls, for example in terms of impact.[43] To our knowledge, there exist no studies on remote detection of seizure-related falls, which may differ from acted falls, or falls not resulting from seizures.

We propose an algorithm for remote detection of high impact falls, aiming for universality and robustness to new data and real-life epilepsy-related falls. We used half of a benchmark dataset with acted falls and normal activities to train the algorithm. We then tested the trained algorithm's performance and robustness on the other half of the benchmark dataset and on a newly collected test set containing real-life falls resulting from seizures.

METHODS

Rule-based analysis of video and audio signals, as deduced from a physical model of a fall, provides us with possible fall events. Unlike most fall detection algorithms that convert the entire time series to signal features for machine learning-based classification, our algorithm uses only signal features at the time of possible fall events. In this way we use prior knowledge of event properties to reduce the quantity of observations to classify. The algorithm does not require segmentation of the person in the video frames, but uses vertical velocity and acceleration analysis from optical flow outputs, complemented with a sound amplitude feature for increased detector specificity. The algorithm consists of three parts: preprocessing of the video and audio signals; event detection; and event classification. An overview of the algorithm is shown in Figure 6.1. All calculations were performed in Matlab (version 2017a, Mathworks Inc., Natic, USA).

DATASETS

We used two datasets for the development and validation of the fall detection algorithm; the publicly available Le2i fall detection database[26] and the SEIN fall database, a video database of recordings of genuine falls from people with epilepsy, collected at our center. Recording information is summarized in Table 6.1.

The Le2i database contains 221 videos simulated by actors, with falls in all directions, various normal activities and challenges such as variable illumination and occlusions. Videos were recorded from four room settings; 'coffee room', 'home', 'lecture room', and 'office', with 320 x 240 pixel resolution and a frame rate of 25 frames per second (fps). A subset of this database ('Office2') was excluded for lack of an audio stream in the video files. The remaining dataset (190 videos) was split in two randomly, providing a training set and a test set.

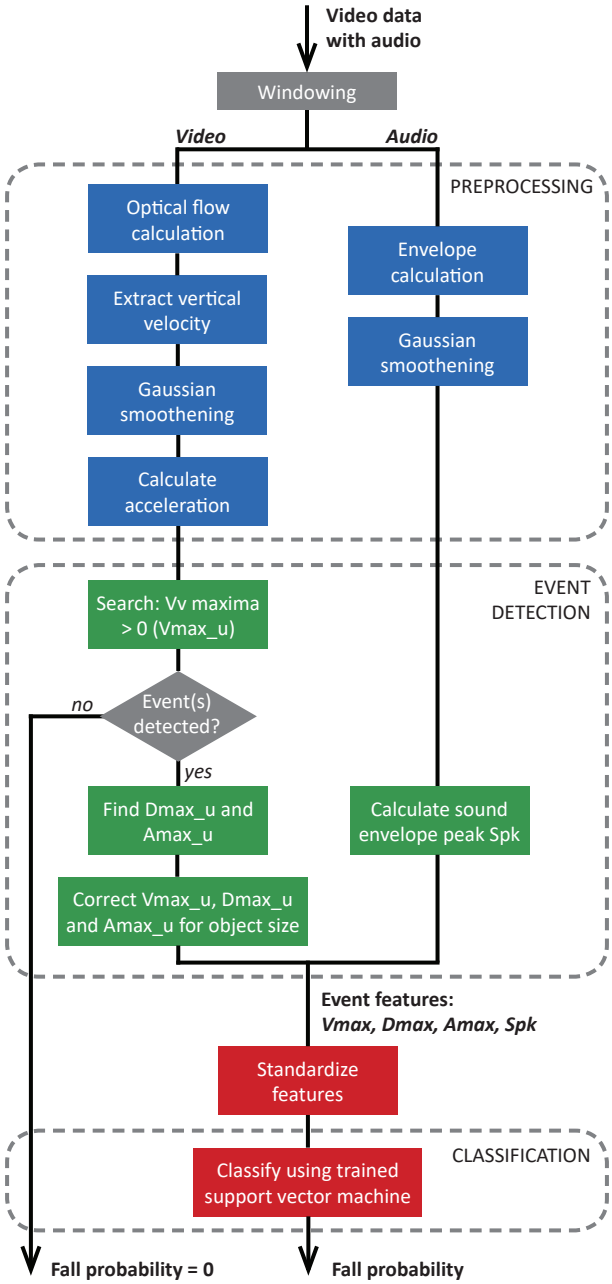


FIGURE 6.1.
Fall detection algorithm workflow. Video and audio streams are analyzed separately to provide signal features for events. Event features are only calculated when an event is detected, i.e. when a downward vertical velocity peak is found. Standardization and support vector machine parameters are obtained with algorithm training. $V_{v:}$ smoothed vertical velocity, V_{max}: positive maximum vertical velocity, D_{max}: deceleration maximum, A_{max}: acceleration maximum (_u: uncorrected for moving object size), S_{pk}: sound envelope peak.

TABLE 6.1.
Used datasets for training and testing the fall detection algorithm.

Dataset	Subjects (N)	N videos	N falls	Total video length (h)	Mean video length (s) [range]
Le2i					
- training set	Actors (9)	94	65	0.41	15.7 [5.3-78]
- test set	Actors (9)	96	65	0.44	16.5 [5.6-64]
SEIN test set	Patients (9)	9	9	1.45	578 [11-1600]

The SEIN fall database was selected from a historical video database containing clinical recordings of people who fell because of a seizure. Only videos with the fall happening in view of the camera were included. Five/nine videos were collected at an epilepsy monitoring unit, had a 1008 x 535 pixel resolution and an audio sampling frequency of 48,000 Hz. The remaining four videos came from several neurologists' personal seizure video databases, had a 352 x 288 pixel resolution and an audio sampling frequency of 44,100 Hz. All videos had a frame rate of 25 fps. The need for written informed consent from the individuals in the videos was waived by our institutional ethics committee. All data was handled anonymously.

Fall start and end times in the Le2i dataset were annotated by the makers of the dataset. The time of impact with the floor was annotated for falls in the training set by EG (one of the authors). Fall start and end times in the SEIN fall database were also annotated by EG.

PREPROCESSING

Video-audio data in the test sets is windowed to be able to count detector true negatives and estimate fall detection performance realistically. We applied a three-second calculation window shifted with one-second steps (two-second overlap) to obtain algorithm output every second.

Optical flow and subsequent vertical velocity calculations provide subject velocity estimates without having to first segment the subject from the video frames, promoting computational lightness. Choosing this preprocessing step also allows combining a seizure detection algorithm previously developed by our group[44] with the fall detection algorithm proposed in this paper into one detection system, without adding much extra computational cost.

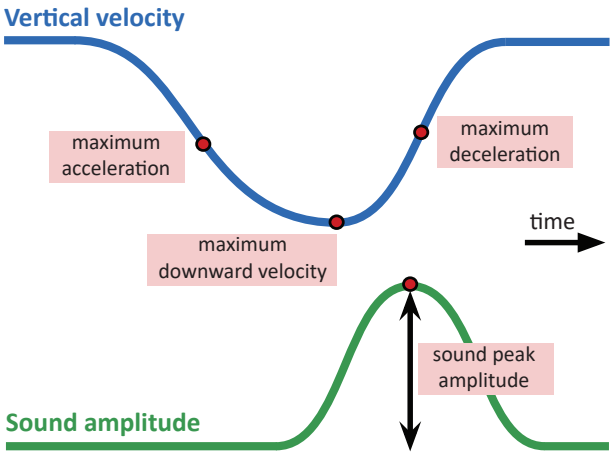


FIGURE 6.2.
Schematic illustration of the fall model used.

Video stream optical flow was estimated using the Horn-Schunck method[45] implemented as standard in Matlab, where the velocity vector field is obtained from the intensity field L(x,y,t) as a function of the 2-D spatial coordinates (x,y) and the time t:

$$L(x,y,t) \rightarrow V(x,y,t) = \{V_x(x,y,t), V_y(x,y,t)\}. \quad (6.1)$$

For further calculations we used only the vertical velocity time signal $V_y(t)$, defined as

$$V_y(t) \equiv \langle V_y(x,y,t) \rangle_{x,y}. \quad (6.2)$$

$V_y(t)$ was smoothed along the time samples, or frames, using a one-dimensional Gaussian filter with its scale parameter set to 4 samples (4/25 second), obtaining the smoothened vertical velocity $V_v(t)$. The scale parameter setting was fine-tuned to this value during classifier training (see "Algorithm settings and classifier training" Section). Additionally, vertical acceleration $A_v(t)$ was calculated by taking the first derivative of the smoothed vertical velocity signal $V_v(t)$. For the purpose of this study we selected the downward direction of V_v to be positive.

Audio was preprocessed to provide a smooth amplitude envelope. From the stereo sound recordings, we used only the first audio channel. The signal envelope was calculated using the Hilbert transform and smoothed using a one-dimensional Gaussian filter with its scale parameter set to 0.1 s. This filter scale setting was also fine-tuned during classifier training (see "Algorithm settings and classifier training" Section).

EVENT DETECTION AND FEATURES

To detect possible fall events, we used the following model of a fall, illustrated schematically in Figure 6.2. Once an individual starts to fall, downward velocity quickly increases and suddenly decreases again when the individual hits the ground. The quick increase and decrease in velocity are reflected in the maximum downward acceleration A_{max} and deceleration D_{max}. Between these two extrema is the point of maximum vertical downward velocity V_{max}. The impact of the person with the floor is accompanied by a sound. An event is detected when an A_{max}, V_{max}, D_{max} sequence, and an accompanying sound, are found within a calculation window.

To calculate the video event features, we first find positive velocity maxima and for each find its leading acceleration maximum and following deceleration maximum. Feature values are subsequently corrected for distance to the camera. Uncorrected maxima (V_{max_u}, A_{max_u}, D_{max_u}) are calculated at corresponding time points (t_V, t_A, t_D) for each calculation window (w). We define for any function f(t) of the discrete time variable t, the set of time points T_f^w of its positive local maxima within w as:

$$T_f^w = \{t \in w | f(t) > max(f(t-1), f(t+1)); f(t) > 0\}, \quad (6.3)$$

where f(t) can be replaced with $V_v(t)$ to obtain T_V, with $A_v(t)$ to obtain T_A, and with $-A_V(t)$ to obtain T_D. Only events with a consecutive $t_A < t_V < t_D$ sequence were considered eligible for a fall. The uncorrected features for each remaining event are:

$$A_{max\_u} \equiv A_v(t_A), \quad (6.4)$$

$$V_{max\_u} \equiv V_v(t_V), \quad (6.5)$$

$$D_{max\_u} \equiv -A_v(t_D). \quad (6.6)$$

The values of our video features depend on both the actual size of the object and its (focal) distance to the camera. Assuming that the only motion signal being observed is that of the subject, we can correct our video features for subject-camera distance or zoom angle changes using the size of the image area representing the velocity field of the subject. We can approximate the size of the image-footprint of a moving object from the optical flow velocity field at each time point t, using the determinant of the second moment tensor in the squared field's centroid. First, we spatially smoothed the magnitude of the velocity field $||V(x,y)||$, using a two-dimensional Gaussian

filter with its scale parameter set to 10 pixels, obtaining the smoothened field magnitude $V_{Ms}(x,y)$. For further calculations we used the squared smoothened magnitude of the velocity field

$$W(x,y) \equiv {V_{Ms}}^2(x,y), \tag{6.7}$$

emphasizing large-amplitude movements that are presumably the subject's. The total amount of movement at a given time, quantified by the sum $N_w = \sum_{x,y} W(x,y)$, will be used as a normalization factor considering (6.7) as a distribution density. The centroid location (x_c,y_c) of the squared velocity field can be found using:

$$x_c \equiv \frac{1}{N_w}\sum_{x,y} xW(x,y); \quad y_c \equiv \frac{1}{N_w}\sum_{x,y} yW(x,y). \tag{6.8}$$

The second moment tensor describes the spatial distribution of the velocity field around the centroid and is defined as

$$M \equiv \frac{1}{N_w}\begin{bmatrix} \sum_{x,y}(x-x_c)(x-x_c)W(x,y) & \sum_{x,y}(x-x_c)(y-y_c)W(x,y) \\ \sum_{x,y}(x-x_c)(y-y_c)W(x,y) & \sum_{x,y}(y-y_c)(y-y_c)W(x,y) \end{bmatrix}. \tag{6.9}$$

We can then approximate the moving object area in the field with

$$Q \equiv \frac{\sqrt{det\,(M)}}{N}, \tag{6.10}$$

where N is the number of pixels. A correction factor was obtained for each event, by averaging the calculated areas over time from the acceleration maximum to the velocity maximum:

$$C_{event} \equiv \langle Q \rangle_{t\in[t_A,t_V]}. \tag{6.11}$$

The three event velocity features can then be corrected as follows:

$$A_{max} \equiv A_{max\_u}/C_{event}, \tag{6.12}$$

$$V_{max} \equiv V_{max\_u}/C_{event}, \tag{6.13}$$

$$D_{max} \equiv D_{max\_u}/C_{event}. \tag{6.14}$$

For time windows with an event we obtained a fourth feature, based on the sound peak during the event. The window's relative peak amplitude (S_{pk}) was calculated using the smoothed envelope's (S_{env}) maximum and minimum:

$$S_{pk}(w) \equiv \frac{max\,(S_{env})}{min(S_{env})}. \tag{6.15}$$

If a window contained multiple events, these received the same sound feature S_{pk}.

CLASSIFICATION

We classified events using a support vector machine (SVM) with a radial basis function (rbf) kernel. The training procedure that provided the SVM hyperparameters and model parameters is described in the "Algorithm settings and classifier training" Section. In each three-second analysis window zero, one, or multiple events may be detected. The V_v fall waveform with corresponding acceleration extrema and sound peak need to occur in the same calculation window for them to constitute a detection. To prevent double detections of a single event owing to overlapping analysis windows we define that t_V is between t = 1s and t = 2s of its corresponding window. If an event was detected, its features were passed to the classifier. The resulting posterior probability was stored as the classifier output for the window. If the window contained multiple events, all were passed separately to the classifier and the maximum posterior probability was stored as output. If no event was present in the analysis window, classifier output was set to zero.

ALGORITHM SETTINGS AND CLASSIFIER TRAINING

Algorithm settings were optimized in the following order: 1) SVM hyperparameters, 2) Gaussian filter scale parameters (used in audio and video preprocessing), 3) feature set standardization parameters. Settings optimization was followed by 4) classifier training, which provided the SVM parameters. In all four steps of this procedure we used only the training set in a ten-fold cross-validation scheme.

The classifier was trained to separate fall events from non-fall events based on event features, aiming for high sensitivity for falls. A sub selection of events was used for training; one event per fall, and a maximum of five non-fall events (selected randomly) per video. The number of non-fall events per video was limited to prevent over-representation of specific situations of a registration in the dataset. To select the right fall event for training (there may be multiple instances of V_{max}), we assumed that the event with the largest velocity within the [impact time -1 s,impact time +1 s] timeframe indicates the fall. In total, the training set contained 370 events, of which 94 were labeled 'fall'.

We tuned the SVM hyperparameters; the soft margin constant, regulating misclassification penalty, and the rbf kernel scale, determining model

flexibility. Tuning aimed to optimize precision and accuracy at 90% sensitivity. To favor high sensitivity over accuracy, misclassification cost for false negative detections was increased until 90% sensitivity (in the cross-validation) was achieved. This process resulted in a subset of suitable hyperparameter settings, as no clear optimum was found. From this subset, we chose the hyperparameter settings that provided the most robust SVM parameters, minimizing variance over the folds.

Next, we fine-tuned the scale parameters for the Gaussian filters used to smooth the vertical velocity and sound envelope signals. The filter scale parameters were originally chosen visually to distinguish signals during falls. Both filter scale parameters were finetuned by optimizing precision and accuracy at 90% sensitivity in the training set.

The feature set was transformed logarithmically and standardized to zero mean and unit standard deviation. The standardization parameters were saved to standardize future observations. Finally, the SVM was trained on the entire standardized training set, again using a 10-fold cross-validation scheme. Other possible sub-selections of features, classification models, and SVM kernels were also explored, but showed worse performance and were thus pursued no further.

PERFORMANCE ANALYSIS

Performance was analyzed in the (so far unused) Le2i test set and the SEIN test set. A fall was detected (true positive, TP) when one of the windows overlapping with fall start-end annotations received supra-threshold classifier output. If none of the windows overlapping with the fall annotation received supra-threshold output, the fall was not detected (false negative, FN). Each non-fall window wrongly classified as a fall provided a false positive detection (FP). Non-fall windows correctly classified provided true negative (TN) detections. Detector sensitivity and specificity were calculated using detection counts N:

$$sensitivity = \frac{N_{TP}}{N_{TP}+N_{FN}} \tag{6.16}$$

$$specificity = \frac{N_{TN}}{N_{FP}+N_{TN}} \tag{6.17}$$

A receiver operating characteristics (ROC) curve for the Le2i test set, plotting sensitivity and specificity for different posterior probability thresholds, provided: ROC area under the curve (AUC), and specificity (SPEC) and positive predictive values (PPV) in different ROC working points. We also

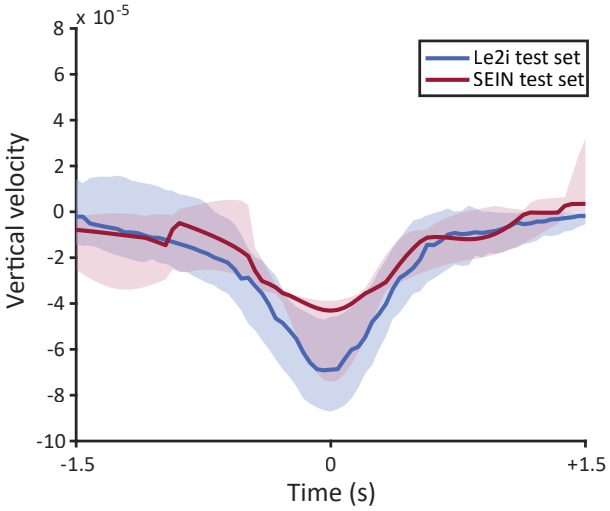


FIGURE 6.3.
Median fall V_V waveforms (solid lines) with interquartile ranges (shaded areas), for the Le2i test set (N=65 falls) and SEIN test set (N=6 falls with fall waveform). Fall waveforms were synchronized by putting all fall V_V minima at t=0.

analyzed the distribution of false positive fraction (FPF) values for individual recordings at 90% sensitivity, defined as the fraction of non-fall windows wrongly classified as a fall. Performance in the SEIN test set was analyzed qualitatively to provide directions for use of the algorithm in practice.

RESULTS

We first compared fall V_v waveforms observed in both test sets to assess the feasibility of applying the trained algorithm to the SEIN test set data, with falls that are possibly different. The results in Figure 6.3 show that fall waveforms in the SEIN test set are similar to those in the Le2i test set. Smaller median V_v amplitudes were observed in the SEIN test set. This could be caused by an overall larger distance of the subjects from the camera. Because differences between amplitudes of both test sets remained after object size correction (results not shown), we surmise that velocities of the real-life falls in the SEIN test set are smaller than those in experimental falls, as has also been described in Reference [43].

Applying the trained classifier to the Le2i test set resulted in the ROC curve shown in Figure 6.4. Results for different ROC curve working points are summed up in Table 6.2. Le2i test set results when using a feature set without sound are also shown, to illustrate the added value of the sound

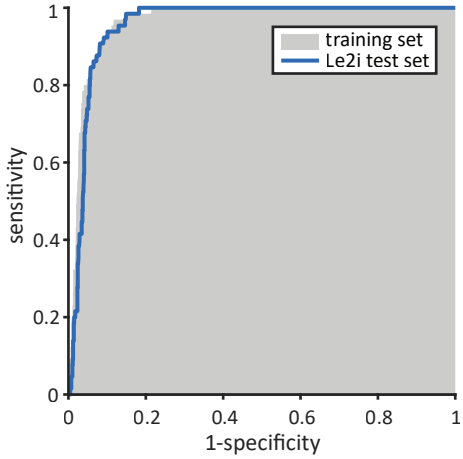


FIGURE 6.4.
Receiver operating characteristics (ROC) curve of fall detection algorithm testing results for the Le2i test set. Training set performance is shown in the background for reference. ROC curves were constructed using different classifier posterior probability thresholds.

TABLE 6.2.
Fall detection performance results for the Le2i test set. Results from using the full feature set, and for only video features are shown. Specificity (SPEC) and positive predictive values (PPV) are given for three working points on the ROC curves chosen according to their sensitivity values. ROC AUC: receiver operating characteristic area under the curve.

	ROC AUC	100% sensitivity		90% sensitivity		80% sensitivity	
		SPEC	PPV	SPEC	PPV	SPEC	PPV
Video & sound	0.957	0.818	0.248	0.919	0.401	0.945	0.468
Only video	0.947	0.799	0.231	0.896	0.345	0.923	0.385

feature. FPF distributions for individual Le2i test set recordings obtained with the posterior probability threshold corresponding to 90% sensitivity, are shown in Figure 6.5. Most of the videos (61%) show zero false positives, encompassing 43% of the total length of the dataset. In 75% of videos, accounting for 71% of total dataset length, a low FPF<0.1 was found. Two relatively short videos had a high FPF>0.3. With a small number of non-fall windows the FP fraction easily becomes high. Most FPs in the Le2i test set occurred when chairs were put down on hard flooring, and video and audio features resembled values of a fall.

Applying the detection threshold that provided 90% sensitivity in the training set in the SEIN test set resulted in detection of six/nine falls. Inspection of the videos and V_v waveforms of undetected falls provided explanations: 1) camera movement (tilt, controlled by staff) cancelled out the

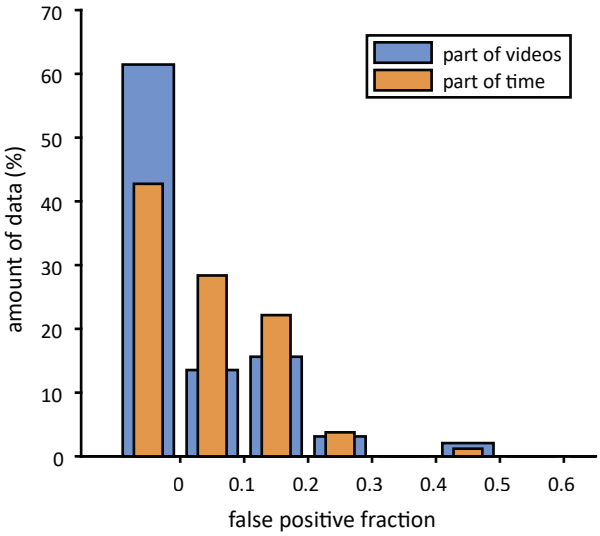


FIGURE 6.5.
False positive fraction (FPF) histograms obtained using the 90% sensitivity posterior probability threshold in the Le2i test set. The dark green histograms show the percentage of video registrations. The amount of time encompassed by the subset of videos within this FPF range is given in light green. The first bin of the histogram holds only videos without any false detections.

fall movement. Static camera positions, as likely in the intended monitoring system, would prevent this. Also, staff-controlled camera movements imply staff observing the fall. 2) A caregiver caught the falling subject before he hit the ground, and 3) the subject loses balance and makes several quick steps towards the back of the room before falling; the negative V_v from the quick upward movement (due to camera placement high on the wall) cancelled out the fall movement. Only situation 3) would realistically result in a missed fall. There were 13 FPs in total, resulting in an overall FPF of 0.0025 (SPEC = 0.997, PPV = 0.32) in the SEIN test set. In four/nine videos there were no false positives. Two FPs were caused by camera movements (controlled by staff), the other 9 FPs occurred during the presence of extra people (caregivers) in view.

DISCUSSION

Our fall detection algorithm trained on a benchmark dataset is able to detect acted and seizure-related falls in data not used for training. High impact falls occurring in any direction with respect to the camera are detected, as long as the vertical component of the fall is visible. In most records, acceptable false positive fractions are found for 90% sensitivity. We consider these results

proof-of-concept of the method, and a promising outcome for eventual use of the method in clinical practice.

Fall detection sensitivity values >90% have been reported with >99% specificity in the Le2i database,[26,28,41] but when algorithms were applied to new datasets with more challenging scenes, more modest performance results were obtained (up to 69% specificity at 60-75% sensitivity).[16,30,41] Whereas the performance of our algorithm in the Le2i database (92% specificity at 90% sensitivity) seems not fully competitive, we used a realistic performance testing scheme which possibly provides less favorable outcomes. We tested performance using shorter detection windows (leaving relatively more non-fall windows) and a random selection of data from all room settings instead of only a subset. We were able to show robustness of our algorithm in a separate test set and in new, challenging data with real-life falls. Our algorithm might also detect non-seizure falls, but was not tested on real-life non-seizure data. As the implementations of published algorithms are not publicly available, we could not test them on our SEIN test set.

Specificity of our algorithm needs to be improved further and validated in ongoing recordings made in real-life environments, where correct detection can be especially challenging.[16] Training the algorithm on real-life movement patterns may improve the algorithm's ability to handle real-life events. Subject segmentation before calculation of the video features could possibly improve detection performance, but could also make the algorithm more sensitive to occlusions. Using a more fall-specific sound feature (e.g. one of the features described in Ref. [37–39]), as opposed to the relatively course amplitude feature in our current algorithm, might also improve performance.

Postprocessing of detection output, for example by waiting after an alarm before generating new alarms, can reduce the false alarm rate when there is some ongoing disturbance. A smart system able to inactivate the detection module when a companion is present could prevent false detections due to multiple individuals in the scene, whereby it is assumed that the companion detects the fall. Similarly, automated inactivation at times of camera movements could prevent false positives. Note that using a 'subject inactivity after a fall' rule, sometimes applied in algorithms to decrease false positives, could decrease sensitivity for seizure-related falls, which may be followed by excess physical (seizure) activity.

Our fall detection algorithm is designed to detect high impact falls with high risk of injury and is less suitable for low impact falls without a clear

moment of impact with the floor. An elderly person more sliding than falling off a chair for example, would likely not result in detection by our algorithm. Falls with soundless impact could in theory also be missed, which raises the question if soft flooring could impede detection. Many falls in the Le2i database involved subjects falling on mattresses, which were detected without problems. Visual inspection of training set sound signals confirmed the presence of peaks at times of mattress-falls. This suggests that soft flooring will probably not impede detection.

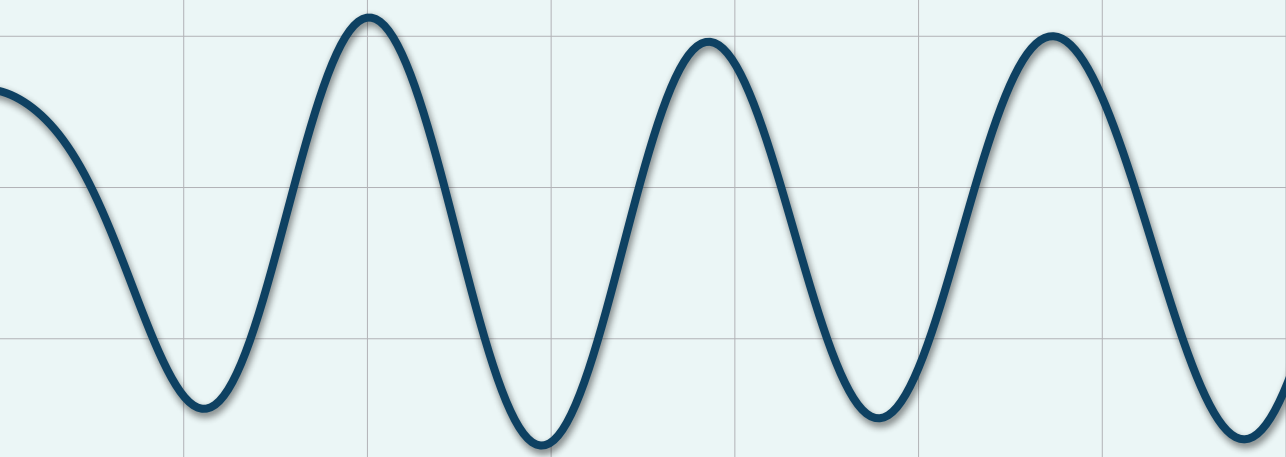
Our video features are amplitude-based and are dependent on camera placement (vertical component needed to detect a fall), camera zoom and subject distance to the camera. The proposed features derived from the physical fall model are, however, generic, and can be applied to other sensors able to quantify vertical velocity. Feature values can be corrected for (focal) distance using moving object size, but this may also introduce errors; for example in case of falling objects. In the SEIN test set, object size correction increased sensitivity (one extra TP) and specificity (two FPs less). Usage of depth sensors could provide more precise distance information and improve performance.[46,47] Although commercially available and affordable, depth cameras have a limited distance range (<5 meters[48]), making them less suitable for application in larger living areas. Depth sensors with a longer range might become more affordable in time.

REFERENCES

1. Rubenstein LZ. Falls in older people: Epidemiology, risk factors and strategies for prevention. *Age Ageing*. 2006;35:ii37-ii41.
2. Krumholz A, Hopp J. Falls give another reason for taking seizures to heart. *Neurology*. 2008;70:1874–5.
3. Russell-Jones DL, Shorvon SD. The frequency and consequences of head injury in epileptic seizures. *J Neurol Neurosurg Psychiatry*. 1989;52:659–62.
4. Mirmahboub B, Samavi S, Karimi N, et al. Automatic monocular system for human fall detection based on variations in silhouette area. *IEEE Trans Biomed Eng*. 2013;60:427–36.
5. Rougier C, Meunier J, St-Arnaud A, et al. Robust video surveillance for fall detection based on human shape deformation. *IEEE Trans Circuits Syst Video Technol*. 2011;21:611–22.
6. Goudelis G, Tsatiris G, Karpouzis K, et al. Fall detection using history triple features. In: *Proc 8th ACM Int Conf PErvasive Technol Relat to Assist Environ*. 2015.
7. Yu M, Yu Y, Rhuma A, et al. An online one class support vector machine-based person-specific fall detection system for monitoring an elderly individual in a room environment. *IEEE J Biomed Heal Informatics*. 2013;17:1002–14.
8. Belshaw M, Taati B, Snoek J, et al. Towards a single sensor passive solution for automated fall detection. In: *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS*. 2011. p. 1773–6.
9. Hazelhoff L, Han J, de With PHN. Video-Based Fall Detection in the Home Using Principal Component Analysis. In: Blanc-Talon J, Bourennane S, Philips W, Popescu D, Scheunders P, editors. Advanced Concepts for Intelligent Vision Systems ACIVS 2008 Lecture Notes in Computer Science, vol 5259. Springer Berlin Heidelberg; 2008. p. 298–309.
10. Tao J, Turjo M, Wong M-F, et al. Fall Incidents Detection for Intelligent Video Surveillance. In: *2005 5th Int Conf Inf Commun Signal Process*. 2005. p. 1590–4.
11. Miaou SG, Sung PH, Huang CY. A customized human fall detection system using omni-camera images and personal information. In: *Conf Proc - 1st Transdiscipl Conf Distrib Diagnosis Home Heal D2H2 2006*. 2006. p. 39–42.
12. Alhimale L, Zedan H, Al-Bayatti A. The implementation of an intelligent and video-based fall detection system using a neural network. *Appl Soft Comput*. 2014;18:59–69.
13. Liu C-L, Lee C-H, Lin P-M. A fall detection system using k-nearest neighbor classifier. *Expert Syst Appl*. 2010;37:7174–81.
14. Vishwakarma V, Mandal C, Sural S. Automatic detection of human fall in video. In: *Int Conf Pattern Recognit Mach Intell*. 2007. p. 616–23.
15. Feng W, Liu R, Zhu M. Fall detection for elderly person care in a vision-based home surveillance environment using a monocular camera. *Signal, Image Video Process*. 2014;8:1129–38.
16. Debard G, Mertens M, Deschodt M, et al. Camera-based fall detection using real-world versus simulated data: How far are we from the solution? *J Ambient Intell Smart Environ*. 2016;8:149–68.
17. Foroughi H, Aski BS, Pourreza H. Intelligent video surveillance for monitoring fall detection of elderly in home environments. In: *Proc 11th Int Conf Comput Inf Technol ICCIT 2008*. 2008. p. 219–24.
18. Rougier C, Meunier J, St-Arnaud A, et al. Monocular 3D head tracking to detect falls of elderly people. In: *Annu Int Conf IEEE Eng Med Biol - Proc*. 2006. p. 6384–7.

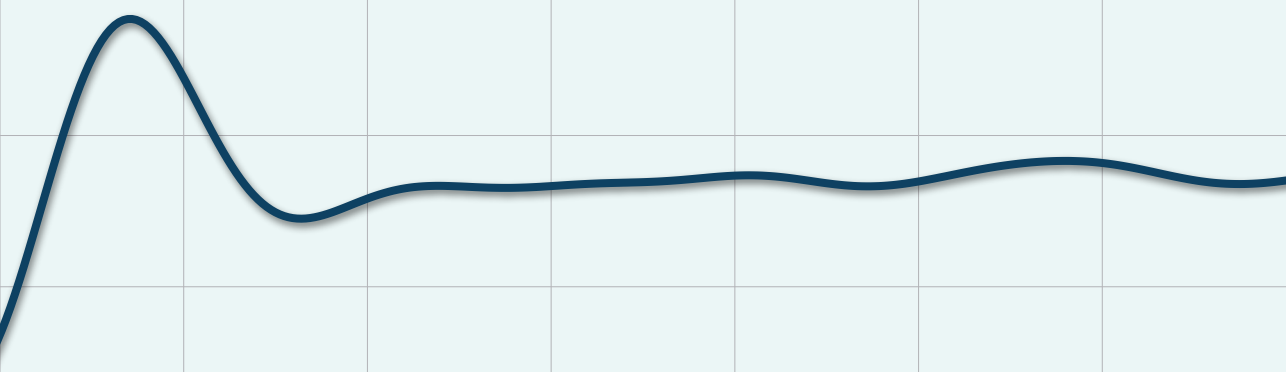
19. Zhang Z, Tong LG, Wang L. Experiments with Computer Vision Methods for Fall Detection. In: *Proc 3rd Int Conf PErvasive Technol Relat to Assist Environ (PETRA '10)*. 2010.
20. Olivieri DN, Gómez Conde I, Vila Sobrino XA. Eigenspace-based fall detection and activity recognition from motion templates and machine learning. *Expert Syst Appl*. 2012;39:5935–45.
21. Khawandi S, Daya B, Chauvet P. Implementation of a monitoring system for fall detection in elderly healthcare. In: *Procedia Comput Sci*. Elsevier; 2011. p. 216–20.
22. Liu T, Yao H, Ji R, et al. Vision-Based Semi-supervised Homecare with Spatial Constraint. In: Advances in Multimedia Information Processing - PCM 2008, Vol 5353 of Lecture Notes in Computer Science. Heidelberg: Springer Berlin; 2008. p. 416–25.
23. Huang C, Chen E, Chung P. Fall Detection Using Modular Neural Networks With Back-Projected Optical Flow. *Biomed Eng Appl Basis Commun*. 2007;19:415–24.
24. Wu G. Distinguishing fall activities from normal activities by velocity characteristics. *J Biomech*. 2000;33:1497–500.
25. Liao YT, Huang C-L, Hsu S-C. Slip and fall event detection using Bayesian Belief Network. *Pattern Recognit*. 2012;45:24–32.
26. Charfi I, Miteran J, Dubois J, et al. Optimized spatio-temporal descriptors for real-time fall detection: comparison of support vector machine and Adaboost-based classification. *J Electron Imaging*. 2013;22:041106.
27. Yun Y, Gu IYH. Human fall detection in videos by fusing statistical features of shape and motion dynamics on Riemannian manifolds. *Neurocomputing*. 2016;207:726–34.
28. Senouci B, Charfi I, Heyrman B, et al. Fast prototyping of a SoC-based smart-camera: a real-time fall detection case study. *J Real-Time Image Process*. 2016;12:649–62.
29. Makantasis K, Protopapadakis E, Doulamis A, et al. 3D measures exploitation for a monocular semi-supervised fall detection system. *Multimed Tools Appl*. 2016;75:15017–49.
30. Debard G, Karsmakers P, Deschodt M, et al. Camera-based fall detection on real world data. In: Dellaert F, Frahm J-M, Pollefeys M, Leal-Taixé L, Rosenhahn B, editors. Outdoor and Large-Scale Real-World Scene Analysis Lecture Notes in Computer Science, vol 7474. Springer, Berlin, Heidelberg; 2012. p. 356-75.
31. Debard G, Baldewijns G, Goedem T, et al. Camera-based fall detection using a particle filter. In: *2015 37th Annu Int Conf IEEE Eng Med Biol Soc*. 2015. p. 6947–50.
32. Zweng A, Zambanini S, Kampel M. Introducing a statistical behavior model into camera-based fall detection. In: Bebis G, Boyle R, Parvin B, Koracin D, Chung R, Hammoud R, et al., editors. Advances in Visual Computing ISVC 2010 Lecture Notes in Computer Science, vol 6453. Springer, Berlin, Heidelberg; 2010. p. 163–72.
33. de Miguel K, Brunete A, Hernando M, et al. Home Camera-Based Fall Detection System for the Elderly. *Sensors*. 2017;17:2864.
34. Nait-Charif H, McKenna S. Activity summarization and fall detection in a supportive home environment. In: *IEEE Int Conf Pattern Recognit*. 2004. p. 20–3.
35. Nater F, Grabner H, Jaeggli T, et al. Tracker trees for unusual event detection. In: *2009 IEEE 12th Int Conf Comput Vis Work ICCV Work*. 2009. p. 1113–20.
36. Zerrouki N, Harrou F, Sun Y, et al. A Data-Driven Monitoring Technique for Enhanced Fall Events Detection. *IFAC-PapersOnLine*. 2016;49:333–8.
37. Li Y, Ho KC, Popescu M. A microphone array system for automatic fall detection. *IEEE Trans Biomed Eng*. 2012;59:1291–301.

38. Popescu M, Li Y, Skubic M, et al. An Acoustic Fall Detector System that Uses Sound Height Information to Reduce the False Alarm Rate. In: *30th Annu Int IEEE EMBS Conf*. 2008. p. 4628–31.
39. Salman Khan M, Yu M, Feng P, et al. An unsupervised acoustic fall detection system using source separation for sound interference suppression. *Signal Processing*. 2015;110:199–210.
40. Toreyin BU, Dedeoglu Y, Cetin AE. HMM Based Falling Person Detection Using Both Audio and Video. In: Sebe N, Lew M, Huang T, editors. Computer Vision in Human-Computer Interaction HCI 2005 Lecture Notes in Computer Science, vol 3766. Springer Berlin Heidelberg; 2005. p. 211–20.
41. Fan Y, Levine MD, Wen G, et al. A deep neural network for real-time detection of falling humans in naturally occurring scenes. *Neurocomputing*. 2017;260:43–58.
42. Wang S, Chen L, Zhou Z, et al. Human fall detection in surveillance video based on PCANet. *Multimed Tools Appl*. 2016;75:11603–13.
43. Kangas M, Vikman I, Nyberg L, et al. Comparison of real-life accidental falls in older people with experimental falls in middle-aged test subjects. *Gait Posture*. 2012;35:500–5.
44. Kalitzin S, Petkov G, Velis D, et al. Automatic segmentation of episodes containing epileptic clonic seizures in video sequences. *IEEE Trans Biomed Eng*. 2012;59:3379–85.
45. Horn BKP, Schunck BG. Determining optical flow. *Artif Intell*. 1981;17:185–203.
46. Ma X, Wang H, Xue B, et al. Depth-based human fall detection via shape features and improved extreme learning machine. *IEEE J Biomed Heal Informatics*. 2014;18:1915–22.
47. Stone EE, Skubic M. Fall detection in homes of older adults using the microsoft kinect. *IEEE J Biomed Heal informatics*. 2015;19:290–301.
48. Horaud R, Hansard M, Evangelidis G, et al. An overview of depth cameras and range scanners based on time-of-flight technologies. *Mach Vis Appl*. 2016;27:1005–20.



CHAPTER 7

AUTOMATED NON-CONTACT DETECTION OF CENTRAL APNEAS USING VIDEO



Based on:
Geertsema EE, Visser GH, Sander JW, Kalitzin SN. Automated non-contact detection of central apneas using video. 2018 [under review].

ABSTRACT

OBJECTIVE

To develop a robust automated non-contact algorithm for real-time detection of central apneas using video cameras.

METHODS

One video registration with simulated apneas and nine with real-life apneas associated with epileptic seizures, each recorded from 3-4 angles, were used to develop the algorithm. Videos were preprocessed using optical flow, from which translation, dilatation and shear rates were extracted. Presence of breathing motions was quantified in the dominant time-frequency spectrum by calculating the relative power in the respiratory range (0.1-1 Hz). Sigmoid modulation was calculated over different scales to quantify sigmoid-like drops in the respiratory range power. Each sigmoid modulation maximum constitutes a possible apnea event. Two event features were calculated to enable distinction between apnea events and movements: modulation maximum amplitude and total spectral power modulation at the time of the event. An ensemble support vector machine was trained to classify events using a bagging procedure and validated in a leave-one-subject-out cross validation procedure.

RESULTS

All apnea episodes were detected in the signals from at least one camera angle. Integrating camera inputs capturing different angles increased overall detection sensitivity (>90%). Specificity of >99% was achieved with individual cameras and integrated camera inputs.

CONCLUSIONS

These results show that it is feasible to detect central apneas automatically in video, using this algorithm. When validated, the algorithm may be used as an online remote apnea detector for safety monitoring.

INTRODUCTION

Central apneas occurring during epileptic seizures are usually self-limiting,[1] but when occurring in the aftermath of a seizure may lead to asystole and could trigger Sudden Unexpected Death in Epilepsy (SUDEP).[2] Observed fatal SUDEP cases occurred within minutes after a convulsive seizure,[2] but not all known cases were preceded by a seizure.[3] Interventions such as repositioning or clearing of airways may have a protective effect preventing SUDEP[4,5] and automated seizure detection may help alert carers. If a seizure was not detected or intervention could not prevent apnea, carers should be alerted to apneas to enable timely resuscitation.

Currently apneas are detected by a respiratory inductance plethysmogram (RIP), an impedance pneumogram, or a blood oxygen saturation measurement. These contact-sensors are, however, not always suitable or tolerated. They would need to be worn continuously, may be uncomfortable and sensors could come loose during seizures. Some populations such as children or people with intellectual disability may not tolerate wearable devices and may try to dislodge them. An alternative solution would be apnea detection using remote sensors.

There are several modes of remote respiratory monitoring using different sensors, but there are few to detect apneas. Algorithms aiming at remote apnea detection use radar,[6] sonar,[7] infrared sensors,[8,9] depth sensors,[10–12] and video.[13–16] Video cameras are suitable for safety monitoring, as they are relatively cheap and sensitive to movement, even at longer distance, provided they have suitable resolution.

Respiration can be quantified in video by using the breathing motions of the trunk or the respiratory modulation of the photoplethysmogram (PPG). PPG techniques are unsuitable for our application, as they need skin regions that may not always be visible (covered or turned away from the camera), and color video, which cannot be used at night. Trunk breathing motions may be quantified, for example, by calculating translation rates with optical flow,[17] tracking image points over time,[18] or comparing breathing motion templates.[13] Most published algorithms aim to detect obstructive sleep apneas (OSA), often using the characteristic sounds and movements that conclude OSA for classification. These algorithms are unsuitable to detect central apneas; in which such sounds and movements are absent.

There is no algorithm available that can detect central apneas in video and is suitable for (postictal) monitoring. This algorithm needs to be 1) able to

work near real-time, providing minimal time delay between apnea onset and its detection, 2) able to distinguish between apnea and gross body motion when breathing motions cease, 3) independent of camera imaging mode to enable daytime and nighttime use, and 4) independent of skin or chest visibility and tracking.

We developed a robust automated algorithm for real-time non-contact detection of central apneas in video registrations. Our algorithm detects cessation of the measured oscillatory movements of breathing, in the absence of gross body motion. Registrations of simulated apneas and real-life, seizure-related central apneas were used to develop and test the algorithm in a leave-one-subject-out cross-validation scheme. We investigated detection performance in a single-camera setup from different angles and propose two different strategies to combine signals from multiple cameras to improve performance.

METHODS

The algorithm consists of three parts: preprocessing of the video signals, event detection and event classification. Video signals from multiple-camera setups were integrated in two ways: early, during preprocessing, or late, after classification. An overview of the algorithm is show in Figure 7.1. An example of algorithm output after different video processing steps is shown in Figure 7.2.

VIDEO REGISTRATIONS

One healthy apnea-simulating subject, and nine subjects with epilepsy (but otherwise healthy), who had seizure-related apneas, were included in this study. The subjects with epilepsy were selected from a database which was prospectively collected (2016-2017) in an epilepsy monitoring unit for the development of seizure detection algorithms. Next to the standard long-term (≥24 h) synchronized electro-encephalogram, video, sound, and electrocardiogram, respiration was quantified for this dataset using RIP belts. The study protocol was approved by the ethics committee of University Medical Center Utrecht and written informed consent was given by all participants or their guardians. All data were handled anonymously.

Subjects were registered with three or four cameras positioned around the room, so the face could be brought in view from at least one angle. Subjects were not always in view of all cameras. Depending on the position of the

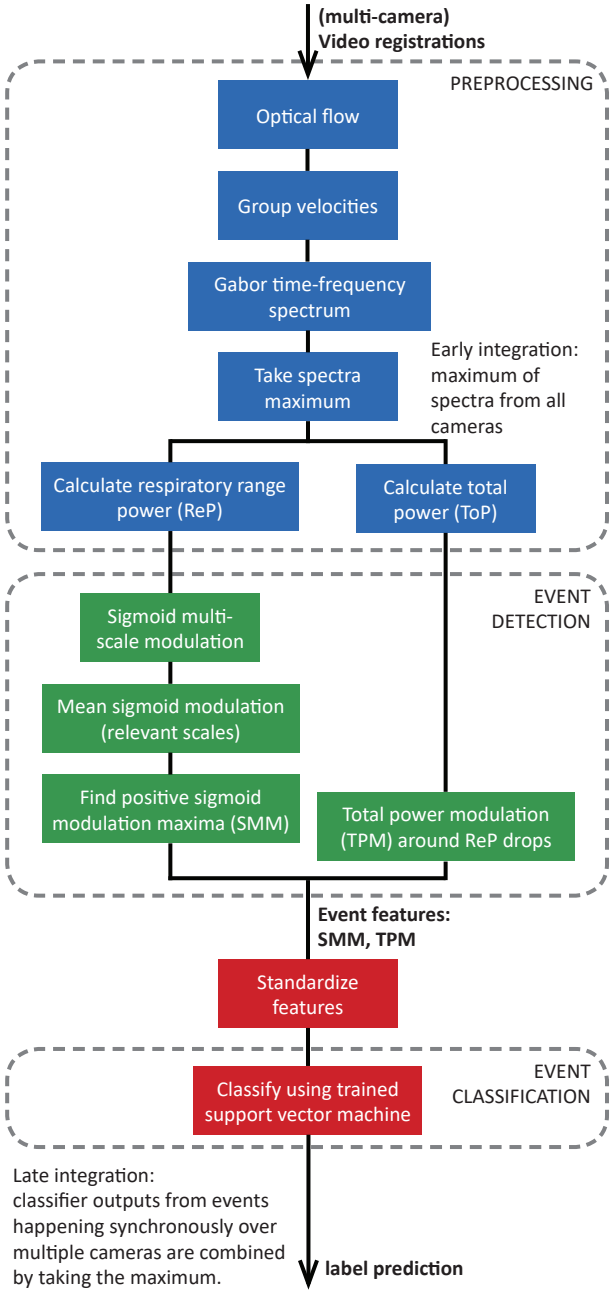


FIGURE 7.1.
Apnea detection algorithm workflow. Total power modulation is only calculated at the time of a detected event, i.e., a positive mean sigmoid modulation maximum.

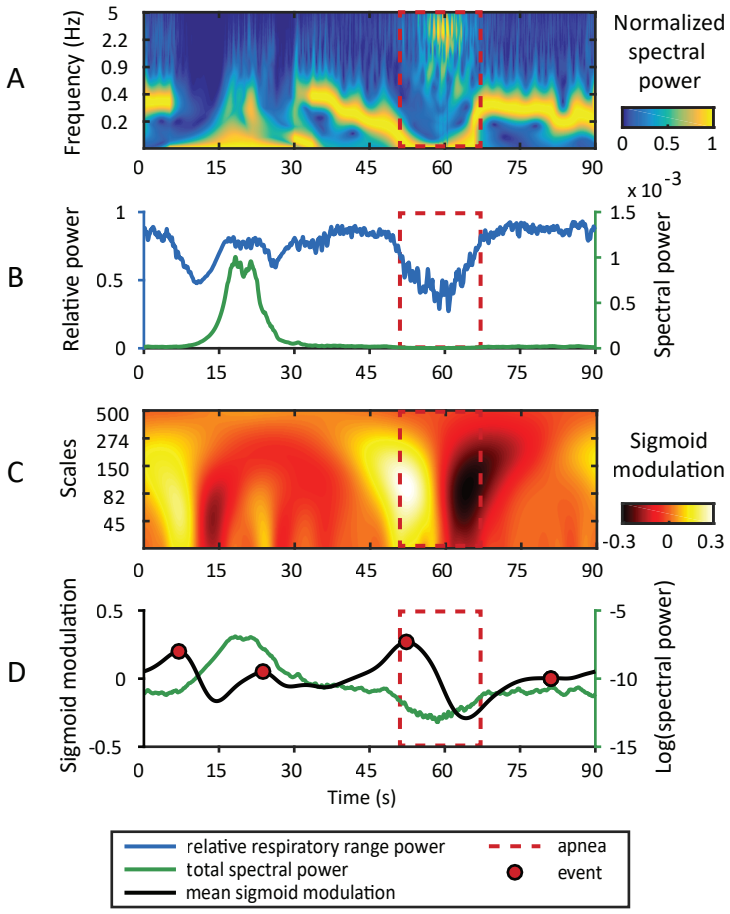


FIGURE 7.2.
Example of the output from different analysis steps in the apnea detection algorithm. A 90-second video fragment with a movement (at t≈15 s) and an apnea (starting at t≈49 s) is shown (subject SIM1, camera 1). (a) Normalized Gabor time-frequency spectrum, showing the dominant component from the five group velocity spectra. (b) Relative power in the respiratory range (Rep) plotted together with the total power. (c) Sigmoid modulation of the ReP, showing high power at significant signal amplitude drops. (d) Mean sigmoid power modulation, calculated over relevant sigmoid scales that fit apneas. Total power is scaled logarithmically here. Events were detected at positive mean sigmoid power modulation maxima.

subject moving freely in the room, subject-to-camera distance ranges 1-5 m. Camera rotation, zoom and pan could be controlled by nurses in a monitoring room. The registration system started a new video automatically after approximately 26 minutes, for storage purposes. Videos were registered using a frame rate of 25 frames per second. Each frame, with a total resolution of 1080x1920 pixels, contains bordered sub-frames from the four cameras, with a 535x959 pixel resolution. Sub-frames were separated after reading for further analyses.

TABLE 7.1.
Dataset information. Relevant apneas were used for training and considered essential to detect. Unique apneas may be observed from different angles in multiple cameras.

ID	Age	Registration length (min)	N cameras	N apneas (unique)		N relevant apneas (unique)	
SIM1	29	16.0	4	36	(9)	30	(8)
PAT1	66	6.00	3	2	(1)	0	(0)
PAT2	45	17.7	3	3	(1)	0	(0)
PAT3	53	26.6	4	4	(1)	0	(0)
PAT4	45	26.6	4	3	(1)	2	(1)
PAT5	18	26.6	3	27	(9)	4	(2)
PAT6	38	26.6	4	3	(1)	2	(1)
PAT7	46	26.6	3	6	(2)	0	(0)
PAT8	27	19.9	4	4	(1)	3	(1)
PAT9	16	26.6	4	4	(1)	2	(1)
Total		219	36	92	(27)	43	(14)

DATA SELECTION

As peri-ictal central apneas are common,[1] the minutes around the seizures of each subject in the database were analyzed for the occurrence of apneas. During this screening, the signals from the respiratory belts were used to locate apneas visually. We presume peri-ictal apneas to be of central origin, as has been described in literature.[1] We did not confirm this with other sensors. For each subject that had peri-ictal apneas, one video was included in the dataset in which one or more apneas occurred. One of the subjects was 16 years old at the time of registration, the remaining subjects were adults. Registration information is shown in Table 7.1.

ANNOTATIONS

To provide ground truth information on the presence and timing of apnea, the beginning and end of each apnea was annotated. Annotations were based on the signals from one, or if available, both RIP belts. For each apnea we noted in which of the cameras the subject was visible, if the subject was moving at apnea onset, and whether there were other persons in view of one of the cameras. Apneas without gross subject movement or other persons in view, and a duration of ≥8 s were considered 'relevant apneas'. Relevant apneas were used during algorithm training and were considered essential to detect during algorithm testing. An apnea may be considered relevant in

a subset of the cameras used in the registration, as a gross body movement or second person may only be visible in a subset of the cameras.

PREPROCESSING

Video registrations were preprocessed in order to extract: 1) respiratory range oscillatory power, which quantifies the presence of a breathing motion signal, and 2) total power, to quantify the presence of gross body movement. The preprocessing procedure used here is similar to the procedure used in a convulsive seizure detection algorithm previously developed by our group.[19] The main difference is the focus on the presence of oscillatory motions in the frequency range of respiration, instead of the frequency range of oscillatory motions in convulsive seizures. A similar preprocessing procedure also allows combining the seizure detection algorithm with the apnea detection algorithm into one detection system, without adding much extra computational cost.

First, subject movement in the video was quantified with optical flow, using the Horn-Schunck method,[20] implemented as standard in Matlab. From the resulting velocity fields five timeseries were derived, representing the spatial transformations, or group velocities; translation rates along the two image axes, dilatation rates and shear rates. Here, we chose to omit rotation, originally the sixth group velocity, for its lack of contribution to the respiratory signal.

Next, the spectral content in each of the five group velocities was calculated using 200 Gabor filters with exponentially spaced central frequencies, v, in the range [0.08,5] Hz. In order to obtain the dominant component of the group velocities, we take the maximum of the five Gabor time-frequency spectra (G_c):

$$Q(t,v) = max_c\big(G_c(t,v)\big), \quad c = 1,\dots,5. \tag{7.1}$$

Here, t is defined as time and c defines the corresponding group velocity. An example of the Q(t,v) spectrum at the time of breathing, movement, and apnea is illustrated in Figure 7.2a. When early integration of video-signals was applied, the maximum in (7.1) was taken over the (15 or 20) group velocities from all cameras.

Relative power in the respiratory frequency band (ReP) was calculated as the power in the 0.1-1 Hz band relative to the total Gabor power (0.08-5 Hz):

$$ReP(t) = \frac{\sum_{v=0.1}^{1} Q(t,v)}{\sum_{v} Q(t,v)} \tag{7.2}$$

In video segments without gross body movement but with the presence of breathing motions, the ReP quantity will presumably be close to one. Note that we do not aim to quantify the respiration frequency with this quantity, but only the presence (and cessation) of respiration. Mean total power (ToP) is calculated over all measured frequencies:

$$ToP(t) = \langle Q(t,v) \rangle_v . \tag{7.3}$$

An example of the ReP(t) and ToP(t) signals at the time of breathing, movement, and apnea is illustrated in Figure 7.2b. Both the ReP(t) and ToP(t) signals were scaled logarithmically before further calculations.

EVENT DETECTION

Apneas manifest as sigmoid-like amplitude drops in the ReP(t) signal. To quantify drop occurrence we calculated sigmoid modulation for different scales to detect drop events of different slopes and amplitudes in ReP(t). This technique is analogous to computing a wavelet time-frequency spectrum. Here, we compute a sigmoid time-scale spectrum instead, using an aperture sigmoid as the generating wavelet template. In this way, the slope (scale) of the drop of ReP(t) and its amplitude (relative to the mean value) can be estimated. The technique is described in detail below.

First, we define a range of 200 scales s_k (k=1,2,...,200), with exponentially spaced values in the range [25,500]. For each scale, an aperture sigmoid template is defined for window τ:

$$S(\tau,k) = N_k^{S^{-1}} \frac{e^{\tau/s_k} - e^{-\tau/s_k}}{e^{\tau/s_k} + e^{-\tau/s_k}} e^{-|\tau|^2/{s_k}^2}, \quad \tau = [-3s_k : 3s_k], \tag{7.4}$$

together with the Gaussian aperture template:

$$G(\tau,k) = N_k^{G^{-1}} e^{-|\tau|^2/{s_k}^2}. \tag{7.5}$$

In Equations (7.4) and (7.5) L2 normalization was applied through the coefficients $N_k^{S,G^{-1}}$, with N_k defined as the squared sum of the k^{th} aperture template. The time window in (7.4) and (7.5) is chosen to be of three scale lengths, as values outside this range are suppressed by the Gaussian aperture factor. Sigmoid time-scale modulation m can then be obtained using the convolutions between the filters and the ReP signal:

$$m(t,k) = \frac{\int_{\tau}^{\infty} S(t-\tau)ReP(\tau)d\tau}{\int_{\tau}^{\infty} G(t-\tau)ReP(\tau)d\tau} \quad (7.6)$$

An example of the sigmoid modulation at the time of breathing, movement, and apnea is illustrated in Figure 7.2c.

To quantify the presence of significant respiratory range power drops, we calculated the mean sigmoid modulation M over the scales that correspond to observed drop times:

$$M(t) = \langle m(t,k)\rangle_{k\in s_{drop}} \quad (7.7)$$

Drop times were observed to be between 4.0 and 8.2 seconds (in the SIM1 registration), and correspond to filters $s_{drop}=[s_{70},s_{129}]$. A positive mean sigmoid modulation maximum corresponds to a significant respiratory range power drop. Its amplitude constitutes our first event feature, to be used in classification. As the M(t) signal fluctuates, many positive maxima and thus events will be detected. High maximum values are however only found when there is a significant respiratory power drop. For the positive local maxima of M we define the set of time points T_M as:

$$T_M = \{M(t) > max(M(t-1), M(t+1)); M(t) > 0\} \quad (7.8)$$

For each event, the sigmoid modulation maximum amplitude is defined as:

$$SMM(t_M) = M(t_M). \quad (7.9)$$

An example of the mean sigmoid modulation and its positive maxima forming the events, is shown in Figure 7.2d.

When an apnea occurs, the total power stays low, or decreases from a low to an even lower value. At the time of events due to gross body movement on the other hand, the total power either increases (movement onset) or drops from a high to a lower value (movement end). An example of the total power change at the time of movement and apnea events is shown in Figure 7.2d. A second classification feature quantifying the change of total power at the time of events may distinguish events due to apneas from events due to gross body movements. For each event we therefore calculated the total power modulation; comparing the two seconds before, to the two seconds after the M maximum:

$$TPM(t_M) = \frac{\langle ToP\rangle_a - \langle ToP\rangle_b}{\langle ToP\rangle_a + \langle ToP\rangle_b}, \quad b = [t_M - 2s, t_M], \quad a = [t_M, t_M + 2s]. \quad (7.10)$$

Presumably, the TPM feature has a small and often negative value for apnea events, and a high value (positive of negative) for gross body movement events.

All events received a ground truth label according to the apnea annotations: 1 (apnea) or 0 (not an apnea). A five second annotation margin allowed detections slightly earlier than the annotated apnea onset. If multiple events were detected within an apnea period, the first event was labelled as the apnea event. Other events during the apnea period were disregarded; an apnea can only be detected once, but an extra detection during an apnea should not be considered a false detection. Relevant apneas were also labelled separately. If no events were found during an apnea, the apnea was missed, but as only events received a ground truth label, such a false negative would not be counted. We therefore corrected performance results by hand to make sure any missed apneas without an event are counted as a false negative.

CLASSIFICATION

To automatically classify detected events as "apnea" or "not an apnea", we used a support vector machine (SVM) with a radial basis function (RBF) kernel. Optimization of the SVM's hyperparameters was followed by training, which provides SVM model parameters. During both hyperparameter and model parameter setting, precision and accuracy were optimized while ensuring 90% sensitivity for apnea events. Each time an SVM model was trained, the misclassification cost for a false negative was increased with 0.1 steps until 90% sensitivity on its training data was achieved.

HYPERPARAMETER OPTIMIZATION

We found the SVM hyperparameters that provided high performance while generalizing well, using a grid-search. The soft margin constant and the RBF kernel scale were tuned to the combination (0.1 and 2, respectively) that performed best in cross validation.

TRAINING AND CROSS VALIDATION

A leave-one-subject-out (LOSO) cross validation procedure was used to train the event classifier and test its performance on new data not used for training. In each of the ten folds the classifier is trained on data from nine/ten subjects and tested on the remaining subject. The training and testing procedure is depicted in Figure 7.3.

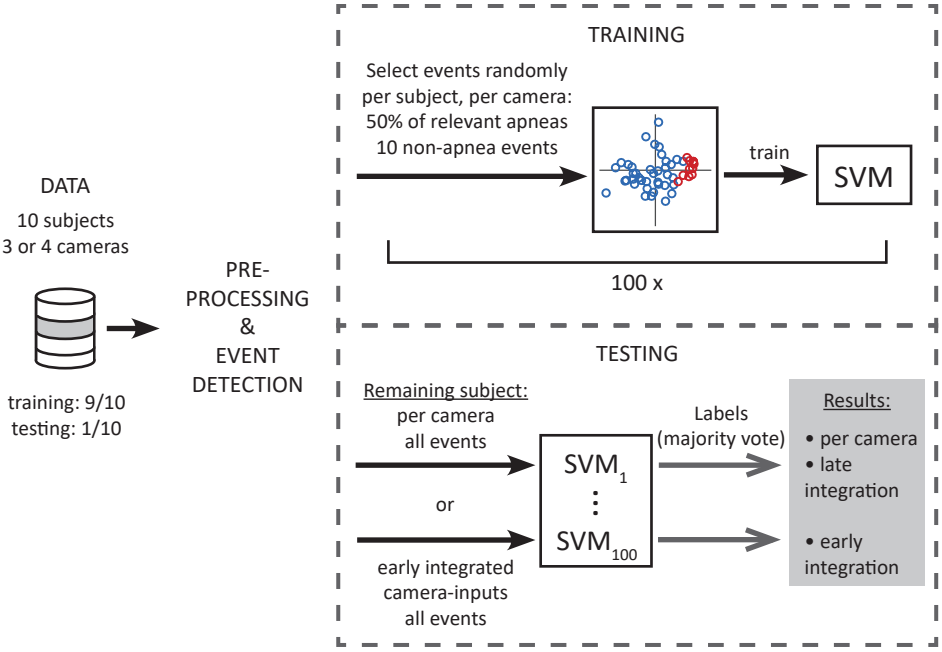


FIGURE 7.3.
Leave-one-subject-out cross validation procedure. In each of the subject-folds training is performed 100 times with a bag of randomly selected training events from 9/10 subjects, resulting in an ensemble of 100 trained SVM models. Testing on the data from the remaining subject is done either 1) on all events per camera, providing predictions per camera, which may be integrated later, or 2) on the events from the early integrated camera inputs.

In each fold, an ensemble classifier was trained using a bagging (i.e. bootstrap aggregating) procedure, for increased classifier stability and accuracy. A sample, with replacement, from the training events (bag) was used to train a classifier. Each bag sample consisted of a random selection of 50% (with a minimum of one) of the relevant apnea events and 10 non-apnea events, selected from each camera registration in the training set. This process was repeated 100 times in each fold to obtain an ensemble classifier.

The classifier is subsequently tested on the registration of the remaining subject of the fold, with two different sets of input events: the events found in the video signals from each individual camera, and the events found in the early integrated video signals. A prediction from an ensemble model was formed by the majority vote of the 100 predicted labels. Late integration of outputs was performed by combining synchronous events, i.e., the maximum output label was obtained for events that occurred within 5 s over two or more cameras. In practice, one camera providing a detection is enough to get a detection from a late integrated camera system.

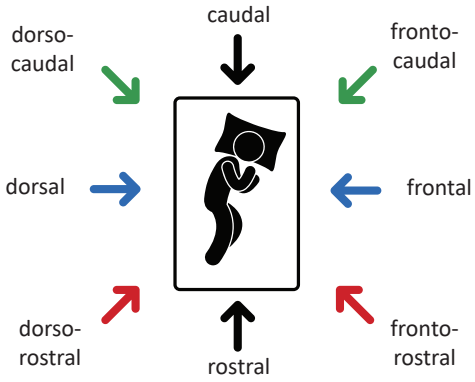


FIGURE 7.4.
Probable camera angles when monitoring a subject. Opposite camera positions form a combination (indicated with an equal color), as they may represent the same view when the subject turns. Adapted from "Sleep by Gan Khoon Lay from the Noun Project".

PERFORMANCE EVALUATION

Comparison of predicted labels with ground truth labels yielded the following counts: true positives (TP); detected apneas, true negatives (TN); events that were correctly classified as not an apnea, false positives (FP); events incorrectly classified as an apnea, and false negatives (FN); events that were incorrectly classified as not an apnea. With these counts we calculated apnea detection sensitivity, false positive rate (FPR), and precision, as follows:

$$Sensitivity = \frac{TP}{TP+FN}, \tag{7.11}$$

$$False\ positive\ rate = \frac{FP}{D}, \tag{7.12}$$

$$precision = \frac{TP}{TP+FP}, \tag{7.13}$$

with D defined as the registration duration in hours. Sensitivity for relevant apneas was analyzed separately. We also inspected algorithm output and video registrations at times of false positives to investigate their causes.

Grouped results were obtained by concatenating all events and comparing the predicted labels with the ground truth. Results from all individual cameras estimate the predicted performance when a random camera angle would be chosen. We also compared found performance with the performance using best and worst camera choices, determined post-hoc per subject according to 1) sensitivity and 2) false positive rate. Thus, we estimate performance when we would have known beforehand the best and

TABLE 7.2.
Grouped algorithm test results, using majority vote labels from the trained ensemble models.

Camera / integration method	Sensitivity (%) relevant apneas (N)		Sensitivity (%) all apneas (N)		False positive rate (/h)	Specificity (%) for events
All individual cameras	83.7	(43)	46.7	(92)	2.17	99.1
best choice	100	(14)	63.0	(27)	1.09	99.5
worst choice	80.0	(10)	37.0	(27)	3.28	98.5
late integrated	100	(14)	66.7	(27)	6.57	97.6
early integrated	92.9	(14)	51.9	(27)	1.64	99.3

worst camera positions to capture apneas. This is done for comparison, and does not pose a realistic situation.

We calculated detection latency for relevant apneas, defined as the time between the annotated start of the apnea and the detection. To estimate real-time latency, Gabor filter windows for respiratory frequencies need to be taken into account. These are calculated in the center of windows of 75-745 frames, effectively adding 1.5 to 15 seconds to real-time latency.

If only a single camera is available, it is important to know beforehand which camera placement gives the highest detection sensitivity. To answer this question, we analyzed the influence of camera perspectives with respect to the subject (as shown in Figure 7.4) on apnea detection sensitivity.

RESULTS

Grouped algorithm test results using predicted labels from the ensemble SVMs are shown in Table 7.2. Detailed results per subject and camera can be found in Supplementary Table 7.S1. All relevant apneas from all subjects were detected in the signals from at least one of the used cameras. In three subjects relevant apneas were not detected in all cameras in which they were annotated. The apnea in the video of PAT9 was detected in camera signals without, but also with a second person in view (who was quietly sleeping in the background). With the best camera angles determined post-hoc, 100% sensitivity would have been achieved with a false positive rate of 1.09 per hour. Mean detection latency for detected relevant apneas was 1.7 s (SD: 1.4 s). Mean online latency is estimated to be in the range of 2.7-16 s.

With late camera input integration 100% sensitivity was achieved, with a FPR of 6.6 per hour. Late integration resulted in a higher FPR than found with randomly chosen cameras, and also higher than found with the post-hoc determined worst camera choices. With early camera input integration, 13/14 unique relevant apneas were detected. Closer inspection of the signals at the time of the missed apnea (PAT6) revealed a slight movement at apnea onset, mainly visible in one of the cameras. For this reason the apnea presumably was detected in the single-camera signals (and in the late integrated outputs), but not in the early integrated video signals, which included the movement as a dominant component.

Best single-camera sensitivity was achieved with angle combination frontocaudal-dorsocaudal (100%). Sensitivity was also high for the combination frontal-dorsal (92%). Three/six frontal view detections involved apneas of awake subjects sitting upright. Sensitivity was lower when the subject was registered from a more rostral view; 84% for the combination dorsorostral-frontorostral, and 50% from a rostral viewpoint. No registrations were made from a caudal view.

Inspection of algorithm output at times of false detections showed several causes for false detections. Coincidental detections, in which a very small movement precedes a larger movement that causes the respiratory power to drop, caused 68% of single-camera false detections. Of the 13 coincidental false detections, 3 were also in the early integrated results. When inspecting the respiratory belt signals, three "false" detections turned out to be short apneas (~5 s) and one was caused by a respiratory frequency drop to 0.1 Hz. In one camera registration providing a closeup of the subject's face during sleep, eyelid myoclonias caused two false detections. These false detections were also seen in the results from the early integrated signal.

DISCUSSION

Our results show that it is feasible to detect apneas automatically in video, using our algorithm. All apneas in all subjects were detected in the signals of at least one camera, with acceptable false detection rates. The algorithm can quickly detect cessation of the breathing motion signal, while distinguishing cessations due to apneas from gross body movements. The algorithm functions with color and intensity video and does not need uncovered body regions in view to quantify breathing motions and detect apneas.

When available, integrating multiple camera signals early and late in the algorithm improves detection sensitivity, as respiration movements can be captured from different angles. Late integration, i.e., combining any synchronous detections over cameras, may, however, induce a high false detection rate. An event leading to a false detection in any camera will cause a false detection in the late integrated results. Early integration, i.e., combining video time-frequency signals before event detection and classification, typically results in less false detections. Coincidental false detections are less likely, because small movements that cause the total power signal to fit the apnea model can be cancelled out during signal integration. Detection sensitivity may however be affected by gross body movement at apnea onset in one of the camera signals. If further study verifies that (near) SUDEP-related apneas do not involve subject movements at apnea onset, early integration may be preferred over late integration.

When only a single camera is used, detection sensitivity depends on the ability to capture the respiratory movements, to detect their arrest. Trunk oscillatory movements of breathing are largest in the craniocaudal and dorsoventral directions.[21] Best perspectives have camera axes perpendicular to these movement directions. Cameras should be placed high enough to get a good overview of the subject. Our results indicate that using a frontocaudal-dorsocaudal camera angle combination likely results in the best performance. Dorsal-frontal and dorsorostral-frontorostral viewpoints may also provide good detection sensitivity when the camera is placed high enough. Rostral views provided only mediocre apnea detection performance, and should be avoided. Closeup views of the subject's face and shoulder region may provide a good respiratory signal, but false detections might be caused by small movements.

Apneas are not necessarily missed when there is a second person in view of the camera, provided that the second person is also showing no gross body movements. The second person's breathing movements add to the respiratory range power, but if only one of the persons stops breathing, the power will drop nonetheless. It will probably help if the subject to monitor is closest to the camera, as the subject's respiration will then influence respiratory range power the most. Four of ten subjects were covered by blankets when their apneas occurred, which did not prevent apnea detection. This suggests that apnea detection is not hampered by being covered. The influence of covers on detection should be further investigated, as this is a likely occurrence in practice.

There are no other central apnea detection algorithms which meet all requirements to be applicable for safety monitoring. Whereas a previously reported algorithm cannot distinguish breathing motion breaks due to movement from actual apneas, it comes close to applicability.[16] Our approaches are similar; quantifying the presence of periodicity to detect apneas (periodicity cessation). In a dataset with 17 apneas, the algorithm detected 90% of apneas with a latency of 30 s and a specificity of 78%, which translates to an FPR of 79 /h with their windowing procedure. The lower specificity compared to our algorithm might be explained by the algorithm's inability to take gross motion into account when classifying apneas.

Our algorithm enables fast detection of apneas, with online detection latencies estimated to be between 3 and 16 seconds, while latencies of >20 s are common in other studies.[15,16] The difference in detection delay can be explained by the respiration-free signal window needed by most other algorithms to measure the absence of periodicity and detect an apnea. Our algorithm detects the moment of cessation of breathing motions, and does not need an large signal window free of breathing motions before it detects an apnea.

A possible limitation of our algorithm is that it can only detect apneas at their onset. If something/someone in view of the camera causes a gross motion during apnea onset, the apnea might be missed. Other movement influences might be prevented by confining camera view as much as possible to the subject. Subject segmentation before quantifying the presence of breathing motions may also improve algorithm performance. Segmentation may, however, also make the algorithm more sensitive to occlusions and heavier computationally.

Obstructive apneas occurring during sleep, which may or may not indicate an obstructive sleep apnea (OSA) disorder, may cause algorithm detections. These apnea detections could be considered true positives, and may help classify undiagnosed OSA cases. OSA is a common disorder, however, and people with the disorder can have frequent apneas throughout the night (severe cases have >30 /h),[22] which may result in alarm fatigue. OSA detection with our algorithm was not studied, and should be further investigated.

A possible solution for high false detection rates due to OSA or other causes might be switching on apnea detection only after detection of a convulsive seizure (automatically, e.g. using,[23] or by a carer). As virtually all known cases of SUDEP followed within 11 minutes after a convulsive seizure,[2] apnea monitoring during the half hour following a convulsive seizure might suffice.

Future work should focus on further improvement and validation of the algorithm. A larger dataset with central apneas captured from good camera angles may improve algorithm performance and enable algorithm validation. Using flexible posterior probability outputs (prediction ranging from 0 to 1) instead of label outputs might further optimize sensitivity, but this may also increase the false detection rate. This might be acceptable if alerts can be easily verified by a caregiver who is close by, monitoring the video stream. Validation data should include central apneas occurring after convulsive seizures, to verify that these apneas indeed do not involve gross body movements. Long-term unselected registrations should be used to demonstrate robustness to the variety of (normal) breathing patterns and real-life false detection rates. Other application areas for the algorithm might also be explored, such as remote neonatal monitoring in the intensive care.

CONCLUSIONS

We present a novel automated algorithm to detect central apneas in video, based on the arrest of oscillatory breathing motions. When fully validated, it may be used as an online remote apnea detector, avoiding the use of contact sensors. Further studies are warranted.

REFERENCES

1. Lacuey N, Zonjy B, Hampson JP, et al. The incidence and significance of periictal apnea in epileptic seizures. *Epilepsia*. 2018;59:573–82.
2. Ryvlin P, Nashef L, Lhatoo SD, et al. Incidence and mechanisms of cardiorespiratory arrests in epilepsy monitoring units (MORTEMUS): A retrospective study. *Lancet Neurol*. 2013;12:966–77.
3. Lhatoo SD, Nei M, Raghavan M, et al. Non-seizure SUDEP: Sudden Unexpected Death in Epilepsy without preceding epileptic seizures. *Epilepsia*. 2016;57:1161–8.
4. Surges R, Thijs RD, Tan HL, et al. Sudden unexpected death in epilepsy: risk factors and potential pathomechanisms. *Nat Rev Neurol*. 2009;5:492–504.
5. Rugg-Gunn F, Duncan J, Hjalgrim H, et al. From unwitnessed fatality to witnessed rescue: Nonpharmacologic interventions in sudden unexpected death in epilepsy. *Epilepsia*. 2016;57:26–34.
6. Zaffaroni A, Kent B, O'Hare E, et al. Assessment of sleep-disordered breathing using a non-contact bio-motion sensor. *J Sleep Res*. 2013;22:231–6.
7. Nandakumar R, Gollakota S, Watson N. Contactless Sleep Apnea Detection on Smartphones. *Proc 13th Annu Int Conf Mob Syst Appl Serv - MobiSys '15*. 2015;45–57.
8. Bani Amer MM, Az-Zaqah R, Aldofash AK, et al. Contactless method for detection of infant sleep apnoea. *J Med Eng Technol*. 2010;34:324–8.
9. Hers V, Corbugy D, Joslet I, et al. New concept using Passive Infrared (PIR) technology for a contactless detection of breathing movement: A pilot study involving a cohort of 169 adult patients. *J Clin Monit Comput*. 2013;27:521–9.
10. Al-Naji A, Gibson K, Lee S-H, et al. Real Time Apnoea Monitoring of Children Using the Microsoft Kinect Sensor: A Pilot Study. *Sensors*. 2017;17:286.
11. Garn H, Kohn B, Wiesmeyr C, et al. 3D detection of the central sleep apnoea syndrome. *Curr Dir Biomed Eng*. 2017;3:829–33.
12. Yang C, Cheung G, Stankovic V, et al. Sleep Apnea Detection via Depth Video and Audio Feature Learning. *IEEE Trans Multimed*. 2017;19:822–35.
13. Wang CW, Hunter A, Gravill N, et al. Unconstrained video monitoring of breathing behavior and application to diagnosis of sleep apnea. *IEEE Trans Biomed Eng*. 2014;61:396–404.
14. Sharma S, Bhattacharyya S, Mukherjee J, et al. Automated detection of newborn sleep apnea using video monitoring system. *2015 Eighth Int Conf Adv Pattern Recognit*. 2015;1–6.
15. Jorge J, Villarroel M, Chaichulee S, et al. Non-Contact Monitoring of Respiration in the Neonatal Intensive Care Unit. *Proc - 12th IEEE Int Conf Autom Face Gesture Recognit*. 2017;286–93.
16. Cattani L, Alinovi D, Ferrari G, et al. Monitoring Infants by Automatic Video Processing: A Unified Approach to Motion Analysis. *Comput Biol Med*. 2017;80:158–65.
17. Koolen N, Decroupet O, Dereymaeker A, et al. Automated Respiration Detection from Neonatal Video Data. *Proc 4th Int Conf Pattern Recognit Appl Methods*. 2015;164–9.
18. Li MH, Yadollahi A, Taati B. Noncontact Vision-Based Cardiopulmonary Monitoring in Different Sleeping Positions. *IEEE J Biomed Heal Informatics*. 2017;21:1367–75.
19. Kalitzin S, Petkov G, Velis D, et al. Automatic segmentation of episodes containing epileptic clonic seizures in video sequences. *IEEE Trans Biomed Eng*. 2012;59:3379–85.
20. Horn BKP, Schunck BG. Determining optical flow. *Artif Intell*. 1981;17:185–203.

21. de Groote A, Wantier M, Cheron G, et al. Chest wall motion during tidal breathing. *J Appl Physiol*. 1997;83:1531–7.
22. Senaratna C V., Perret JL, Lodge CJ, et al. Prevalence of obstructive sleep apnea in the general population: A systematic review. *Sleep Med Rev*. 2017;34:70–81.
23. Geertsema EE, Thijs RD, Gutter T, et al. Automated video-based detection of nocturnal convulsive seizures in a residential care setting. *Epilepsia*. 2018;59:53–60.

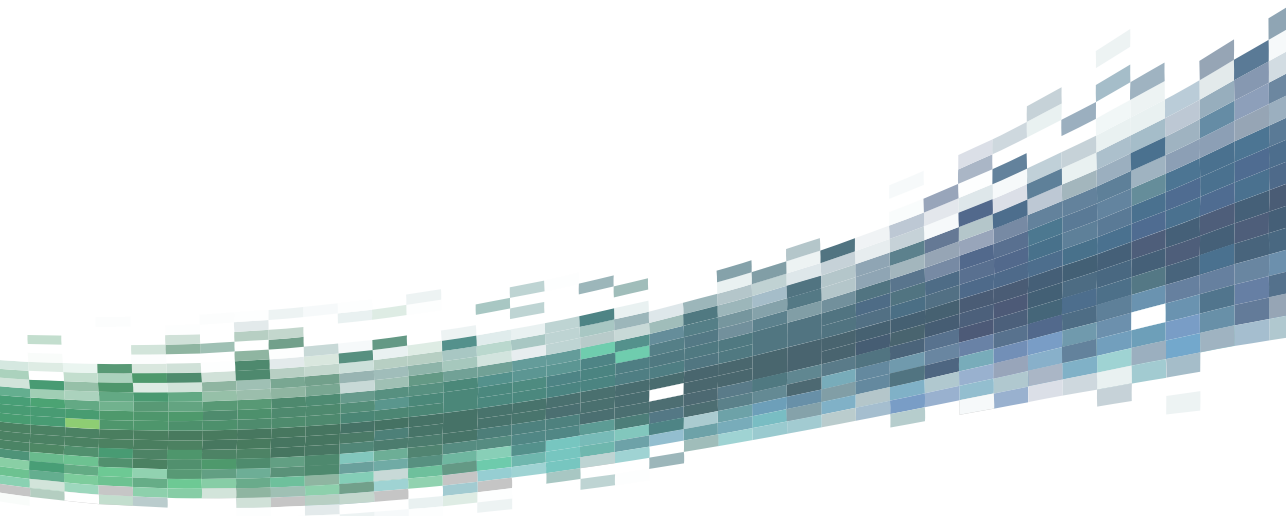
SUPPLEMENTARY MATERIALS

TABLE 7.S1

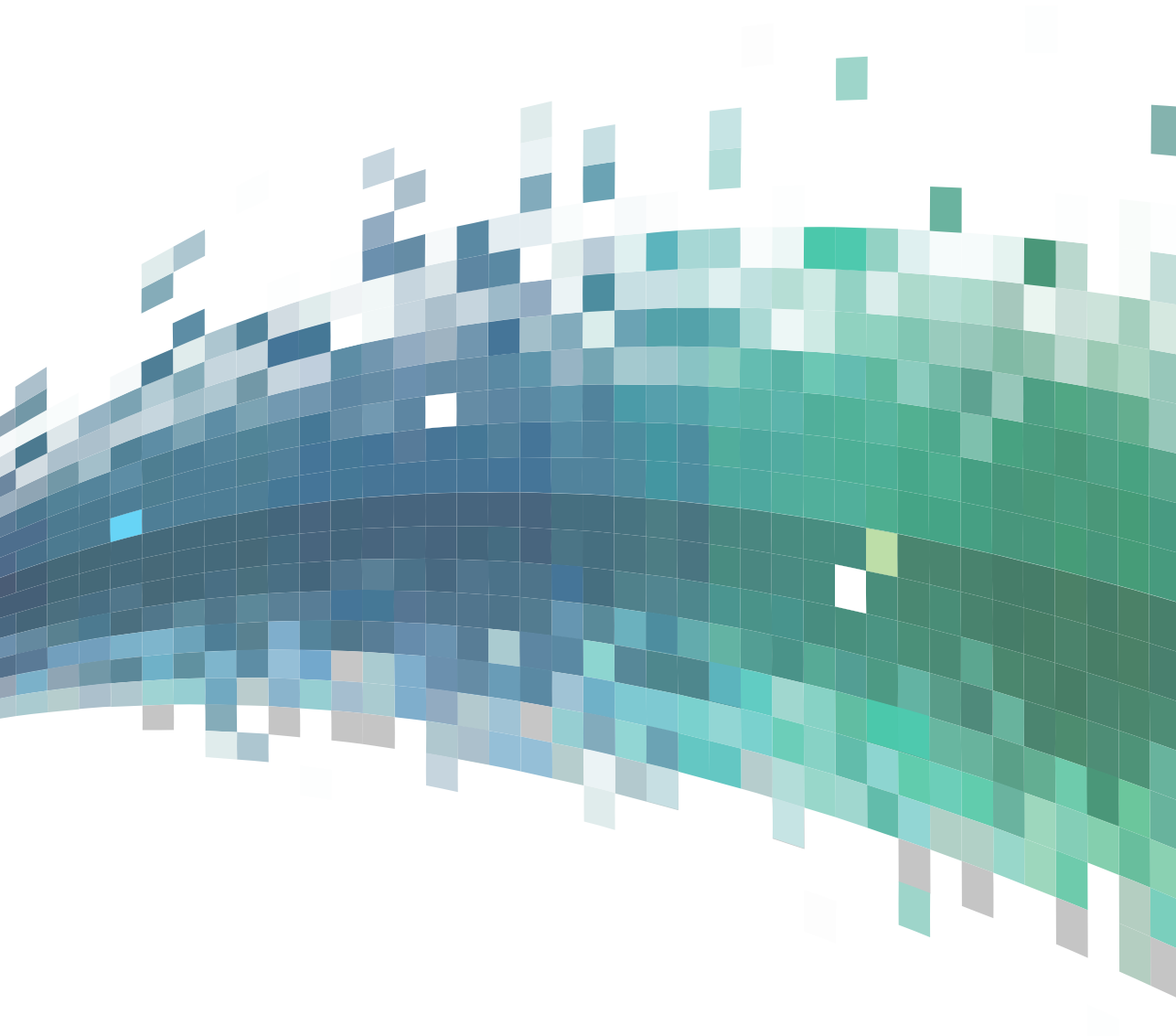
Algorithm test results per camera for each subject. A dash in both sensitivity columns for relevant and all apneas means that the subject was not in view of that camera during the apnea(s). A dash in only the relevant apneas column implies that the apnea(s) registered from that camera angle were non-relevant. Cameras in which performance was best are underlined per subject.

ID	Camera	Sensitivity (%) relevant apneas (N apneas)	Sensitivity (%) all apneas (N apneas)	False positive rate (/h)	Specificity for events (%)
SIM1	1	87.5 (8)	77.8 (9)	0	100
	2	87.5 (8)	77.8 (9)	0	100
	3	100 (8)	88.9 (9)	0	100
	4	33.3 (6)	22.2 (9)	0	100
	late integrated	100 (8)	88.9 (9)	0	100
	early integrated	100 (8)	88.9 (9)	0	100
PAT1	1	- (0)	- (0)	0	100
	2	- (0)	0 (1)	0	100
	3	- (0)	0 (1)	0	100
	late integrated	- (0)	0 (1)	0	100
	early integrated	- (0)	0 (1)	0	100
PAT2	1	- (0)	0 (1)	3.4	98.6
	2	- (0)	0 (1)	3.4	98.7
	3	- (0)	0 (1)	3.4	98.6
	late integrated	- (0)	0 (1)	3.4	98.9
	early integrated	- (0)	0 (1)	3.4	98.7
PAT3	1	- (0)	0 (1)	0	100
	2	- (0)	0 (1)	2.3	99.0
	3	- (0)	0 (1)	0	100
	4	- (0)	0 (1)	2.3	99.0
	late integrated	- (0)	0 (1)	4.5	98.5
	early integrated	- (0)	0 (1)	2.3	99.0
PAT4	1	100 (1)	100 (1)	6.8	97.2
	2	0 (1)	0 (1)	2.3	99.1
	3	- (0)	0 (1)	9.0	96.5
	4	- (0)	- (0)	6.8	97.1
	late integrated	100 (1)	100 (1)	22.6	92.5
	early integrated	100 (1)	100 (1)	0	100

PAT5	1	100	(2)	33.3	(9)	2.3	99.0
	2	100	(2)	33.3	(9)	0	100
	3	-	(0)	0	(9)	2.3	99.0
	late integrated	100	(2)	44.4	(9)	4.5	98.3
	early integrated	100	(2)	33.3	(9)	2.3	98.9
PAT6	1	-	(0)	-	(0)	0	100
	2	100	(1)	100	(1)	0	100
	3	100	(1)	100	(1)	0	100
	4	-	(0)	0	(1)	0	100
	late integrated	100	(1)	100	(1)	0	100
	early integrated	0	(1)	0	(1)	0	100
PAT7	1	-	(0)	0	(2)	0	100
	2	-	(0)	100	(2)	0	100
	3	-	(0)	100	(2)	6.8	97.0
	late integrated	-	(0)	100	(2)	6.8	97.4
	early integrated	-	(0)	0	(2)	0	100
PAT8	1	100	(1)	100	(1)	9.0	96.2
	2	100	(1)	100	(1)	0	100
	3	100	(1)	100	(1)	0	100
	4	-	(0)	0	(1)	0	100
	late integrated	100	(1)	100	(1)	9.0	96.7
	early integrated	100	(1)	100	(1)	9.0	96.2
PAT9	1	100	(1)	100	(1)	0	100
	2	-	(0)	0	(1)	4.5	98.2
	3	-	(0)	100	(1)	4.5	98.1
	4	100	(1)	100	(1)	2.3	99.1
	late integrated	100	(1)	100	(1)	6.8	97.7
	early integrated	100	(1)	100	(1)	0	100
Grouped	all individual cameras	83.7	(43)	46.7	(92)	2.2	99.1
	best choice	100	(14)	63.0	(27)	1.1	99.5
	worst choice	80.0	(10)	37.0	(27)	3.3	98.5
	late integrated	100	(14)	66.7	(27)	6.6	97.6
	early integrated	92.9	(14)	51.9	(27)	1.6	99.3



SUMMARY, DISCUSSION AND APPENDICES



CHAPTER 8

SUMMARY AND DISCUSSION

SUMMARY

In specialized epilepsy care, data streams are often monitored and analyzed real-time by a human observer to detect events. Such events can be ictal (i.e. seizures) or interictal (e.g. epileptic spikes or high-frequency oscillations) transient signals produced by the brain, which usually occur unexpectedly. In the epilepsy monitoring unit (EMU) for example, seizures need to be detected by staff so they can go to the patient to reduce risks arising from seizures (e.g., treat injuries) and to perform testing during the seizure to aid diagnosis. In the operating room, the electrocorticogram (ECoG) is monitored for spikes or high-frequency oscillations when a tailored surgical epilepsy resection is performed. The locations and frequencies of these markers help delineate the area of brain tissue that needs to be resected for the patient to become seizure free. And in homes of people with epilepsy, monitoring for (results of) dangerous seizures can improve safety by indicating whether someone is in need of assistance and at risk of sudden death.

Knowledge about epilepsy-related events can aid diagnosis or direct treatment, and can also indicate the need of immediate assistance. It might be of vital importance that these events are noted directly upon occurrence. The visual observation usually performed to detect these events is time-consuming, subjective, and sensitive to distractions. Consequently, important events could be missed, which might impair safety and the quality of diagnosis and treatment. Automated markers can help detect occurrences and characteristics (e.g. timing and location) of events in the data streams, and in addition might be able to identify data streams (e.g. epochs on specific channels) that are likely to contain events.

The aim of this thesis was to improve situations of real-time data monitoring for event detection in epilepsy, by constructing and validating automated algorithms to detect markers of epilepsy.

PART I: AUTOMATED MARKERS TO ENHANCE DIAGNOSIS

In the epilepsy monitoring unit (EMU), seizures need to be detected by staff so they can go to the patient to reduce risks arising from seizures (e.g. treat injuries) and to perform tests during the seizure to aid diagnosis. Online seizure detection algorithms might help detect seizures that could have otherwise been missed or recognized too late. The added value of seizure

detection algorithms to the detection by already present staff is, however, unclear.

In **Chapter 2** we investigated the added value of applying algorithms for online seizure detection in the electroencephalogram (EEG). We retrospectively analyzed a representative sample of the EEG-video recordings encountered in the EMU (with and without seizures), using two commercially available seizure detection algorithms. We found that EEG seizure detection algorithms may improve the response to seizures on the EMU, since they increase both the total number of seizures detected and the speed of detection.

PART II: AUTOMATED MARKERS TO ENHANCE TREATMENT

Invasive presurgical or intraoperative EEG can be monitored for interictal events to help delineate the brain tissue that needs to be resected, in order to prevent seizures. Surgical removal of cortex showing interictal high-frequency oscillations (HFOs), in particular fast ripples (250-500 Hz), has been associated with post-operative seizure freedom. Visual analysis to find these events is, however, time consuming and, when performed for surgery tailoring, requires the presence of expert reviewers during surgery. An automated algorithm is needed that can independently, reliably, and reproducibly delineate the epileptogenic zone to be resected during surgery.

In **Chapter 3** we presented a novel algorithm - autoregressive model residual variation (ARR) - for this purpose. ARR reflects the amount of non-harmonicity in the signal's high-frequency components. We tested the ability of ARR to approximate the seizure onset zone (SOZ), a surrogate marker for the epileptogenic zone, in interictal intracranial (depth) electroencephalograms that were obtained presurgically. We found that ARR values were higher inside SOZ areas than in channels outside the SOZ. We concluded that ARR may be applied to identify channels in the SOZ automatically in interictal intracranial EEGs, possibly providing a new way to delineate the epileptogenic zone.

In **Chapter 4**, we adjusted the ARR algorithm to reduce the influence of artefacts and tested the potential of the new algorithm to identify epileptogenic tissue during surgery. The autoregressive model residual modulation (ARRm) is less sensitive to common, but visually hard to spot, peri-surgical artefacts. We found that high ARRm values measured in the

post-resection ECoG were associated with poor postsurgical outcome. We concluded that the ARRm algorithm might enable intra-operative delineation of epileptogenic tissue by providing 'on demand' interpretation per electrode about the need to remove underlying tissue to optimize the chance of seizure freedom.

PART III: AUTOMATED MARKERS TO ENHANCE SAFETY MONITORING

In homes of people with epilepsy, monitoring for (results of) dangerous seizures can improve safety by indicating whether someone is in need of assistance and at risk of sudden death. Automated real-time seizure detection systems can help alert caregivers, but popular wearable sensors are not always tolerated. We aimed to design a remote monitoring system with three modules to detect convulsive seizures, falls, and apneas.

In **Chapter 5** we established performance of a non-contact convulsive seizure detection algorithm, by determining a detection threshold and by investigating detection performance as a function of several variables. The algorithm calculates power in the 2-6 Hz range (convulsive seizure spectral footprint) relative to the total spectral power in group velocity signals derived from video sequence optical flow. With the detection threshold determined in a training set, all convulsive seizures were detected in the test set of new subjects (100% sensitivity), with an acceptable false detection rate. This algorithm could improve safety unobtrusively by automated real-time detection of convulsive seizures in video registrations.

In **Chapter 6** we presented an automated algorithm for remote detection of falls, based on a physical model of a fall, aiming at universality and robustness. The algorithm uses vertical velocity and acceleration features from optical flow outputs, corrected for distance from the camera using moving object size estimation. A sound amplitude feature was used to increase detector specificity. Applying the trained algorithm to an acted dataset and real life data with seizure-related falls resulted in high sensitivity and specificity for detection of falls. These results reflect the algorithm's robustness and confirms the feasibility of detecting falls using this algorithm.

In **Chapter 7** we presented a novel automated algorithm for real-time detection of apnea events in video, aiming at fast detection when the subject is immobile. Our algorithm is based on detecting a cessation of the measured oscillatory movements of breathing, in the absence of gross body

motion. We used registrations of simulated apneas and real-life peri-ictal central apneas, each recorded from multiple angles, to develop and test the algorithm. All apnea episodes were detected in the signals of at least one camera. Integrating camera inputs capturing different angles increased detection sensitivity (>90%). Specificity of >99% was achieved with both individual cameras and integrated camera inputs. These results show that it is feasible to detect central apneas automatically in video, using the proposed algorithm.

CONCLUSION

Automated markers can enhance epilepsy diagnosis, treatment, and safety monitoring by providing real-time, objective, and fast identification of epilepsy-specific events, or data streams likely to contain events. The work in this thesis contributed new algorithms to detect tissue to be removed during epilepsy surgery, and to remotely detect falls and apneas using video. This thesis also adds knowledge about the validity and added value of automated seizure detection algorithms in epilepsy monitoring units and at home. With the construction and validation of these algorithms to detect markers of epilepsy, we brought automated markers a step further towards clinical and domestic practice.

DISCUSSION

In this section we will view the work of this thesis in a broader perspective. We will discuss what we have learned and what challenges are yet to be overcome. We will also discuss how future use of the automated markers could look, and what studies are needed to get there.

CHALLENGES OF AUTOMATED MARKER PERFORMANCE TESTING

The performance of automated markers determines their applicability in clinical practice. Obtaining realistic and usable information on the performance can, however, be challenging. The ultimate test, implementing the marker online and comparing its use with the current situation in terms of outcome, is often infeasible; real-time functionality may be lacking and a large amount of time and data is needed to collect enough performance information (e.g., wait for seizures). Therefore, marker performance is usually estimated retrospectively, using offline data. The development

phase of a marker determines the level of evidence needed. Showing cross-validation performance results in the selected, 'clean' training data might suffice for a proof of principle, whereas new, unselected, and continuous registrations are needed to show what real-life performance of a marker would be.[1]

The definition of 'good performance' is determined by the marker's intended purpose and study objectives. In seizure detection, sensitivity is usually favored over specificity, as the damage of a missed seizure is greater than the cost of a false alarm. For this reason we tuned our detection algorithms for convulsive seizures, falls, and apneas to have high sensitivity. In contrast, high specificity is favored when delineating the epileptogenic zone; the surgeon wants to be sure of the non-epileptogenicity of remaining tissue after resection. Also, the performance of the combination user and automated marker may be more relevant than the performance of the automated marker alone. It is improbable, for example, that automated seizure detection algorithms would be applied on an EMU without nurses present to co-monitor the patient's EEG-video stream. Furthermore, detection latency may be decisive when assessing feasibility of online marker usage. Usage of ARRm to delineate the epileptogenic zone during surgery is only feasible when it can provide analysis results within seconds. In addition, early intervention after a convulsive seizure or subsequent postictal apnea might be life-saving in cases of near-SUDEP, thus short detection latencies are required.[2,3] The latency of the entire system encompassing detection and intervention is, however, determined by its weakest link; this may well be caregiver travel time when a person with epilepsy lives alone.

Another challenge when testing automated markers is unavailability or incorrectness of a ground truth needed to define true and false detections. The ground truth is unavailable in studies focused on delineating the epileptogenic zone, i.e. the minimal resection to result in seizure freedom. When a patient became seizure free after surgery, the resected tissue presumably included the epileptogenic zone. The only plausible assumption with this information, however, is that marker detections outside the resected area are false positive detections; the resected tissue need not be entirely epileptogenic. Resections that did not result in seizure freedom provide no reliable information about epileptogenic zone location. In practice, researchers often use approximations of the epileptogenic zone, like the seizure onset zone, or they make comparisons with other markers like spikes and high-frequency oscillations. Similarly, comparison of automated seizure detector output with the current gold standard for seizure detection, synchronized video-EEG registration, is infeasible when collecting real-life

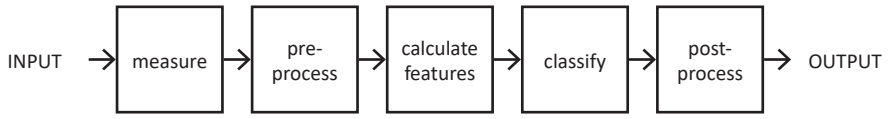


FIGURE 8.1.
Automated marker algorithm design.

data. Researchers are therefore dependent on achievable reference standards like seizure diaries or video recordings.

Flexible handling of data labels using individual case-knowledge enables correction when the ground truth is unavailable or flawed and permits, ultimately, for markers to transcend a flawed reference standard. Alleged false detections might be questioned in some cases: epileptic events are not rarely missed by patients themselves[4] or even during monitoring by trained staff.[5,6] Also, automated markers can reveal subtle events virtually invisible with the naked eye, in video (e.g., subtle seizure onsets[7] or apneas[8]) and electroencephalographic registrations (e.g., interictal non-linear or high-frequency activity[9,10] or seizures[11]). Careful review by experts of recordings at times of marker detections allowed us to improve reference standard quality; some artefacts were relabeled HFO in **Chapter 4**, some alleged convulsive seizures were relabeled as hyperkinetic seizures in **Chapter 5**, and some false detections were in reality short apneas in **Chapter 7**.

IMPROVING AUTOMATED MARKERS

There are many ways to improve an automated marker, and each of its typical components (see Figure 8.1) could provide opportunities. Investigating which components of the algorithm impact the performance parameter to be improved the most, can help find those opportunities.

Increasing the reach of sensor sampling might improve marker sensitivity. Detecting nothing can mean three things; either nothing is happening, the algorithm is not detecting an event (false negative), or there is a spatial or temporal sampling problem (technically also a false negative, but not the fault of the algorithm). When marker sensitivity is low, checking whether the input signal(s) can pick up events to be detected may be worthwhile. Seizures or interictal events to be detected using EEG, for example, may involve brain tissue that is not sampled adequately; e.g., epileptic activity in tissue too far away from the electrodes (inside a sulcus or deeper structures), or the sampling frequency or electrode density is too low to pick up subtle signals. Or perhaps the event you are trying to detect with a video camera is happening outside the field of view. Having a sound hypothesis on where to

place sensors, increasing the spatial reach and having an adequate sampling frequency can potentially help solve these issues. Also, integrating signals from cameras capturing a subtle event from different angles can increase detection sensitivity, as we showed for apneas in **Chapter 7**.

Understanding of the dynamics of events can lead to model-based features, which allow rapid advancement of detection strategies, as no lengthy trial-and-error attempts are required. Also, using model-based features provides insight in algorithm functioning, which facilitates tracking down potential sources of error if necessary. In fact, all features included in the automated markers of this thesis are based on models (or model hypotheses) of the events to detect.

Adding extra input signals or features might improve marker sensitivity, specificity and/or latency, depending on how they are implemented in an algorithm: early integration, by adding new inputs and features to the existing algorithm, or late integration, in which additional classifiers or classifier outputs are combined (e.g. using a voting scheme). New signals or features can be especially interesting when they add information not available before. Complementing seizure detection based on EEG by adding a heart rate feature, for example, might improve sensitivity and latency, by capturing seizure activity not detectable in EEG.[12] In our fall detection algorithm, we improved detector specificity by adding an extra sound level feature, using early integration. Even features that are calculated using the same source signal can add information and improve marker performance. In one example of a seizure detection algorithm, as many as 2974 EEG-features were calculated in time-, frequency-, and time-frequency domains, from which a feature subset was selected as input for the classifier.[13] Features should be carefully selected (or extracted) to provide a suitable feature set for the automated marker's classification task. While adding new signals or features to an automated marker may improve one performance parameter, this improvement may go at the cost of other performance parameters.

Post-processing of classifier output could also improve automated marker performance. In a detection/alarming system, for example, waiting after an alarm before generating new alarms can reduce the false alarm rate when there is some ongoing disturbance.

Personalization of automated markers might improve performance in some cases, and is usually achieved by changes in the classifier of the algorithm. Generic algorithms provide a plug-and-play solution, with algorithm settings based on a typically large and representative group of

subjects. Every person is different however, and when generic algorithms fail, performance improvement by personalization may be attempted. To personalize or tune a detection system, sufficient data (e.g., seizures) and expert annotations are needed. Waiting to collect enough seizures for algorithm tuning might take weeks or months when the subject has a low seizure frequency. A self-learning algorithm could render waiting for events unnecessary, because it can tune the algorithm with each new observation. Subjects may also label each seizure-event after it has (or has not) happened, after which the classifier might be updated.[14] Personalization processes can, however, make the algorithm vulnerable; incorrect labelling and outliers can decrease overall performance. Therefore, personalization should only be attempted when needed, and by professionals in a controlled setting, with the possibility to check and analyze missed events and false detections.

FUTURE PERSPECTIVES

Let us look ahead 10 years. If we assume that epilepsy will not yet have been eliminated by that time, automated markers have the potential to make life of people with epilepsy, their caregivers, and healthcare professionals a bit easier. How would the use of these automated markers look, and what is needed to achieve this objective?

The quality of **diagnostics** and patient **safety** on the EMU will likely be improved by using software that automatically detects epileptic seizures. EMU staff and automated algorithms complement each other; together they detect seizures more often and earlier than either staff or algorithms alone. In this way, staff response to seizures can be improved, which helps ensure patient safety, and improves the quality of the diagnostic process because of timely assessment of patient consciousness and cognition. The current state of art regarding performance allows implementation and online use, which should be the focus of future research. Because seizure detection software needs access to the online, real-time EEG, a software-specific coupling with the EEG registration software needs to be established. It might be interesting to also investigate the added value of different levels of alarming; as algorithms may have various levels of certainty of seizure occurrence. A high level of detection certainty warrants high level (action calling) alarming, but a lower level of detection certainty should perhaps only alert staff to pay close attention to the subject. Multi-level alarming might reduce the chance of alarm fatigue and consequent desensitization or disabling of alarms by staff.[15]

Automated identification of epileptogenic tissue may improve **therapeutic** outcome in certain cases of epilepsy. Epilepsy surgery tailoring, for example, may be guided by automated markers that delineate the epileptogenic zone using ECoG. A device with electrodes might measure, calculate and directly point out tissue areas – under specific electrodes - to be removed, providing also the level of prediction certainty. Resection tailoring with successful markers could increase the chance of seizure freedom after epilepsy surgery, and decrease the chance of negative side effects, like loss of function, thanks to a smaller resection. While a lot of research is being performed on many different electrophysiological markers, there is no consensus on what is the best marker to delineate the epileptogenic zone. Marker validation studies typically show many cases in which the marker was successful, but also exceptions in which the epileptogenic zone is missed or presumably non-epileptogenic areas are detected.[16–21] It is possible that a combination of markers, measuring different signal characteristics, is needed to adequately point out tissue areas to be removed. Future work might therefore focus on a machine learning approach on a big dataset. Different features with shown classification power (such as ARRm, spikes, ripples and fast ripples) and relevant patient and epilepsy data might be integrated for improved classification of tissue epileptogenicity.

When therapeutic approaches have been unable to provide cure, those suffering from therapy-resistant seizures may be **safeguarded** by automated systems. People with epilepsy may be monitored remotely in their home by a system automatically detecting the occurrence of convulsive seizures, falls, and apneas using video and audio, and subsequently alarming a caregiver. Timely detection of these events might decrease risks of sudden death and injury by enabling care to be given quickly. More patients would have access to automated seizure detection, since patients who could not be monitored using currently available bed-attached or worn sensors might be monitored using our remote detection system. Further developments toward this monitoring system should include improving the specificity of the fall detection algorithm. Using a fall-specific sound feature instead of the currently used sound amplitude feature might decrease the number of false positives. Our apnea detection algorithm should be further improved by training it on a large dataset with central apneas captured from good camera angles. Offline validation of the fall and apnea detection algorithms should follow, using real-life uninterrupted data streams. The three validated algorithms may be combined in a smart monitoring system that detects convulsive seizures, falls, and apneas. This system might automatically activate the right detection module when needed, preventing false detections that could occur when all modules run continuously. The apnea detection

module might for example be activated after detecting a convulsive seizure, and the monitoring system might automatically be put in 'sleep mode' when a companion is present in the same room as the subject.

REFERENCES

1. Beniczky S, Ryvlin P. Standards for testing and clinical validation of seizure detection devices. *Epilepsia*. 2018;59:9–13.
2. Ryvlin P, Nashef L, Lhatoo SD, et al. Incidence and mechanisms of cardiorespiratory arrests in epilepsy monitoring units (MORTEMUS): A retrospective study. *Lancet Neurol*. 2013;12:966–77.
3. Rugg-Gunn F, Duncan J, Hjalgrim H, et al. From unwitnessed fatality to witnessed rescue: Nonpharmacologic interventions in sudden unexpected death in epilepsy. *Epilepsia*. 2016;57:26–34.
4. Elger CE, Hoppe C. Diagnostic challenges in epilepsy: seizure under-reporting and seizure detection. *Lancet Neurol*. 2018;17:279–88.
5. Atkinson M, Hari K, Schaefer K, et al. Improving safety outcomes in the epilepsy monitoring unit. *Seizure*. 2012;21:124–7.
6. van der Lende M, Cox FME, Visser GH, et al. Value of video monitoring for nocturnal seizure detection in a residential setting. 2016;1748–53.
7. Geertsema EE, Thijs RD, Gutter T, et al. Automated video-based detection of nocturnal convulsive seizures in a residential care setting. *Epilepsia*. 2018;59:53–60.
8. Sharma S, Bhattacharyya S, Mukherjee J, et al. Automated detection of newborn sleep apnea using video monitoring system. *2015 Eighth Int Conf Adv Pattern Recognit*. 2015;1–6.
9. Andrzejak RG, David O, Gnatkovsky V, et al. Localization of Epileptogenic Zone on Pre-surgical Intracranial EEG Recordings: Toward a Validation of Quantitative Signal Analysis Approaches. *Brain Topogr*. 2015;28:832–7.
10. Bartolomei F, Lagarde S, Wendling F, et al. Defining epileptogenic networks: Contribution of SEEG and signal analysis. *Epilepsia*. 2017;58:1131–47.
11. Acharya UR, Vinitha Sree S, Swapna G, et al. Automated EEG analysis of epilepsy: A review. *Knowledge-Based Syst*. 2013;45:147–65.
12. Leutmezer F, Schernthaner C, Lurger S, et al. Electrocardiographic changes at the onset of epileptic seizures. *Epilepsia*. 2003;44:348–54.
13. Zhang Y, Yang S, Liu Y, et al. Integration of 24 Feature Types to Accurately Detect and Predict Seizures Using Scalp EEG Signals. *Sensors*. 2018;18:1372.
14. De Cooman T, Kjær TW, Van Huffel S, et al. Adaptive heart rate-based epileptic seizure detection using real-time user feedback. *Physiol Meas*. 2018;39:014005.
15. Cvach M. Monitor alarm fatigue: An integrative review. *Biomed Instrum Technol*. 2012;46:268–77.
16. van 't Klooster MA, Van Klink NEC, Leijten FSS, et al. Residual fast ripples in the intraoperative corticogram predict epilepsy surgery outcome. *Neurology*. 2015;85:120–8.
17. Cho JR, Koo DL, Joo EY, et al. Resection of individually identified high-rate high-frequency oscillations region is associated with favorable outcome in neocortical epilepsy. *Epilepsia*. 2014;1–12.
18. Rummel C, Abela E, Andrzejak RG, et al. Resected brain tissue, seizure onset zone and quantitative EEG measures: Towards prediction of post-surgical seizure control. *PLoS One*. 2015;10:e0141023.
19. Fujiwara H, Greiner HM, Lee KH, et al. Resection of ictal high-frequency oscillations leads to favorable surgical outcome in pediatric epilepsy. *Epilepsia*. 2012;53:1607–17.

20. van 't Klooster MA, van Klink NEC, Zweiphenning WJEM, et al. Tailoring epilepsy surgery with fast ripples in the intra-operative electrocorticogram. *Ann Neurol*. 2017;4.
21. Grinenko O, Li J, Mosher JC, et al. A fingerprint of the epileptogenic zone in human epilepsies. *Brain*. 2018;141:117–31.

APPENDIX

NEDERLANDSE SAMENVATTING | DUTCH SUMMARY

PUBLICATION LIST

DANKWOORD | ACKNOWLEDGMENTS

BIOGRAPHY

NEDERLANDSE SAMENVATTING | DUTCH SUMMARY

In de specialistische epilepsiezorg monitoren en analyseren mensen vaak informatiestromen om belangrijke events te vinden. Dit zijn kortdurende gebeurtenissen die we kunnen meten bij mensen met epilepsie. Een event kan bijvoorbeeld een aanval zijn, of een korte signaalpiek gemeten in het hersenfilmpje (EEG) van iemand met epilepsie. Het moment of de locatie van het optreden van events verschaft informatie over de juiste diagnose, behandeling of over de noodzaak van onmiddellijke (medische) hulp als iemand een epileptische aanval heeft.

Om events te detecteren, of op te merken, moeten mensen vaak meerdere signalen of videobeelden van patiënten tegelijk in de gaten houden. Deze observatie is tijdrovend, subjectief en gevoelig voor afleiding. Hierdoor kunnen belangrijke gebeurtenissen gemist worden, wat de veiligheid en de kwaliteit van diagnose en behandeling in gevaar zou kunnen brengen. Automatische markers kunnen hierbij helpen, door de events voor ons te detecteren. Een algoritme, of computerprogramma, geeft automatisch een schatting van de locatie of het moment van optreden van events.

In dit proefschrift proberen we met automatische markers het monitoren van informatiestromen om events te detecteren makkelijker te maken. We maakten zelf algoritmes, en onderzochten hoe goed ze werken. Op deze manier proberen we de diagnose, behandeling en veiligheidsmonitoring bij mensen met epilepsie te verbeteren.

In de epilepsie monitoring unit worden mensen met een vermoedelijke epileptische aandoening langdurig onderzocht met EEG en video. Medewerkers moeten aanvallen tijdig opmerken, zodat ze naar de patiënt kunnen gaan. Medewerkers kunnen dan risico's bij een aanval verkleinen door bijvoorbeeld verwondingen te behandelen en tests tijdens de aanval uitvoeren om een diagnose te ondersteunen. Automatische algoritmes zouden aanvallen kunnen helpen detecteren welke anders gemist of te laat opgemerkt zouden zijn. We onderzochten daarom de toegevoegde waarde van algoritmes die automatisch aanvallen detecteren in het EEG. We vonden dat de algoritmes de reactie op aanvallen kunnen verbeteren, omdat ze het totale aantal gedetecteerde aanvallen en de detectiesnelheid vergrootten.

Als het niet lukt om aanvallen met medicijnen onder controle te krijgen, wordt soms epilepsie chirurgie overwogen. De operatie heeft als doel precies

en alleen dat weefsel te verwijderen waarmee de patiënt aanvalsvrij wordt; we noemen dit weefsel epileptogeen. Om dit stukje weefsel af te grenzen kan het nodig zijn om voorafgaand aan, of tijdens de operatie, een invasief EEG - een meting in of op de hersenen – te maken. Er is een automatisch algoritme nodig om met het invasieve EEG het epileptogene weefsel af te kunnen grenzen. We maakten een algoritme met dit doel. We zagen dat het algoritme hogere waarden mat in weefsel dat waarschijnlijk epileptogeen was. Patiënten waarbij er na de operatie nog steeds zulke hoge waarden gemeten werden, waren bovendien vaak niet aanvalsvrij. We concludeerden dat ons algoritme bruikbaar zou kunnen zijn om epileptogeen weefsel af te grenzen, om mogelijk de kans op aanvalsvrijheid na een operatie te vergoten.

Thuis monitoren op gevaarlijke aanvallen kan de veiligheid vergroten door aan te geven of iemand hulp nodig heeft en risico loopt op plotse dood bij een aanval. Automatische aanvalsdetectie-systemen kunnen helpen zorgverleners te waarschuwen, maar populaire draagbare sensoren (wearables) worden niet altijd verdragen. Ons doel was daarom om een contactloos monitoring systeem te ontwerpen, dat gevaarlijke aanvallen, vallen en ademstops kan detecteren in videoregistraties. Eerst onderzochten we hoe goed een detectie-algoritme convulsieve aanvallen, het meest risicovolle aanvalstype, kan opmerken. Het algoritme detecteerde alle convulsieve aanvallen in een grote test-dataset, met een acceptabel aantal valse alarmen, en zou gebruikt kunnen worden om de veiligheid te vergroten. Daarnaast maakten en testten we een algoritme om vallen te detecteren. We maakten hiervoor gebruik van videoregistraties met geacteerde vallen en echte vallen door een aanval. Het algoritme detecteerde veel van de vallen, met weinig valse alarmen. Deze resultaten tonen de haalbaarheid van het detecteren van vallen met ons algoritme. Verder ontwikkelden en testten we een algoritme dat ademstops detecteert. Hiervoor gebruikten we videoregistraties vanuit verschillende camerahoeken met gesimuleerde en aanvals-gerelateerde ademstops. De haalbaarheid van het detecteren van ademstops met het gemaakte algoritme was zichtbaar in de resultaten: ademstops werden altijd gedetecteerd vanuit minstens één camerahoek, en er waren weinig valse alarmen.

Concluderend draagt het werk in dit proefschrift bij aan de verbetering van de diagnose, behandeling en veiligheidsmonitoring van epilepsie door onderzoek naar automatische markers voor epilepsie. Ons onderzoek verschafte 1) nieuwe algoritmes om te verwijderen epileptogeen weefsel af te grenzen en contactloos vallen en ademstops te detecteren en 2) kennis over de werking en toegevoegde waarde van aanvalsdetectie-algoritmes.

PUBLICATION LIST

THIS THESIS

Geertsema EE, Visser GH, Velis DN, Claus SP, Zijlmans M*, Kalitzin SN*. Automated Seizure Onset Zone Approximation Based on Nonharmonic High-Frequency Oscillations in Human Interictal Intracranial EEGs. International Journal of Neural Systems. 2015;25:1550015.

Geertsema EE*, van 't Klooster MA*, van Klink NEC, Leijten FSS, van Rijen PC, Visser GH, Kalitzin SN, Zijlmans M. Non-harmonicity in high-frequency components of the intra-operative corticogram to delineate epileptogenic tissue during surgery. Clinical Neurophysiology 2017;128:153–64.

Geertsema EE, Thijs RD, Gutter T, Vledder B, Arends JB, Leijten FS, Visser GH, Kalitzin SN. Automated video-based detection of convulsive seizures in a residential care setting. Epilepsia 2018; 59(S1):53-60.

Rommens N, **Geertsema E**, Jansen Holleboom L, Cox F, Visser G. Improved staff response to seizures on the EMU with online EEG seizure detection algorithms. Epilepsy & Behavior 2018; 84: 99-104.

Geertsema EE, Visser GH, Viergever MA, Kalitzin SN. Automated remote fall detection using impact features from video and audio. 2018. [under review]

Geertsema EE, Visser GH, Sander JW, Kalitzin SN. Automated non-contact detection of central apneas using video. 2018. [under review]

* Both authors contributed equally

OTHER

Bennis FC, **Geertsema EE**, Velis DN, Reus EEM, Visser GH. The use of single bipolar scalp derivation for the detection of ictal events during long-term EEG monitoring. Epileptic Disorders 2017;19(3):307-14.

Kalitzin S, **Geertsema E**, Petkov G. Scale-iterative optical flow reconstruction from multi-channel image sequences. In: Petkov N, Strisciuglio N, Travieso-Gonzalez C, eds. Proceedings of the 1st International Conference on Applications of Intelligent Systems, APPIS 2018, Frontiers in Artificial Intelligence and Applications. Amsterdam: IOS press. [accepted]

Kalitzin S, **Geertsema E**, Petkov G. Optical flow group-parameter reconstruction from multi-channel image sequences. In: Petkov N, Strisciuglio N, Travieso-Gonzalez C, eds. Proceedings of the 1st International Conference on Applications of Intelligent Systems, APPIS 2018, Frontiers in Artificial Intelligence and Applications. Amsterdam: IOS press. [accepted]

DANKWOORD | ACKNOWLEDGMENTS

Promotieonderzoek doen is een beetje alsof je voor de eerste keer een berg beklimt. Vanaf een afstand denk je: poeh, dat is ver en hoog! Ik moet nog maar zien hoe ik ooit boven ga komen. Het eerste stuk lijkt nooit af te komen, maar gelukkig word je er daarna steeds een beetje beter in. Een aantal stukken hebben een makkelijk pad, waarbij je alleen maar de bordjes hoeft te volgen. Op andere stukken moet je zelf een goede route kiezen. Heel soms zie je bij het klimmen even niet meer waar je je voeten neer kunt zetten. Maar als je samen met anderen de berg beklimt, is er altijd iemand in de buurt die je een voetsteun kan aanwijzen.

Er zijn een boel fijne mensen die mijn leven als PhD student de afgelopen jaren een stuk makkelijker en leuker hebben gemaakt. Een aantal mensen wil ik hieronder in het bijzonder bedanken.

Prof. Viergever, beste Max. Je durfde het aan: een technisch geneeskundige promovenda, die denkt ook technisch onderzoek te kunnen doen. Dank voor je vertrouwen. In je begeleiding ben je zowel kritisch als flexibel op de juiste momenten, en dat is een prettige mix. Toen je voor het eerst een stuk van me ging lezen, waarschuwde je me nog dat je heel veel commentaar gaf, alsof dat iets naars zou zijn. Ik waardeerde al jouw feedback enorm! Je flexibiliteit, vooral in de afrondingsfase wanneer alles volgens mijn nogal strakke planning moest gebeuren, was erg prettig. Ook buiten de inhoud om was je betrokken. Als er iets voor me geregeld moest worden bij het UMCU, had je het vaak al voor elkaar nog voor ik zelf actie kon ondernemen. Bedankt dat je mijn promotor wilde zijn.

Prof. Sander, dear Ley. You were of course already involved in my work, but I was happy that you agreed to also formally be my second promotor during the last year. I really appreciated your quality-check of all my papers, and your critical feedback. Even if it meant that an in my eyes publication-ready paper needed more work. It was a pleasure to learn from you, about being a researcher, scientific writing, science politics, and epilepsy.

Dr. Visser, beste Gerhard. Er zijn veel mensen belangrijk geweest voor mijn proefschrift, maar in dit geval weet ik zeker: zonder jou was dit boekje er niet geweest. Toen ik in 2012 mijn afstudeerproject bij jou startte, had ik niet gedacht dat ik daarna nog jaren zou mogen blijven als onderzoeker. Dank je wel dat je mij deze kans hebt gegeven! De combinatie van de vrijheid die ik van je kreeg om mijn eigen weg te vinden en je altijd open deur om even

iets te overleggen, vond ik erg prettig. Ook waardeer ik de ruimte die je gaf om me ook op andere vlakken naast het onderzoek te ontwikkelen, enorm.

Dr. Kalitzin, beste Stiliyan. Wat heb ik veel van je geleerd. Onder jouw begeleiding ging ik in een paar jaar van "ik weet eigenlijk niet zo goed wat ik aan het doen ben" naar het ontwerpen en programmeren van mijn eigen algoritmes. Ik waardeer je betrokkenheid bij alle projecten, ook je bezoekjes aan afdeling Research om te vragen "of het al af is". Je was altijd bereid om even mee te denken of om samen wat data te martelen.

Gerhard en Stiliyan, waar harde oppervlakken botsen ontstaan soms vonken. Uit jullie vonken ontstaan gave epilepsie-innovaties en onderzoeken, en bijna alle hoofdstukken van dit boekje zijn daar gevolgen van. Het was me een waar genoegen om met jullie te werken aan zinvolle innovaties in de epilepsie!

Dr. Huiskamp, beste Geertjan. In Utrecht maakte ik kennis met de wereld van de KNF en epilepsie, en mocht ik profiteren van jouw kennis, bereidheid te helpen, en humor. Wat leuk dat jij mijn one-man "aiobegeleidingscommissie" wilde zijn, en we elkaar op deze manier af en toe nog eens tegenkwamen.

Beste leden van de beoordelingscommissie. Ik wil jullie van harte bedanken voor de tijd en moeite die jullie gestoken hebben en nog gaan steken in de beoordeling van mijn proefschrift.

Beste neurologie- en epilepsie-inspiratiebronnen Roland Thijs, Maeike Zijlmans, Frans Leijten, Ben Vledder, Rob Klahn en Thea Gutter. Jullie hebben een groot aandeel gehad in mijn enthousiasme om juist voor deze doelgroep aan de slag te gaan. Roland, ik heb veel geleerd van het samen scoren van aanvalsvideo's bij de "is het een TC, ja of nee" quiz. Hopelijk zullen we met ons gezamenlijke onderzoek in de toekomst SUDEP gevallen voorkomen. We zullen misschien nooit weten of het gelukt is. Maeike, het is inspirerend om bij jou te zien hoeveel je als jonge onderzoeker kunt bereiken. Het was me een genoegen om met jou aan HFO onderzoek te mogen werken! Ik heb het nooit helemaal kunnen loslaten (getuige hoofdstuk 4). Frans, jij introduceerde me in de wereld van epilepsie en epilepsiechirurgie. Het was mooi om wat van "kneepjes van het vak" bij je educatieve presentaties af te kijken. Ben, Rob en Thea, jullie passie voor aanvalsdetectie werkte aanstekelijk! Jullie lieten me zien voor wie we het doen, hoe we vroeger aanvallen detecteerden (mooi speelgoed) en dat het beter kan.

Ik wil iedereen met wie ik samen heb gewerkt aan de onderzoeken in dit proefschrift en daarbuiten, van harte bedanken. Het was prettig werken met de korte lijntjes binnen de epilepsie-wereld, en ik heb veel geleerd van jullie tips, aanvullingen en feedback. Maryse, ik heb genoten van onze samenwerking in onze joint effort om ARRm en HFOs samen te brengen. Ik bewonder je enthousiasme en energieke manier van werken, en heb me vaak afgevraagd hoe je toch al die verschillende registraties en patiënten uit elkaar kon houden. Stiekem staat onze brain-cake ook op mijn lijstje met lifetime achievements. Lisanne en Nicole, ik wilde jullie eerst bij de studenten noemen, maar jullie horen eigenlijk tussen de volwaardig onderzoekers. Wat fijn dat ik een flink stuk onderzoek aan jullie over kon laten, om er later samen ook nog een mooi paper over te schrijven. Zonder jullie was er geen hoofdstuk 2!

Beste EMU verpleging, veel dank voor jullie hulp bij alle metingen. Ik wist wel dat ik jullie eigenlijk niet mocht afleiden, maar het was veel te gezellig bij jullie.

Best KNF laboranten, of het nu om EEG-vragen ging, of het opstarten van metingen, jullie waren altijd bereid te helpen. Ik heb veel bewondering voor jullie kennis en jullie bereidheid die te delen. Fia, heel veel dank voor je hulp bij het scoren van de aanvallen!

Beste Holger en Dennis, dank voor alle technische support en dat jullie af en toe mijn hulplijn wilden zijn.

Lieve Research collega's en oud-collega's. Leuke collega's maken ieder werk beter. Borrels, BBQ's, spelletjesavonden, uitjes, congressen, en een schrijf-retreat, waren samen met jullie een feestje! Bij jullie op kantoor binnenstappen was voor mij als thuiskomen. Ik genoot van de gezellige koffie en lunch momentjes, waar ik het liefst de hele middag bleef hangen. Al mijn verhalen, zowel met als zonder clue, kon ik bij jullie kwijt. Er was altijd wel iemand met een antwoord op welke vraag dan ook, en wanneer nodig werd er acuut tijd gemaakt voor een peptalk of een luisterend oor. Heel veel dank voor de mooie tijd!

Lieve paranimfen Anouk en Robert, ik kan me geen betere mensen wensen om me te helpen bij alles wat nog komen gaat. Robert, jij was onderdeel van mijn SEIN-avontuur vanaf dag 1. Ik mocht alleen bij SEIN afstuderen als jij een korte stage kwam doen, wat je geloof ik een beetje opgedrongen was. Wat blijkbaar goed beviel, want je bent er nog steeds! Je bent een rots in de branding; jij houdt het hoofd altijd koel en hebt een oplossing voor ieder probleem. Anouk, jou ken ik iets minder lang, maar ik kon me al snel geen

SEIN zonder jou meer voorstellen. Als een soort telepathie voelde jij, door de computerschermen heen, altijd feilloos aan wanneer er iets hilarisch was of er even geklaagd moest worden. Wat fijn dat jij mijn bureau-buddy was!

Lieve vrienden, dank voor alle leuke uitjes, etentjes, reisjes en logeerpartijtjes, ik liet me altijd met veel plezier door jullie uit mijn promotie-kluizenaarsschap halen! Ideale broertjes en zusjes Anique, Martijn, Remy en Wieke, al vanaf de intro voor onze studie, en nog steeds, een team. Ik geniet altijd enorm van onze leuke activiteiten, en er zullen er ongetwijfeld nog veel volgen. Lieve Wieke, ik heb zo'n vermoeden dat wij nog veel mooie avonturen tijdens onze toekomstige (snoep)reisjes gaan beleven! Ik zie ons nog wel als oude vrouwtjes samen de wereld over reizen. Lieve Tianne, ik leerde jou goed kennen bij Neverland, waar we als partners in crime samen wel eens even de boel op stelten zouden zetten. Wij zitten voor mijn gevoel op dezelfde golflengte. Which reminds me, ik heb hier nog een surfles met jouw naam erop ;) Lieve Robin, van al deze mensen, afgezien van mijn familie, ken ik jou het langst. Ik heb even gerekend, mijn schatting zou 14 jaar zijn (bijna de helft van mijn leven). Met jou is het altijd onverminderd gezellig!

Lieve Jos en Hanny, Vicky, Silke en Robbie, bedankt voor jullie interesse en de altijd gezellige weekendjes in zonnig Zeeland als welkome intermezzo's.

Lieve Joris, Marianne en Thijs, ik geniet altijd erg van onze avonden gevuld met biertjes, hapjes, spelletjes, en veel hard lachen. Opdat er nog maar veel mogen volgen!

Lieve Pap en Mam, bedankt voor jullie oneindige steun en vertrouwen de afgelopen jaren. Jullie staan altijd voor me klaar, en dat waardeer ik enorm. Ik bof maar met ouders zoals jullie.

Lieve Joren, het was een avontuur de afgelopen jaren, maar gelukkig ging jij met me mee. Thuis komen bij jou en Poes is een van mijn favoriete onderdelen van de dag. Bedankt voor al je vertrouwen, relativeringsvermogen en geduld, vooral de afgelopen maanden. Met jou samen is geen berg te hoog, dus op naar nieuwe avonturen!

BIOGRAPHY

Evelien Geertsema was born on July 4th, 1988 in Hengelo, the Netherlands. Interested in medicine, but also wanting to work with physics and mathematics, she started her bachelor's study in technical medicine at Twente University in 2006. During this study, she was trained to improve healthcare by applying medical technology. After receiving her bachelor degree in 2009, she started her master's with a specialization in medical signal analysis. Next to her studies, Evelien took an interest in teaching, and took several jobs as a teaching assistant. After four clinical internships in different hospitals 2011 and 2012, she started a one-year graduation internship at Stichting Epilepsie Instellingen Nederland (SEIN) in Heemstede. During this internship, she worked on automated algorithms to delineate tissue that should be removed during epilepsy surgery. She received her technical medicine master's degree from Twente University in 2014.

Enthusiastic about epilepsy, research and algorithm design, she started a PhD study at SEIN in 2014. She worked there under the supervision of Prof.dr.ir. M.A. Viergever, Prof.dr. J.W. Sander, Dr. G.H. Visser, and Dr. S.N. Kalitzin. During the following years, she validated an algorithm to detect tissue to be removed during epilepsy surgery, and designed algorithms to remotely detect falls and apneas using video. She also investigated the validity and added value of automated seizure detection algorithms in epilepsy monitoring units and at home. Her work resulted in several publications, and this thesis. She presented her research at several conferences and symposia, and gave a number of lectures and workshops about seizure detection.