EEG in delirium

TIANNE NUMAN

Cover design Kiki Spoelstra
Lay-out Ton Loesink
Printed by Ridderprint BV
ISBN 978-90-393-6828-2

For consistency some terms and graphical presentation of Tables and Figures have been standardized throughout this thesis. Therefore, there may be some differences with the articles that have been published.

Copyright © 2017 T. Numan

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage or retrieval system, without permission in writing from the author. The copyright of articles that have been published or accepted for publication has been transferred to the respective journals.

Financial support for printing this thesis was kindly provided by University Medical Center Utrecht and Brain Centre Rudolf Magnus.

The studies described in this thesis were performed at the Department of Intensive Care Medicine, University Medical Center Utrecht, the Netherlands.

EEG in delirium

EEG bij delirium

(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. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op donderdag 21 september 2017 des middags te 4.15 uur

door

Trijntje Numan

geboren op 27 april 1988 te Sneek

Promotoren: Prof. dr. A.J.C. Slooter
Prof. dr. C.J. Stam

Copromotor: Dr. E. van Dellen

TABLE OF CONTENTS

Chapter 1 General introduction 7

Part I Resting state EEG during hypoactive delirium and light levels of sedation

Chapter 2 Opportunities and methodological challenges in EEG and MEG resting state functional brain network research 19

Chapter 3 Functional connectivity and network analysis during hypoactive delirium and recovery from anesthesia 53

Chapter 4 Resting state electroencephalography characteristics during sedation with midazolam and propofol in older patients 87

Part II Validation of an EEG-based delirium monitor

Chapter 5 Recognition of delirium in postoperative elderly patients: a multicenter study 109

Chapter 6 Delirium detection based on a brief, single-channel quantitative electroencephalography recording in older postoperative patients: a prospective, multicenter validation study 127

Chapter 7 Predicting delirium with a single-channel EEG recording 149

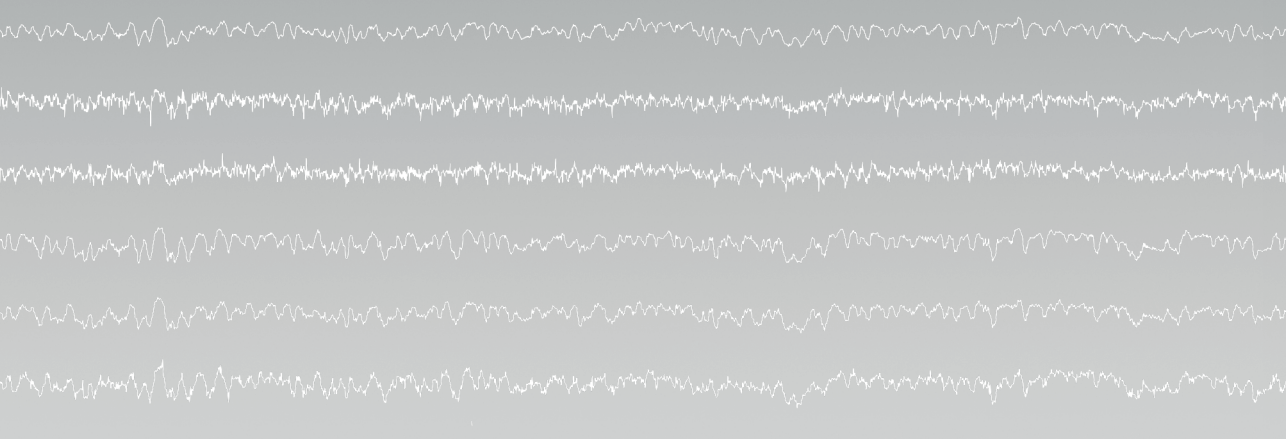
Part III Summary and general discussion

Chapter 8 Summary and general discussion 169

A Nederlandse samenvatting (Dutch summary) 189

B Dankwoord (Acknowledgements) 195

C About the author 201



1

GENERAL INTRODUCTION

BACKGROUND

Case description

Mrs. R, an 83 years old patient, was admitted to the hospital for cardiac surgery. After surgery she was admitted to the ICU according to the protocol and she started to wake up around midnight and could be extubated. The morning after surgery, she was drowsy and did not sleep very well, but she was not in pain. During the morning she slept and she was transmitted to the medium care in the afternoon, still drowsy, but feeling okay. During twilight, the nurse was warned by the alarms of the monitors and found the patient being confused, disorientated, and agitated. Mrs. R. had unplugged herself of the ECG monitor and pulled off her IV line. The physician was consulted and Mrs. R. was diagnosed with delirium.

At the moment of delirium diagnosis, Mrs. R. was confused and did not understand what happened. It was frightening for Mrs. R. not knowing where she was and for what reason. Her family was asked to come over so Mrs. R. had familiar people around her, who could also help her explain what was happening. The family told the nurse that Mrs. R. saw a red truck driving into here room, which of course was not really there. They were explained that such experiences are visual hallucinations, which can be frightening for the patient. The nurse also helped in reorientation, with pictures of her family and a clock in front of here. Moreover, the physician prescribed haloperidol to reduce the symptoms for that moment. Despite the good recovery of her surgery, she stayed longer in the hospital than on average. Within a year after surgery, she suffered from cognitive problems and she was admitted to a nursing home.

Box 1.1

DELIRIUM: DEFINITION AND DIAGNOSIS

Delirium is a neurocognitive disorder with an acute decline in cognition and attention as a consequence of an underlying physical problem[1]. It is a serious health care problem, especially in critically ill and elderly patients[2,3]. Consequences of delirium, besides the suffering of the current symptoms by the patient, are prolonged hospital and intensive care unit (ICU) stay, long-term cognitive problems, institutionalization after the hospital stay, and increased mortality and health care costs[4–9]. Therefore, delirium is a significant problem for the individual patient, as described in the case description in Box 1.1, and for health care in general.

Criteria for delirium as described in the Diagnostic and Statistical Manual fifth edition (DSM-5)[1] include a disturbance of attention and awareness which develops in a short period of hours to a few days, and additional cognitive disturbances. The disturbances of attention, awareness, and cognition cannot be explained by already existing neurocognitive disorders and there is always a physical underlying cause. Three motor subtypes of delirium are defined: hyperactive, hypoactive, and mixed type. Hyperactive delirium is characterized by restlessness, agitation, and hyper vigilance, whereas hypoactive delirium is characterized by sleepiness, inactivity, and slow responses to questions. Patients showing features of both hypoactive and hyperactive delirium have the mixed type. In elderly and critically ill patients, the hypoactive form of delirium occurs most frequently[10,11].

Ideally, delirium is diagnosed by delirium experts: geriatricians, psychiatrists, and neurologists with expertise in geriatric neuropsychiatry. However, approximately 50% of the ICU patients and elderly patients admitted to the hospital suffer from delirium during their ICU stay[2,5]. Thus, the high incidence of delirium and the increasing number of elderly patients admitted to the hospital make it unfeasible for experts to evaluate the presence of delirium in all patients. Moreover, due to the fluctuating course of delirium a once a day visit might miss the delirious episode. Furthermore, it was shown that physicians at the ICU often do not recognize delirium in their patients[12]. Therefore, several tools for clinical nurses and physicians have been developed to improve recognition. A broad range of different assessment tools are described and validated in the literature. At the general wards, the confusion assessment method (CAM) is a frequently used detection tool[13]. It is a questionnaire consisting of four features, which are displayed in Box 1.2. A patient is recognized as having delirium when feature 1 and 2 and either 3 and/or 4 are positive.

Features of the confusion assessment method (CAM)[13]

1. Acute onset and fluctuating course
2. Inattention
3. Disorganized thinking
4. Altered level of consciousness

Box 1.2

For ICU patients, the most frequently used tool is the CAM-ICU, one of the adapted versions of the CAM[14,15]. The CAM-ICU contains the same features as the CAM, but is adjusted to be applicable in patients who are intubated and thus unable to speak. In a research setting, both tools showed high feasibility and reliability. However, in

clinical practice the delirium detection rate was insufficient with a sensitivity of 47% for the CAM-ICU[16] and 25% for the CAM[17]. The lowest detection rates were found in the hypoactive delirium cases. In the clinical setting, interpretation of subjective features can be difficult. For example, the first feature of the CAM(-ICU) may be difficult to interpret if one has only the patient record as reference and one has not seen the patient before. Therefore, there is a clear need for an objective and reliable tool for delirium screening by clinical nurses.

THE BRAIN AS A COMPLEX NETWORK

The brain is a complex organ with billions of neurons[18]. Processing and transferring information by neurons occurs by electrical signals. This electrical activity can be recorded as field potentials using electroencephalography (EEG). These EEG signals are recorded using electrodes, which are placed on different locations on the skull, and can therefore measure brain activity from different areas of the brain. For each signal recorded at a site of the brain, it is possible to calculate which frequencies this signal contains. Moreover, combining all these signals, it is possible to assess the global function of the brain, or the function of the brain as a complex network. Evaluation of the degree of interaction or synchronization between all EEG signals, gives a surrogate measure of information exchange between these EEG signals (i.e., brain areas). Since EEG electrodes record field potentials at the scalp and not individual action potentials of neurons, it is not possible to record the process of individual neurons using EEG. However, EEG based network analysis provides insights in pathology and improved diagnosis of several conditions, such as Alzheimer's disease[19] and epilepsy[20]. Moreover, disturbances of consciousness can be studied using functional brain network analysis.

Consciousness is a broadly studied process of human function, but still not fully understood. Previous studies proposed that consciousness is the result of information processing and transferring over a highly integrated brain network[21]. Sedation, a broadly applied model to study reduced levels of consciousness, was suggested to disrupt these highly integrated brain networks, described as cognitive unbinding[22,23]. Reduced functional connectivity (i.e., reduced synchronization) during reduced levels of consciousness due to sedatives was found, indicating a less- or less efficient-transferring of information between brain areas[21,24]. Moreover, the connectivity directed from frontal to parietal regions of the brain was found to be disrupted during loss of consciousness[24,25]. It was suggested that these forward and backward connections are related to higher-order processing[21], and thus for being consciousness[25].

Delirium is a diffuse cerebral dysfunction and therefore EEG can be used to assess the disturbance of the electrical activity during delirium. Since decades, it is known that the EEG recording of delirious patients shows slowing of the background pattern compared to non-delirious subjects[26]. Specifically, increased delta (i.e., frequencies lower than 4 Hz) and theta (i.e., frequencies between 4 and 8 Hz) band activity were found to be related to delirium[27]. Moreover, network analyses were performed in delirious patients compared to non-delirious patients using EEG and functional magnetic resonance imaging (fMRI)[28,29]. EEG based analysis showed reduced functional connectivity, disruption of the directed connectivity, and disturbance of the functional brain network[28]. Although not formally tested, these results seemed similar to the effects of sedatives[24], which may not be surprising since the disturbance of attention is one of the main features of delirium[1]. The results of the fMRI study showed a negative correlation of the fMRI time series between parietal (i.e., posterior cingulate cortex) and frontal (i.e., dorsolateral prefrontal cortex) brain regions in the control subject[29]. This negative correlation can be explained by the alternating activity of these two brain areas. The dorsolateral prefrontal cortex is part of the executive network, which are brain regions that are activated during a task. The posterior cingulate cortex is part of the default mode network, in which brain regions which are activated at rest. During the delirium episode, the correlation between these brain areas was positive instead of negative. The authors suggested that this may contribute to the disturbance of awareness during delirium[29].

Even though the results of the EEG and fMRI functional network studies are not directly comparable, due to the differences between the used modalities, they both suggested that delirium showed disturbed functional brain network, specifically the disturbance in the parietal to frontal connectivity.

THE CLINICAL SETTING

In the ICU setting, several patients show reduced levels of consciousness, due to either hypoactive delirium or the effects of sedatives. As symptoms are similar[30], it may be difficult to distinguish both conditions. However, it is clinically important to know which patients suffer from delirium and may need an intervention. The neurophysiology of these two states of reduced consciousness may be different. This can be evaluated using power spectral analysis and functional network analysis based on EEG recordings. It is unclear if the functional brain network changes differently during reduced levels of consciousness due to sedatives and reduced levels of consciousness due to hypoactive delirium.

For delirium, it was shown that relative delta power (i.e., frequencies lower than 4 Hz) based on a bipolar EEG recording can distinguish between delirious and definite non-delirium patients after cardiac surgery[31]. Based on this study, an EEG-based delirium monitor was developed, having only four electrodes making it easy to apply and feasible for clinical practice. This EEG-based delirium monitor needs to be validated in an unselected patient population, including the entire spectrum from definite no delirium towards definite delirium. The goal is to find a simple tool which can objectively detect delirium and monitor the course of delirium during hospital stay. To be able to implement such a tool in clinical practice, other effects influencing the EEG have to be explored. One important factor in this context of the ICU setting is the use of sedatives. Sedatives are frequently prescribed in ICU patients for comfort[32]. Combining the results of the validation study and the effects of sedatives on the EEG may help in recognizing delirium based on the bipolar EEG-based delirium monitor, specifically for ICU patients. Moreover, this EEG-based method to recognize delirium may be able to detect delirium in an earlier stage than the current diagnosis. The patient example, Mrs. R. described in Box 1.1, already showed some symptoms of delirium in the morning after surgery (i.e., being drowsy), however she was not diagnosed as delirious until the end of the day, when she fulfilled all DSM-5 criteria. This may indicate that the delirious episode has already started in the morning, which might be an opportunity to detect the ongoing process of delirium using EEG in earlier stage compared to the expert diagnosis based on symptoms.

OBJECTIVE OF THIS THESIS

This thesis has two main objectives. The first objective is to improve our knowledge of reduced levels of consciousness during hypoactive delirium, the most common delirium subtype. For this purpose, the functional brain network during hypoactive delirium was assessed using EEG recordings. In addition, the functional brain network during similar states of reduced levels of consciousness, induced by sedatives, was studied to assess differences compared to hypoactive delirium.

The second objective of this thesis is to improve delirium detection for routine clinical practice using EEG. First, we studied the current quality of delirium diagnosis in clinical practice. Thereafter, the EEG-based delirium monitor was validated in postoperative elderly patients, assessing the possibility to detect delirium in clinical practice based on

a simple EEG recording. Finally, improved delirium detection could also be achieved by early detection or prediction of delirium prior to the clinical diagnosis. Therefore, the possibility to predict delirium based on EEG was evaluated.

OUTLINE OF THIS THESIS

Part I: Resting state EEG during hypoactive delirium and light levels of sedation.
In the first part of this thesis, the changes in functional network during reduced levels of consciousness due to either hypoactive delirium or the resulting effects of sedation were evaluated. **Chapter 2** is a review describing the analysis pipeline of functional connectivity and network analysis, including the choices and challenges for each step to consider. With the increase of interest in functional brain networks and the available functional network measures, comparison between studies becomes more challenging. This review describes the advantages and limitations of the currently used measures for functional connectivity. In **chapter 3** the results of the functional connectivity and network analysis performed in patients with hypoactive delirium, patients recovering from anesthesia, and non-delirious patients are presented. In **chapter 4** the effects of two commonly used sedatives, midazolam and propofol, are assessed using EEG-based functional network analysis in patients undergoing light sedation.

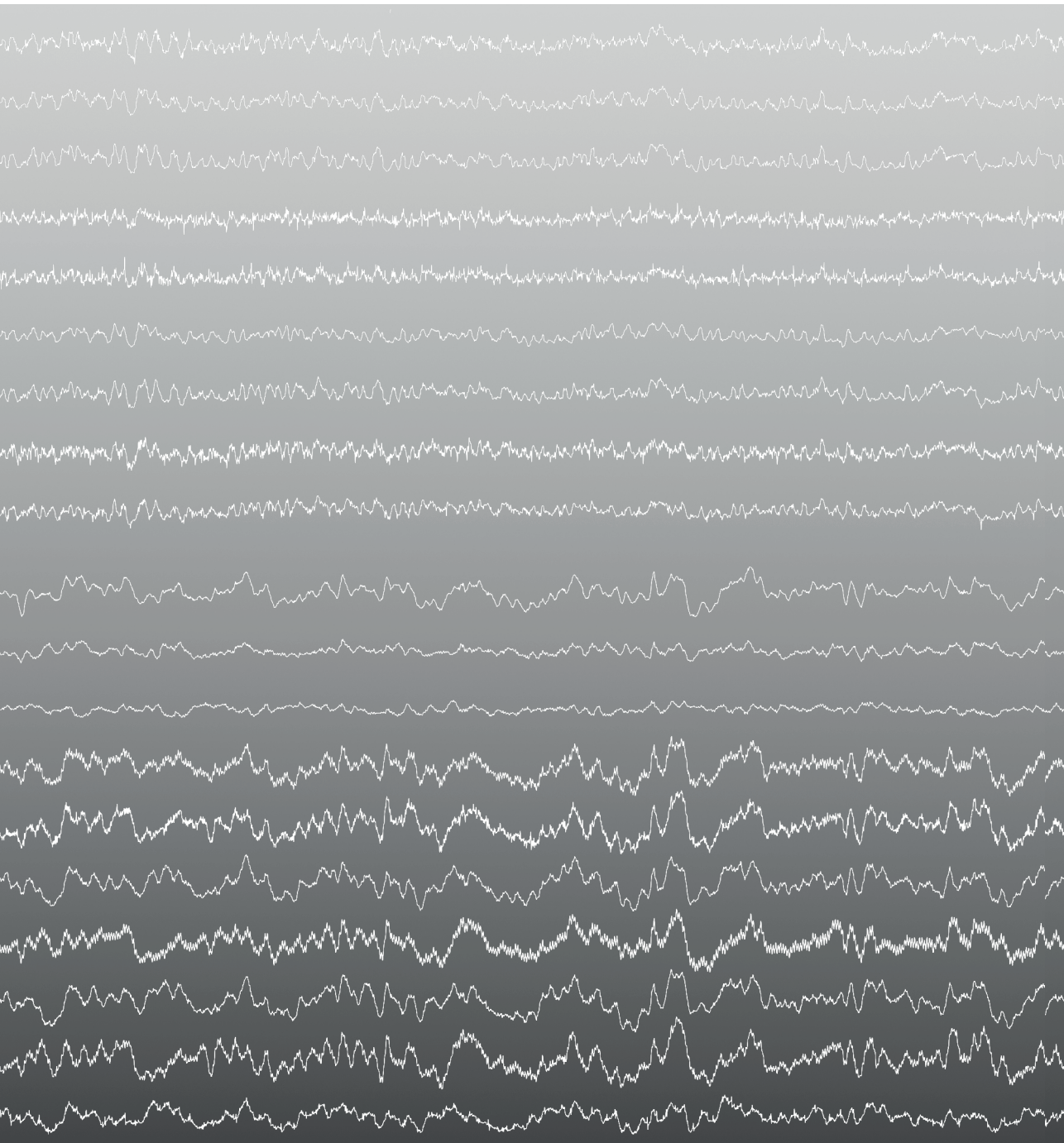
Part II: Validation of an EEG-based delirium monitor.
In the second part of this thesis, the results of the validation of the EEG-based delirium monitor are described. In **chapter 5** the currently used methods of diagnosing delirium are evaluated, in both delirium experts and clinical nurses using delirium screening tools. **Chapter 6** describes the results of the validation study of an EEG-based delirium monitor in postoperative older patients. In **chapter 7** the possibility of using this EEG-based delirium monitor to predict delirium prior to clinical diagnosis is evaluated.

Part III: Summary and general discussion.
Finally, the main results are discussed and a summary of this thesis is given. The results are discussed in relation to the two main objectives, and how these results can help to improve delirium detection. Suggestions for future perspective of delirium detection in clinical practice and research setting are considered.

REFERENCES

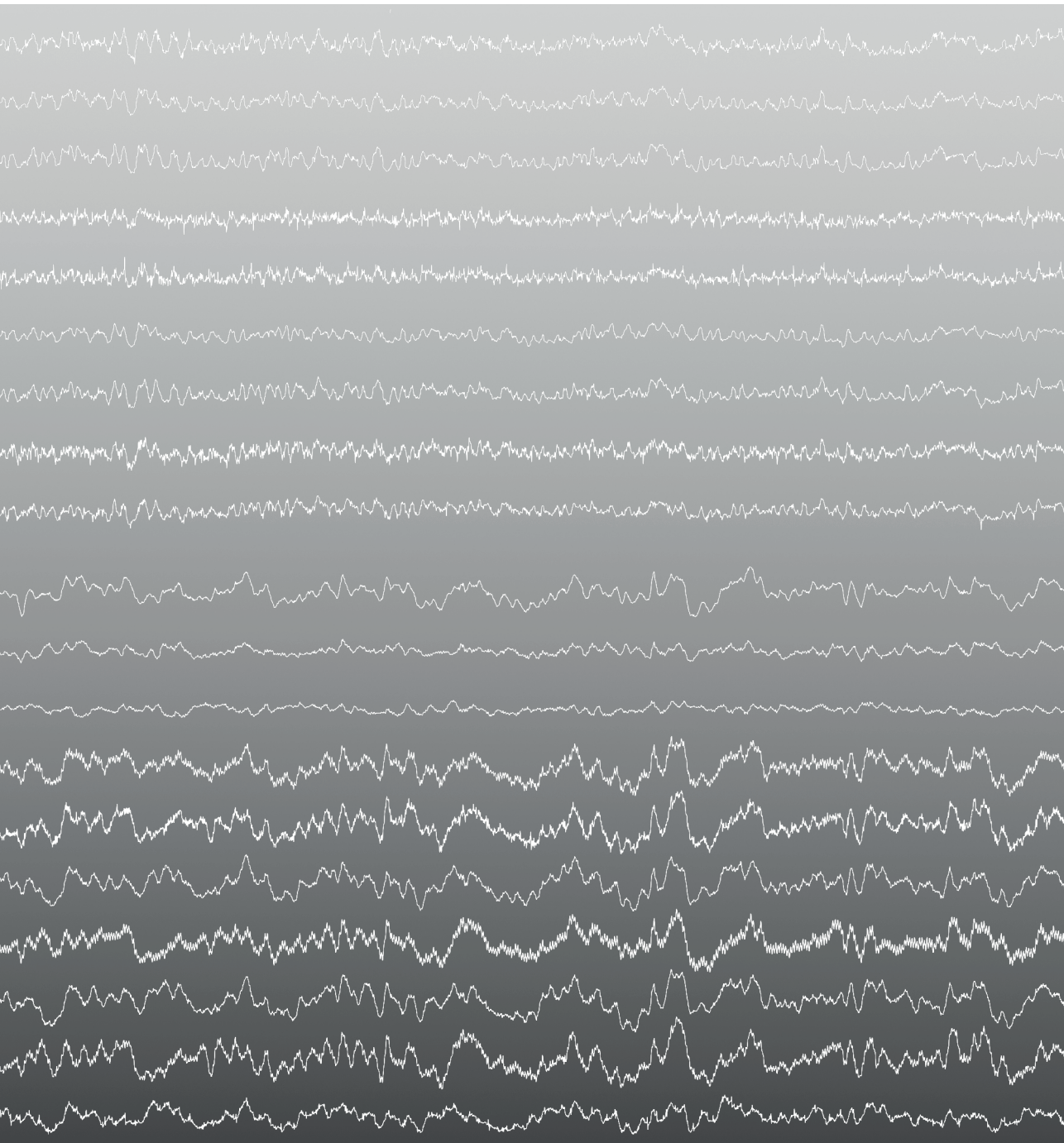
1 American Psychiatric Association. Diagnostics and statistical manual of mental disorders. 5th edition. (2013).
2 Inouye, S. K., Westendorp, R. G. & Saczynski, J. S. Delirium in elderly people. Lancet 383, 911–922 (2014).
3 Wolters, A. E., van Dijk, D., Pasma, W., et al. Long-term outcome of delirium during intensive care unit stay in survivors of critical illness: a prospective cohort study. Crit. Care 18, R125 (2014).
4 Gleason, L. J., Schmitt, E. M., Kosar, C. M., et al. Effect of delirium and other major complications on outcomes after elective surgery in older adults. JAMA Surg. 150, 1134-1140 (2015).
5 Salluh, J. I. F., Wang, H., Schneider, E. B., et al. Outcome of delirium in critically ill patients: systematic review and meta-analysis. BMJ 350: h2538 (2015).
6 Klein Klouwenberg, P. M. C., Zaal, I. J., Spitoni, C., et al. The attributable mortality of delirium in critically ill patients: prospective cohort study. BMJ 349: g6652, 1–10 (2014).
7 Pandharipande, P. P., Girard, T. D., Jackson, J. C., et al. Long-term cognitive impairment after critical illness. N. Engl. J. Med. 369, 1306–1316 (2013).
8 Khan, B. A., Zawahiri, M., Campbell, N. L., et al. Delirium in hospitalized patients: Implications of current evidence on clinical practice and future avenues for research-A systematic evidence review. J. Hosp. Med. 7, 580–589 (2012).
9 Leslie, D. L., Marcantonio, E. R., Zhang, Y., et al. One-year health care costs associated with delirium in the elderly population. Arch. Intern. Med. 168, 27–32 (2008).
10 Peterson, J. F., Pun, B. T., Dittus, R. S., et al. Delirium and its motoric subtypes: a study of 614 critically ill patients. J. Am. Geriatr. Soc. 54, 479–484 (2006).
11 Fong, T. G., Tulebaev, S. R. & Inouye, S. K. Delirium in elderly adults: diagnosis, prevention and treatment. Nat. Rev. Neurol. 5, 210–220 (2009).
12 Eijk, M. M. J. van, Marum, R. J. van, Klijn, I. A. M., et al. Comparison of delirium assessment tools in a mixed intensive care unit. Crit. Care Med. 37, 1881–1885 (2009).
13 Nouye, S. K., Van Dyck, C. H., Alessi, C. A., et al. Clarifying confusion: the confusion assessment method: a new method for detection of delirium. Ann. Intern. Med. 113, 941–948 (1990).
14 Ely, E. W., Margolin, R., Francis, J., et al. Evaluation of delirium in critically ill patients: validation of the confusion assessment method for the intensive care unit (CAM-ICU). Crit. Care Med. 29, 1370–1379 (2001).
15 Neufeld, K. J., Nelliot, A., Inouye, S. K., et al. Delirium diagnosis methodology used in research: a survey-based Study. Am. J. Geriatr. Psychiatry 22, 1513–1521 (2014).
16 Eijk, M. M. van, Boogaard, M. van den, Marum, R. J. van, et al. Routine use of the confusion assessment method for the intensive care unit: a multicenter study. Am. J. Respir. Crit. Care Med. 184, 340–344 (2011).
17 Rice, K. L., Bennett, M., Gomez, M., et al. Nurses' recognition of delirium in the hospitalized older adult. Clin. Nurse Spec. 25, 299–311 (2011).
18 Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10, 186–198 (2009).
19 Tijms, B. M., Wink, A. M., de Haan, W., et al. Alzheimer's disease: connecting findings from graph theoretical studies of brain networks. Neurobiol. Aging 34, 2023-2036 (2013).

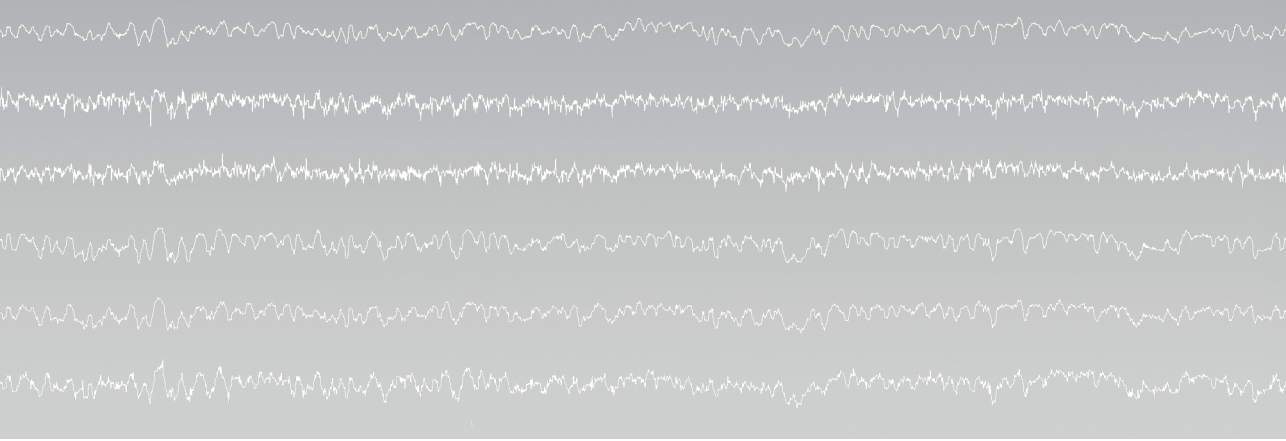
20 Van Diessen, E., Otte, W. M., Braun, K. P. J., et al. Improved diagnosis in children with partial epilepsy using a multivariable prediction model based on EEG network characteristics. PLoS One 8, e59764 (2013).
21 Hudetz, A. G. General anesthesia and human brain connectivity. Brain Connect. 2, 291–302 (2012).
22 Mashour, G. A. Cognitive unbinding in sleep and anesthesia. Science 310, 1768b–1769b (2005).
23 Mashour, G. A. Cognitive unbinding: a neuroscientific paradigm of general anesthesia and related states of unconsciousness. Neurosci. Biobehav. Rev. 37, 2751–2759 (2013).
24 Lee, H., Mashour, G. A., Noh, G.-J., et al. Reconfiguration of network hub structure after propofol-induced unconsciousness. Anesthesiology 119, 1347–1359 (2013).
25 Boly, M., Moran, R., Murphy, M., et al. Connectivity changes underlying spectral EEG changes during propofol-induced loss of consciousness. J. Neurosci. 32, 7082–7090 (2012).
26 Engel, G. L. & Romano, J. Delirium, a syndrome of cerebral insufficiency. J. Chronic Dis. 9, 260–277 (1959).
27 Van der Kooi, A. W., Leijten, F. S. S., van der Wekken, R. J., et al. What are the opportunities for EEG-based monitoring of delirium in the ICU? J. Neuropsychiatry Clin. Neurosci. 24, 472–477 (2012).
28 Van Dellen, E., van der Kooi, A. W., Numan, T., et al. Decreased functional connectivity and disturbed directionality of information flow in the electroencephalography of intensive care unit patients with delirium after cardiac surgery. Anesthesiology 121, 328–335 (2014).
29 Choi, S.-H., Lee, H., Chung, T.-S., et al. Neural network functional connectivity during and after an episode of delirium. Am. J. Psychiatry 169, 498–507 (2012).
30 Zaal, I. J. & Slooter, A. J. C. Light levels of sedation and DSM-5 criteria for delirium. Intensive Care Med. 40, 300–301 (2014).
31 Van der Kooi, A. W., Zaal, I. J., Klijn, F. A., et al. Delirium detection using EEG. Chest 147, 94–101 (2015).
32 Barr, J., Fraser, G. L., Puntillo, K., et al. Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit. Crit. Care Med. 41, 263–306 (2013).

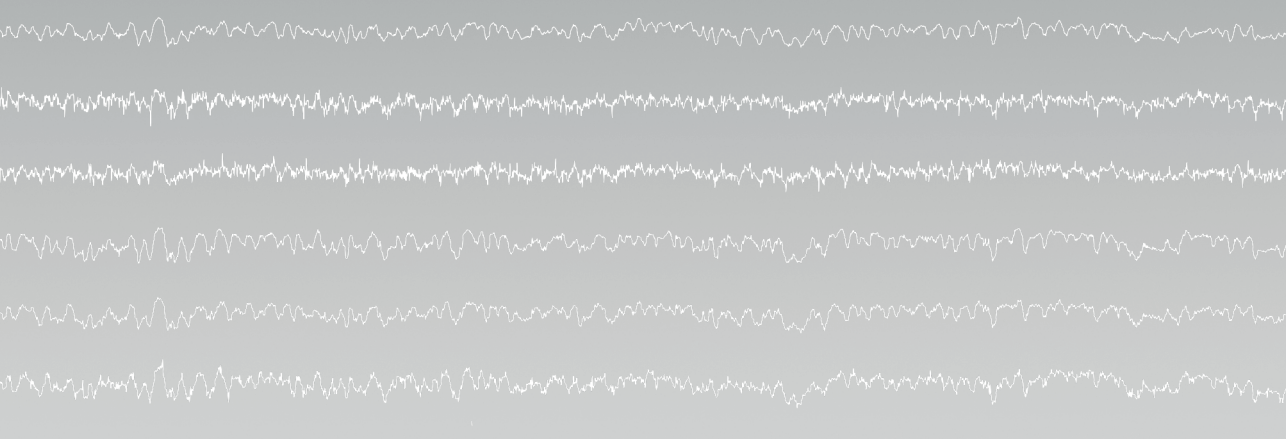


PART 1

RESTING STATE EEG DURING HYPOACTIVE DELIRIUM AND LIGHT LEVELS OF SEDATION







2

OPPORTUNITIES AND METHODOLOGICAL CHALLENGES IN EEG AND MEG RESTING STATE FUNCTIONAL BRAIN NETWORK RESEARCH

E. van Diessen*
T. Numan*
E. van Dellen
A.W. van der Kooi
M. Boersma
D. Hofman
R. van Lutterveld
B.W. van Dijk
E.C.W. van Straaten
A. Hillebrand
C.J. Stam

** Authors contributed equally*

Clinical Neurophysiology 2015; 126(8): 1468-1481

ABSTRACT

Electroencephalogram (EEG) and magnetoencephalogram (MEG) recordings during resting state are increasingly used to study functional connectivity and network topology. Moreover, the number of different analysis approaches is expanding along with the rising interest in this research area. The comparison between studies can therefore be challenging and discussion is needed to underscore methodological opportunities and pitfalls in functional connectivity and network studies. In this overview we discuss methodological considerations throughout the analysis pipeline of recording and analyzing resting state EEG and MEG data, with a focus on functional connectivity and network analysis. We summarize current common practices with their advantages and disadvantages; provide practical tips, and suggestions for future research. Finally, we discuss how methodological choices in resting state research can affect the construction of functional networks. When taking advantage of current best practices and avoid the most obvious pitfalls, functional connectivity and network studies can be improved and enable a more accurate interpretation and comparison between studies.

INTRODUCTION AND RATIONALE

In recent years, there has been a growing interest in characterizing the functional network of the brain 'at rest'. This so-called 'resting state' paradigm is believed to reflect intrinsic activity of the brain, which may reveal valuable information on how different brain areas communicate[1–3]. It has linked spontaneous – task independent – fluctuations in neural activity to diseases, cognitive decline, and disturbances in consciousness[4–7].

This interest in the 'resting state' is associated with several breakthroughs in functional magnetic resonance imaging (fMRI) research[8]. The claim, however, that valuable information on communication between brain areas can be inferred from intrinsic activity – obtained with neurophysiological techniques – is much older (for a comprehensive overview see[9,10]). The high spatial resolution might be a favorable feature of fMRI; still this technique only provides an indirect measurement of brain activity and has a limited temporal resolution. Information processing in the brain, however, acts on multiple time-scales, depending on the specific cognitive or behavioral function[11]. A considerable part of the information processed in the brain at rest is encoded on time scales from milliseconds to seconds[12], a time scale that better suits techniques such as electroencephalography (EEG) and magnetoencephalography (MEG).

In the last decade, EEG and MEG connectivity and functional brain network studies have gained considerable interest resulting in a yearly growing number of published studies on this subject (Figure 2.1). These studies have provided valuable information on the deviant organization in the diseased brain, such as in Alzheimer's disease[13,14], epilepsy[15,16], schizophrenia[17,18], multiple sclerosis[19,20], Parkinson's disease[21], as well as in

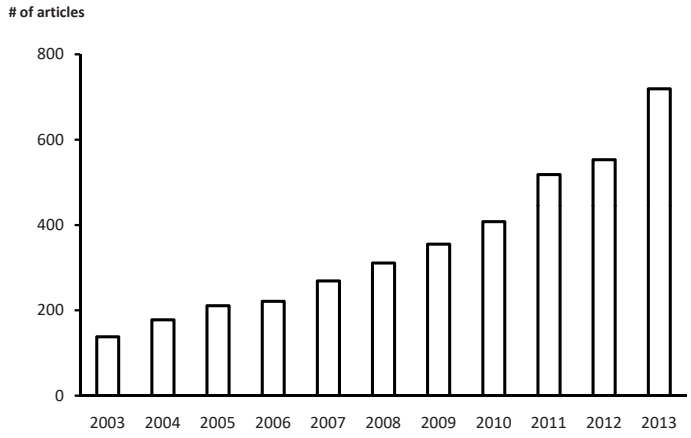


Figure 2.1. Number of articles per year from Pubmed search using keywords "(EEG OR MEG) AND (connectivity OR brain networks OR functional networks OR graph theory OR network analysis)" in the period 2003-2013.

the healthy brain on topics as aging[22,23], gender differences[24] and a healthy lifestyle[25]. Furthermore, connectivity and functional brain network studies can be used in the clinical setting. For example, in epilepsy it has been shown to prompt early diagnosis[26] and to improve accuracy of epilepsy surgery by removing aberrant network nodes[27]. In Alzheimer's disease, EEG connectivity studies were used to monitor the success of novel interventions[28]. Similarly, progression of cognitive deficits in Parkinson's disease was correlated with functional brain network changes[29]. Together these examples clearly underline the importance and additional value of connectivity and brain network analyses in EEG and MEG research.

When performing these analyses, one makes several assumptions and choices that may influence the eventual results. Moreover, the literature on functional connectivity and functional network studies is rapidly evolving, with an increasing number of analysis methods becoming available. Discussion is needed to obtain uniformity and comparability between different studies[30,31]. The present paper therefore aims to highlight challenges, problems, and opportunities that are encountered when performing this type of research. As there are only few methodological studies that address these issues systematically, our review can be seen as a reflection of the current state of the field. We provide an overview of the methodological issues that should be considered when performing functional connectivity and network studies with EEG or MEG, and highlight the advantages and disadvantages of different approaches. Although we specifically focus on resting state EEG and MEG studies, most of the information provided is also applicable to task-related studies and other imaging techniques such as fMRI.

We start with subject-related methodological issues that are of interest when conducting a resting state study. What is resting state, and how does the heterogeneous experience of subjects affect resting state studies? Furthermore, we explain how the state of vigilance and an eyes open versus eyes closed condition might influence the resting state recording. We then summarize analysis-related methodological choices that can strongly influence the eventual results of a functional connectivity or network study, namely: choice of EEG reference, source space analysis, artifact handling and filtering, epoch selection, choice of frequency bands, and test-retest reliability. For each methodological choice we introduce current common practice, explain why these choices are important for the eventual data analysis and summarize the advantages and disadvantages of each

choice. Furthermore, we discuss issues that are still unresolved or subject of debate and give some general recommendations. Subsequently, a concise overview of currently used functional connectivity measures is provided and how these connectivity measures relate to various methodological choices and assumptions. Finally, we discuss current challenges in functional network analysis based on EEG and MEG recordings.

SUBJECT-RELATED METHODOLOGICAL ISSUES

WHAT IS 'RESTING STATE' AND HOW DOES IT AFFECT THE RECORDING?

Resting state is the state in which a subject is awake and not performing an explicit mental or physical task. Traditionally, the 'resting state' condition was commonly used in EEG research – besides event-related potential studies – to study patterns of brain activity, whereas fMRI research was mainly focused on alterations in activity during task performance. Early EEG studies, including the first EEG recordings performed by Berger[32], already provided evidence for patterns of brain activity when subjects were not performing a task[33]. Ironically, it was not until Biswal and colleagues revealed a distinct fMRI pattern of interacting brain regions when not performing a task that the resting state condition became a research paradigm for the study of interconnectivity of brain regions[34]. Since then, many studies have identified sets of brain regions that share a common activation pattern during the resting state[1,35] including the 'default mode network'[36,37] and other so-called resting state networks[38]. These resting state networks have been replicated and validated both in neuroimaging and neurophysiological studies[39–41], suggesting that resting state patterns of connectivity are the result of robust and specific intrinsic neural activity.

In contrast to task related activity, it is difficult to control the behavior of subjects during a resting state experiment. Differences in the specific instructions for the resting state period may at least partly influence the activity of the default mode network[42]. Commonly, subjects are asked to close their eyes and not to fall asleep. As a result, thoughts are drifting and thereby creating stimulus independent thoughts[43,44].

A recent study investigated what kind of thoughts patients experienced during a resting state recording and found several phenotypes of resting state cognition[45]. Even when

similar instructions were given to subjects, the subjective experience during a resting state recording varied greatly, thereby potentially confounding the results. In addition, the cognitive state before the recording could also influence resting state dynamics[46]. These findings underline the heterogeneous experience of subjects during the resting state condition. Instructing a resting state condition might thus not be as straightforward as it seems. Controlling spontaneous thoughts is difficult, however, it might not be necessary as the experienced random episodic spontaneous thoughts seems to activate similar brain regions as resting state networks[43]. Furthermore, consistent activation patterns of resting state networks in healthy controls among studies suggest only a limited influence[35].
To reduce the externally induced heterogeneity of the resting state condition, we recommend the use of a priori defined instructions throughout a study, which should be reported in the method section of the eventual manuscript. Furthermore, subjects should have a similar pre-experimental procedure, to avoid introducing variance in cognitive state. Therefore, researchers should consider the order of neuropsychological testing and recording of resting state. We suggest to first record a resting state condition after which an experimental session can be performed to minimize disturbances of the resting state condition. Finally, longer registration increases the stability of resting state networks, as with time it becomes more likely that the complete repertoire of resting state networks has been activated[47], but this also increases the risk of drowsiness in subjects when no precautions are taken. The exact relation between duration of recording and stability in EEG and MEG resting state functional connectivity remains complex and is discussed further in section 'Epoch selection'.

STATE OF VIGILANCE

During the day, the brain is constantly shifting between different levels of activation, also called ‘states of vigilance’. In EEG and MEG research, we distinguish several vigilance states. These different states can be identified by visual inspection or spectral analysis of the EEG[48,49]. The three most studied vigilance states are wake, sleep, and sleep deprived. The state ‘drowsiness’ is often avoided in resting state research, because alertness or wakefulness is reduced during a drowsiness state and may vary and influence measurements greatly. However, recognizing drowsiness in resting state recordings can be difficult and requires a systematic approach[50,51]. Each state of vigilance has specific characteristics that contribute to differences in spectral power[52,53] and functional connectivity[54,55]. Several factors have been identified that influence state of vigilance: circadian rhythm[54], task performance before the recording[56], including neuropsychological testing, medication use, or even caffeine intake[57–59]. Body posture

and recording environment may also affect vigilance and, consequently, functional connectivity measures. For example, drowsiness is more likely to occur in a dimly lit and sound attenuated room with the subject in supine position, compared to a noisy environment with the subject sitting in daylight. Also, the amplitude of the EEG recording changes as a result of different body postures due to shifts in cerebrospinal fluid layer thickness[60]. The effect size of these possible confounders in resting state studies is unknown. It is therefore recommended to eliminate or record and correct for these possible confounders as much as possible.

EYES OPEN VERSUS EYES CLOSED

Whether a recording is performed with eyes open or eyes closed influences the resting state condition greatly. Evidence from fMRI, MEG and EEG studies has revealed differences between eyes open and eyes closed conditions for functional connectivity measures and functional networks[61–64]. Irrespective of the condition, eye movements affect neurophysiological recordings, particularly the frontal channels[65,66] and are thus a potential confounder in connectivity analyses. Both eyes open and eyes closed are associated with specific eye movements. For example, eye blinks are more prevalent during eyes open condition, whereas rolling of the eyes might particularly influence the eyes closed condition. Rolling of the eyes during eyes closed condition is often due to drowsiness, which is normally not a state of vigilance where resting state studies are interested in and, as a result, discarded from further analysis. Furthermore, the eyes closed condition is more stable over sessions when quantifying EEG parameters than the eyes open condition[67] and standardization of the procedure is relatively straightforward even in subjects who are difficult to instruct such as children, or patients with behavioral or cognitive problems. Together with the robust topographic effect posteriorly in the alpha frequency band when eyes are closed, thereby giving good guidance for selecting resting state epochs, we advocate to use the eyes closed condition during resting state recordings. Methods to automatically remove eye movements from the EEG and MEG recordings are discussed in section 'Artifact handling and filtering'.

ANALYSIS-RELATED METHODOLOGICAL CHOICES

CHOICE OF REFERENCE

In contrast to MEG, the electric potentials measured by EEG electrodes are defined with respect to a reference. Besides bipolar recordings, in which EEG activity is defined by

the electric potential difference between two electrodes, EEG recordings often use a single common reference such as auricular, mastoid or central electrode as reference. These conventional reference montages are confounded by brain activity that will eventually affect further analysis. As a result, recordings are often re-referenced offline to compute reference montages that are electrophysiologically more silent[68–71]. The common average reference has previously been suggested as a practical compromise to reduce the confounding effect of brain activity that is picked up by the reference[69]. The advantage of the common average reference of approximating a zero sum reference is, however, increasingly lost in low-density EEG recordings[72,73]. To this end, several new other methods have been proposed. These include the infinity reference, which tries to estimate a time-varying constant that is removed from the recorded data[74], and the surface Laplacian, which represents a truly reference-free transformation[75,76]. Similar to the conventional and average common reference, the infinity reference is reversible to the original reference scheme, whereas the surface Laplacian involves the estimation of radial current flow at the scalp, which cannot be undone[76]. Both methods, in particular surface Laplacian, have been used empirically in various basic and applied contexts[73]. Recently, a statistically robust method has been proposed to adequately mitigate the influence of neural activity in the common average reference[77]. More investigations are needed to further explore the performance of this theoretically appealing method.

A related question is to what extent reference choice will affect the computation of connectivity measures. For example, Qin and colleagues demonstrated that infinity reference has a superior performance compared to other reference montages when estimating functional connectivity by means of coherence[78]. Correlation measures such as coherence, however, are increasingly abandoned in connectivity studies, as they fail to include information on the intrinsic nonlinearity of brain activity and it is currently unclear whether the superior effect of the infinity reference is maintained when using nonlinear connectivity measures. We await future studies that critically evaluate possible biases due to reference as was illustrated for effective connectivity measures[79].

In the light of the ongoing discussion on reference choice[71,80] and its effect on connectivity measures[78,79] we encourage researchers to explore the effects of different types of references when computing connectivity measures.

SIGNAL VERSUS SOURCE SPACE

Many resting state EEG and MEG studies use the activity at the electrode level to infer how brain regions are (functionally) interconnected. These analyses are performed in so-called 'signal space' as neural activity is directly inferred from signals measured at the EEG electrode or MEG sensor. When performing connectivity analysis in signal space, several factors should be considered. Firstly, multiple electrodes pick up activity from a single source due to the nature of the signal, also called 'field spread'[81]. A second problem is related to volume conduction: the 'blurring' effect due to the electrical conduction properties of the human head[82]. Together, these factors can result in an erroneous estimate of the actual connectivity between brain areas. To obtain more reliable information on the communication between brain areas, studies project the activity measured at the electrode or sensor (signal space) back to the underlying sources, the so-called 'source space'.

The mapping from signal space to source space is known as the inverse problem[52]. Unfortunately, no unique solution exists to this problem[83], unless constraints and assumptions are made. These assumptions concern, for example, the number of possible sources or the non-linearity between sources[84,85]. Furthermore, analyzing neurophysiological signals in source space does not completely overcome the problems of field spread and volume conduction, and it has therefore been suggested to combine source space analysis with a robust connectivity measure[86]. These robust measures include, for example, imaginary coherence, phase slope index or phase lag index (section 'Connectivity measures'). Secondly, the mixture of signals arising from spatially separated sources at a single electrode also hampers the interpretation of connectivity estimate in signal space[73,87]. Source space analysis could be helpful in demixing signals[85]. Various approaches for source space analysis have been offered[88], including low resolution brain electromagnetic tomography (LORETA) in EEG research[89] and the beamforming approach for EEG and MEG recordings[84,90]. In general, the accuracy of source localization increases with the number of recording sites[91] if the signal quality remains constant, although a noiseless recording condition may also allow source localization with a standard 10-20 system[92]. A high channel density recording might even become a disadvantage for source reconstruction approaches that rely on an accurate description of the lead fields. Lead field is defined as the electrode or sensor signal that is produced by a source of unitary strength. Incorrect source and head models lead to deviations from the 'true' lead fields, and subsequently to source reconstruction errors[93,94]. These errors in the lead fields are

more discernible when using good quality data. Counterintuitively, recordings with higher signal-to-noise ratio and higher density may therefore degrade source reconstructions in the presence of these modeling errors. Finally, increased channel density can result in bridging, the incidental spread of electrolyte gel between adjacent electrodes, thereby negatively influencing the recording and subsequent inverse solutions and/or connectivity estimates[95].

ARTIFACT HANDLING AND FILTERING

To minimize the influence of artifacts on the results, visual inspection and automatic detection of artifacts are often used to remove artifacts or to select artifact-free data segments. Many attempts have been made to reject or mitigate eye movement artifacts, to reduce interobserver variability, and to improve efficiency in visible inspection of the data[96,97]. Blind source separation, such as independent component analysis, is increasingly used for the detection and removal of ocular artifacts[98,99], although it is unknown to what extent these artifact reduction methods influence functional connectivity and network metrics. Furthermore, a number of techniques is available for automatic removal of muscle artifacts; however, none of them guarantees muscle artifact free data[100]. Improvement of software-based muscle artifact recognition is therefore needed[101,102]. Currently, most studies use visual inspection to eliminate epochs with myogenic or eye-movement artifacts, although the precise procedure is often not described. Some studies mention the removal of frontopolar and auricular channels to reduce the influence of artifacts[103,104]. Although these artifacts will also affect other channels due to volume conduction, this influence is reduced when excluding frontopolar and auricular channels prior to offline re-referencing. It is, however, preferable to select artifact free epochs. In addition, visual recognition of any EEG artifact depends on the chosen reference montage. This means that a possible artifact on the reference channel could influence all other recording sites and could lead to the undesirable rejection of an epoch. Although this problem is relatively easy resolved by replacing (e.g., interpolating) the original reference, it underlines the importance that visual inspection should be done by a well-trained researcher.

In EEG and MEG recordings analogue filtering is needed to prevent from aliasing and eliminate the direct current (DC) component. Besides the analogue filtering, software programs have digital filtering options that can be useful to improve inspection and selection of epochs. It is important to realize, however, that this digital filtering may affect amplitude and phase of EEG and MEG recordings. It is essential to know how the

data is exported from the recording device and whether a digital filter is contaminating the signal. To avoid any possible influences of the digital filters on the recorded signal, it is therefore important to consider elimination of software based filtering or to use zero-phase filtering, for example by a forward and reversed filtering approach.

EPOCH SELECTION

Standardization and interobserver variability in epoch selection

Selection of epochs based on visual inspection is a subjective approach, which may result in inter-observer variability. The effect of epoch selection on functional connectivity has never been investigated systematically. A few studies, however, assessed the stability of their outcome measures by repeating the analysis with a different number and selection of epochs, showing minimal changes between conditions[29,104,105]. This may indicate a modest subjective influence on epoch selection when sufficient epochs are selected. Still, automated analysis could be helpful for researchers and clinicians in evaluating and improving epoch selection in EEG and MEG recordings[106]. The current complexity and limited transparency of automated detection systems demotivate researchers to use it on a larger scale[107] although recent advances are promising[108]. Also, it is recommended to define the selection criteria prior to the epoch selection. Selection by two or more experienced researchers or clinicians could improve the reliability of the epoch selection. For example, when a researcher selects epochs according to predefined criteria, a second researcher can be asked to independently evaluate the selected epochs. When both researchers agree on the quality, the epoch can be included. Epochs without consensus can be replaced by new epochs. Another option is the use of an automatic software-based rejection procedure or a random selection among all selected artifact-free epochs[108]. Beside the elimination of artifacts in the included data, it is important to define the time segments in the recording from which epochs are selected. As the variance of vigilance increases with the length of the EEG recording[109], we recommend selecting the first artifact-free epochs of sufficient quality after the start of the resting state recording. In this way, selection bias of EEG epochs will be minimalized.

Number and length of epochs and sample frequency

Different epoch lengths and number of epochs are currently used in resting state functional connectivity studies, ranging from one second[110,111] to a few minutes[112] or even a day[54] for epoch length, and from one epoch[113] to over 100 epochs containing the entire EEG recording[110]. Previous studies have investigated epoch length in relation

to connectivity stability[111,114,115] and showed that the length of epochs to obtain connectivity measures is highly dependent on the type of connectivity measure. Recent studies have used a more pragmatic approach and extended the original analysis to investigate connectivity stability for the included subjects, for example, by determining the minimum number of epochs that are needed using a leave-one-out analysis[105]. Similarly, varying the epoch length revealed that a longer epoch length does not automatically imply a more stable connectivity value. For phase synchronization measures, longer epochs could result in lower connectivity values based on the asymmetrical distribution of the phase difference[104]. From this perspective, we recommend to use epochs of an identical length within a study, as epoch length can influence the connectivity measure[104], and to choose an epoch length in accordance with the connectivity measure of choice. Furthermore, we recommend computing connectivity values per epoch and consequently an average value (over epochs) per subject to increase the stability of connectivity values. In some studies, it might be particularly interesting to investigate temporal dynamics of functional connectivity and networks. The use of sequential short epochs might be useful to study these dynamical properties[116].

An often neglected issue when choosing epoch length and number is the sample frequency. Increasing sample frequency will result in a higher temporal resolution and, consequently, a higher number of samples in one epoch. It is unclear how sample frequency influences connectivity measures exactly, although it is reasonable to assume that a lower sample frequency leads to a reduced sensitivity and a larger variability over epochs for detection of coupling between signals. Sample frequencies commonly used in functional connectivity EEG and MEG studies range between 250 and 512 Hz. Analysis of connectivity measures in higher frequency bands requires higher sample frequencies to fulfill the Nyquist-Shannon sampling theorem[117]. Otherwise, the spectral resolution is only dependent on the length of the epochs. We recommend comparing epochs with a similar sample rate even if this would mean down sampling for some epochs.

CHOICE OF FREQUENCY BANDS

Factor analysis revealed that the classification of EEG recordings into distinct frequency bands, namely delta (0.5 – 4 Hz), theta (4 – 8 Hz), alpha (8 – 13 Hz), beta (13 – 30 Hz), and gamma (30 – 90 Hz), is statistically sound[118]. Furthermore, each frequency band is associated with distinct cognitive functions[119.] Nevertheless, considerable disagreement exists whether traditional frequency bands are indeed fixed entities. For example,

spectral band limits seems to depend on factors such as age[24,120]. In addition, it has been argued that the lower (8 – 10 Hz) and higher (10 – 13 Hz) alpha frequency band are involved in different cognitive processes[121]. The reliability of the gamma frequency band is under debate as it remains questionable whether gamma oscillations can be reliably inferred from surface recordings[11]. Although various research fields have consistently identified gamma activity being related to tasks (e.g., visual attention, perception and memory[11]) or diseases (e.g., epilepsy, autism, ADHD and schizophrenia[122]), studies have suggested that EEG oscillations >20 Hz from surface recordings reflect myogenic artifact[102,123]. These artifacts can, however, be removed using surface transformations of the EEG[124]. Besides myogenic artifacts, saccadic spike potentials are also known to affect the gamma frequency band[125,126]. Finally, various diseases of the brain have been associated with abnormal slowing in specific frequency bands and thereby acting as disease specific modifiers, for review see[127]. These disease specific spectral changes might become a problem when evaluating connectivity measures in these patients, particularly when connectivity measures are strongly dependent on amplitude. Connectivity measures based on phase differences are therefore recommended, as they are not or less sensitive to amplitude differences. In addition, we suggest performing a power spectral analysis along with connectivity and functional network analyses to disentangle disease specific spectral changes from disease related connectivity and network changes in the frequency bands[26,103]. To overcome previously described frequency band-related problems, some studies have used a broadband approach. This could be an option when exploring an undirected hypothesis and to avoid inflated type I error rates due to multiple comparisons across frequency bands, but it will probably fail to disentangle simultaneous (opposing) changes in different frequency bands.

TEST RE-TEST RELIABILITY

To evaluate changes in resting state EEG recordings over time or the effects of an intervention, it is important to understand the effects of multiple recordings over time on the outcome measures. Previous literature showed a moderate to high intra-individual correlation in EEG spectral analysis, with correlation coefficients ranging between 0.4 and 0.9[128–132]. These correlations do not seem to depend on the time interval between the recordings[29,133]. When the interval between two tests are separated by a few months or years, one should also be aware of the effects of aging, especially in pediatric studies[22,24].

CONNECTIVITY MEASURES

To investigate functional interactions between brain regions, EEG and MEG studies have used different connectivity measures, for an overview see[134–136]. The quantification of interacting brain regions can be subdivided into functional and effective connectivity measures[137,138]. Connectivity measures are based on statistical interdependencies between signals[139]. The extent to which brain regions are connected is defined by the strength or consistency of this statistical interdependency[140], also called synchronization. In dynamical systems the term synchronization generally refers to (phase) coupling of two or more harmonic oscillators[141], but it is currently used in a more liberal way in brain connectivity analyses[142]. Thus, a stronger synchronization, often reflected by a high coupling or consistency of oscillating systems, leads to a stronger connection. In contrast to functional connectivity measures, which only give information about the temporal correlation, effective connectivity measures also provide information on causal and therefore directed information flow between brain regions[138]. Increasing evidence suggests that information processing in the brain follows a complex, directional pattern between brain regions[143]. Besides effective connectivity, directed functional connectivity measures can reveal disturbances in the normal directionality of information flow in the diseased brain, for example, in dementia[79], epilepsy[144], or when consciousness is disturbed[145]. These directed functional measures do not provide information about the causal relation between brain areas.

Deciding on the most appropriate connectivity measure can be arduous, as several issues should be considered. This includes the consideration of linear or nonlinear relations, analysis in time or frequency domain, using an amplitude or phase-based measure, obtaining directed or undirected information, and whether to include indirect relations or not (i.e., multivariate or bivariate). Linear correlations to investigate connectivity of the brain have been used for several decades and are relatively straightforward in terms of computation and interpretation, for a review see[146]. Since these linear methods are not able to take into account the intrinsic nonlinearity of neuronal activity, various nonlinear connectivity measures have been introduced[134,147–149], including several phase-based connectivity measures, such as the phase locking value[150] and phase lag index[151]. Besides considering nonlinearity, novel multivariate connectivity measures are aiming to differentiate between direct and indirect interrelations[138]. Whereas bivariate measures disregard the influence of other signals when computing the interaction between two signals of interest, multivariate measures try to disentangle this information in a

Table 2.1. Overview of different connectivity measures used in EEG and MEG research, with their advantages and disadvantages.

Connectivity measure	Property measured	Advantage(s)	Disadvantage(s)	Key reference(s)
Correlation	The linear relation between the amplitude of two signals in the time domain	Commonly used and straightforward method	Nonlinearity not considered Not possible to make a distinction between direct and indirect relations Sensitive to volume conduction	162
Coherence	The linear relation between the amplitude of two signals in the frequency domain	Commonly used and straightforward method	Nonlinearity not considered Not possible to make a distinction between direct and indirect relations Sensitive to volume conduction	163
Granger causality	The future of signal X can be predicted more precisely when the past of signal Y is included and vice versa	Estimates causal interaction, and therefore directionality is assessed Well established and widely used in many fields of research	Nonlinearity not considered Methodological choices (e.g., choice of reference in EEG), as well as other confounders (e.g., volume conduction) could interfere with the actual causality	164-166
Directed coherence	The directed linear relation between two signal in the frequency domain based on the Granger causality principle	Directionality of information flow	Nonlinearity not considered Not possible to make a distinction between direct and indirect relations Sensitive to volume conduction	167
Directed transfer function	Gives the causal relation between the outflow of node X towards node Y in the frequency domain based on Granger causality principle, normalized by all inflows towards node Y	Directionality of information flow Distinction between common source and interconnectedness Insensitive to volume conduction	Nonlinearity not considered Not possible to make a distinction between direct and indirect relations Noisy channels affect the directionality Difficult to estimate an optimal order for the multivariate model	168

Table 2.1 (Continued). Overview of different connectivity measures used in EEG and MEG research, with their advantages and disadvantages.

Connectivity measure	Property measured	Advantage(s)	Disadvantage(s)	Key reference(s)
Partial directed coherence	Gives the causal relation between the outflow of node X towards node Y in the frequency domain based on Granger causality principle, normalized by all outflows from node X	Directionality of information flow Insensitive to volume conduction	Nonlinearity not considered Not possible to make a distinction between direct and indirect relations No conclusion about the strength of coupling, due to normalization Difficult to estimate an optimal order for the multivariate model	169
Imaginary part of coherency	Based on coherency*, but excluded the influence of volume conduction by including only the imaginary part of the coherency	Less sensitive to volume conduction	Nonlinearity not considered Imaginary part is mostly small, thereby risking to miss meaningful interactions Not possible to make a distinction between direct and indirect relations	153
Mutual information	Gives the amount of information in signal X that can be explained by signal Y and vice versa, based on the probability distribution of X and Y, and the joint probability distribution of X and Y	Mutual information is sensitive in narrow-frequency band analysis	No directionality of the interaction Weak coupling could be missed Complicated computational measure to obtain from experimental time series Not possible to make a distinction between direct and indirect relations	170
Synchronization likelihood	Describes the normalized strength of the mutual information between signal X with signal Y in state space	Adequately deals with complexity caused by interacting systems Sensitive to nonlinear relations	Sensitive to volume conduction Not possible to make a distinction between direct and indirect relations	171
Phase locking value	Gives the modulus of the averaged instantaneous phase differences between two time series	Nonlinearity is taken into account	No directionality of interaction Sensitive to volume conduction The size of the instantaneous phase difference is included, however, there is no evidence that the size of the phase difference is important for the coupling strength Not possible to make a distinction between direct and indirect relations	150

**Coherency between two time signals is the lineair relation at a specific frequency, an imaginary-valued measure containing information about the magnitude and phase between the signals. Coherence is the absolute value of coherency, containing only information about the magnitude*

Phase slope index	Estimates the direction of information flow, based on the slope of the phase difference of the cross spectral density between signal X and Y	Directionality of information flow Weighs the contribution of different time series Not affected by mixture of independent sources (e.g., background activity)	Not possible to make a distinction between direct and indirect relations Complicated computational method	154
Phase lag index	The asymmetry of the distribution of phase differences between two signals	Less sensitive to volume conduction, common sources, and montage	The risk to miss linear but functionally meaningful interactions The instantaneous phase differences are binarized, therefore, small phase differences may also be missed under noisy conditions Not possible to make a distinction between direct and indirect relations	151
Weighted phase lag index	The contribution of the observed phase leads and lags is weighted by the magnitude of the imaginary part of the coherency	Reduced sensitivity to noise (cf. PLI) Improved detection of phase synchronization changes (cf. PLI)	The size of the instantaneous phase difference is included, however, there is no evidence that the size of the phase difference is important for the coupling strength Relative insensitive to phase differences around 0 and 180 degrees Mixes information about consistency and magnitude of phase differences, hampering interpretation Not possible to make a distinction between direct and indirect relations	172
Directed phase lag index	The probability that the instantaneous phase of signal X was smaller than the phase of signal Y (modulo π) over time	Directionality of information flow Less sensitive to volume conduction and common sources (cf. PLI)	Directionality can be ambiguous as leading with a small differences is similar to lagging with a large phase difference Not possible to make a distinction between direct and indirect relations	173

meaningful manner. Obviously, all these different connectivity measures have their unique advantages and disadvantages. In Table 2.1 we provided an overview of the currently most often used measures, including their main advantages and disadvantages. In addition to computational differences between connectivity measures, it is important to consider related methodological issues that could be encountered when choosing, and interpreting the results of, a connectivity measure. This includes, for example, the problem of field spread, volume conduction and specific reference montages, which for many measures leads to an erroneously high estimate of connectivity between two recording sites[69,87,151]. It is possible, however, to remove these biases prior to computing connectivity measures[41,152] or to estimate the influence of the bias on the connectivity measure through simulations[40]. A more straightforward approach is to use phase-based connectivity measures that are less sensitive to these spurious interactions. Typically, these phase-based measures, such as the imaginary part of coherence[153], phase-slope index[154] and phase lag index[151] ignore the zero phase interaction that are the result of volume conduction/field spread (at the expense of ignoring true zero phase interactions). The phase lag index has the additional advantage that it does not depend directly on the amplitude of the signal[155].

FUNCTIONAL NETWORKS

Resting state EEG and MEG data can be used to construct connectivity matrices and, consequently, functional networks by using network analysis[6,156,157]. In contrast to connectivity measures, which only provide information on how pairs of different brain regions are (functionally) connected, network analysis characterizes the organization of networks[158]. Complex network analysis, a branch of graph theory, reduces the brain into a collection of ‘nodes’ and ‘edges’ and allows quantitative characterization of these networks. In EEG and MEG research, nodes correspond to the recording sites (electrodes or sensors), or specific brain regions when using a source space analysis. Edges are connections between nodes and represent (functional) connectivity values. Together, nodes and edges form the basic elements of a network, and from these elements various global and local network measures can be inferred, for an overview see[159] and section 'Global and local network metrics'. Providing an overview on the vast literature on functional networks and its statistical challenges[160] is beyond the scope of this paper; rather, we discuss relevant issues that need to be considered when preparing and analyzing a resting state EEG or MEG study.

NETWORK CONSTRUCTION

Functional networks are based on the strength or consistency of functional interactions between the network nodes. In a weighted network, the strength of this interaction is taken into account, whereas in an unweighted network only the existence or absence of an interaction is taken into account. Such a binary network is obtained by setting a threshold for the functional connectivity, above which a functional connection is considered to be present. A motivation to use a binary network could be to discard spurious connections that are potentially influenced by, for example, noise[6]. Selecting the value for the threshold is, however, arbitrary and may vary between individuals and groups[161]. Although a weighted network overcomes the problem of this subjective factor and provides a more realistic representation of functional networks, spurious weak connections are also taken into account, potentially influencing network metrics. Besides weighted or unweighted, a network can be directed or undirected. In order to construct a directed network, one should use an effective connectivity measure to infer information on the directionality of communication. Although this could potentially provide useful additional information on network functioning, most studies use undirected networks[161].

NETWORK DENSITY

Network density refers to the number of connections in a network and is influenced by the size of the network. When comparing networks between subjects, the number of nodes should therefore be equal, as it will directly influence network density and various network metrics[161]. A straightforward approach to correct for network density is by using a binary network. This, however, will lead to a data reduction wherein valuable information of the network is not taken into account[161]. Another way to reduce the effect of network density on network metrics is by using weighted networks in combination with normalization procedures that use network metrics based on surrogate data for comparison[158,159]. Although these steps reduce the influence of network density on the eventual network metrics, it is still difficult to make an unbiased comparison between networks[161,174]. Furthermore, each of these steps can potentially influence the computation of network metrics and should therefore be included in the methodological section of the study.

MINIMUM SPANNING TREE

We have briefly mentioned several strategies to improve comparison between networks. Some of these strategies involve specific methodological choices, such as

epoch selection, artifact handling, specific frequency bands, and connectivity measures. Other solutions include normalization procedures and the use of weighted networks[161]. Despite these efforts, traditional network metrics will remain sensitive to the effect of network density. A new approach to overcome the problem of network density in network analytical studies involves computation of the minimum spanning tree (MST). The MST is a unique acyclic subgraph that contains most of the strongest connections of the original undirected, weighted graph, for review see[174]. As the communication in the original graph follows the most efficient paths[175], the MST can be considered the backbone of the functional brain network[176,177]. During the process of constructing the MST graph, connections that will lead to loops in the network will be excluded from the eventual network. By doing so, the number of connections in an MST graph will always correspond to the number of nodes minus one. As a result, MST networks with the same number of nodes will automatically have the same number of connections, thereby facilitating the comparison of networks. Moreover, normalization procedures with surrogate networks are not necessary anymore. A possible disadvantage of the MST approach is that it may miss information about the network topology that is contained in the weaker connections of the network. An increasing number of studies have evaluated the practical utility of the MST approach for network analysis in resting state EEG[51,178–181] and MEG[29,182,183] recordings.

GLOBAL AND LOCAL NETWORK METRICS

Several metrics exist to characterize the organization of networks[159]. The two most commonly used network metrics are the average path length, a measure of global integration of the network, and the average clustering coefficient that defines local segregation of a network. Both average path length and average clustering coefficient are considered to be global network metrics as they provide global information of the network and are commonly used to describe the network topology. An optimal network organization is characterized by a short average shortest path length and a high average clustering coefficient, also called a ‘small-world’ configuration[184]. Although these global network metrics are appealing to use, as they have been widely used in network analytical studies, they fail to explain the diversity found at node level[6,158].

To explain this diversity at node level, local network metrics are used. Network metrics like degree, betweenness centrality, and eigenvector centrality are used to specify the level of importance of a specific node in a network[159]. Nodes with many connections and a central position within the network are considered 'hubs'. Removal of a hub-node will have a considerable impact on the network[6,158]. Often, a network is built out of smaller subnetworks, also called 'modules'. These modules are formed by groups of nodes that are highly connected to each other[185], but much less so to other nodes in the network. Provincial hubs are the most highly connected nodes within a module. Connector hubs are the important nodes connecting one module with other modules in the network[186]. Each local network metric captures specific information about the network topology, making it appropriate for certain analysis, depending on the specific research question. For example, the eigenvector centrality gives a more accurate estimation of the centrality of a specific node in the network than degree[187], and has lower computational costs than betweenness centrality[188], it is less sensitive for the detection of hubs in modules[189] and has specific normalization problems[190].

When making choices for specific network metrics, it is recommendable to take into account the chosen connectivity measure. Connectivity measures that depend on linear correlations or the level of synchronization are more susceptible to volume conduction (Table 2.1). As a result, particularly the interaction between two nearby network nodes will be overestimated[191], which might result in an overestimation of (local) clustering. Another important issue to consider when choosing network metrics, is the influence of network density[161]. Although it is possible to reduce the effect of network density, the MST approach offers an elegant manner to control for differences in network density. Like in conventional networks, MST metrics are inferred from MST graphs. For example, MST diameter and leaf number are global metrics that provide information on network integration and segregation, similarly to average path length and average clustering coefficient. Furthermore, various other local MST metrics can be computed[174,181,182]. Finally, it is important to realize that some network metrics are highly correlated[26,192], which means that some combinations of network metrics are redundant. It is difficult to recommend specific network metrics, as this will eventually depend on the specific research question.

CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

We have summarized several problems and challenges by reviewing current practice in resting state functional connectivity EEG and MEG research. First, performing a resting state recording might not be as straightforward as it seems; behavior during, and perception of, a stimulus independent condition may vary greatly between subjects despite similar instructions[45]. In our overview, we differentiated subject-related from measurement-related methodological issues. For future research we suggest to explicitly mention the instructions given to the subject and to control for factors that might influence the state of vigilance of subjects. Second, we mentioned technical issues that are important to consider when collecting data from resting state EEG and MEG recordings (Figure 2.2 for summary). Since the current literature is too diverse to provide a uniform methodological guideline, we suggest including different methodological approaches in resting state studies to better understand the influence of these approaches on study results. Some recommendations are, however, useful irrespective of the chosen approach[193]. We summarized these recommendations in Table 2.2.

Since resting state EEG and MEG recordings are increasingly used for a network analytical approach[6,158], we briefly introduced some relevant topics. It is recommendable to decide on the connectivity measure and network metrics simultaneously as they are mutually dependent. We provided suggestions to overcome some limitations that are inherent to conventional network analysis[161] and offered a new approach to overcome the influence of network density on network metrics: the minimum spanning tree[174].

Finally, methodological studies are needed to systematically investigate the influence of various choices that researchers have to make when conducting a functional connectivity resting state experiment. Particularly the process of selecting data and artifact handling needs a more evidence-based approach. Regarding the choice of an appropriate connectivity measure, more information on the advantages and disadvantage is

available[134,147–149]. Increasing evidence exist that some connectivity measures are more vulnerable to volume conduction (see Table 2.1), leading to unreliable connectivity values, and consequently, unreliable network estimations. Future studies should use this knowledge to make appropriate decisions. We await methodological studies wherein different methodological issues are investigated systematically. This would be invaluable to the field of functional connectivity and network studies.

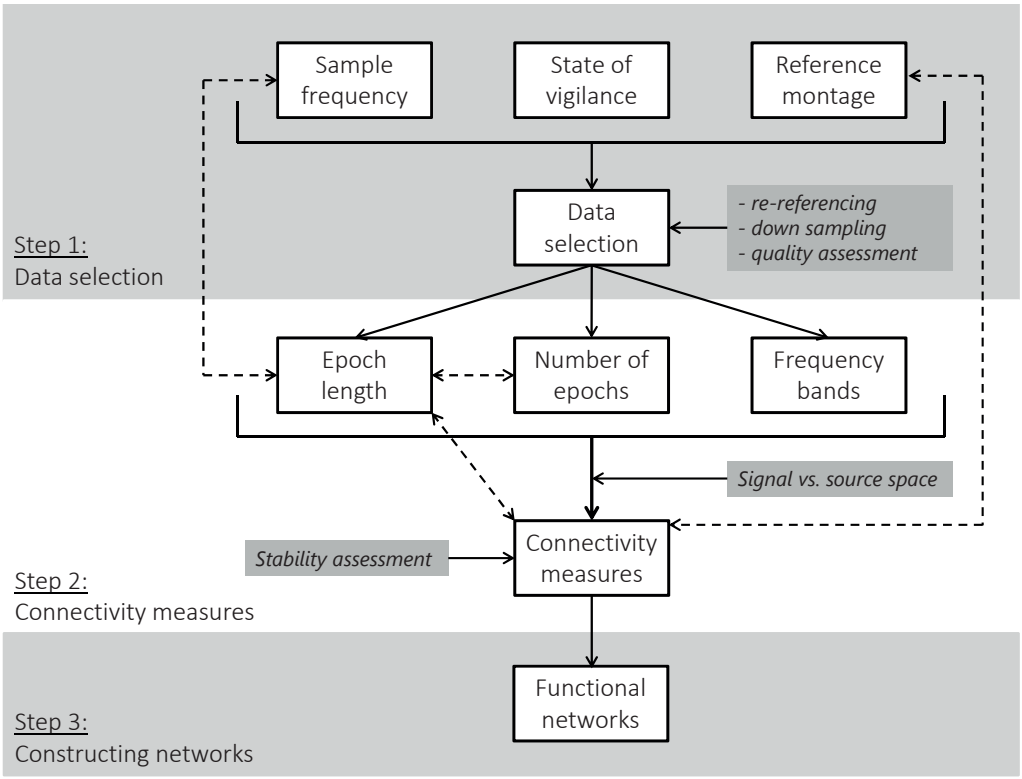


Figure 2.2. Schematic illustration how various methodological choices will influence resting state data and, consequently, the construction of functional networks.
Note that choice of connectivity measures is highly depended on methodological choices. Furthermore, several methodological options exist during step 3. These options are discussed in section 'Functional networks'.

Table 2.2. General recommendations for methodology of functional connectivity resting state EEG and MEG research.

Subject-related methodological issues
Heterogeneity of resting state condition - Use a priori defined instructions throughout a study to reduce the external induced heterogeneity - Eliminate or control for possible confounders such as time of the day, intake of caffeine or medication, task or physical performance prior to recording and state of vigilance
Measurement-related methodological issues
Choice of reference - Conventional reference montages will influence EEG measures, both a reference-free and (robust) common average reference are reasonable choices - Various reference montages will influence the estimated strength and directionality of information flow and thereby the outcome of functional and effective connectivity Epoch selection - Select the first artifact free epochs after the start of the resting state recording to minimize any potential selection bias - Re-inspect selected epochs by an independent researcher to avoid any selection or systematic bias - Include extra epochs to perform a leave-one-out analysis to investigate stability of connectivity and network measures - Length and sample frequency of epochs should be equal and appropriate for the connectivity measure that is used Filtering and artifact handling - Select epochs without eye-movement or muscle artifacts. If not possible, use an automatic artifact reduction approach and describe these procedures or exclude affected channels (but maintain an equal number of channels per subject) - Use zero-phase filtering to eliminate the phase shift of digital filters on the signal Frequency bands - Differentiate neurophysiological signals in separate frequency bands when an effect is expected in specific bands. Broadband analysis could be used when testing an undirected hypothesis - Avoid gamma band when impossible to control for myogenic influences - Perform a spectral power analysis along with your connectivity analysis
Connectivity measures
Volume conduction - Avoid connectivity measures that are susceptible for volume conduction - Compute connectivity values per epoch and consequently an average value per subject to increase the stability of connectivity values
Functional networks
Network construction - Use weighted networks to avoid subjectivity in unweighted network analysis or approaches not influenced by network density (e.g., minimum spanning tree) - Use directed networks in combination with effective connectivity measures when information is needed on the directionality within the network Network density - When using traditional network metrics use a combination of weighted networks and normalization procedures with surrogate data to correct for the influence of density or use the minimum spanning tree approach

REFERENCES

1 Greicius, M., Krasnow, B., Reiss, A, et al. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc. Natl. Acad. Sci. U. S. A. 100, 253-258 (2003).

2 Deco, G., Jirsa, V. K. & McIntosh, A. R. Emerging concepts for the dynamical organization of resting-state activity in the brain. Nat. Rev. Neurosci. 12, 43–56 (2011).

3 Birn, R. M. The role of physiological noise in resting-state functional connectivity. Neuroimage 62, 864–870 (2012).

4 Greicius, M. Resting-state functional connectivity in neuropsychiatric disorders. Curr. Opin. Neurol. 24, 424–430 (2008).

5 Bassett, D. S. & Bullmore, E. T. Human brain networks in health and disease. Curr. Opin. Neurol. 22, 340–347 (2009).

6 Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10, 186–198 (2009).

7 Stam, C. J. Modern network science of neurological disorders. Nat. Rev. Neurosci. 15, 683–695 (2014).

8 Raichle, M. E. A Paradigm shift in functional brain imaging. J. Neurosci. 29, 12729–12734 (2009).

9 Pinneo, L. R. On noise in the nervous system. Psychol. Rev. 73, 242–247 (1966).

10 Snyder, A. Z. & Raichle, M. E. A brief history of the resting state: The Washington University perspective. Neuroimage 62, 902–910 (2012).

11 Lopes da Silva, F. EEG and MEG: relevance to neuroscience. Neuron. 80, 1112–1128 (2013).

12 Koenig, T., Studer, D., Hubl, D., et al. Brain connectivity at different time-scales measured with EEG. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 360, 1015–1023 (2005).

13 Stam, C. J., Jones, B. F., Nolte, G., et al. Small-world networks and functional connectivity in Alzheimer's disease. Cereb. cortex 17, 92–99 (2007).

14 Dubovik, S., Bouzerda-Wahlen, A., Nahum, L., et al. Adaptive reorganization of cortical networks in Alzheimer's disease. Clin. Neurophysiol. 124, 35–43 (2013).

15 Bartolomei, F., Bosma, I., Klein, M., et al. Disturbed functional connectivity in brain tumour patients: Evaluation by graph analysis of synchronization matrices. Clin. Neurophysiol. 117, 2039–2049 (2006).

16 Ibrahim, G. M., Anderson, R., Akiyama, T., et al. Neocortical pathological high-frequency oscillations are associated with frequency-dependent alterations in functional network topology. J. Neurophysiol. 110, 2475–2483 (2013).

17 Hinkley, L. Cognitive impairments in schizophrenia as assessed through activation and connectivity measures of magnetoencephalography (MEG) data. Front. Hum. Neurosci. 3, (2010).

18 Siebenhühner, F., Weiss, S. A., Coppola, R., et al. Intra- and inter-frequency brain network structure in health and schizophrenia. PLoS One 8, e72351 (2013).

19 Schoonheim, M. M., Geurts, J. J. G., Landi, D., et al. Functional connectivity changes in multiple sclerosis patients : a graph analytical study of MEG resting state data. Hum. Brain Mapp. 61, 52–61 (2013).

20 Van Schependom, J., Gielen, J., Laton, J., et al. Graph theoretical analysis indicates cognitive impairment in MS stems from neural disconnection. NeuroImage Clin. 4, 403–410 (2014).

21 Fogelson, N., Li, L., Li, Y., et al. Functional connectivity abnormalities during contextual processing in schizophrenia and in Parkinson's disease. Brain Cogn. 82, 243–253 (2013).

22 Smit, D. J. A, Boersma, M., Schnack, H. G., et al. The brain matures with stronger functional connectivity and decreased randomness of its network. PLoS One 7, e36896 (2012).

23 Vecchio, F., Miraglia, F., Bramanti, P., et al. Human brain networks in physiological aging: A graph theoretical analysis of cortical connectivity from EEG data. J. Alzheimer's Dis. 41, 1239–1249 (2014).
24 Boersma, M., Smit, D. J. A, de Bie, H. M. A, et al. Network analysis of resting state EEG in the developing young brain: structure comes with maturation. Hum. Brain Mapp. 32, 413–425 (2011).
25 Douw, L., Nieboer, D., van Dijk, B. W., et al. A healthy brain in a healthy body: brain network correlates of physical and mental fitness. PLoS One 9, e88202 (2014).
26 Van Diessen, E., Otte, W. M., Braun, K. P. J., et al. Improved diagnosis in children with partial epilepsy using a multivariable prediction model based on EEG network characteristics. PLoS One 8, e59764 (2013).
27 Wilke, C., Worrell, G. & He, B. Graph analysis of epileptogenic networks in human partial epilepsy. Epilepsia 52, 84–93 (2011).
28 De Waal, H., Stam, C. J., Lansbergen, M. M., et al. The effect of Souvenaid on functional brain network organisation in patients with mild Alzheimer's disease: a randomised controlled study. PLoS One 9, e86558 (2014).
29 Olde Dubbelink, K. T. E., Hillebrand, A., Stoffers, D., et al. Disrupted brain network topology in Parkinson's disease: a longitudinal magnetoencephalography study. Brain 137, 197–207 (2014).
30 Duncan, N. & Northoff, G. Overview of potential procedural and participant- related confounds for neuroimaging of the resting state. J. Psychiatry Neurosci. 38, 84–96 (2013).
31 Gross, J. Analytical methods and experimental approaches for electrophysiological studies of brain oscillations. J. Neurosci. Methods 228, 57–66 (2014).
32 Berger, H. Über das elektrenkephalogramm des menschen. Eur. Arch. Psychiatry Clin. Neurosci. 87, 527–570 (1929).
33 Collura, T. F. History and evoluation of electroencephalographic instruments and techniques. J. Clin. Neurophysiol. 10, 476–504 (1993).
34 Biswal, B., Yetkin, F. Z., Haughton, V. M., et al. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn. Reson. Imaging 34, 537–541 (1995).
35 Damoiseaux, J. S., Rombouts, S. A. R. B., Barkhof, F., et al. Consistent resting-state networks across healthy subjects. Proc. Natl. Acad. Sci. 103, 13848–13853 (2006).
36 Shulman, G. L., Corbetta, M., Buckner, R. L., et al. Common blood flow changes across visual tasks: I. Increases in subcortical structures and cerebellum but not in nonvisual cortex. J. Cogn. Neurosci. 9, 624–647 (1997).
37 Raichle, M. E., MacLeod, A. M., Snyder, A. Z., et al. A default mode of brain function. Proc. Natl. Acad. Sci. U. S. A. 98, 676–682 (2001).
38 Rosazza, C. & Minati, L. Resting-state brain networks: literature review and clinical applications. Neurol. Sci. 32, 773–785 (2011).
39 Miller, K. J., Weaver, K. E. & Ojemann, J. G. Direct electrophysiological measurement of human default network areas. Proc. Natl. Acad. Sci. 106, 12174–12177 (2009).
40 Brookes, M. J., Woolrich, M., Luckhoo, H., et al. Investigating the electrophysiological basis of resting state networks using magnetoencephalography. Proc. Natl. Acad. Sci. 108, 16783–16788 (2011).
41 Hipp, J. F., Hawellek, D. J., Corbetta, M., et al. Large-scale cortical correlation structure of spontaneous oscillatory activity. Nat. Neurosci. 15, 884–890 (2012).
42 Benjamin, C., Lieberman, D. A., Chang, M., et al. The influence of rest period instructions on the default mode network. Front. Hum. Neurosci. 4, 1–9 (2010).
43 Andreasen, N. C., O'Leary, D. S., Cizadlo, T., et al. Remembering the past: two facets of

episodic memory explored with positron emission tomography. Am. J. Psychiatry 152, 1576–1585 (1995).
44 Teasdale, J. D., Dritschel, B. H., Taylor, M. J., et al. Stimulus-independent thought depends on central executive resources. Mem. Cognit.23, 551–559 (1995).
45 Diaz, B. A., Van Der Sluis, S., Moens, S., et al. The Amsterdam resting-state questionnaire reveals multiple phenotypes of resting-state cognition. Front. Hum. Neurosci. 7, 446 (2013).
46 López Zunini, R. A, Thivierge, J.-P., Kousaie, S., et al. Alterations in resting-state activity relate to performance in a verbal recognition task. PLoS One 8, e65608 (2013).
47 Honey, C. J., Thesen, T., Donner, T. H., et al. Slow cortical dynamics and the accumulation of information over long timescales. Neuron 76, 423–434 (2012).
48 Olbrich, S., Mulert, C., Karch, S., et al. EEG-vigilance and BOLD effect during simultaneous EEG/fMRI measurement. Neuroimage 45, 319–332 (2009).
49 Minkwitz, J., Trenner, M. U., Sander, C., et al. Prestimulus vigilance predicts response speed in an easy visual discrimination task. Behav. brain Funct. 7, 1–8 (2011).
50 Koenis, M. M. G., Romeijn, N., Piantoni, G., et al. Does sleep restore the topology of functional brain networks? Hum. Brain Mapp. 34, 487–500 (2013).
51 Van Diessen, E., Otte, W. M., Braun, K. P. J., et al. Does sleep deprivation alter functional EEG networks in children with focal epilepsy? Front. Syst. Neurosci. 8, 67 (2014).
52 Niedermeyer, E. & Lopes da Silva, F. Electroencephalography, basic principles, clinical applications, and related fields. (Urban & Schwarzenberg, 1987).
53 Cantero, J. L., Atienza, M., Salas, R. M., et al. Alpha EEG coherence in different brain states: an electrophysiological index of the arousal level in human subjects. Neurosci. Lett. 271, 159–162 (1999).
54 Kuhnert, M.-T., Elger, C. E. & Lehnertz, K. Long-term variability of global statistical properties of epileptic brain networks. Chaos 20, 043126 (2010).
55 Piantoni, G., Cheung, B. L. P., Van Veen, B. D., et al. Disrupted directed connectivity along the cingulate cortex determines vigilance after sleep deprivation. Neuroimage 79, 213–222 (2013).
56 Klimesch, W., Doppelmayr, M., Schwaiger, J., et al. 'Paradoxical' alpha synchronization in a memory task. Cogn. Brain Res. 7, 493–501 (1999).
57 Siepmann, M. & Kirch, W. Effects of caffeine on topographic quantitative EEG. Neuropsychobiology 45, 161–166 (2002).
58 Barry, R. J., Clarke, A. R. & Johnstone, S. J. Caffeine and opening the eyes have additive effects on resting arousal measures. Clin. Neurophysiol. 122, 2010–2015 (2011).
59 Tal, O., Diwakar, M., Wong, C.-W., et al. Caffeine-induced global reductions in resting-state BOLD connectivity reflect widespread decreases in MEG connectivity. Front. Hum. Neurosci. 7, 63 (2013).
60 Rice, J. K., Rorden, C., Little, J. S., et al. Subject position affects EEG magnitudes. Neuroimage 64, 476–484 (2013).
61 Horstmann, M.-T., Bialonski, S., Noennig, N., et al. State dependent properties of epileptic brain networks: comparative graph-theoretical analyses of simultaneously recorded EEG and MEG. Clin. Neurophysiol. 121, 172–185 (2010).
62 Tan, B., Kong, X., Yang, P., et al. The difference of brain functional connectivity between eyes-closed and eyes-open using graph theoretical analysis. Comput. Math. Methods Med. 2013, 976365 (2013).
63 Jin, S.-H., Jeong, W., Lee, D.-S., et al. Preserved high-centrality hubs but efficient network reorganization during eyes-open state compared with eyes-closed resting state: an MEG study. J. Neurophysiol. 111, 1455–1465 (2014).

64 Xu, P., Huang, R., Wang, J., et al. Different topological organization of human brain functional networks with eyes open versus eyes closed. Neuroimage 90, 246–255 (2014).
65 Davidson, R. J. EEG measures of cerebral asymmetry: conceptual and methodological issues. Int. J. Neurosci. 39, 71–89 (1988).
66 Allen, J. J. B., Coan, J. A. & Nazarian, M. Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion. Biol. Psychol. 67, 183–218 (2004).
67 Corsi-Cabrera, M., Galindo-Vilchis, L., del-Río-Portilla, Y., et al. Within-subject reliability and inter-session stability of EEG power and coherent activity in women evaluated monthly over nine months. Clin. Neurophysiol. 118, 9–21 (2007).
68 Pivik, R. T., Broughton, R. J., Coppola, R., et al. Guidelines for the recording and quantitative analysis of electroencephalographic activity in research contexts. Psychophysiology 30, 547–558 (1993).
69 Nunez, P. L., Srinivasan, R., Westdorp, A. F., et al. EEG coherency. I: Statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales. Electroencephalogr. Clin. Neurophysiol. 103, 499–515 (1997).
70 Hu, S., Stead, M., Dai, Q., et al. On the recording reference contribution to EEG correlation, phase synchorony, and coherence. IEEE Trans. Syst. Man, Cybern. Part B 40, 1294–1304 (2010).
71 Kayser, J. & Tenke, C. E. In search of the Rosetta Stone for scalp EEG: converging on reference-free techniques. Clinical Neurophysiology 121, 1973–1975 (2010).
72 Schiff, S. J. Dangerous phase. Neuroinformatics 3, 315–318 (2005).
73 Nunez, P. L. & Srinivasan, R. Electric field of the brain. The neurophysics of EEG. (Oxford University Press, 2006).
74 Yao, D. A method to standardize a reference of scalp EEG recordings to a point at infinity. Physiol. Meas. 22, 693–711 (2001).
75 Hjorth, B. An on-line transformation of EEG scalp potentials into orthogonal source derivations. Electroencephalogr. Clin. Neurophysiol. 39, 526–530 (1975).
76 Tenke, C. E. & Kayser, J. Generator localization by current source density (CSD): Implications of volume conduction and field closure at intracranial and scalp resolutions. Clin. Neurophysiol. 123, 2328–2345 (2012).
77 Lepage, K. Q., Kramer, M. A.& Chu, C. J. A statistically robust EEG re-referencing procedure to mitigate reference effect. J. Neurosci. Methods 235C, 101–116 (2014).
78 Qin, Y., Xu, P. & Yao, D. A comparative study of different references for EEG default mode network: the use of the infinity reference. Clin. Neurophysiol. 121, 1981–1991 (2010).
79 Van Straaten, E. C. W., den Haan, J., de Waal, H., et al. Disturbed phase relations in white matter hyperintensity based vascular dementia: An EEG directed connectivity study. Clin. Neurophysiol. 126, 497–504 (2015).
80 Nunez, P. L. REST: A good idea but not the gold standard. Clin. Neurophysiol. 121, 2177–2180 (2010).
81 Sarvas, J. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. Phys. Med. Biol. 32, 11–22 (1987).
82 Van den Broek, S. P., Reinders, F., Donderwinkel, M., et al. Volume conduction effects in EEG and MEG. Electroencephalogr. Clin. Neurophysiol. 106, 522–534 (1998).
83 Helmholtz, H. Ueber einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern mit Anwendung auf die thierisch-elektrischen Versuche. Ann. der Phys. 165, 211–233 (1853).
84 Hillebrand, A. & Barnes, G. R. Beamformer analysis of MEG data. Int. Rev. Neurobiol. 68, 149–171 (2005).
85 Michel, C. M. & Murray, M. M. Towards the utilization of EEG as a brain imaging tool. Neuroimage 61, 371–385 (2012).

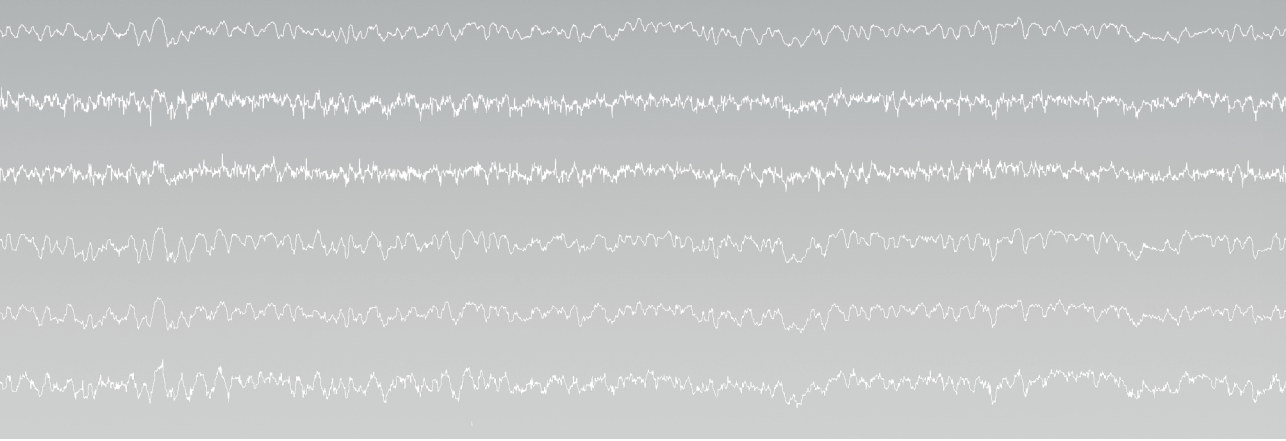
86 Hillebrand, A., Barnes, G. R., Bosboom, J. L., et al. Frequency-dependent functional connectivity within resting-state networks: an atlas-based MEG beamformer solution. Neuroimage 59, 3909–3921 (2012).
87 Schoffelen, J.-M. & Gross, J. Source connectivity analysis with MEG and EEG. Hum. Brain Mapp. 30, 1857–1865 (2009).
88 Baillet, S., Mosher, J. C. & Leahy, R. M. Electromagnetic brain mapping. IEEE Signal Process. Mag. 18, 14–30 (2001).
89 Pascual-Marqui, R. D., Michel, C. M. & Lehmann, D. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18, 49–65 (1994).
90 Van Drongelen, W., Yuchtman, M., Van Veen, B., et al. A spatial filtering technique to detect and localize multiple sources in the brain. Brain Topogr. 9, 39–49 (1996).
91 Lantz, G., Grave de Peralta, R., Spinelli, L., et al. Epileptic source localization with high density EEG: how many electrodes are needed? Clin. Neurophysiol. 114, 63–69 (2003).
92 Laarne, P. H., Tenhunen-Eskelinen, M. L., Hyttinen, J. K., et al. Effect of EEG electrode density on dipole localization accuracy using two realistically shaped skull resistivity models. Brain Topogr. 12, 249–254 (2000).
93 Hillebrand, A. & Barnes, G. R. The use of anatomical constraints with MEG beamformers. Neuroimage 20, 2302–2313 (2003).
94 Hillebrand, A. & Barnes, G. R. Practical constraints on estimation of source extent with MEG beamformers. Neuroimage 54, 2732–2740 (2011).
95 Alschuler, D. M., Tenke, C. E., Bruder, G. E., et al. Identifying electrode bridging from electrical distance distributions: A survey of publicly-available EEG data using a new method. Clin. Neurophysiol. 125, 484–490 (2014).
96 Croft, R. J. & Barry, R. J. Removal of ocular artifact from the EEG: A review. Neurophysiologie Clinique 30, 5–19 (2000).
97 Cassani, R., Falk, T. H., Fraga, F. J., et al. The effects of automated artifact removal algorithms on electroencephalography-based Alzheimer's disease diagnosis. Front. Aging Neurosci. 6, 1–13 (2014).
98 LeVan, P., Urrestarazu, E. & Gotman, J. A system for automatic artifact removal in ictal scalp EEG based on independent component analysis and Bayesian classification. Clin. Neurophysiol. 117, 912–927 (2006).
99 Gao, J. F., Yang, Y., Lin, P., et al. Automatic removal of eye-movement and blink artifacts from EEG signals. Brain Topogr. 23, 105–114 (2010).
100 Muthukumaraswamy, S. D. High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations. Front. Hum. Neurosci. 7, 138 (2013).
101 Delorme, A., Sejnowski, T. & Makeig, S. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. Neuroimage 34, 1443–1449 (2007).
102 Whitham, E. M., Pope, K. J., Fitzgibbon, S. P., et al. Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. Clin. Neurophysiol. 118, 1877–1888 (2007).
103 De Haan, W., Pijnenburg, Y. AL, Strijers, R. L., et al. Functional neural network analysis in frontotemporal dementia and Alzheimer's disease using EEG and graph theory. BMC Neurosci. 10, 101 (2009).
104 Van Dellen, E., van der Kooi, A. W., Numan, T., et al. Decreased functional connectivity and disturbed directionality of information flow in the electroencephalography of intensive care unit patients with delirium after cardiac surgery. Anesthesiology 121, 328–335 (2014).

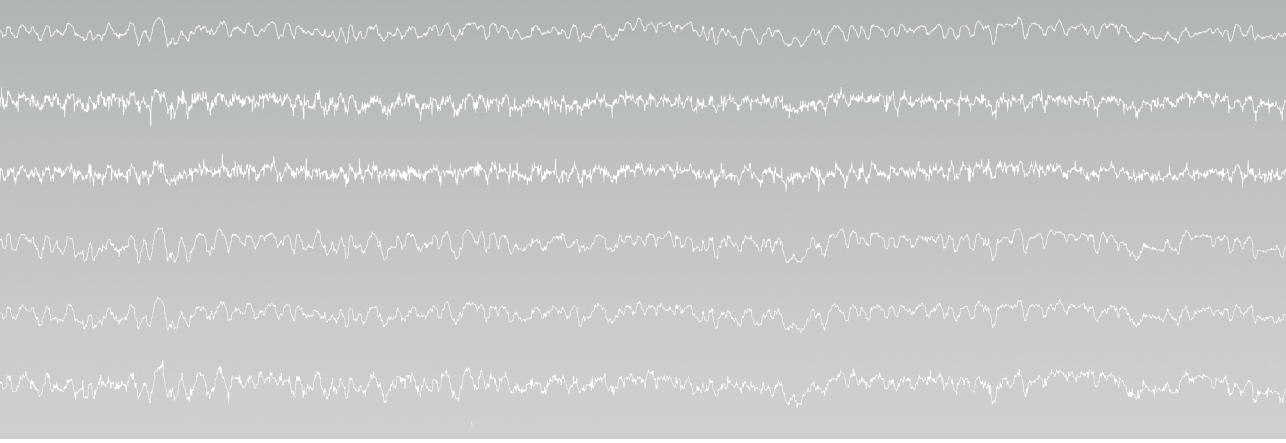
105 Douw, L., de Groot, M., van Dellen, E., et al. Local MEG networks: the missing link between protein expression and epilepsy in glioma patients? Neuroimage 75, 203–211 (2013).
106 Lodder, S. S., Askamp, J. & van Putten, M. J. A. M. Computer-assisted interpretation of the EEG background pattern: a clinical evaluation. PLoS One 9, e85966 (2014).
107 Anderson, N. R. & Doolittle, L. M. Automated analysis of EEG: opportunities and pitfalls. J. Clin. Neurophysiol. 27, 453–457 (2010).
108 Shibasaki, H., Nakamura, M., Sugi, T., et al. Automatic interpretation and writing report of the adult waking electroencephalogram. Clin. Neurophysiol. 125, 1081–1094 (2014).
109 Maltez, J., Hyllienmark, L., Nikulin, V. V., et al. Time course and variability of power in different frequency bands of EEG during resting conditions. Neurophysiol. Clin. Neurophysiol. 34, 195–202 (2004).
110 Knyazeva, M. G., Jalili, M., Brioschi, A., et al. Topography of EEG multivariate phase synchronization in early Alzheimer's disease. Neurobiol. Aging 31, 1132–1144 (2010).
111 Chu, C. J., Kramer, M. A., Pathmanathan, J., et al. Emergence of stable functional networks in long-term human electroencephalography. J. Neurosci. 32, 2703–2713 (2012).
112 Tahaei, M. S., Jalili, M. & Knyazeva, M. G. Synchronizability of EEG-based functional networks in early Alzheimer 's disease. IEEE Trans. neural Syst. Rehabil. Eng. 20, 636–641 (2012).
113 Ahmadlou, M. & Adeli, H. Functional community analysis of brain: a new approach for EEG-based investigation of the brain pathology. Neuroimage 58, 401–408 (2011).
114 David, O., Cosmelli, D. & Friston, K. J. Evaluation of different measures of functional connectivity using a neural mass model. Neuroimage 21, 659–673 (2004).
115 Honey, C. J., Kötter, R., Breakspear, M., et al. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. Proc. Natl. Acad. Sci. U. S. A. 104, 10240–10245 (2007).
116 Singer, W. Cortical dynamics revisited. Trends Cogn. Sci. 17, 616–626 (2013).
117 Candes, E. J. Compressive sampling. Proc. Int. Congr. Math. Madrid, Spain 52, 1433–1452 (2006).
118 Lopes da Silva, F. in Electroencephalography: basic principles, clinical applications and related fields 1164–1189 (1998).
119 Başar, E., Başar-Eroglu, C., Karakaş, S., et al. Gamma, alpha, delta, and theta oscillations govern cognitive processes. Int. J. Psychophysiol. 39, 241–248 (2001).
120 Aurlien, H., Gjerde, I. O., Aarseth, J. H., et al. EEG background activity described by a large computerized database. Clin. Neurophysiol. 115, 665–673 (2004).
121 Klimesch, W. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Res. Brain Res. Rev. 29, 169–195 (1999).
122 Hermann, C. & Demiralp, T. Human EEG gamma oscillations in neuropsychiatric disorders. Clin. Neurophysiol. 116, 2719–2733 (2005).
123 Pope, K. J., Fitzgibbon, S. P., Lewis, T. W., et al. Relation of gamma oscillations in scalp recordings to muscular activity. Brain Topogr. 22, 13–17 (2009).
124 Fitzgibbon, S. P., Lewis, T. W., Powers, D. M. W., et al. Surface Laplacian of central scalp electrical signals is insensitive to muscle contamination. IEEE Trans. Biomed. Eng. 60, 4–9 (2013).
125 Yuval-Greenberg, S., Tomer, O., Keren, A. S., et al. Transient induced gamma-band response in EEG as a manifestation of miniature saccades. Neuron 58, 429–441 (2008).
126 Yuval-Greenberg, S. & Deouell, L. Y. The broadband-transient induced gamma-band response in scalp EEG reflects the execution of saccades. Brain Topogr. 22, 3–6 (2009).
127 Kaplan, P. W. & Rossetti, A. O. EEG patterns and imaging correlations in encephalopathy: encephalopathy part II. J. Clin. Neurophysiol. 28, 233–251 (2011).
128 Gasser, T., Bacher, P. & Steinberg, H. Test-retest reliability of spectral parameters of the EEG. Electroencephalogr. Clin. Neurophysiol. 60, 312–319 (1985).

129 Kondacs, a & Szabó, M. Long-term intra-individual variability of the background EEG in normals. Clin. Neurophysiol. 110, 1708–1716 (1999).
130 Fingelkurts, A. A., Fingelkurts, A. A., Ermolaev, V. A., et al. Stability, reliability and consistency of the compositions of brain oscillations. Int. J. Psychophysiol. 59, 116–126 (2006).
131 Gudmundsson, S., Runarsson, T. P., Sigurdsson, S., et al. Reliability of quantitative EEG features. Clin. Neurophysiol. 118, 2162–2171 (2007).
132 Näpflin, M., Wildi, M. & Sarnthein, J. Test-retest reliability of resting EEG spectra validates a statistical signature of persons. Clin. Neurophysiol. 118, 2519–2524 (2007).
133 Salinsky, M. C., Oken, B. S. & Morehead, L. Test-retest reliability in EEG frequency analysis. Electroencephalogr. Clin. Neurophysiol. 79, 382–392 (1991).
134 Pereda, E., Quiroga, R. Q. & Bhattacharya, J. Nonlinear multivariate analysis of neurophysiological signals. Prog. Neurobiol. 77, 1–37 (2005).
135 Stam, C. J. Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. Clin. Neurophysiol. 116, 2266–2301 (2005).
136 Bonita, J. D., Ambolode, L. C. C., Rosenberg, B. M., et al. Time domain measures of inter-channel EEG correlations: a comparison of linear, nonparametric and nonlinear measures. Cogn. Neurodyn. 8, 1–15 (2014).
137 Friston, K. J. Functional and effective connectivity in neuroimaging : a synthesis. Hum. Brain Mapp. 2, 56–78 (1994).
138 Friston, K. J. Functional and effective connectivity: a review. Brain Connect. 1, 13–36 (2011).
139 Aertsen, A. M. H. J., Gerstein, G. L., Habib, M. K., et al. Dynamics of neuronal firing correlation: modulation of 'effective connectivity'. J. Neurophysiol. 61, 900–917 (1989).
140 Varela, F., Lachaux, J., Rodriguez, E., et al. The brainweb: phase synchronization and large-scaleintegration. Nat. Rev. Neurosci. 2, 229–239 (2001).
141 Boccaletti, S., Kurths, J., Osipov, G., et al. The synchronization of chaotic systems. Phys. Rep. 366, 1–101 (2002).
142 Daffertshofer, A. & van Wijk, B. C. M. On the influence of amplitude on the connectivity between phases. Front. Neuroinform. 5, 6 (2011).
143 Stephan, K. E. & Roebroeck, A. A short history of causal modeling of fMRI data. Neuroimage 62, 856–863 (2012).
144 Korzeniewska, A., Cervenka, M. C., Jouny, C. C., et al. Ictal propagation of high frequency activity is recapitulated in interictal recordings: Effective connectivity of epileptogenic networks recorded with intracranial EEG. Neuroimage 101, 96–113 (2014).
145 Lee, H., Mashour, G. A., Noh, G.-J., et al. Reconfiguration of network hub structure after propofol-induced unconsciousness. Anesthesiology 119, 1347–1359 (2013).
146 Shaw, J. C. Correlation and coherence analysis of the EEG: A selective tutorial review. Int. J. Psychophysiol. 1, 255–266 (1984).
147 Wendling, F., Ansari-Asl, K., Bartolomei, F., et al. From EEG signals to brain connectivity: A model-based evaluation of interdependence measures. J. Neurosci. Methods 183, 9–18 (2009).
148 Van Mierlo, P., Papadopoulou, M., Carrette, E., et al. Functional brain connectivity from EEG in epilepsy: Seizure prediction and epileptogenic focus localization. Prog. Neurobiol. 121, 19-35 (2014).
149 Vindiola, M., Vettel, J. M., Gordon, S. M., et al. Applying EEG phase synchronization measures to non-linearly coupled neural mass models. J. Neurosci. Methods 226, 1–14 (2014).
150 Lachaux, J. P., Rodriguez, E., Martinerie, J., et al. Measuring phase synchrony in brain signals. Hum. Brain Mapp. 8, 194–208 (1999).

151 Stam, C. J., Nolte, G. & Daffertshofer, A. Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. Hum. Brain Mapp. 28, 1178–1193 (2007).
152 Brookes, M. J., Woolrich, M. W. & Barnes, G. R. Measuring functional connectivity in MEG: A multivariate approach insensitive to linear source leakage. Neuroimage 63, 910–920 (2012).
153 Nolte, G., Bai, O., Wheaton, L., et al. Identifying true brain interaction from EEG data using the imaginary part of coherency. Clin. Neurophysiol. 115, 2292–2307 (2004).
154 Nolte, G., Ziehe, A., Nikulin, V., et al. Robustly estimating the flow direction of information in complex physical systems. Phys. Rev. Lett. 100, 234101 (2008).
155 Muthukumaraswamy, S. D. & Singh, K. D. A cautionary note on the interpretation of phase-locking estimates with concurrent changes in power. Clinical Neurophysiology 122, 2324–2325 (2011).
156 Sporns, O., Chialvo, D., Kaiser, M., et al. Organization, development and function of complex brain networks. Trends Cogn. Sci. 8, 418–425 (2004).
157 Stam, C. J. Characterization of anatomical and functional connectivity in the brain: a complex networks perspective. Int. J. Psychophysiol. 77, 186–194 (2010).
158 Stam, C. J. & van Straaten, E. C. W. The organization of physiological brain networks. Clin. Neurophysiol. 123, 1067–1087 (2012).
159 Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52, 1059–1069 (2010).
160 Zalesky, A., Fornito, A. & Bullmore, E. T. Network-based statistic: identifying differences in brain networks. Neuroimage 53, 1197–1207 (2010).
161 Van Wijk, B. C. M., Stam, C. J. & Daffertshofer, A. Comparing brain networks of different size and connectivity density using graph theory. PLoS One 5, e13701 (2010).
162 Brazier, M. A. B. & Barlow, J. S. Some applications of correlation analysis to clinical problems in electroencephalography. Electroencephalogr. Clin. Neurophysiol. 8, 325–331 (1956).
163 Adey, W., Elul, R., Walter, R., et al. The cooperative behavior of neuronal populations during sleep anad mental tasks. Electroencephalogr. Clin. Neurophysiol. 23, 88 (1967).
164 Granger, C. W. J. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37, 424 (1969).
165 Hesse, W., Möller, E., Arnold, M., et al. The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies. J. Neurosci. Methods 124, 27–44 (2003).
166 Bressler, S. L. & Seth, A. K. Wiener-Granger causality: a well established methodology. Neuroimage 58, 323–329 (2011).
167 Wang, G. & Takigawa, M. Directed coherence as a measure of interhemispheric correlation of EEG. Int. J. Psychophysiol. 13, 119–128 (1992).
168 Kaminski, M. J. & Blinowska, K. J. A new method of the description of the information flow in the brain structures. Biol. Cybern. 65, 203–210 (1991).
169 Baccalá, L. A. & Sameshima, K. Partial directed coherence: a new concept in neural structure determination. Biol. Cybern. 84, 463–474 (2001).
170 Fraser, A. M. & Swinney, H. L. Independent coordinates for strange attractors from mutual information. Phys. Rev. A 33, 1134–1141 (1986).
171 Stam, C. J. & van Dijk, B. W. Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets. Phys. D Nonlinear Phenom. 163, 236–251 (2002).
172 Vinck, M., Oostenveld, R., van Wingerden, M., et al. An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias. Neuroimage 55, 1548–1565 (2011).

173 Stam, C. J. & van Straaten, E. C. W. Go with the flow: use of a directed phase lag index (dPLI) to characterize patterns of phase relations in a large-scale model of brain dynamics. Neuroimage 62, 1415–1428 (2012).
174 Stam, C. J. The trees and the forest: characterization of complex brain networks with minimum spanning trees. Int. J. Psychophysiol. 92, 129-138 (2014).
175 Van Mieghem, P. & van Langen, S. Influence of the link weight structure on the shortest path. Phys. Rev. E 71, 056113 (2005).
176 Van Mieghem, P. & Magdalena, S. M. Phase transition in the link weight structure of networks. Phys. Rev. E 72, 056138 (2005).
177 Wang, H., Hernandez, J. M. & Van Mieghem, P. Betweenness centrality in a weighted network. Phys. Rev. E 77, 046105 (2008).
178 Ortega, G. J., Sola, R. G. & Pastor, J. Complex network analysis of human ECoG data. Neurosci. Lett. 447, 129–133 (2008).
179 Lee, U., Oh, G., Kim, S., et al. Brain networks maintain a scale-free organization across consciousness, anesthesia and recovery. Anesthesiology 113, 1081–1091 (2010).
180 Schoen, W., Chang, J. S., Lee, U., et al. The temporal organization of functional brain connectivity is abnormal in schizophrenia but does not correlate with symptomatology. Conscious. Cogn. 20, 1050–1054 (2011).
181 Boersma, M., Smit, D. J. A., Boomsma, D. I., et al. Growing trees in child brains: graph theoretical analysis of electroencephalography-derived minimum spanning tree in 5- and 7-year-old children reflects brain maturation. Brain Connect. 3, 50–60 (2013).
182 Tewarie, P., Hillebrand, A., Schoonheim, M. M., et al. Functional brain network analysis using minimum spanning trees in Multiple Sclerosis: An MEG source-space study. Neuroimage 88, 308–318 (2014).
183 Van Dellen, E., Douw, L., Hillebrand, A., et al. Epilepsy surgery outcome and functional network alterations in longitudinal MEG: a minimum spanning tree analysis. Neuroimage 86, 354–363 (2014).
184 Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. Nature 393, 440–442 (1998).
185 Newman, M. E. J. Modularity and community structure in networks. (2006).
186 Van den Heuvel, M. P. & Sporns, O. Network hubs in the human brain. Trends Cogn. Sci. 17, 683–696 (2013).
187 Batool, K. & Niazi, M. A. Towards a methodology for validation of centrality measures in complex networks. PLoS One 9, e90283 (2014).
188 Lohmann, G., Margulies, D. S., Horstmann, A., et al. Eigenvector centrality mapping for analyzing connectivity patterns in fMRI data of the human brain. PLoS One 5, e10232 (2010).
189 Joyce, K. E., Laurienti, P. J., Burdette, J. H., et al. A new measure of centrality for brain networks. PLoS One 5, e12200 (2010).
190 Ruhnau, B. Eigenvector-centrality — a node-centrality? Soc. Networks 22, 357–365 (2000).
191 Stam, C. J., Nolte, G. & Daffertshofer, A. Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. Hum. Brain Mapp. 28, 1178–1193 (2007).
192 Li, C., Wang, H., de Haan, W., et al. The correlation of metrics in complex networks with applications in functional brain networks. J. Stat. Mech. Theory Exp. 2011, P11018 (2011).
193 Gross, J., Baillet, S., Barnes, G. R., et al. Good practice for conducting and reporting MEG research. Neuroimage 65, 349–363 (2013).





3

FUNCTIONAL CONNECTIVITY AND NETWORK ANALYSIS DURING HYPOACTIVE DELIRIUM AND RECOVERY FROM ANESTHESIA

T. Numan
A.J.C. Slooter
A.W. van der Kooi
A.M.L. Hoekman
W.J.L. Suyker
C.J. Stam
E. van Dellen

Clinical Neurophysiology 2017; 128(6): 914-924

ABSTRACT

OBJECTIVE: To gain insight in the underlying mechanism of reduced levels of consciousness due to hypoactive delirium versus recovery from anesthesia, we studied functional connectivity and network topology using electroencephalography (EEG).

METHODS: EEG recordings were performed in age and sex-matched patients with hypoactive delirium (n=18), patients recovering from anesthesia (n=20), and non-delirious control patients (n=20), all after cardiac surgery. Functional and directed connectivity were studied with phase lag index and directed phase transfer entropy. Network topology was characterized using the minimum spanning tree (MST). A random forest classifier was calculated based on all measures to obtain discriminative ability between the three groups.

RESULTS: Non-delirious control subjects showed a back-to-front information flow, which was lost during hypoactive delirium ($p=0.01$) and recovery from anesthesia ($p<0.01$). The recovery from anesthesia group had a more integrated network in the delta band compared to non-delirious controls. In contrast, hypoactive delirium showed a less integrated network in the alpha band. High accuracy for discrimination between hypoactive delirious patients and controls (86%) and recovery from anesthesia and controls (85%) were found. Accuracy for discrimination between hypoactive delirium and recovery from anesthesia was 73%.

CONCLUSION: Loss of functional and directed connectivity were observed in both hypoactive delirium and recovery from anesthesia, which might be related to the reduced level of consciousness in both states. These states could be distinguished in topology, which was a less integrated network during hypoactive delirium.

SIGNIFICANCE: Functional and directed connectivity are similarly disturbed during a reduced level of consciousness due to hypoactive delirium and sedatives, however topology was differently affected.

INTRODUCTION

Although disturbed consciousness is common in intensive care unit (ICU) patients[1–3], understanding of different mechanisms leading to reduced levels of consciousness is limited. Consciousness is a complex state of wakefulness and awareness. (Disturbance of) consciousness has been studied with various techniques, and there is compelling evidence that thalamocoritcal interactions and reduced (cortical) activity are signatures of reduced consciousness[4,5]. Disturbance of consciousness may therefore be the result of disruption of interactions between brain areas[6]. Several methods can be used to characterize these interactions: the (undirected) functional connectivity strength reflects the statistical interdependence of activity of remote brain areas. Directed connectivity could provide additional insight in patterns of the direction of information flow. Finally, the global organization of functional brain networks could be characterized using graph theory.

Two important causes for disturbed consciousness in the ICU are the use of sedatives and hypoactive delirium, which is an acute disturbance of awareness, attention, and cognition that fluctuates within hours due to acute brain dysfunction[7]. Whereas the clinical expression of hypoactive delirium and sedation is similar[8], the underlying cause of the disturbance of consciousness is different. Therefore, connectivity patterns might be different as well. Reduced functional connectivity[9,10] was found during both delirium and recovery from anesthesia in the alpha frequency band. The disturbance of directed connectivity band seemed similar between delirium and loss of consciousness, although in different frequency bands, being the delta and alpha frequency band, respectively. Moreover, during recovery of consciousness the directed connectivity in the alpha frequency band showed an intermediate state between feedforward and feedback patterns[10]. This recovery of consciousness period is more similar to the level of consciousness in hypoactive delirious patients. However, connectivity alterations during delirium and recovery from anesthesia have not been compared directly. Delirium was studied in patients with advanced age after cardiac surgery[9], while propofol-induced loss of consciousness was studied in healthy young subjects[10]. Moreover, recent insights in methodology of EEG connectivity analysis suggest that the interpretation of these previous studies may be problematic. The directed phase lag index (dPLI) was used to characterize directed connectivity in these studies, which measures the consistency of the phase difference of two EEG time series. This phase difference is dependent on the EEG reference and cannot be interpreted unequivocally in terms of the direction

of information flow[11]. The phase transfer entropy (PTE), an information theory-based measure of directed connectivity, has been introduced to solve these problems[11–13].

Both delirium and sedatives also appear to cause changes in functional brain network organization of alpha band (8-13Hz) oscillations[9,10,14]. A more random network topology was found during these states of reduced level of consciousness, which might reflect an inability to properly integrate information[15]. However, the interpretation of these findings is problematic due to methodological issues[16], and altered network topology may merely reflect a general decrease in connectivity strength[17]. The minimum spanning tree (MST) has been developed as an unbiased method to characterize network topology across conditions and groups. The MST can be seen as the backbone structure of the network, containing only the strongest connections without forming loops. Based on the MST, network measures can be calculated, which can be interpreted along the lines of conventional graph measures, while characteristics remain unaffected by differences in connectivity strength (Figure 3.1)[18].

The aim of this study was to compare functional and directional connectivity as well as network topology between the episodes of hypoactive delirium and recovery from anesthesia, and non-delirious controls. Based on previous research we hypothesized that

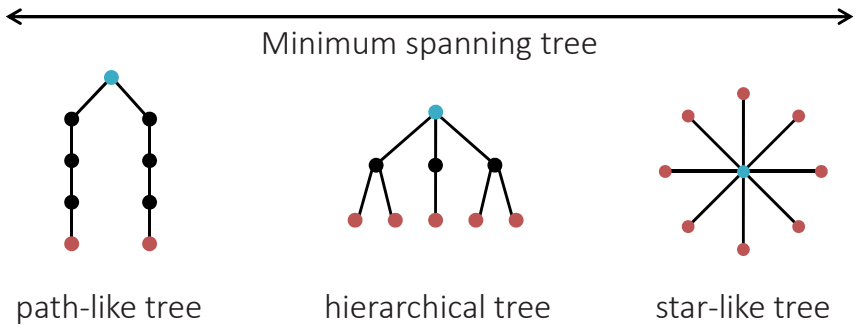


Figure 3.1. Schematic presentation of the MST.
MST structures can range between a path-like tree (i.e., less integrated network) to a star-like tree (i.e., more integrated network). A path-like tree has two leaf nodes (red), all other nodes are connected to their neighbors, which has the disadvantage of being inefficient. A path-like tree has a low leaf number and a high diameter. The other end of the spectrum is a star-like tree, which has one central node (blue) and all other nodes are leaf nodes. It thus has a high leaf number and a small diameter. Information can spread easily across the network, but the central node in the star-like tree might suffer from overloading of information. A hierarchical tree is a hypothesized optimal topology, combining a relatively low diameter (which indicates high efficiency) while the relatively low maximum number of connections per node or maximum degree prevents overload of central hub regions. This figure was made based on Figure 2 in van Dellen et al. (2014)[19].

functional connectivity is reduced during reduced levels of consciousness, due to both hypoactive delirium and recovery from anesthesia[9,10]. However, we expected that the direction of information flow and the underlying brain network topology are different depending on the underlying cause of reduced levels of consciousness.

METHODS

STUDY POPULATION AND PROCEDURES

The study population consisted of three groups of cardiac surgery patients. Two groups were described before[9]. From this study, we included all patients with hypoactive delirium and age- and sex matched non-delirious control patients. In summary, patients were preoperatively included after obtaining informed consent for participation. During the first five consecutive days after cardiac surgery, patients were screened for delirium and examined by a geriatrician, neurologist or psychiatrist who made the classification delirious or non-delirious according to the Diagnostic and Statistical Manual of Mental Disorders DSM-IV-R criteria[20]. The Richmond agitation-sedation scale (RASS)[21] was used to determine the type of delirium, with negative RASS scores indicating hypoactive delirium. This study was approved by the local ethics committee (protocol number 11-073).

A third group of patients was included to study recovery from anesthesia, and EEG registrations were made at the ICU directly after cardiac surgery. These patients were age- and sex matched with the hypoactive delirious and non-delirious control patients on group level. Exclusion criteria for the recovery from anesthesia group were cerebral damage prior to surgery based on chart review, postoperative signs of delirium, hallucinations without delirium, transient ischemic attack, stroke, or any other type of perioperative cerebral damage. The local medical research ethics committee waived the need for informed consent for EEG recordings in the recovery from anesthesia group (protocol number 13-011), although written approval was obtained preoperatively.

There was a time delay of approximately two years between the recording of the recovery from anesthesia group and the hypoactive delirium and non-delirious control groups, however, since patient population and inclusion, recording setup, and (pre-) processing steps were similar, we do not expect that this time delay will affect the results.

Anesthesia protocols in all patients were as follows: no premedication was used, and anesthesia was initiated using 0.1 mg/kg midazolam and 1 µg/kg sufentanil. Pancuronium 0.1 mg/kg was used to facilitate orotracheal intubation. Inhalation anesthesia with sevoflurane or isoflurane was applied together with 0.5 µg/kg/h sufentanil i.v. for maintenance.

CLINICAL ASSESSMENT AND DATA COLLECTION

During EEG recordings, the RASS score was evaluated to obtain the level of consciousness. For the patients in the recovery from anesthesia group, the recording was completed after the patient reached the level of RASS -2 (eye contact in response to voice for less than ten seconds) or RASS -1 (eye contact in response to voice for more than ten seconds without being fully alert), and maintained this level of consciousness for 10 minutes. The two following days, researchers performed a chart review, used the confusion assessment method for ICU patients (CAM-ICU)[1], and asked patients for the presence of hallucinations. Age, sex, type of cardiac surgery, acute physiology and chronic health evaluation (APACHE) IV score, duration of surgery and bypass-time were collected for all included patients.

EEG RECORDINGS, SELECTION AND PREPROCESSING

Twenty-one electrodes were applied immediately after ICU admission, according to the 10-20 system (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2, A1, A2) for the EEG registration. The registration was performed with a Micromed EEG system (Trevisio, Italy), using a sample frequency of 512 Hz and analogue filter settings of 0.15 Hz high pass and 70 Hz low pass filter.

Epoch selection was performed using Micromed. Time series of electrodes A1, A2, Fp1 and Fp2 were excluded because of eye movement and muscle artifacts. The first four artifact-free epochs of eight seconds with eyes closed[22] were selected within three minutes after a RASS score of -2 or -1 was obtained in the recovery from anesthesia group. In the hypoactive delirium and non-delirious control groups, the first four artifact-free epochs during the eyes closed condition were selected[9,22].

SPECTRAL DENSITY

Power spectral density analysis was performed in MATLAB R2014b (The MathWorks Inc., Natick, MA, 2000) to obtain an impression of the EEG data based on standard frequency analysis. All epochs were high pass (0.5 Hz) and low pass (50 Hz) filtered by a second

order Butterworth filter. The power spectral densities per epoch and per electrode were obtained by the Welch method, using a window length of 4 seconds with 50% overlap. After normalization, by the sum of power between 0.5 and 20 Hz, all power spectral densities were averaged and plotted for the three groups.

FUNCTIONAL CONNECTIVITY

Data was re-referenced towards a source reference, as approximation of the Laplacian[23], and filtered in the four frequency bands, delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), and beta (13-20 Hz) for functional and directed connectivity analyses. Frequencies above 20 Hz were excluded to avoid artifacts due to muscle contractions[24].

The phase lag index (PLI) was used to calculate the functional connectivity strength[25]. The PLI is based on the Hilbert transformed instantaneous phase differences (Δϕ), capturing the asymmetry in phase leading and lagging between two signals:

$$PLI = |\langle sign[\sin(\Delta\varphi(t))]\rangle|$$

The sign[Δϕ(t)] is 1 for all positive phase differences and -1 for all negative phase differences, which was averaged over an epoch. PLI values range between 0 and 1; 1 indicates complete phase-locking and values reaching 0 mean no phase synchronization or equal in leading and lagging over the epoch. An average over all PLI values between all channel pairs was calculated and compared between the three groups.

DIRECTED CONNECTIVITY

The direction of information flow was estimated with the phase transfer entropy (PTE)[11–13]. Transfer entropy is a measure of the direction of causal influence between two signals. A causal relation between signal X and signal Y is assumed, when the prediction of the future of signal X is improved by adding not only the information of the past of signal X, but also the past of signal Y. The phase transfer entropy estimates the strength of causality between two signals based on the instantaneous phase differences, calculated with the Hilbert transform. This measure is relatively robust to noise and linear mixing and is independent of the selected reference[11,13].

To normalize the PTE, the directed PTE (dPTE) is calculated, since there is no meaningful upper bound of the PTE:

$$dPTE_{x\to y} = \frac{PTE_{x\to y}}{PTE_{x\to y} + PTE_{y\to x}}$$

The dPTE ranges between 0 and 1. $dPTE_{xy}$ values above 0.5 and below 0.5 indicate an information flow preferentially from X to Y and from Y to X, respectively. There is no preferential information flow when the dPTE value is 0.5.

An anterior-posterior (AP) ratio was calculated to evaluate the presence of a consistent direction of information flow, which has been reported in previous work[9,10,26,27]. The anterior dPTE values (channels F7, F3, Fz, F4 and F8) were averaged and divided by the averaged posterior dPTE values (channels T5, P3, Pz, P4, T6, O1 and O2).

MST ANALYSIS

To evaluate the organization of the functional networks of the three groups, MST networks were calculated for all frequency bands, based on all connectivity matrices for each epoch using Kruskal's algorithm[28]. The PLI values of the connectivity matrix were ranked starting with the highest PLI value, which was included as the first MST connection. Then the second highest PLI value was added as an MST connection, and so on. If adding a connection would result in a loop or triangle, this connection was discarded and the following PLI value was evaluated. This procedure was continued until all nodes (i.e., EEG channels) were connected. In this study 17 EEG channels were included for analyses, resulting in a MST network containing 17 nodes with 16 edges for each epoch. The connectivity strength within the MST was calculated by averaging PLI values of included connections in the MST.

Using MATLAB a MST-reference was calculated for each frequency band to test whether the MSTs between the three groups were different[29,30]. To obtain this reference MST, the functional connectivity matrices of the controls subjects were averaged. Based on the averaged PLI matrix, the MST-reference was calculated. The number of overlapping MST edges in each epoch with the MST-reference was counted for all subjects and for all four frequency bands. MST overlap was defined as the number of overlapping edges divided by the total number of edges, ranging between 0 (i.e., no overlapping edges) and 1 (i.e., exactly similar MST) and averaged over four epochs for each patient. Moreover, the functional connectivity within the MST was calculated for all four frequency bands, including only the PLI values of the included edges within the MST. When there was a significant difference between the three groups in MST overlap, MST network measures were calculated and averaged over 4 epochs per patient in that frequency band (Table 3.1). Connectivity and network analyses were performed in BrainWave (Version 0.9.152.2.7 freely available on http://home.kpn.nl/stam7883/brainwave.html).

STATISTICAL ANALYSIS

Patients' characteristics were compared between the three groups. Continuous parameters were tested for normality using the Shapiro-Wilk test. Normally distributed variables were presented as mean and standard deviation (SD) and compared between the three groups using a one-way ANOVA test. Post-hoc tests between groups were performed in case of overall statistically significant findings (i.e., $p<0.05$), using independent samples t-tests. Data that did not follow a Gaussian distribution were presented as median (Interquartile range (IQR); 1st quartile-3rd quartile) and compared using Kruskal Wallis test, while Mann-Whitney U tests were used for post-hoc analyses. Categorical data were compared using the chi-square test or Fisher's exact test, where appropriate.

Table 3.1. Minimum spanning tree measures.

MST measures	Definition – node (local measures)	Definition - MST (global measures)
Degree	The number of edges for each node divided by the maximum number of edges possible (m = n-1).	The highest degree within the MST is given.
Leaf fraction (Lf)	-	A leaf node (L) is a node with only one edge. The leaf fraction is the ratio between the number of leaf nodes divided by the total number of nodes
Diameter (D)	-	The diameter is the longest distance between any two nodes in the MST network.
Betweenness centrality (BC)	The number of paths between all other nodes in the MST crossing the node of interest, divided by the total number of paths in the MST. The central node in a star-like network has a high BC (BC=1), the BC of a leaf node is 0.	The BC of the MST corresponds to the highest BC value in the MST and is therefore called maximum betweenness centrality (BC_{max}).
Eccentricity (Ecc)	The maximum distance (i.e., number of edges) between a node and any other node in the MST. The Ecc will be low in a central node (i.e., all other nodes are 'nearby') and will be higher more peripheral node.	The Ecc of the MST represents a range, which is calculated by subtracting the Ecc of the most peripheral node from Ecc of the most central node.
Tree hierarchy (Th)	-	The Th is given by $Th = Lf/(2mBC_{max})$ and defines the hierarchy of the MST organization, the optimal topology. It considers the leaf fraction Lf, the BC_{max} and the maximum diameter m of the tree (m = n-1). For a path-like MST, Th decreases towards 0. In a star-like MST, Th = 0.5. MST structures between path and star-like can reach Th values up to 1.

PLI, AP-ratios, MST overlap and MST measures were compared between the three groups in the four frequency bands using the Kruskal-Wallis test. Level of significance was adjusted using the Bonferroni correction for the four frequency bands, to correct for multiple testing, p-values <0.0125 were assumed to indicate statistical significance (uncorrected p-values are reported throughout the manuscript). Post-hoc Mann-Whitney U tests were performed for the frequency bands showing significant differences for PLI and AP-ratios. When the MST overlap in a frequency band was significantly different between the groups, subsequent tests were performed to explore possible differences in other MST measures. These tests were considered exploratory post-hoc tests, and therefore no correction for multiple testing was performed. Statistical analyses were performed with SPSS, version 21 (SPSS Inc., Chicago, IL).

Finally, we used a random forest classifier[31] to study the included EEG network characteristics (i.e., the 36 calculated measures, see Table S3.7 in supplementary material) in relation to group assignment (non-delirious controls vs. hypoactive delirium, non-delirious controls vs. recovery from anesthesia, and hypoactive delirium vs. recovery from anesthesia)[32]. Random forest is a machine learning method for classification, using multiple decision trees created by bootstrap samples of the data (see Dauwan et al. 2016 for full description of random forest algorithm[33]). The number of trees included in the random forest classifier was set to 500. Moreover, the number of variables calculated at each split of each individual tree was set to 6 (i.e., the square root of 36 variables). However, the set parameters are not highly influencing the classification outcome[34]. The discriminative ability of the resulting classifier based on all features was evaluated based on accuracy (i.e., ratio between correctly classified subjects and total number of subjects), sensitivity, and specificity. Moreover, variable importance (VIMP), a value between 0 and 1, was calculated for each included feature.

RESULTS

CHARACTERISTICS OF THE STUDY POPULATION

We included 18 hypoactive delirious patients and 20 non-delirious patients from our previous study[9]. Further, we included 34 patients in the recovery from anesthesia group. From this group, patients were excluded because of signs of delirium (n=6), hallucinations (n=2), or stroke (n=2) on the two consecutive days after the EEG recording. In four other patients it was not feasible to select artifact free epochs. This resulted in

20 patients recovering from anesthesia who were age and sex matched on group level to the hypoactive delirious patients and non-delirious control patients (Table 3.2). The APACHE IV score, the duration of surgery, and bypass time were significantly higher in the delirium group than in the other two groups.

All non-delirious control patients had a RASS score of 0 during the EEG recording. Patients in the recovery from anesthesia group did not deviate in RASS score compared to the hypoactive delirious group (p=0.19). The course of regaining consciousness in patients recovering from anesthesia was highly variable. The average time between ICU admission and the moment of the selected epochs was 163 minutes (SD=50), with a range of 85-268 minutes (depending on the time needed to regain consciousness, which was highly variable between patients).

Table 3.2. Patient characteristics.

	Non-delirious controls (n=20)	Hypoactive delirium (n=18)	Recovery from anesthesia (n=20)	p-value‡
Age (years)	76 ± 7	76 ± 5	74 ± 7	0.50
Gender (M/F)	11/9	7/11	10/10	0.58
Duration of				
Surgery (min.)	173 (129-208)	217 (179-314)	174 (141-239)	0.03
Bypass (min.)	88 (77-132)	129 (101-180)	88 (66-141)	0.03
Type of surgery (n)				0.06
CABG	3 (15%)	2 (11%)	6 (30%)	
Valve	8 (40%)	5 (28%)	11 (55%)	
CABG + Valve	5 (25%)	9 (50%)	3 (15%)	
Other	4 (20%)	2 (11%)	-	
RASS during EEG				0.19*
0	20 (100%)	-	-	
-1	-	11 (61%)	7 (35%)	
-2	-	7 (39%)	13 (65%)	
APACHE IV score	41 ± 8	55 ± 15	43 ± 7	<0.01
# days between surgery & EEG	3 (2-4)	3 (2-4)	-	0.68
# minutes between ICU admission & epoch selection	-	-	163 ± 50	-

Data is presented as mean ± standard deviation (SD), median (interquartile range), or numbers (%). APACHE = acute physiology and chronic health evaluation, CABG = coronary artery bypass grafting, M/F = male/female, n = number.
*‡The p-values represent the comparison between the three groups.*Fisher's Exact test is performed between hypoactive delirium and recovery from anesthesia.*

SPECTRAL DENSITY

Figure 3.2 shows the mean spectral power density over all 17 EEG channels of the three groups. Both hypoactive delirium and recovery from anesthesia were associated with slowing of background activity as compared to the non-delirious controls. The relative delta power was higher and alpha power was lower in hypoactive delirium and recovery from anesthesia compared to non-delirious controls (Table S3.1). Post-hoc comparison showed a lower alpha power in the hypoactive delirious patients (median, 0.07; IQR, 0.04-0.11) compared to recovery from anesthesia (median, 0.10; IQR, 0.07-0.15, p=0.019).

FUNCTIONAL CONNECTIVITY

The average PLI was significantly lower in hypoactive delirium (median, 0.122; IQR, 0.116-0.130, p<0.01) and during recovery from anesthesia (median, 0.125; IQR, 0.116-0.130, p<0.01) compared to the non-delirious controls (median, 0.136; IQR, 0.122-0.161) (Table 3.3). PLI values in hypoactive delirium did not significantly differ from PLI values during recovery from anesthesia (p=0.98). After correction for multiple testing there were no significant differences between the three groups in the other frequency bands.

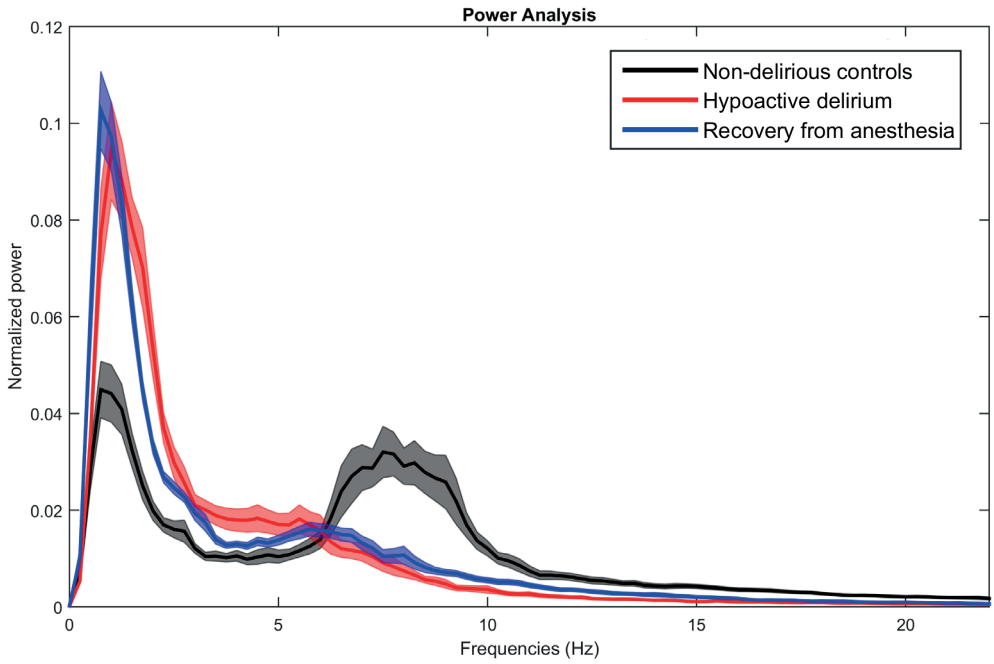


Figure 3.2. Normalized power spectral density of non-delirious controls, hypoactive delirium, and recovering from anesthesia after cardiac surgery.
Thicker line represents the mean power spectrum for all three groups, the filled area represents the standard error of the mean.

Table 3.3. Phase lag index in the four frequency bands.

	Non-delirious controls (n=20)	Hypoactive delirium (n=18)	Recovery from anesthesia (n=20)	p-value
Delta	0.152 (0.143-0.160)	0.157 (0.142-0.172)	0.167 (0.159-0.184)	0.03
Theta	0.144 (0.136-0.160)	0.148 (0.137-0.159)	0.143 (0.132-0.160)	0.87
Alpha	0.136 (0.122-0.161)	0.122 (0.116-0.130)	0.125 (0.116-0.130)	<0.01*
Beta	0.098 (0.094-0.104)	0.095 (0.091-0.103)	0.098 (0.095-0.100)	0.71

*P-values <0.0125 were assumed significant based on the Bonferroni correction for the four frequency bands. *Significant difference between the three groups. Post-hoc Mann Whitney U test showed a significant reduction of alpha PLI in delirium and recovery from anesthesia compared to non-delirious controls (p<0.01 and p<0.01 respectively).*

DIRECTED CONNECTIVITY

In the alpha frequency band, both the hypoactive delirium group and the recovery from anesthesia group showed a loss of posterior-anterior directionality, compared to the observed posterior-anterior connectivity in the non-delirious control patients (Figure 3.3 and 3.4). Post-hoc analyses revealed a significantly higher AP-ratio in delirium patients (median, 1.00; IQR, 0.98-1.00, p=0.01) and recovery from anesthesia patients (median, 1.00; IQR, 0.99-1.01, p<0.01) compared to non-delirious controls (median, 0.98; IQR, 0.96-0.99), indicating a loss of back-to-front pattern. There was no significant difference in AP-ratio between hypoactive delirium and recovery from anesthesia patients (p=0.07).

MST ANALYSIS

Functional connectivity within the MST was significantly reduced in the alpha frequency band in hypoactive delirium (p=0.002) and recovery from anesthesia (p=0.003) compared to the non-delirious control group (Table 3.4). The overlap of the MSTs with the MST-reference was compared between the three groups to study overall differences in network topology. Overall, significant differences of this overlap were observed in the delta, alpha and beta frequency bands (Table 3.4). Post-hoc comparisons showed significantly lower overlap in patients recovering from anesthesia compared to non-delirious controls in the delta (p=0.004), alpha (p=0.013), and beta bands (p=0.006). Secondly, lower overlap was found in the hypoactive delirium group compared to non-delirious controls in the delta (p=0.003), alpha (p<0.001), and beta bands (p=0.028).

Thirdly, lower overlap was observed in patients with hypoactive delirium compared to the recovery from anesthesia group in de alpha frequency band only (p=0.012). See for the visualization of the averaged MST topology Figure S3.3.

Based on these findings, we first compared MST measures between the groups in the frequency bands in which differences were observed (Figure 3.5 and Table S3.2). Patients recovering from anesthesia had, compared to non-delirious controls, an increased delta band degree (p=0.033) and leaf fraction (p=0.028) and decreased diameter (p=0.033) and eccentricity (p=0.010). This indicates a more star-like topology and therefore more integrated network organization during recovery from anesthesia. No differences in MST measures were observed in the alpha and beta bands between these groups (Figure 3.5). Secondly, a comparison of MST measures between hypoactive delirium patients and non-delirious controls showed a decrease in degree (p=0.003), leaf fraction (p=0.033) and maximum betweenness centrality (p=0.016) in the alpha band during hypoactive delirium, which indicates a shift towards a more path-like topology and a less integrated network organization. No group differences were found in MST measures in the delta and beta bands (Figure 3.5).

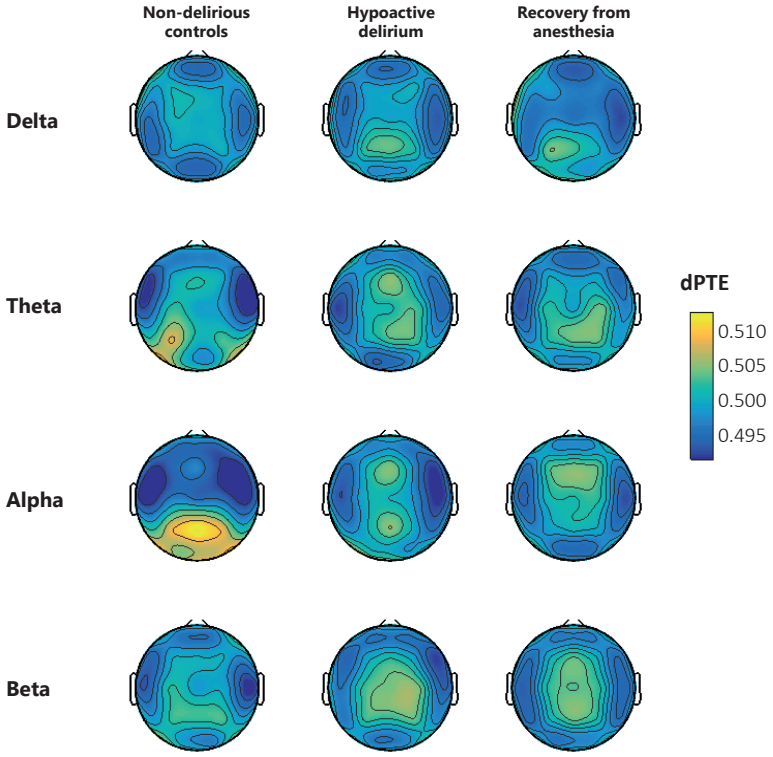


Figure 3.3. Mean directed phase transfer entropy results for all electrodes.
dPTE values above 0.5 indicate and information flow from that electrodes, whereas dPTE values below 0.5 indicate information flow towards that electrode. A1, A2, Fp1, Fp2 were excluded for analysis.

Table 3.4. Functional connectivity strength within MST and Overlap with the reference MST.

	Frequency band	Non-delirious controls (n=20)	Hypoactive delirium (n=18)	Recovery from anesthesia (n=20)	P-value‡
PLI within MST	Delta	0.344 (0.335-0.369)	0.358 (0.331-0.372)	0.373 (0.357-0.401)	0.079
	Theta	0.334 (0.317-0.369)	0.333 (0.315-0.363)	0.321 (0.305-0.356)	0.505
	Alpha	0.320 (0.291-0.358)	0.287 (0.278-0.297)	0.286 (0.269-0.295)	0.002[a,b]
	Beta	0.228 (0.221-0.245)	0.236 (0.216-0.243)	0.231 (0.225-0.236)	0.748
Overlap with MST-reference	Delta	0.156 (0.141-0.184)	0.117 (0.094-0.145)	0.125 (0.094-0.141)	0.003[a,b]
	Theta	0.188 (0.109-0.230)	0.125 (0.109-0.156)	0.141 (0.109-0.172)	0.078
	Alpha	0.188 (0.156-0.219)	0.125 (0.078-0.141)	0.156 (0.125-0.188)	<0.001[a,b,c]
	Beta	0.156 (0.141-0.184)	0.141 (0.109-0.156)	0.125 (0.109-0.152)	0.012[a,b]

MST = minimum spanning tree, PLI = phase lag index. ‡ The p-values represent the comparison between the three groups, [a] significant difference between recovery from anesthesia and controls, [b] significant difference between hypoactive delirium and controls, [c] significant difference between hypoactive delirium and recovery from anesthesia.

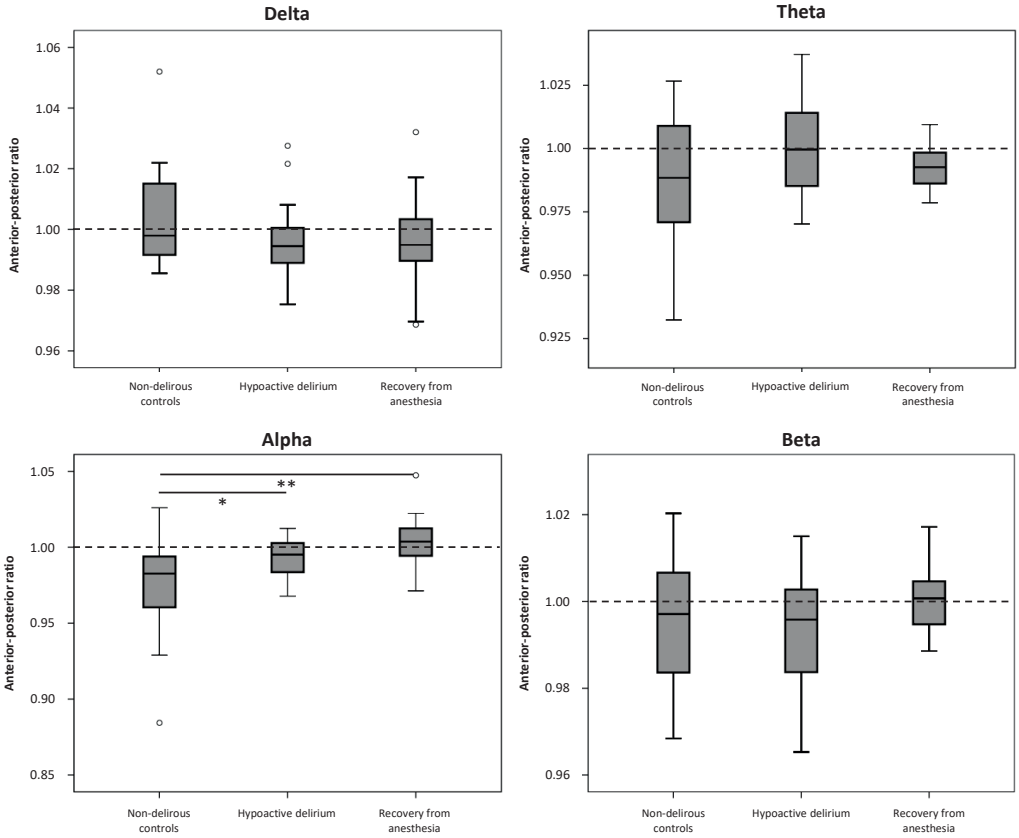


Figure 3.4. Anterior-posterior ratios based on directed phase transfer entropy in the four frequency bands.
*Striped line indicates an anterior-posterior (AP) ratio of 1, indicating no preferential directionality. AP-ratios >1 indicate higher dPTE values in the anterior regions compared to posterior regions and therefore an information flow from anterior towards posterior. AP-ratios <1 suggest a posterior leading of information flow. *p<0.05, **p<0.01.*

Thirdly, we compared hypoactive delirium with recovery from anesthesia with regard to functional network topology in the alpha band. Maximum betweenness centrality was significantly lower in hypoactive delirium compared to recovery from anesthesia (p=0.007). Other MST measures did not differ between the three groups. To correct for any difference in level of consciousness, we performed an additional analysis between the hypoactive delirium and recovery from anesthesia groups with adjustment for the RASS score. In the adjusted analysis, the differences in maximum betweenness centrality remained significant, while differences in degree, leaf fraction, diameter and eccentricity reached statistical significance (see supplementary material additional analysis – correction for RASS score).

Finally, within the recovery from anesthesia group, we compared connectivity measures during unconsciousness (RASS -5 or -4) and regaining consciousness (RASS -2 or -1), aiming to study the effects of the level of consciousness on our outcome measures (PLI, dPTE and MST measures). Data were available in 13 patients. We found higher a higher PLI in the theta band during regaining consciousness (median, 0.142; IQR, 0.134-0.155) compared to unconsciousness (median 0.131; IQR, 0.121-0.136, p=0.046). No differences were found in the directed connectivity. In the delta band, a more integrated network was found during regaining consciousness compared to unconsciousness based on the leaf fraction (p=0.011), diameter (p=0.003), and eccentricity (0.004). In the alpha band, the tree hierarchy was lower during regaining consciousness compared to the state of unconsciousness (p=0.023) (see supplementary material additional analysis – unconsciousness vs. regaining consciousness).

RANDOM FOREST – CONTRIBUTION IN CLASSIFICATION

The results of random forest analysis are shown in supplementary material (additional analysis – random forest classification). Discrimination between hypoactive delirium and controls based on all included features resulted in an accuracy of 86%, sensitivity and specificity were 77% and 95%, respectively. Relative delta, alpha, and beta power were found to have the highest VIMP values.

The classifier for discrimination between recovery from anesthesia and controls had an accuracy, sensitivity, and specificity of 85%. The relative delta power was most essential, based on highest VIMP values.

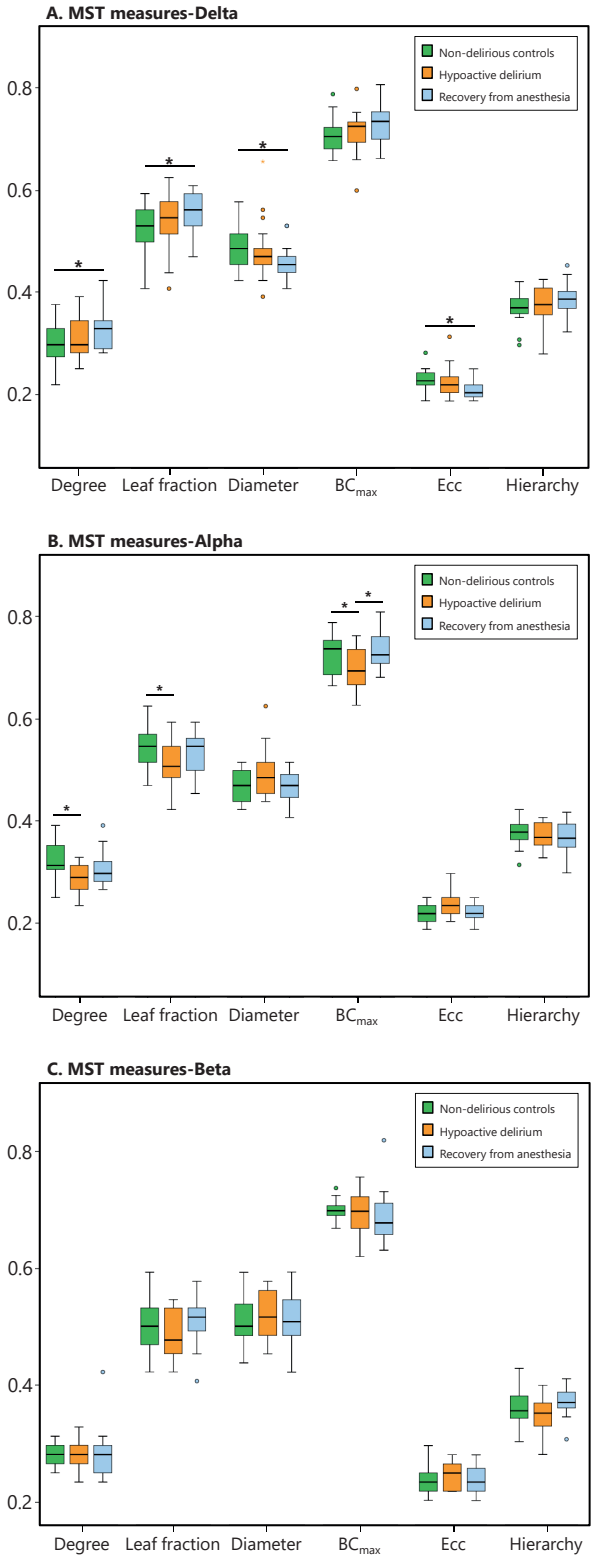


Figure 3.5. MST measures in non-delirious controls, hypoactive delirium and recovery from anesthesia patients in the delta (A), alpha (B) and beta (C) frequency bands. *In the delta frequency band (A) significant differences were found between recovery from anesthesia and non-delirious controls in degree, leaf fraction, diameter, and eccentricity, indicating a more star-like topology in recovery from anesthesia (Figure 3.1). In the alpha frequency band (B) hypoactive delirium had lower degree, leaf fraction and maximum betweenness centrality compared to non-delirious controls, indicating a more path-like topology. The maximum betweenness centrality in the alpha frequency band (B) was significantly lower in the hypoactive delirium group compared to recovery from anesthesia. No differences were found in the beta band (C). The circles with corresponding colors indicate outliers of the boxplot. BC_{max} = maximum betweenness centrality, Ecc = eccentricity. *With horizontal line denotes significant difference between groups.*

Finally, the classifier for distinction between hypoactive delirium and recovery from anesthesia showed an accuracy of 73%, sensitivity of 78% and specificity of 68% using all features. Features with highest VIMP values were i) maximum betweenness centrality in the alpha frequency band and ii) relative theta power, and iii) eccentricity in the alpha frequency band. As previously shown maximum betweenness centrality was significantly lower in hypoactive delirium compared to recovery from anesthesia. Additionally, evaluation of betweenness centrality in the alpha frequency band as single classifier had an accuracy of 71%.

DISCUSSION

This study shows a similar disturbance of functional and directed connectivity during hypoactive delirium and recovery from anesthesia compared to non-delirious controls, while network organization appeared to differ between the three groups. We found a reduction of functional connectivity and a disturbance of the back-to-front directed connectivity pattern in the alpha frequency band during both hypoactive delirium and recovery from anesthesia. Minimum spanning tree analysis showed a more star-like organization in the delta frequency band in patients recovering from anesthesia compared to non-delirious control subjects, which is assumed to be more efficient (i.e., faster) and enables integration of information from a graph theoretical perspective. In contrast, patients with hypoactive delirium showed a shift towards a more path-like MST in the alpha band compared to non-delirious controls and recovery from anesthesia, which indicates a less integrated network (i.e., less efficient information exchange) and loss of functional hubs. Patients with hypoactive delirium therefore had a distinct deviation of their network topology compared to recovering from anesthesia, despite a similar clinical expression including a similar level of consciousness.

Consciousness and disturbed consciousness are broadly studied phenomena. Previous work has considered consciousness to be the ability of processing information[35]. Interactions between brain areas, especially the thalamocortical coupling and frontal-parietal networks are thought to be crucial for consciousness information processing[4,35,36]. Focusing on loss of consciousness after sedation, the thalamus was suggested to work as a conscious switch, since anesthetics are related with reduced thalamic blood flow at loss of consciousness[37]. In the current study, using scalp EEG, we were not able to assess the thalamocortical interactions. However, our finding of

reduced functional connectivity and reduced information processing are in line with the previous findings by Lee et al.[10]

Reduced functional connectivity was previously reported both in delirium[9] and in loss of consciousness by propofol[10,38,39]. It has been suggested that a delirious episode results in cognitive deficits due to this reduction of functional interactions between brain regions[9] and that administration of anesthetics results in cognitive unbinding[40]. Reduced functional connectivity was also found during a minimally conscious state, a vegetative state/unresponsive state[41], and in dementia[30,42,43]. Therefore, the reduced alpha band functional connectivity that we observed in both delirium and in recovery from anesthesia may be a sensitive but non-specific correlate of altered cognitive status.

Directionality of connectivity has been investigated before, both during delirium and loss of consciousness due to propofol[9,10]. However, these previous studies used the dPLI. The dPLI is a reliable measure to characterize a systematic phase difference and is sensitive to affected brain networks, but it is not possible to interpret results unambiguously as directed causal influences. Moreover, the dPLI depends on the EEG reference. The dPTE overcomes these problems and is therefore considered a more appropriate measure for the direction of causal influence[11]. In the current study, the dPTE scalp plots of hypoactive delirium and recovery from anesthesia showed loss of the parietal-to-frontal orientation in the alpha frequency band, which was clearly pronounced in the non-delirious control patients. This loss of directionality was similar to previous studies evaluating delirium[9] and propofol[10], however dPLI showed a frontal-to-parietal pattern in the non-delirious and conscious state.

Neuroimaging studies in patients with a disturbance of consciousness, other than induced by anesthesia or delirium, showed a reduced activity in the default mode network[44]. The frontoparietal network is assumed to be essential for consciousness[45–47]. Disturbance of consciousness may therefore result from reduced information transfer between areas in the default mode network and the frontoparietal network[14,48]. Our findings suggest that this information transfer is normally directed from parieto-occipital areas to frontal areas, while this direction is lost during hypoactive delirium and recovery from anesthesia compared to non-delirious controls. However, it remains to be investigated whether other disorders of consciousness are also associated with dPTE changes.

During recovery from anesthesia we found a more integrated brain network topology compared to non-delirious control patients, which is assumed to be more efficient and

facilitates integration of information. As patients recovering from anesthesia had reduced level of consciousness, we did not expect to find a more integrated network. A speculative explanation is that the brain network topology is optimized to regain consciousness during recovery from anesthesia. In an additional comparison in the recovery from anesthesia patients, we found also a more integrated network in the delta band during regaining of consciousness (RASS -2/-1) compared to unconsciousness (RASS -5/-4). Interestingly, this more integrated network topology in recovery of consciousness compared to unconsciousness was found in the delta frequency band. Increased delta band oscillatory brain activity is associated with reduced level of consciousness and found in slow-wave phase of sleep, sedation[9,49], and during delirium[50,51]. Our findings suggest that during recovery from anesthesia, the functional network topology in the delta frequency band is optimized towards a more star-like (i.e., more integrated) and therefore more efficient network.

Hypoactive delirium is characterized by reduced level of consciousness with additional cognitive problems, in particular inattention. In this study, hypoactive delirious patients showed a decrease in degree, leaf fraction, and maximum betweenness centrality in the alpha band compared to non-delirious controls (i.e., a more path-like network), which indicates a less integrated and less efficient organization of the functional network. In an EEG study with dementia with Lewy bodies patients, a similar shift towards a path-like MST structure was found in the alpha frequency band[30]. Dementia with Lewy bodies is characterized by dementia with additional fluctuations of cognition and recurrent delirium[52,53]. Dementia with Lewy bodies and delirium share key-symptoms and thus seem to have a similar disturbance of the functional brain networks. Interestingly, in patients with Alzheimer's disease, MST measures also indicated a less efficient network because of a reduced leaf fraction[54], which is similar to the results in delirium in this study. The common feature in dementia with Lewy bodies, Alzheimer's disease, and delirium is the presence of cognitive deficits, and therefore a shift of MST topology towards a less efficient network might underlie cognitive problems in these three groups. Moreover, delirium is associated with long-term cognitive decline and dementia predisposes to delirium[55,56]. Functional connectivity and network analysis may provide further insight in the relation between an episode of delirium and long-term cognitive decline.

This is the first study in which a direct comparison could be made between the functional network characteristics of hypoactive delirium and recovery from anesthesia. The maximum betweenness centrality in the alpha frequency band was significantly higher during recovery from anesthesia compared to hypoactive delirium. This finding was

supported by the additional analysis in which adjustments were made for RASS scores, showing stronger differences in MST topology between hypoactive delirium and recovery from anesthesia. Therefore, differences in network topology seem to be related to the cause of reduced level of consciousness and not to the level of consciousness. We performed our study using EEG recordings obtained in clinical setting, which is especially time consuming and complicated in delirious patients. We were therefore unable to measure the complete trajectory of recovery from anesthesia and hypoactive delirium towards total recovery, which is an interesting topic of future research.

The discriminative accuracy of the EEG features was high, with high importance values for relative delta, alpha, and beta power to distinguish between hypoactive delirium and controls. Where these results were based on relative powers over all EEG channels, a previous approach showed that it is possible to detect delirium based on a bipolar derivation using relative delta power[51], which is more feasible in clinical practice compared to full EEG in all patients at risk for delirium. However, distinction between hypoactive delirious patients and patients who are sedated is mainly based on betweenness centrality in the alpha frequency band. Therefore, functional network measures contribute to the distinction between hypoactive delirium and recovery from anesthesia. The results of random classifier algorithm give an indication of which parameters are important, however, the accuracy levels should be interpreted with caution for other populations, because of the relative small sample size. Moreover, the randomization procedure, which is part of the classification algorithm, may lead to slightly different results in a replication experiment.

Recovery from anesthesia was used as a model for light sedation, which cannot be distinguished clinically from hypoactive delirium[8]. Light sedation in ICU patients is difficult to study in isolation, as these patients could have concomitant delirium as well. However, during recovery from anesthesia after cardiac surgery, other factors may affect the level of consciousness including inflammation, cerebral edema, and micro-emboli[57]. To increase our understanding of underlying network changes, we compared non-delirious patients recovering from anesthesia with non-sedated patients with hypoactive delirium. Our findings suggest that these conditions can be distinguished with EEG, but we acknowledge that it is not practical to use full EEG in all patients in a daily clinical setting[51]. The interaction between altered network topology across different frequency bands in hypoactive delirium and recovery from anesthesia remains unclear. In future research, a multilayer network approach, analyzing the different frequency bands as layers, might reveal a complete overview of the global brain networks including all frequency bands[58].

A strength of this study compared to previous EEG network studies in delirium and reduced levels of consciousness is the use of reliable measures, such as the dPTE to evaluate directed connectivity and the MST to explore the topology of the functional networks. The MST excludes spurious influences of connectivity strength, which might be a problem in conventional network measures, and provides a unique sub-network with the most important connections of the connectivity matrix without arbitrary choices[16,18]. Another strength is the direct comparison of delirium with recovery from anesthesia within the same homogeneous population (i.e., elderly patients after cardiac surgery). Moreover, the clinical assessment of the delirious patients was performed by experts, and classification of the level of consciousness was based on the RASS score, a frequently used, validated assessment tool[21]. A limitation is that the sample size of our study population was relatively small, partly because of the difficulty to record artifact free data in patients who have difficulty in following instructions. Our results should therefore be validated in a different, preferably larger population. Additionally, the recordings were performed with a limited number of electrodes; high density EEG might enable more specific and local network change. Finally, the EEG systems were slightly different between the recovery from anesthesia group and the delirious and non-delirious control group, i.e., an online vs. an off-line Micromed EEG system, although the hardware settings were similar. These differences in recording system might have a subtle influence on the EEG signals. However, it is unlikely that our observations rely on these minor differences in recording system, and the recording software was the same.

CONCLUSION

In conclusion, a similar loss of functional connectivity and disturbance of the back-to-front pattern of information flow was found in hypoactive delirium and recovery from anesthesia compared to non-delirious controls, which might be related to the reduced state of consciousness. However, the organization of the functional brain networks was different, with hypoactive delirious patients having a less integrated network with a less central role for hubs. During recovery from anesthesia a more integrated organization of the network was found. A more integrated network might reflect the process of rapidly regaining consciousness during recovery from anesthesia, while a loss of efficiency may reflect the inability to do so during hypoactive delirium.

REFERENCES

1 Ely, E. W., Margolin, R., Francis, J., et al. Evaluation of delirium in critically ill patients: validation of the confusion assessment method for the intensive care unit (CAM-ICU). Crit. Care Med. 29, 1370–1379 (2001).

2 Eijk, M. M. van, Boogaard, M. van den, Marum, R. J. van, et al.Routine use of the confusion assessment method for the intensive care unit: a multicenter study. Am. J. Respir. Crit. Care Med. 184, 340–344 (2011).

3 Barr, J., Fraser, G. L., Puntillo, K., et al. Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit. Crit. Care Med. 41, 263–306 (2013).

4 Alkire, M. T., Haier, R. J. & Fallon, J. H. Toward a unified theory of narcosis: brain imaging evidence for a thalamocortical switch as the neurophysiologic basis of anesthetic-induced unconsciousness. Conscious. Cogn. 9, 370–386 (2000).

5 Schröter, M. S., Spoormaker, V. I., Schorer, A., et al. Spatiotemporal reconfiguration of large-scale brain functional networks during propofol-induced loss of consciousness. J. Neurosci. 32, 12832–12840 (2012).

6 Moon, J.-Y., Lee, U., Blain-Moraes, S., et al. General relationship of global topology, local dynamics, and directionality in large-scale brain networks. PLOS Comput. Biol. 11, e1004225 (2015).

7 American Psychiatric Association. Diagnostics and statistical manual of mental disorders. 5th edition. (2013).

8 Zaal, I. J. & Slooter, A. J. C. Light levels of sedation and DSM-5 criteria for delirium. Intensive Care Med. 40, 300–301 (2014).

9 Van Dellen, E., van der Kooi, A. W., Numan, T., et al. Decreased functional connectivity and disturbed directionality of information flow in the electroencephalography of intensive care unit patients with delirium after cardiac surgery. Anesthesiology 121, 328–335 (2014).

10 Lee, H., Mashour, G. A., Noh, G.-J., et al. Reconfiguration of network hub structure after propofol-induced unconsciousness. Anesthesiology 119, 1347–1359 (2013).

11 Hillebrand, A., Tewarie, P., van Dellen, E., et al. Direction of information flow in large-scale resting-state networks is frequency-dependent. Proc. Natl. Acad. Sci. 113, 3867–3872 (2016).

12 Paluš, M. & Vejmelka, M. Directionality of coupling from bivariate time series: How to avoid false causalities and missed connections. Phys. Rev. E 75, 056211 (2007).

13 Lobier, M., Siebenhühner, F., Palva, S., et al. Phase transfer entropy: a novel phase-based measure for directed connectivity in networks coupled by oscillatory interactions. Neuroimage 85 Part 2, 853–872 (2014).

14 Di Perri, C., Stender, J., Laureys, S., et al. Functional neuroanatomy of disorders of consciousness. Epilepsy Behav. 30, 28–32 (2014).

15 Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10, 186–198 (2009).

16 Van Wijk, B. C. M., Stam, C. J. & Daffertshofer, A. Comparing brain networks of different size and connectivity density using graph theory. PLoS One 5, e13701 (2010).

17 Numan, T., Stam, C. J., Slooter, A. J. C., et al. Being conscious of methodological pitfalls in functional brain network analysis. Anesthesiology 123, 484–485 (2015).

18 Tewarie, P., van Dellen, E., Hillebrand, A., et al. The minimum spanning tree: an unbiased method for brain network analysis. Neuroimage 104, 177–188 (2015).

19 Van Dellen, E., Douw, L., Hillebrand, A., et al. Epilepsy surgery outcome and functional network alterations in longitudinal MEG: a minimum spanning tree analysis. Neuroimage 86, 354–363 (2014).

20 American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 135–147 (2000).
21 Sessler, C. N., Gosnell, M. S., Grap, M. J., et al. The Richmond agitation–sedation scale. Am. J. Respir. Crit. Care Med. 166, 1338–1344 (2002).
22 Van Diessen, E., Numan, T., van Dellen, E., et al. Opportunities and methodological challenges in EEG and MEG resting state functional brain network research. Clin. Neurophysiol. 126, 1468–1481 (2015).
23 Hjorth, B. An on-line transformation of EEG scalp potentials into orthogonal source derivations. Electroencephalogr. Clin. Neurophysiol. 39, 526–530 (1975).
24 Whitham, E. M., Pope, K. J., Fitzgibbon, S. P., et al. Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. Clin. Neurophysiol. 118, 1877–1888 (2007).
25 Stam, C. J., Nolte, G. & Daffertshofer, A. Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. Hum. Brain Mapp. 28, 1178–1193 (2007).
26 Nolte, G., Ziehe, A., Nikulin, V., et al. Robustly estimating the flow direction of information in complex physical systems. Phys. Rev. Lett. 100, 234101 (2008).
27 Stam, C. J. & van Straaten, E. C. W. Go with the flow: use of a directed phase lag index (dPLI) to characterize patterns of phase relations in a large-scale model of brain dynamics. Neuroimage 62, 1415–1428 (2012).
28 Kruskal, J. B. On the shortest spanning subtree of a graph and the traveling salesman problem. Am. Math. Soiety 7, 48–50 (1956).
29 Van Dellen, E., Bohlken, M. M., Draaisma, L., et al. Structural brain network disturbances in the psychosis spectrum. Schizophr. Bull. 42, 782–789 (2016).
30 Van Dellen, E., de Waal, H., van der Flier, W. M., et al. Loss of EEG network efficiency is related to cognitive impairment in dementia with Lewy bodies. Mov. Disord. 30, 1785–1793 (2015).
31 Breiman, L. Random forest. Mach. Learn. 45, 5–32 (2001).
32 Van Diessen, E., Otte, W. M., Braun, K. P. J., et al. Improved diagnosis in children with partial epilepsy using a multivariable prediction model based on EEG network characteristics. PLoS One 8, e59764 (2013).
33 Dauwan, M., van der Zande, J. J., van Dellen, E., et al. Random forest to differentiate dementia with Lewy bodies from Alzheimer's disease. Alzheimer's Dement. Diagnosis, Assess. Dis. Monit. 4, 99–106 (2016).
34 Liaw, A. & Wiener, M. Classification and regression by random Forest. R news 2, 18–22 (2002).
35 Dehaene, S. & Changeux, J.-P. Experimental and theoretical approaches to conscious processing. Neuron 70, 200–227 (2011).
36 Engel, A. K., Fries, P., König, P., et al. Temporal Binding, Binocular Rivalry, and Consciousness. Conscious. Cogn. 8, 128–151 (1999).
37 Alkire, M. T., Hudetz, A. G. & Tononi, G. Consciousness and anesthesia. Science. 322, 876–880 (2008).
38 Lee, U., Kim, S., Noh, G.-J., et al. The directionality and functional organization of frontoparietal connectivity during consciousness and anesthesia in humans. Conscious. Cogn. 18, 1069–1078 (2009).
39 Maksimow, A., Silfverhuth, M., Långsjö, J., et al. Directional connectivity between frontal and posterior brain regions is altered with increasing concentrations of propofol. PLoS One 9, e113616 (2014).

40 Mashour, G. A. Integrating the science of consciousness and anesthesia. Anesth. Analg. 103, 975–982 (2006).
41 Lehembre, R., Bruno, M.-A., Vanhaudenhuyse, A., et al. Resting-state EEG study of comatose patients: a connectivity and frequency analysis to find differences between vegetative and minimally conscious states. Funct. Neurol. 27, 41–47 (2012).
42 Stam, C. J., de Haan, W., Daffertshofer, A., et al. Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease. Brain 132, 213–224 (2008).
43 Engels, M. M., Stam, C. J., van der Flier, W. M., et al. Declining functional connectivity and changing hub locations in Alzheimer's disease: an EEG study. BMC Neurol. 15, 145 (2015).
44 Hannawi, Y., Lindquist, M. A., Caffo, B. S., et al. Resting brain activity in disorders of consciousness: A systematic review and meta-analysis. Neurology 84, 1272–1280 (2015).
45 Hudetz, A. G. Suppressing consciousness: Mechanisms of general anesthesia. Semin. Anesth. Perioper. Med. Pain 25, 196–204 (2006).
46 Marino, S., Bonanno, L. & Giorgio, A. Functional connectivity in disorders of consciousness: methodological aspects and clinical relevance. Brain Imaging Behav. 10, 604–608 (2016).
47 Hudetz, A. G. & Mashour, G. A. Disconnecting consciousness. Anesth. Analg. 123, 1228–1240 (2016).
48 Jordan, D., Ilg, R., Riedl, V., et al. Simultaneous electroencephalographic and functional magnetic resonance imaging indicate impaired cortical top-down processing in association with anesthetic-induced unconsciousness. Anesthesiology 119, 1031–1042 (2013).
49 Traast, H. S. & Kalkman, C. J. Electroencephalographic characteristics of emergence from propofol/sufentanil total intravenous anesthesia. Anesth. Analg. 81, 366–371 (1995).
50 Engel, G. L. & Romano, J. Delirium, a syndrome of cerebral insufficiency. J. Chronic Dis. 9, 260–277 (1959).
51 Van der Kooi, A. W., Zaal, I. J., Klijn, F. A., et al. Delirium detection using EEG. Chest 147, 94–101 (2015).
52 McKeith, I. G., Dickson, D. W., Lowe, J., et al. Diagnosis and management of dementia with Lewy bodies: Third report of the DLB consortium. Neurology 65, 1863–1872 (2005).
53 Gunther, M. L., Morandi, A., Krauskopf, E., et al. The association between brain volumes, delirium duration, and cognitive outcomes in intensive care unit survivors. Crit. Care Med. 40, 2022–2032 (2012).
54 Olde Dubbelink, K. T. E., Hillebrand, A., Stoffers, D., et al. Disrupted brain network topology in Parkinson's disease: a longitudinal magnetoencephalography study. Brain 137, 197–207 (2014).
55 Gore, R. L., Vardy, E. R. L. C. & O'Brien, J. T. Delirium and dementia with Lewy bodies: distinct diagnoses or part of the same spectrum? J. Neurol. Neurosurg. Psychiatry 86, 50–59 (2015).
56 Wolters, A. E., van Dijk, D., Pasma, W., et al. Long-term outcome of delirium during intensive care unit stay in survivors of critical illness: a prospective cohort study. Crit. Care 18, R125 (2014).
57 Goto, T. & Maekawa, K. Cerebral dysfunction after coronary artery bypass surgery. J. Anesth. 28, 242–248 (2014).
58 Tewarie, P., Hillebrand, A., van Dijk, B. W., et al. Integrating cross-frequency and within band functional networks in resting-state MEG: A multi-layer network approach. Neuroimage 142, 324–336 (2016).

SUPPLEMENTARY MATERIAL

Table S3.1. Relative powers of the three groups.

	Relative powers				Post-hoc comparison p -values		
Relative power	**Non-delirious controls (n=20)**	**Hypoactive delirium (n=18)**	**Recovery from anesthesia (n=20)**	**P-value (overall)**	**Hypoactive delirium vs. non-delirious controls**	**Hypoactive delirium vs. Recovery from anesthesia**	**Recovery from anesthesia vs. non-delirious controls**
Delta	0.29 (0.24-0.38)	0.68 (0.54-0.74)	0.59 (0.53-0.72)	<0.001	<0.001	0.346	<0.001
Theta	0.29 (0.20-0.43)	0.24 (0.19-0.34)	0.24 (0.18-0.29)	0.087	-	-	-
Alpha	0.22 (0.15-0.37)	0.07 (0.04-0.11)	0.10 (0.07-0.15)	<0.001	<0.001	0.019	<0.001
Beta	0.08 (0.06-0.14)	0.02 (0.02-0.05)	0.04 (0.03-0.06)	<0.001	<0.001	0.072	<0.001

Post-hoc analyses were performed when overall significant difference was found between the three groups.

Table S3.2. Post-hoc comparison of MST measures in non-delirious controls, hypoactive delirium and recovery from anesthesia.

MST measure	Non-delirious controls (n=20)	Hypoactive delirium (n=18)	Recovery from anesthesia (n=20)	p-value non-delirious controls vs. delirium	p-value non-delirious controls vs. recovery from anesthesia	p-value hypoactive delirium vs. recovery from anesthesia
Delta						
Degree	0.299 (0.270-0.329)	0.305 (0.278-0.344)	0.331 (0.285-0.344)	0.784	0.033*	-
Leaf fraction	0.522 (0.500-0.563)	0.535 (0.504-0.579)	0.558 (0.532-0.594)	0.443	0.028*	-
Diameter	0.486 (0.446-0.516)	0.481 (0.449-0.493)	0.457 (0.438-0.469)	0.409	0.033*	-
BC_{max}	0.707 (0.680-0.725)	0.714 (0.693-0.735)	0.730 (0.700-0.754)	0.290	0.102	-
Ecc	0.227 (0.219-0.246)	0.226 (0.203-0.235)	0.210 (0.192-0.219)	0.534	0.010*	-
Th	0.371 (0.356-0.388)	0.376 (0.355-0.408)	0.385 (0.366-0.402)	0.443	0.149	-
Alpha						
Degree	0.323 (0.301-0.356)	0.288 (0.266-0.313)	0.307 (0.282-0.325)	0.003*	0.086	0.093
Leaf fraction	0.545 (0.516-0.574)	0.513 (0.485-0.547)	0.534 (0.493-0.563)	0.033*	0.602	0.099
Diameter	0.469 (0.438-0.500)	0.497 (0.454-0.516)	0.467 (0.442-0.496)	0.105	0.779	0.126
BC_{max}	0.726 (0.682-0.754)	0.697 (0.666-0.736)	0.730 (0.708-0.761)	0.016*	0.820	0.007*
Ecc	0.217 (0.203-0.235)	0.234 (0.215-0.250)	0.219 (0.207-0.235)	0.082	0.925	0.067
Th	0.377 (0.363-0.393)	0.370 (0.350-0.398)	0.368 (0.346-0.394)	0.478	0.383	0.965
Beta						
Degree	0.285 (0.266-0.297)	0.282 (0.266-0.297)	0.279 (0.250-0.297)	0.806	0.201	-
Leaf fraction	0.501 (0.469-0.532)	0.485 (0.450-0.531)	0.511 (0.488-0.532)	0.251	0.429	-
Diameter	0.510 (0.485-0.543)	0.522 (0.485-0.563)	0.512 (0.485-0.547)	0.331	0.820	-
BC_{max}	0.699 (0.690-0.708)	0.695 (0.668-0.723)	0.689 (0.657-0.713)	0.806	0.221	-
Ecc	0.241 (0.219-0.250)	0.245 (0.219-0.266)	0.240 (0.219-0.262)	0.426	0.989	-
Th	0.360 (0.343-0.382)	0.351 (0.330-0.370)	0.372 (0.360-0.389)	0.478	0.121	-

*Post-hoc comparisons of MST measures were only performed when the MST overlap was significantly different between the three groups. Only in the alpha frequency band a significant difference of MST overlap was found in the hypoactive delirium and recovery from anesthesia patients. BC_{max} = betweenness centrality, Ecc = eccentricity, Th = tree hierarchy, *p<0.05.*

ADDITIONAL ANALYSIS – CORRECTION FOR RASS SCORE

The averaged RASS score of the hypoactive delirium group (mean,-1.4; SD, 0.5) was slightly higher compared to the recovery from anesthesia group (mean,-1.7; SD, 0.5), although not significantly different (p=0.19). To study MST network measures in the alpha frequency band independent of the level of consciousness, we performed a general linear model analysis with the RASS score as covariate.

In Table S3.3 the estimated mean of MST measures corrected for the RASS score are presented. Maximum betweenness centrality was significantly higher in recovery from anesthesia compared to hypoactive delirium in the crude analysis. In the adjusted analysis, this difference remained significant. Moreover, degree and leaf fraction were significantly higher and diameter and eccentricity were significantly lower in recovery from anesthesia compared to hypoactive delirium, after adjustment for level of consciousness. This indicates a deviant shift towards a path-like MST structure for hypoactive delirium and a star-like structure during recovery from anesthesia.

Table S3.3. Estimated mean MST measures (95% confidence interval) in the alpha frequency band; RASS included as covariate.

MST measures	Hypoactive delirium (n=18)	Recovery from anesthesia (n=20)	P-value (unadjusted)
Degree	0.284 (0.271-0.298)	0.309 (0.297-0.322)	0.009
Leaf fraction	0.508 (0.491-0.525)	0.539 (0.523-0.555)	0.012
Diameter	0.502 (0.483-0.520)	0.463 (0.446-0.480)	0.004
BCmax	0.694 (0.678-0.711)	0.732 (0.716-0.748)	0.002
Ecc	0.237 (0.227-0.247)	0.217 (0.207-0.226)	0.007
Th	0.368 (0.354-0.381)	0.371 (0.358-0.383)	0.749

BC_{max} = maximum betweenness centrality, Ecc = eccentricity, Th = tree hierarchy

ADDITIONAL ANALYSIS – UNCONSCIOUSNESS VS. REGAINING CONSCIOUSNESS

In the main analysis reduced consciousness due to recovery from anesthesia and hypoactive delirium were compared. The effect of level of consciousness was previously studied by Lee et al. (2013), using the directed phase lag index (dPLI) as measure for directed connectivity[10]. The directed phase transfer entropy (dPTE) was shown to be a

more reliable measure for directed connectivity[11]. Therefore, we repeated this study in a subset of our recovery from anesthesia group using the PLI, dPTE and minimum spanning tree.

In the recovery from anesthesia group the EEG electrodes were directly applied when possible after ICU admission, to be able to start recording before the patient was fully alert (RASS=0). The recording was continued until the patient maintained the level of RASS -2/-1 for 10 minutes. In the main analysis, the selected EEG epochs were during a reduced state of consciousness with RASS scores of -2 or -1. A broad range of variation between patients was found in this process of regaining consciousness. In 13 patients we were able to select 4 epochs of 8 seconds during a period of unconsciousness (RASS -4 or-5) and compared these with the period of regaining consciousness (RASS -2 or -1) using the same measures as the main study.

In Table S3.4 the PLI results are presented. The uncorrected p-values showed a significant higher PLI in the theta frequency band during regaining of consciousness compared to unconsciousness (p=0.046). Lee et al. found a reduction of the PLI during the process of regaining consciousness in the alpha frequency band, in this subanalysis no significant differences in the alpha band PLI were found[10]. The directed connectivity (dPTE) did not show a significant difference between unconsciousness and regaining consciousness, see Figure S3.1 and Table S3.5. Based on the dPLI Lee et al. found a flip of the directed connectivity (i.e., feedforward instead of feedbackward) during loss of consciousness, which was significantly restored during recovery of consciousness[10]. Based on our dPTE results, we were not able to reproduce this difference between the levels of reduced consciousness. However, in the main analysis, we did find the loss of directionality in the recovery from anesthesia group (i.e., RASS -1/-2) compared to the non-delirious control subjects.

Table S3.4. Phase lag index (PLI) in the four frequency bands (median (interquartile range)) between unconsciousness (RASS-4 or -5) and regaining consciousness (RASS -2 or -1).

PLI	Unconsciousness	Regaining consciousness	p-value
Delta	0.165 (0.153-0.177)	0.163 (0.152-0.182)	0.917
Theta	0.131 (0.121-0.136)	0.142 (0.134-0.155)	0.046
Alpha	0.129 (0.122-0.151)	0.126 (0.116-0.133)	0.087
Beta	0.097 (0.092-0.102)	0.098 (0.086-0.100)	0.552

RASS = Richmond agitation-sedation scale

Table S3.5. Anterior-posterior (AP) ratios based on the directed phase transfer entropy in the four frequency bands during unconsciousness (RASS -4 or -5) and regaining consciousness (RASS -2 or -1).

AP-ratio	Unconsciousness	Regaining consciousness	p-value
Delta	0.998 ± 0.008	0.996 ± 0.017	0.668
Theta	0.997 ± 0.007	0.992 ± 0.008	0.113
Alpha	1.009 ± 0.019	1.004 ± 0.016	0.211
Beta	1.000 ± 0.007	1.000 ± 0.007	0.246

AP-ratios >1 indicate an information flow from anterior to posterior, AP-ratio <1 indicate a posterior-anterior information flow. Data are presented as mean ± standard deviation.

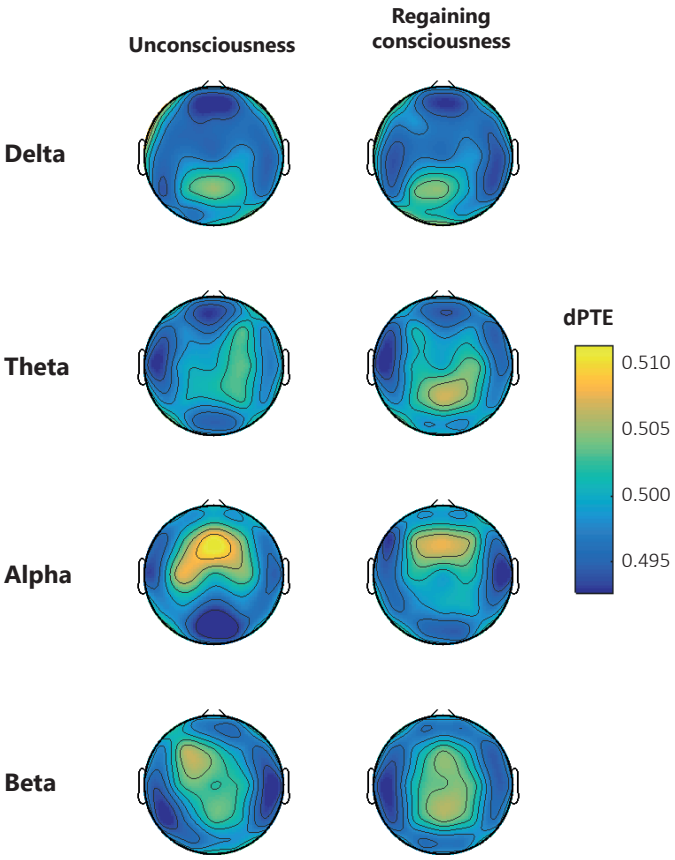


Figure S3.1. Mean directed phase transfer entropy results during unconsciousness (RASS -4 or -5) and regaining consciousness (RASS -2 or -1).
dPTE values above 0.5 indicate and information flow from that electrodes, whereas dPTE values below 0.5 indicate information flow towards that electrode. A1, A2, Fp1, Fp2 were excluded for analysis.

The MST measures were compared for the delta, alpha and beta frequency band, because in the main analysis a difference in topology was found in these frequency bands between recovery from anesthesia and non-delirious control subjects using the MST overlap with MST-reference. In the delta frequency band, an increase in leaf fraction (p=0.011) and decrease in diameter (p=0.003) and eccentricity (p=0.004) was found in the process of regaining consciousness (Table S3.6). This indicates that the functional organization of the brain is more efficient when patients are regaining consciousness compared to the unconscious state in the same patients. In the alpha frequency band the tree hierarchy measure was found to be significantly lower in patients regaining consciousness compared to unconscious state. It may be counterintuitive that the unconscious brain has a more efficient organization than the brain regaining consciousness. However, this possible difference in MST topology is not confirmed by other MST parameters in the alpha band. Moreover, the presented results were obtained in a small group of 13 subjects, no correction for multiple testing was applied and all other MST parameters in the alpha band were not different. Therefore, these results should be interpreted with caution.

Table S3.6. MST measures during unconsciousness and regaining consciousness.

MST measure	Unconsciousness	Regaining consciousness	p-value
Delta			
Degree	0.313 (0.282-0.352)	0.329 (0.301-0.275)	0.055
Leaf fraction	0.532 (0.500-0.547)	0.579 (0.555-0.602)	0.011*
Diameter	0.485 (0.469-0.516)	0.453 (0.422-0.454)	0.003*
BC_{max}	0.696 (0.681-0.753)	0.746 (0.687-0.753)	0.345
Ecc	0.234 (0.219-0.250)	0.203 (0.188-0.219)	0.004*
Th	0.367 (0.349-0.400)	0.390 (0.378-0.423)	0.055
Alpha			
Degree	0.313 (0.281-0.383)	0.313 (0.297-0.336)	0.959
Leaf fraction	0.579 (0.532-0.594)	0.547 (0.500-0.563)	0.050
Diameter	0.438 (0.422-0.492)	0.485 (0.430-0.493)	0.277
BC_{max}	0.702 (0.689-0.763)	0.727 (0.707-0.765)	0.279
Ecc	0.219 (0.188-0.234)	0.219 (0.203-0.234)	0.545
Th	0.394 (0.368-0.416)	0.358 (0.343-0.388)	0.023*
Beta			
Degree	0.282 (0.251-0.271)	0.282 (0.250-0.297)	0.964
Leaf fraction	0.500 (0.469-0.516)	0.501 (0.492-0.532)	0.278
Diameter	0.516 (0.500-0.532)	0.532 (0.485-0.555)	0.327
BC_{max}	0.696 (0.671-0.708)	0.675 (0.658-0.712)	0.600
Ecc	0.250 (0.219-0.250)	0.250 (0.234-0.266)	0.207
Th	0.362 (0.342-0.376)	0.369 (0.358-0.384)	0.196

*BC_{max} = maximum betweenness centrality, Ecc = eccentricity, Th = tree hierarchy, *p<0.05*

In conclusion, the differences in level of consciousness can be captured in the MST delta band measures, with the regaining of consciousness showing a more integrated network compared to unconsciousness.

ADDITIONAL ANALYSIS – RANDOM FOREST CLASSIFICATION

Random forest classification was performed in three datasets: 1) hypoactive delirium versus controls, 2) recovery from anesthesia versus controls, and 3) hypoactive delirium versus recovery from anesthesia. The variables are listed in Table S3.7. Variable importance results are presented in Figure S3.2.

Table S3.7. Features included in random forest classifier analysis.

Number	Feature name	Number	Feature name
1	Relative delta power	19	MST – Degree theta
2	Relative theta power	20	MST – Leaf fraction theta
3	Relative alpha power	21	MST – Diameter theta
4	Relative beta power	22	MST – BC_{max} theta
5	PLI delta	23	MST – Ecc theta
6	PLI theta	24	MST – Tree hierarchy theta
7	PLI alpha	25	MST – Degree alpha
8	PLI beta	26	MST – Leaf fraction alpha
9	PTE ratio delta	27	MST – Diameter alpha
10	PTE ratio theta	28	MST – BC_{max} alpha
11	PTE ratio alpha	29	MST – Ecc alpha
12	PTE ratio beta	30	MST – Tree hierarchy alpha
13	MST – Degree delta	31	MST – Degree beta
14	MST – Leaf fraction delta	32	MST – Leaf fraction beta
15	MST – Diameter delta	33	MST – Diameter beta
16	MST – BC_{max} delta	34	MST – BC_{max} beta
17	MST – Ecc delta	35	MST – Ecc beta
18	MST – Tree hierarchy delta	36	MST – Tree hierarchy beta

BC_{max} = maximum betweenness centrality, Ecc = eccentricity, MST = minimum spanning tree, PLI = phase lag index.

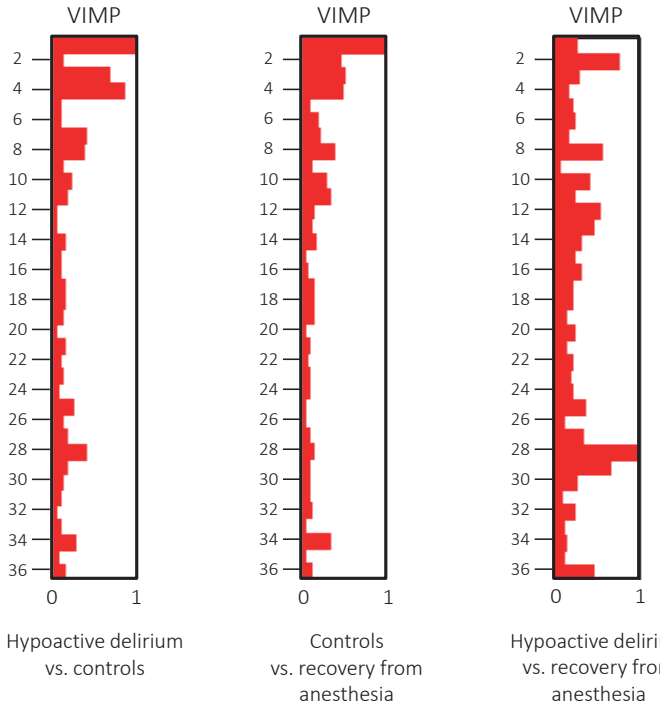


Figure S3.2. Variable importance (VIMP) results of random forest classifier for three datasets.
Feature numbers relate to Table S3.7.

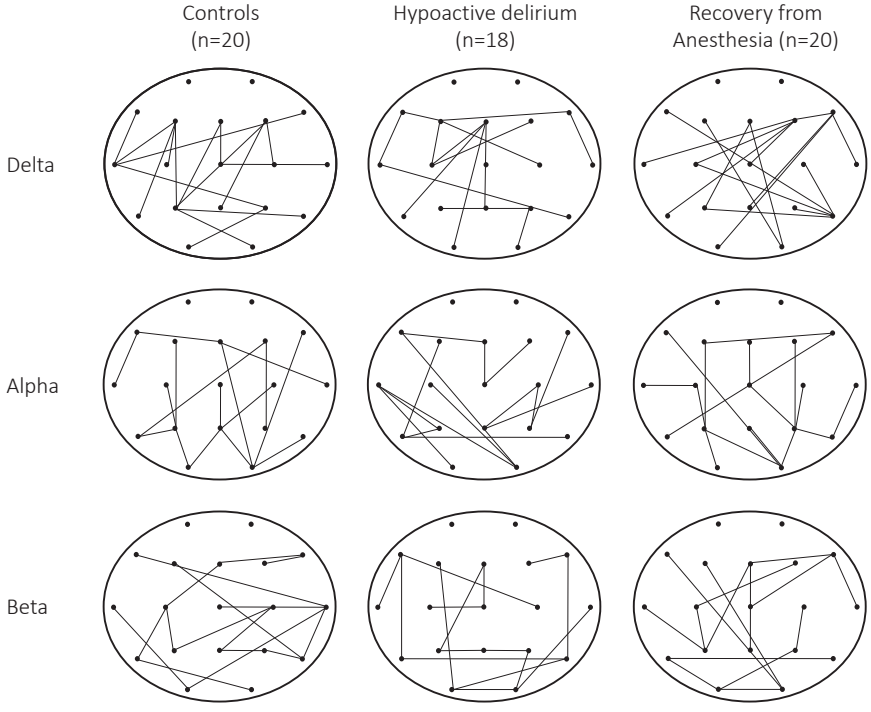
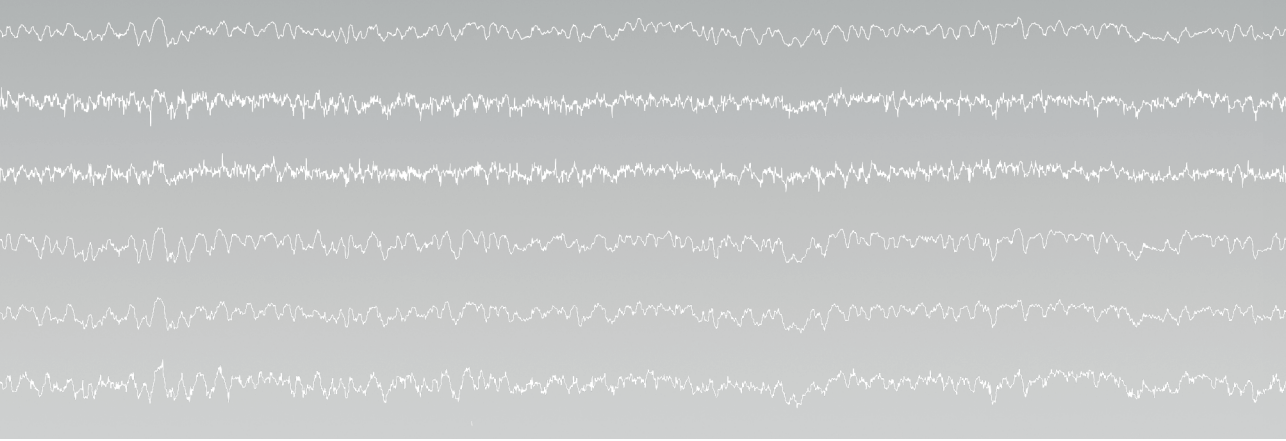


Figure S3.3. Visualization plots of the averaged MST topology in the three groups.
The circle represents the head with the nose on top.



4

RESTING STATE ELECTROENCEPHALOGRAPHY CHARACTERISTICS DURING SEDATION WITH MIDAZOLAM AND PROPOFOL IN OLDER PATIENTS

T. Numan
E. van Dellen
F.P. Vleggaar
P. van Vlieberghe
C.J. Stam
A.J.C. Slooter

Submitted

ABSTRACT

BACKGROUND: Little is known about the neurophysiological effects of midazolam and propofol despite widespread use. Sedative-induced unconsciousness alters functional brain networks. However, these effects were found in young subjects, whereas the effects of sedatives in older subjects might be different. Sedatives are frequently administered to older individuals. However, elderly may have reduced reserves for these agents that affect functional connectivity. The aim of this study was to assess the effects of midazolam and propofol on a variety of EEG measures in older subjects.

METHODS: In patients (≥60 years without neuropsychiatric disease such as delirium) scheduled for endoscopy two electroencephalography (EEG) recordings were performed, before and after administration of either midazolam (n=22) or propofol (n=26). Power spectrum, functional connectivity, directed connectivity, and network topology based on the minimum spanning tree were compared within subjects (i.e., baseline versus light sedation with either midazolam or propofol) and between subjects (i.e., midazolam versus propofol).

RESULTS: Both midazolam and propofol altered power spectra and reduced functional connectivity. The back-to-front pattern of information flow in the alpha band during the awake state was lost during midazolam and propofol sedation, but network topology was not affected. Propofol showed a larger reduction of functional connectivity and stronger disruption of the directed connectivity compared to midazolam.

CONCLUSION: Midazolam and propofol had a distinct neurophysiological signature for light sedation with regard to relative powers, functional connectivity, and directed connectivity. These effects are more pronounced during sedation with propofol than with midazolam. No clear effects were found on network topology.

INTRODUCTION

Every day, midazolam and propofol are administered to numerous individuals for general anesthesia, procedural sedation or to comfort patients in the intensive care unit (ICU). Yet, little is known about the neurophysiological effects of both sedatives, particularly in elderly individuals. In a healthy, awake resting state, electroencephalography (EEG) shows an alpha rhythm (8-13 Hz) in posterior derivations when the eyes are closed. Moreover, in a resting state, brain areas are functionally connected with a dominant posterior to anterior flow of information[1–4]. Sedative agents alter these characteristics, and these alterations coincide with the level of consciousness. These effects have mostly been studied for propofol[5,6], while literature on other sedatives, such as midazolam and other benzodiazepines, is relatively scarce.

A previous study using EEG coherence suggested that light sedation with benzodiazepines results in functional uncoupling of frontal brain regions[7]. However, coherence as connectivity measure in EEG is difficult to interpret as it can be influenced by the signal power and the effect of volume conduction[8,9]. Functional magnetic resonance imaging (fMRI) studies on midazolam-induced light levels of sedation showed disruption of networks related to higher order brain functions with preservation of lower-order networks[10], and reduced functional connectivity within the default mode network[11]. Propofol-induced loss of consciousness showed similar effects: decreased cortico-cortical and thalamo-cortical connectivity[6,12] was found using fMRI, and EEG studies showed reduced functional connectivity, a less efficient brain network, and disruption of the directed feedforward connectivity[1,2,12].

It should be noted that previous studies on functional connectivity and light levels of sedation were performed in younger subjects. In clinical practice however, particularly older patients receive sedation, for example during endoscopy[13,14]. The functional brain network changes with aging, in older subjects reduced functional connectivity and less efficient network topology were found[15]. Therefore, as the functional brain network of elderly is already affected, the effects of sedatives on functional connectivity may be different in elderly individuals versus younger persons. This may be related to frailty as defined by a decreased reserve and increased vulnerability to adverse outcomes[16]. Moreover, the application of sedation can lead to subsequent delirium[17,18] with associated long-term cognitive impairment[19,20]. Therefore, the primary aim of this study was to assess the individual effects of midazolam and propofol on the functional brain

network in older subjects. Moreover, we aimed to compare the effects of midazolam and propofol on EEG connectivity. We hypothesized that light levels of sedation in older patients, due to midazolam and propofol, result in an altered power spectrum, reduced functional connectivity, disturbed directed connectivity and altered network topology.

METHOD

STUDY POPULATION

This study was performed from December 2014 to December 2016 in patients scheduled for a gastrointestinal endoscopic procedure in the University Medical Center Utrecht (UMCU). The ethical committee of UMCU waived the need for approval (number 14-466). Inclusion criteria were age above 60 years and planned endoscopy with light sedation with either midazolam or propofol. Exclusion criteria were (previous) delirium, previous stroke, epilepsy, and any psychiatric disease. Patients were included when they gave written informed consent. Patients who received midazolam were age- and sex-matched on group level to patients who received propofol.

STUDY PROCEDURES AND DATA COLLECTION

The following patient and procedural characteristics were collected: age, sex, type and duration of endoscopy, type of sedation, and type of analgesia. Selection of sedative type was done by the gastroenterologist performing the endoscopy based on physical states (i.e., ASA classification), invasiveness and risk of the procedure, the need of lying still by the patient, and previous experience by the patient. One hour prior their endoscopic procedure patients arrived at the endoscopy unit for the preparation of the set-up. An EEG cap with 32 electrodes (Brain products GmbH, Munich, Germany) was placed according to the 10-20 system. A reference electrode was located between Fz and Cz. Electrode impedances were kept below 20 kOhm and sample frequency was 5000 Hz.

For the baseline EEG recording, patients were lying on their back and were instructed to relax and keep their eyes closed for 5 minutes, while the researcher ensured that the patient was not falling asleep. The patient was subsequently transported to the endoscopy room where impedances were checked again, and the EEG cap was adjusted when necessary. The sedation EEG recording was started and the level of consciousness was assessed using the Richmond agitation-sedation scale (RASS) every 10 minutes[21].

EEG ANALYSIS

Data was down sampled using spline interpolation to 512 Hz with the Brain Vision Analyzer software (version 2.0.4.368 Brain products GmbH, Munich, Germany) to enable comparison with previous studies[22]. From all baseline recordings, four epochs of eight seconds were selected, as this was previously shown to be sufficient for stable results[22,23]. From each sedation recording (with either midazolam or propofol), four epochs of eight seconds were selected within three minutes after the first determination of the reduced level of consciousness having artifact-free EEG data. We assumed stable levels of consciousness within these three minutes as the administration of sedatives was unchanged in this timeframe. For all analyses, the following 17 electrodes were included: F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1 and O2 to enable comparison with previous literature[9,24]. A source reference was used for connectivity and topology analysis, as approximation of the Laplacian[25] eliminating the influence of active reference in the analysis.

Spectral analysis

First data was filtered in forward and backward arrangement, excluding phase disturbances by filtering, using a Butterworth filter with a high pass of 0.5 Hz and low pass of 50 Hz and an order of 2. Subsequently, for each epoch the power spectral density was calculated with the Welch method and normalized according to the power between 0.5 and 20 Hz. Relative delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), and beta (13-20 Hz) power were calculated for each epoch and averaged over the four epochs. Frequencies above 20 Hz were excluded, as these represented particularly muscle activity due to the clinical procedures[26,27].

Functional connectivity

The phase lag index (PLI) was calculated to assess functional connectivity, a measure of synchronization between electrode time series[28]:

$$PLI = |\langle sign[\sin(\Delta\varphi(t))]\rangle|$$

For each sample (t) the instantaneous phase was obtained using the Hilbert transformation, and the instantaneous phase difference between two times series ($\Delta\phi$) was calculated, ranging between $-\pi$ and π. The sign-function is 1 for all positive phase differences and -1 for al negative phase differences, which are averaged over the epoch. Calculating the absolute value results in the PLI, a value ranging between 0 (no phase

synchronization or zero lag phase synchronization) and 1 (complete, non-zero phase-locking over the epoch). The PLI was averaged over all combinations of electrodes.

Directed connectivity

Besides the functional connectivity, we also assessed the direction of the information flow using the phase transfer entropy (PTE)[29–31]. Transfer entropy assesses the direction of the causal influence between two signals. A causal influence is assumed when prediction of the future of signal 1 can be improved by adding current information of signal 2. The PTE estimates the extent of this directed causality, based on the phase differences between the two signals. The advantages of the PTE are the robustness to noise and the independence of the choice of reference[30,31].

To ensure that the directed connectivity measures are comparable between subjects, we normalized the PTE towards the directed PTE (dPTE) as follows:

$$dPTE_{x\to y} = \frac{PTE_{x\to y}}{PTE_{x\to y} + PTE_{y\to x}}$$

After normalization, dPTE values range between 0 and 1, with 0.5 indicating no favored directionality. Values above 0.5 indicate a preferred directionality from X to Y, values below 0.5 indicate the opposite directionality, from Y to X.

To assess the global directionality, the anterior-posterior (AP) ratio was calculated by dividing the average of the anterior dPTE values (channels F3, F4, F7, F8, and Fz) by the average of the posterior dPTE values (channels P3, P4, Pz, T5, T6, O1, and O2). An AP-ratio above 1 indicates a front-to-back information flow, whereas an AP-ratio below 1 means a back-to-front pattern.

Minimum spanning tree

The topology of the functional network was assessed using the minimum spanning tree (MST). Based on the PLI connectivity matrix, the MST was obtained using Kruskal's algorithm, including the strongest connections without forming loops[32]. This resulted in a network of 17 nodes (i.e., electrodes) and 16 connections (i.e., PLI values).

First, to assess overall differences in topology, an MST-reference was calculated. Based on all baseline recordings, separately for the midazolam and the propofol group, an averaged PLI matrix was calculated over all epochs. This averaged PLI matrix was used

to create an MST-reference, an approximation of the averaged baseline topography. The number of overlapping connections of the MST of each epoch with the MST-reference was counted. The ratio of overlap (MST-overlap) was calculated by dividing the number of overlapping connections by the total number of connections within the MST (i.e., 16 connections). Subsequently the MST-overlap was averaged over the four epochs. To overcome bias due to repeated measurements in the same individuals, we used as MST-reference for the midazolam group the averaged PLI matrix over all propofol baseline recordings, and as MST-reference for the propofol group the averaged PLI matrix over all midazolam baseline recordings. MST measures were calculated when a significant difference in MST-overlap was found. Degree, diameter, and leaf fraction were calculated to characterize global integration of information, betweenness centrality and eccentricity quantified the centrality of nodes, and tree hierarchy gives a global measure of hierarchical organization of the MST topology (see supplementary material Table S4.1 for the included MST measures).

Functional connectivity, AP-ratio as directed connectivity measure, and MST-overlap were calculated for each frequency band, and averaged over the four epochs. Connectivity and network analysis were performed in Brainwave (Version 0.9.152.4.1 freely available on http://home.kpn.nl/stam7883/brainwave.html) and MATLAB (Version 2015a, The MathWorks, Inc.) using the connectivity toolbox[33].

STATISTICAL ANALYSIS

Patient characteristics were tested for normality and presented as mean with standard deviation (SD), median with interquartile range (IQR), or number (%) as appropriate. Patient characteristics were compared between the midazolam and propofol groups, using t-test, Mann-Whitney U test, or Chi-square test where appropriate. Statistical significance was defined as p-value <0.05.

To study the difference in EEG characteristics between baseline and sedation with either midazolam or propofol, Student's paired t-test was used in case of a normal distribution, and the Wilcoxon signed rank test was used for a skewed distribution.

To address differences between the effect of midazolam and propofol, general linear model analyses were performed, with the difference in the included measure (e.g., ΔPLI in alpha frequency band by sedation PLI – baseline PLI) as dependent variable, type of sedation (i.e., midazolam or propofol) as fixed factor, and with baseline of the included

measure (e.g., baseline PLI in alpha frequency band) and level of consciousness during sedation (i.e., RASS score) as covariates. As sensitivity analysis assessing the reliability of the results, we re-evaluated the difference between the effect of midazolam and propofol without correction for the RASS score.

As the aim of this study was to explore different effects of sedatives, and no prior information was available, we chose not to correct for multiple comparisons to avoid type II statistical errors. Statistical analyses were performed with SPSS, version 21 (SPSS Inc., Chicago, IL).

RESULTS

STUDY POPULATION

In total, 85 patients signed informed consent, however 37 patients had to be excluded because of the following reasons: no sedation during endoscopy (n=1), meningioma or stroke in the medical history (n=2), no EEG recording possible due to logistic reasons (n=4), no eyes closed segments (n=1), no EEG signal on one or more of the included channels (n=6), or no artifact free epoch selection possible due to the noisy environment of the experimental setup (n=23). In the remaining 48 patients, 22 received midazolam sedation and 26 propofol sedation. In Table 4.1 the patient characteristics are described. The type of analgesia was different between the midazolam and the propofol groups, in accordance with the protocol at the endoscopy unit. Furthermore, the patients receiving propofol reached a lower level of consciousness as assessed by RASS ($p<0.001$).

SPECTRAL ANALYSIS

Table 4.2 shows that midazolam sedation resulted in a significant increase in relative delta and beta power, and a significant decrease of relative theta and alpha power compared to the baseline recording. Propofol sedation increased the relative delta power and decreased the relative theta and alpha power (Table 4.2). No significant effect on beta power was observed during sedation with propofol. When we compared the effects on relative powers between midazolam sedation and propofol sedation, no differences were found (Table 4.3).

FUNCTIONAL CONNECTIVITY

Sedation with midazolam resulted in a decrease of the PLI in the alpha frequency band (median, 0.135; IQR, 0.115-0.187) compared to baseline recording (median, 0.168; IQR, 0.136-0.214, p=0.033). In the other frequency bands, no differences were found in PLI in recordings during midazolam sedation compared to baseline. Propofol sedation resulted in a significantly increased PLI in the delta and beta frequency bands, and a significantly decreased PLI in the theta and alpha frequency bands (Table 4.2). Compared to midazolam sedation, sedation with propofol resulted in a larger decrease in PLI in the alpha band PLI (p=0.037), see Table 4.3.

DIRECTED CONNECTIVITY

At baseline, before administration of midazolam or propofol, the AP-ratio in the alpha frequency band was below 1, indicating a back-to-front information flow. Midazolam sedation increased this AP-ratio and resulted therefore in loss of directionality. Sedation with propofol resulted in an AP-ratio in the alpha frequency band above 1, indicating a front-to-back pattern and therefore a flip of directionality (Table 4.2).

Table 4.1. Patient characteristics.

	Midazolam (n=22)	Propofol (n=26)	P-value (unadjusted)
Mean Age (SD)	72.7 (5.9)	71.5 (6.4)	0.481
Male sex (%)	13 (59.1%)	17 (65.4%)	0.768
Endoscopic procedure			0.239
Coloscopy	11 (50.0%)	9 (34.6%)	
Gastroscopy	11 (50.0%)	17 (65.4%)	
Planned duration of endoscopic procedure (min.)	33.0 (23.8-41.3)	35.0 (25.0-77.5)	0.189
Analgesia			<0.001*
None	5 (22.7%)	-	
Alfentanil	1 (4.5%)	25 (96.2%)	
Fentanyl	16 (72.7%)	-	
Remifentanil	-	1 (3.8%)	
RASS			<0.001*
0	2 (9.1%)	-	
-1	5 (22.7%)	-	
-2	10 (45.5%)	4 (15.4%)	
-3	4 (18.2%)	5 (19.2%)	
-4	1 (4.5%)	17 (65.4%)	

*RASS = Richmond agitation-sedation scale, SD = standard deviation. *Significantly different (p<0.05).*

Table 4.2. Electroencephalography effects of light sedation with midazolam or propofol.

	Midazolam (n=22)			Propofol (n=26)		
Measure	**Baseline**	**Sedation**	**p-value**	**Baseline**	**Sedation**	**p-value**
Relative power						
Delta	0.188 (0.107-0.264)	0.365 (0.274-0.429)	<0.001*	0.205 (0.159-0.300)	0.443 (0.296-0.640)	<0.001*
Theta	0.176 (0.130-0.213)	0.127 (0.106-0.156)	0.036*	0.206 (0.164-0.282)	0.104 (0.079-0.128)	<0.001*
Alpha	0.484 (0.312-0.575)	0.235 (0.194-0.324)	<0.001*	0.330 (0.218-0.501)	0.208 (0.140-0.292)	<0.001*
Beta	0.160 (0.118-0.226)	0.245 (0.164-0.289)	0.003*	0.157 (0.115-0.191)	0.188 (0.107-0.251)	0.501
Functional connectivity, PLI						
Delta	0.158 (0.147-0.170)	0.175 (0.157-0.182)	0.074	0.160 (0.152-0.180)	0.191 (0.173-0.211)	0.003*
Theta	0.133 (0.117-0.142)	0.132 (0.121-0.146)	0.987	0.136 (0.121-0.153)	0.124 (0.119-0.136)	0.020*
Alpha	0.168 (0.136-0.214)	0.135 (0.115-0.187)	0.033*	0.160 (0.131-0.210)	0.128 (0.118-0.138)	0.001*
Beta	0.102 (0.093-0.106)	0.104 (0.098-0.110)	0.445	0.099 (0.096-0.105)	0.101 (0.093-0.109)	0.001*
Directed connectivity, AP-ratio						
Delta	1.013 (0.999-1.018)	1.007 (0.997-1.020)	0.833	1.010 (0.991-1.021)	1.001 (0.986-1.008)	0.124
Theta	0.999 (0.989-1.006)	0.992 (0.983-1.011)	0.808	0.998 (0.983-1.007)	0.100 (0.997-1.013)	0.118
Alpha	0.959 (0.937-0.978)	0.991 (0.977-0.999)	0.002*	0.967 (0.940-0.986)	1.012 (0.992-1.016)	<0.001*
Beta	0.991 (0.986-1.004)	0.994 (0.989-1.003)	0.709	0.994 (0.984-1.002)	1.001 (0.995-1.006)	0.016*
Topology, MST-overlap						
Delta	0.109 (0.094-0.148)	0.125 (0.090-0.141)	0.888	0.109 (0.094-0.145)	0.117 (0.078-0.141)	0.392
Theta	0.141 (0.106-0.160)	0.125 (0.106-0.156)	0.393	0.125 (0.094-0.145)	0.109 (0.094-0.145)	0.972
Alpha	0.164 (0.125-0.207)	0.188 (0.125-0.234)	0.485	0.141 (0.109-0.176)	0.125 (0.094-0.176)	0.184
Beta	0.141 (0.102-0.188)	0.141 (0.106-0.188)	0.952	0.156 (0.137-0.176)	0.109 (0.094-0.156)	0.012*

*Relative delta power, functional connectivity (phase lag index, PLI) directed connectivity (anterior-posterior (AP)-ratio based on directed phase transfer entropy) and MST-overlap were presented as median (interquartile range). *Significantly different (p<0.05).*

Moreover, propofol increased the AP-ratio in the beta frequency band (p=0.016, Table 4.2). When midazolam and propofol sedation were directly compared, we found a difference in the adjusted change in directed connectivity in the delta frequency band (p=0.006, see Table 4.3). The dPTE pattern in the delta frequency band is shown in Figure 4.1 in more detail at baseline and during sedation with midazolam or propofol. Whereas recordings at baseline and during midazolam sedation did not have a preferred directed connectivity, propofol sedation resulted in a more back-to-front pattern.

Table 4.3. Comparison of electroencephalography effects of midazolam versus propofol.

Measure	Effect of Midazolam (n=22)	Effect of Propofol (n=26)	F-test	p-value
Δ Relative power				
Delta	0.138 (0.003 to 0.243)	0.262 (0.164 to 0.360)	F(1,40) = 3.045	0.089
Theta	-0.079 (-0.122 to -0.037)	-0.101 (-0.133 to 0.069)	F(1,40) = 0.671	0.418
Alpha	-0.147 (-0.198 to -0.095)	0.173 (-0.219 to -0.128)	F(1,40) = 0.607	0.441
Beta	0.045 (-0.008 to 0.098)	0.017 (-0.033 to 0.067)	F(1,40) = 0.609	0.440
Δ Functional connectivity, Δ PLI				
Delta	0.010 (-0.003 to 0.024)	0.023 (0.009 to 0.036)	F(1,40) = 1.626	0.210
Theta	-0.006 (-0.019 to 0.006)	-0.015 (-0.023 to -0.006)	F(1,40) = 1.278	0.265
Alpha	-0.015 (-0.032 to 0.003)	-0.041 (-0.059 to -0.024)	F(1,40) = 4.656	0.037*
Beta	0.001 (-0.007 to 0.008)	-0.003 (-0.010 to 0.004)	F(1,40) = 0.425	0.518
Δ Directed connectivity, Δ AP-ratio				
Delta	0.008 (-0.005 to 0.021)	-0.017 (-0.029 to 0.005)	F(1,40) = 8.315	0.006*
Theta	0.004 (-0.008 to 0.016)	0.010 (-0.001 to 0.021)	F(1,40) = 0.504	0.482
Alpha	0.022 (0.010 to 0.034)	0.035 (0.023 to 0.046)	F(1,40) = 2.457	0.125
Beta	0.007 (0.000 to 0.015)	0.008 (0.001 to 0.016)	F(1,40) = 0.043	0.837
Δ Topology, Δ MST-overlap				
Delta	-0.004 (-0.027 to 0.018)	-0.011 (-0.032 to 0.011)	F(1,40) = 0.156	0.695
Theta	0.003 (-0.020 to 0.026)	-0.012 (-0.033 to 0.010)	F(1,40) = 0.905	0.347
Alpha	0.017 (-0.015 to 0.048)	-0.010 (-0.044 to 0.024)	F(1,40) = 1.333	0.255
Beta	-0.033 (-0.065 to 0.001)	-0.006 (-0.034 to 0.023)	F(1,40) = 1.625	0.210

*Data are presented as estimated marginal means (95% confidence interval). Difference between effect of midazolam and effect of propofol were evaluated with a linear regression model with sedative type as fixed factor and baseline measure and RASS score as covariates. AP-ratio = anterior-posterior ratio, MST = minimum spanning tree, PLI = phase lag index. *Significantly different (p<0.05).*

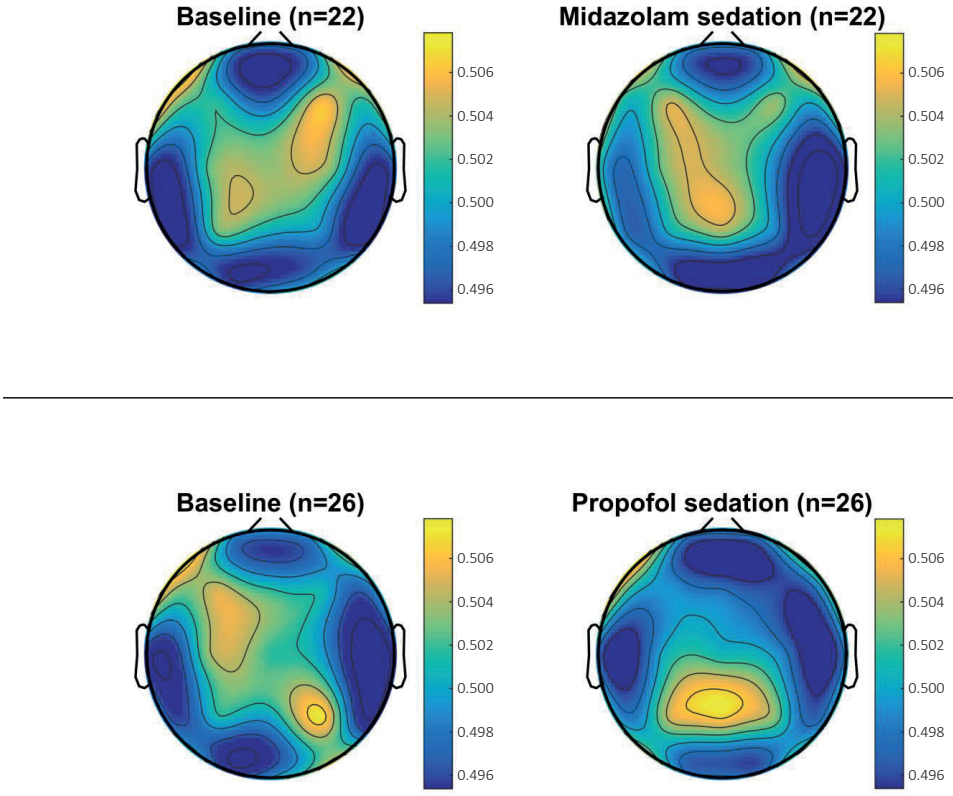


Figure 4.1. The directed connectivity patterns in the delta frequency band (0.5-4 Hz) based on the dPTE before and after administration of midazolam (upper panels) or propofol (lower panels).
Whereas recordings at baseline and during midazolam sedation did not have a preferred directed connectivity, propofol sedation resulted in a more back-to-front pattern.

Table 4.4. MST measures in the beta frequency band of the baseline and propofol-sedation recordings.

MST measures	Baseline (n=26)	Propofol sedation (n=26)	p-value
Degree	0.305 (0.282-0.332)	0.313 (0.282-0.332)	0.808
Leaf fraction	0.532 (0.500-0.563)	0.547 (0.500-0.563)	0.427
Diameter	0.485 (0.449-0.500)	0.477 (0.453-0.500)	0.977
BC_{max}	0.718 (0.690-0.741)	0.706 (0.667-0.746)	0.328
Eccentricity	0.381 (0.357-0.403)	0.381 (0.360-0.395)	0.839
Tree hierarcy	0.373 (0.356-0.393)	0.389 (0.358-0.407)	0.218

BC_{max} = maximum betweenness centrality.

TOPOLOGY OF THE FUNCTIONAL BRAIN NETWORKS

To study network topology, MST-references were created and MST-overlap was calculated and compared between the groups. Midazolam sedation did not affect the MST (Table 4.2). By contrast, propofol sedation resulted in a significant reduction of MST-overlap in the beta frequency band (p=0.012), suggesting an alteration of functional network topology. To explore this in more detail, we calculated MST measures in the beta frequency band before and after administration of propofol. This did not reveal significant differences for any MST-measure between baseline and propofol sedation, see Table 4.4. When we compared the change in network topology between midazolam and propofol sedation, we did not detect significant differences in any of the four frequency bands, see Table 4.3.

SENSITIVITY ANALYSIS

Because of the difference in RASS score between patients undergoing midazolam and propofol sedation, all previous comparisons were corrected for the RASS score. To assess the effect of this approach, we repeated these analyses without correction for RASS score and detected larger differences between the effect of midazolam and the effect of propofol (see supplementary material Table S4.2).

DISCUSSION

In this study, we found that midazolam sedation increased relative delta and beta power and decreased relative theta and alpha power. The functional connectivity strength decreased in the alpha frequency band with midazolam, and the back-to-front pattern in the alpha band disappeared, without an alteration of network topology. Propofol sedation increased the relative delta power and decreased the relative theta and alpha power. Functional connectivity increased in delta and beta band after propofol administration, and significantly decreased in the theta and alpha band. The back-to-front pattern in the alpha band was flipped during propofol sedation, and a more front-to-back pattern was found in the beta band. When we compared the effects of midazolam with the effects of propofol, and corrected for differences in level of consciousness, we found that propofol particularly decreased the alpha band functional connectivity. The directed connectivity pattern in the delta frequency band was also more affected during propofol sedation than midazolam sedation. Together, these results indicate different neurophysiological signatures for light sedation with midazolam versus propofol.

Our findings on the effects of midazolam and propofol on relative power are consistent with most previous literature[34–36], that described a slowing of background activity and an increase in beta power due to midazolam. However, previous neurophysiological studies on midazolam and propofol were all performed in younger individuals, and mainly in subjects who lost consciousness or regained consciousness from anesthesia. In the current study, midazolam or propofol were administered to comfort patients during an endoscopic procedure, and not to induce general anesthesia. Therefore, the results of the current study cannot directly be compared with these previous investigations.

We found that the alpha band PLI was significantly reduced during light levels of sedation with midazolam and particularly propofol. The observed differences in PLI between midazolam and propofol may be the result of different working mechanisms. It is less likely that this different effect of midazolam compared to propofol results from differences in RASS score between both groups, as this difference remained after adjusting for the RASS score and a previous study suggested that functional connectivity in the alpha frequency band is not directly related to the level of consciousness[2].

Frontal-parietal communication has been described to be essential for conscious perception[1,37]. Moreover, this was found as common mediator for loss of consciousness in general anesthesia, independent of the type of administered anesthetics[1]. We found that light levels of sedation with midazolam or propofol affected frontal-parietal information flow in the alpha frequency band. These findings therefore suggest that frontal-parietal communication is already affected during light levels of sedation.

We found that propofol affected network topology in the beta frequency band, but not the MST-measures. This suggests that propofol may alter the connections forming the MST, but not the global network organization. A previous study suggested that loss of consciousness and return of consciousness due to propofol is associated with a less efficient network in all four frequency bands[2]. It is unclear whether differences in age or level of consciousness between this study and the previous study may have influenced the different results. Increasing age (above around 57 years) may result in a significant reduction of functional connectivity and a less efficient minimum spanning tree, compared to younger subjects[15]. Aging and loss of consciousness appeared to result in similar topology changes (i.e., towards a less efficient network). The transition towards a less efficient functional brain network in older patients, may therefore explain why older individuals need lower amounts of sedatives to reach the same level of reduced consciousness compared to younger subjects[13,14,38].

Since use of benzodiazepines such as midazolam is a risk factor for delirium[39,40], it can be hypothesized that administration of benzodiazepines has similar effects on the functional brain network as delirium. Indeed, we previously showed that delirium is associated with a less efficient network, similar to findings in the current study on midazolam. It should however be noted that propofol had a stronger effect on functional connectivity and directed connectivity than midazolam. This suggests that both midazolam and propofol have similar effects on the functional brain networks as delirium, which can be the result of disturbed consciousness in both states.

This is the first study on EEG characteristics of light levels of sedation with midazolam versus propofol in older subjects. We applied advanced and state of the art techniques to unravel the mechanisms of action of drugs that are administered every day to numerous individuals. However, some limitations of the current study need to be addressed. First, the number of excluded patients was high, due to the clinical setting of this study. As patients were mainly excluded for logistic and technical reasons such as environmental artifacts, it is unlikely that this would have affected our findings. Secondly, the type of analgesia during the endoscopic procedure was significantly different between patients who received midazolam and those who received propofol, with unclear consequences. Thirdly, no correction for multiple testing was applied due to the explorative nature of this study. Finally, it is unclear how the difference in RASS score affected the results. However, when we adjusted for the RASS score in multivariate analyses, our findings remained.

In conclusion, this is the first EEG study on the effects of light sedation with midazolam or propofol in older subjects. Both agents had a distinct neurophysiological signature for light levels of sedation with regard to relative powers, functional connectivity and directed connectivity whereas no clear differences were found in network topology.

REFERENCES

1 Lee, U., Ku, S., Noh, G., et al. Disruption of frontal–parietal communication by ketamine, propofol, and sevoflurane. Anesthesiology 118, 1264–1275 (2013).
2 Lee, H., Mashour, G. A., Noh, G.-J., et al. Reconfiguration of network hub structure after propofol-induced unconsciousness. Anesthesiology 119, 1347–1359 (2013).
3 Vecchio, F., Miraglia, F., Bramanti, P., et al. Human brain networks in physiological aging: A graph theoretical analysis of cortical connectivity from EEG data. J. Alzheimer's Dis. 41, 1239–1249 (2014).
4 Stam, C. J. The trees and the forest: characterization of complex brain networks with minimum spanning trees. Int. J. Psychophysiol. 92, 129–138 (2014).
5 Hudetz, A. G. General anesthesia and human brain connectivity. Brain Connect. 2, 291–302 (2012).
6 Tang, C. Y. & Ramani, R. Functional Connectivity and Anesthesia. Int. Anesthesiol. Clin. 54, 143–155 (2016).
7 Alonso, J. F., Mañanas, M. A., Romero, S., et al. Cross-conditional entropy and coherence analysis of pharmaco-EEG changes induced by alprazolam. Psychopharmacology. 221, 397–406 (2012).
8 Numan, T., Stam, C. J., Slooter, A. J. C., et al. Being conscious of methodological pitfalls in functional brain network analysis. Anesthesiology 123, 484–485 (2015).
9 Van Diessen, E., Numan, T., van Dellen, E., et al. Opportunities and methodological challenges in EEG and MEG resting state functional brain network research. Clin. Neurophysiol. 126, 1468–1481 (2015).
10 Liang, P., Zhang, H., Xu, Y., et al. Disruption of cortical integration during midazolam-induced light sedation. Hum. Brain Mapp. 36, 4247–4261 (2015).
11 Greicius, M. D., Kiviniemi, V., Tervonen, O., et al. Persistent default mode network connectivity during light sedation. Hum. Brain Mapp. 29, 839–847 (2008).
12 Jordan, D., Ilg, R., Riedl, V., et al. Simultaneous electroencephalographic and functional magnetic resonance imaging indicate impaired cortical top-down processing in association with anesthetic-induced unconsciousness. Anesthesiology 119, 1031–1042 (2013).
13 Travis, A. C., Pievsky, D. & Saltzman, J. R. Endoscopy in the elderly. Am. J. Gastroenterol. 107, 1495–1501 (2012).
14 Triantafillidis, J. K. Sedation in gastrointestinal endoscopy: current issues. World J. Gastroenterol. 19, 463 (2013).
15 Smit, D. J. A., Geus, E. J. C. De, Boersma, M., et al. Life-span development of brain network integration assessed with Phase Lag Index connectivity and minimum spanning tree graphs. Brain Connect. 8, 1–38 (2016).
16 Fried, L., Tangen, C., Walston, J., et al. Frailty in older adults: evidence for a phenotype. J Gerontol. 56A, M146–M156 (2001).
17 Shafi, M. M., Santarnecchi, E., Fong, T. G., et al. Advancing the neurophysiological understanding of delirium. J. Am. Geriatr. Soc. 65, 1114-1118 (2017).
18 Inouye, S. K., Westendorp, R. G. & Saczynski, J. S. Delirium in elderly people. Lancet 383, 911–922 (2014).
19 Rudolph, J. L. & Marcantonio, E. R. Review articles: postoperative delirium: acute change with long-term implications. Anesth. Analg. 112, 1202–1211 (2011).
20 Koster, S., Hensens, A. G., Schuurmans, M. J., et al. Risk factors of delirium after cardiac surgery. Eur. J. Cardiovasc. Nurs. 10, 197–204 (2011).
21 Sessler, C. N., Gosnell, M. S., Grap, M. J., et al. The Richmond agitation–sedation scale. Am. J. Respir. Crit. Care Med. 166, 1338–1344 (2002).

22 Van Dellen, E., van der Kooi, A. W., Numan, T., et al. Decreased functional connectivity and disturbed directionality of information flow in the electroencephalography of intensive care unit patients with delirium after cardiac surgery. Anesthesiology 121, 328–335 (2014).
23 Fraschini, M., Demuru, M., Crobe, A., et al. The effect of epoch length on estimated EEG functional connectivity and brain network organisation. J. Neural Eng. 13, 036015 (2016).
24 Van Wijk, B. C. M., Stam, C. J. & Daffertshofer, A. Comparing brain networks of different size and connectivity density Uuing graph Theory. PLoS One 5, e13701 (2010).
25 Hjorth, B. An on-line transformation of EEG scalp potentials into orthogonal source derivations. Electroencephalogr. Clin. Neurophysiol. 39, 526–530 (1975).
26 Whitham, E. M., Pope, K. J., Fitzgibbon, S. P., et al. Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. Clin. Neurophysiol. 118, 1877–1888 (2007).
27 Pope, K. J., Fitzgibbon, S. P., Lewis, T. W., et al. Relation of gamma oscillations in scalp recordings to muscular activity. Brain Topogr. 22, 13–17 (2009).
28 Stam, C. J., Nolte, G. & Daffertshofer, A. Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. Hum. Brain Mapp. 28, 1178–1193 (2007).
29 Paluš, M. & Vejmelka, M. Directionality of coupling from bivariate time series: How to avoid false causalities and missed connections. Phys. Rev. E 75, 056211 (2007).
30 Lobier, M., Siebenhühner, F., Palva, S., et al. Phase transfer entropy: a novel phase-based measure for directed connectivity in networks coupled by oscillatory interactions. Neuroimage 85 Part 2, 853–872 (2014).
31 Hillebrand, A., Tewarie, P., van Dellen, E., et al. Direction of information flow in large-scale resting-state networks is frequency-dependent. Proc. Natl. Acad. Sci. 113, 3867–3872 (2016).
32 Kruskal, J. B. On the shortest spanning subtree of a graph and the traveling salesman problem. Am. Math. Soiety 7, 48–50 (1956).
33 Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52, 1059–1069 (2010).
34 Hotz, M. a, Ritz, R., Linder, L., et al. Auditory and electroencephalographic effects of midazolam and alpha-hydroxy-midazolam in healthy subjects. Br. J. Clin. Pharmacol. 49, 72–79 (2000).
35 Nishida, M., Zestos, M. M. & Asano, E. Spatial–temporal patterns of electrocorticographic spectral changes during midazolam sedation. Clin. Neurophysiol. 127, 1223–1232 (2016).
36 Breimer, L. T. M., Hennis, P. J., Burm, A. G. L., et al. Quantification of the EEG Effect of Midazolam by Aperiodic Analysis in Volunteers. Clin. Pharmacokinet. 18, 245–253 (1990).
37 Boveroux, P., Vanhaudenhuyse, A., Bruno, M.-A., et al. Breakdown of within- and between-network resting state during propofol-induced loss of consciousness. Am. Soc. Anesthesiol. 113, 1038–1053 (2010).
38 Albrecht, S., Ihmsen, H., Hering, W., et al. The effect of age on the pharmacokinetics and pharmacodynamics of midazolam. Clin. Pharmacol. Ther. 65, 630–639 (1999).
39 Zaal, I. J. & Slooter, A. J. C. Delirium in critically ill patients: epidemiology, pathophysiology, diagnosis and management. Drugs 72, 1457–1471 (2012).
40 Zaal, I. J., Devlin, J. W., Hazelbag, M., et al. Benzodiazepine-associated delirium in critically ill adults. Intensive Care Med. 41, 2130–2137 (2015).

SUPPLEMENTARY MATERIAL

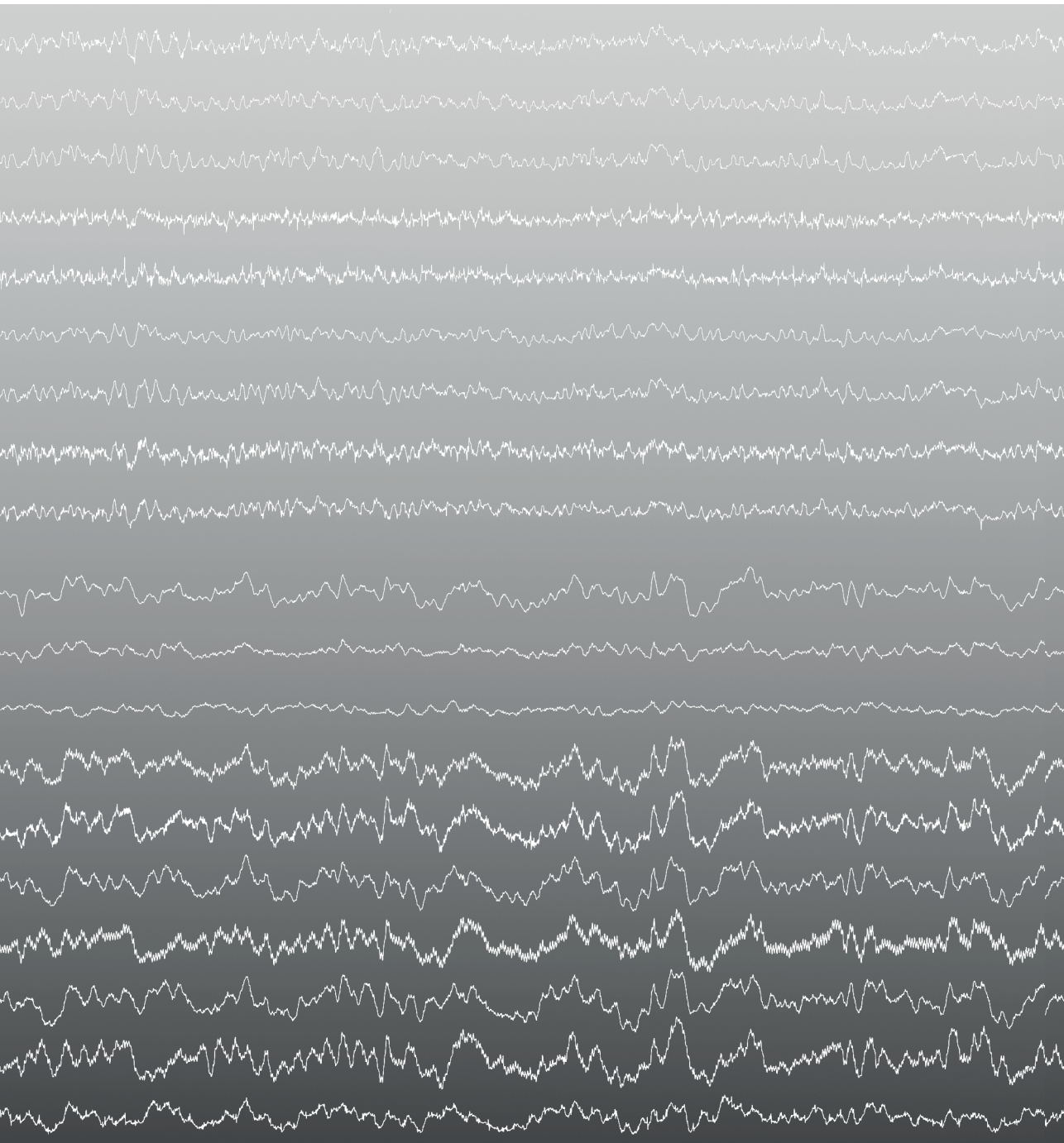
Table S4.1. Minimum spanning tree measure and definitions.

MST measure	Definition
Degree	The degree of each node is calculated as the ratio between number of connections divided between the maximum number of connections (i.e., 16 connections). The global degree is defined as the highest degree within the MST.
Leaf fraction (Lf)	The leaf fraction is a global measure defined as the number of leaf nodes (i.e., nodes with only 1 connection) divided by the maximum number of nodes within the MST (i.e., 17 nodes).
Diameter (D)	The diameter is a global measure and defined as the longest distance between any two nodes in the network.
Betweenness centrality (BC)	The number of paths crossing this node, divided by the total number of paths within the MST (i.e., connections). The maximum BC is used global MST measure. BC value of 0 indicates a leaf node, maximum BC is 1 which is the central node in a star-like network.
Eccentricity (Ecc)	The maximum distance between a node an any other node in the MST network. The global eccentricity is the range between the highest Ecc and lowest Ecc.
Tree hierarchy (Th)	The tree hierarchy is a global measure ($Th=Lf/(2mBC_{max})$) and describes the extent of hierarchical organization of the MST. A star-like tree has an Th of 0.5, for a path-like tree the Th value decreases towards 0. MST structures with a hierarchical organization can have a Th value up to 1. m is number paths in MST (i.e., 16 connections).

Table S4.2. Difference between effect of midazolam and effect of propofol.

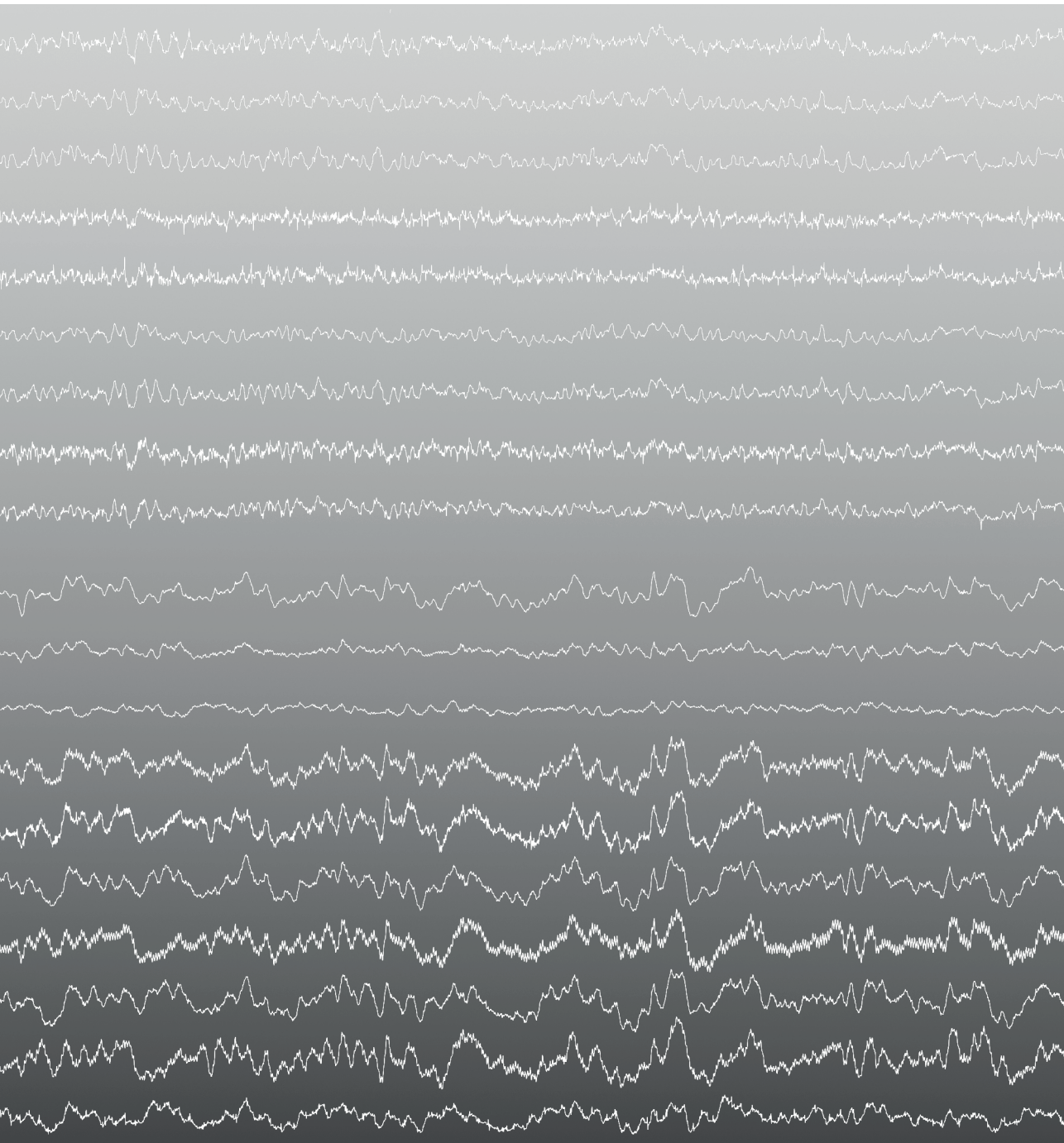
Measure	Effect of Midazolam (n=22)	Effect of Propofol (n=26)	F-test	p-value
Δ Relative power				
Delta	0.167 ± 0.128	0.252 ± 0.216	F(1,45) = 5.10	0.029*
Theta	-0.037 ± 0.081	-0.131 ± 0.144	F(1,45) = 2.01	0.163
Alpha	-0.183 ± 0.099	-0.144 ± 0.142	F(1,45) = 0.34	0.562
Beta	0.052 ± 0.074	0.023 ± 0.106	F(1,45) = 3.37	0.073
Δ Functional connectivity, Δ PLI				
Delta	0.009 ± 0.025	0.025 ± 0.035	F(1,45) = 4.88	0.032*
Theta	0.001 ± 0.024	-0.019 ± 0.042	F(1,45) = 3.12	0.084
Alpha	-0.027 ± 0.058	-0.042 ± 0.058	F(1,45) = 4.72	0.035*
Beta	-0.001 ± 0.024	0.001 ± 0.013	F(1,45) = 1.03	0.316
Δ Directed connectivity, Δ AP-ratio				
Delta	-0.001 ± 0.026	-0.008 ± 0.031	F(1,45) = 4.16	0.047*
Theta	-0.002 ± 0.028	0.009 ± 0.026	F(1,45) = 3.09	0.086
Alpha	0.028 ± 0.035	0.038 ± 0.034	F(1,45) = 8.24	0.006*
Beta	0.002 ± 0.018	0.010 ± 0.020	F(1,45) = 3.25	0.162
Δ Topology, Δ MST-overlap				
Delta	-0.001 ± 0.055	-0.010 ± 0.059	F(1,45) = 0.19	0.664
Theta	-0.015 ± 0.069	-0.018 ± 0.053	F(1,45) = 0.68	0.415
Alpha	0.010 ± 0.076	-0.022 ± 0.081	F(1,45) = 9.55	0.003*
Beta	-0.009 ± 0.094	-0.032 ± 0.060	F(1,45) = 1.52	0.224

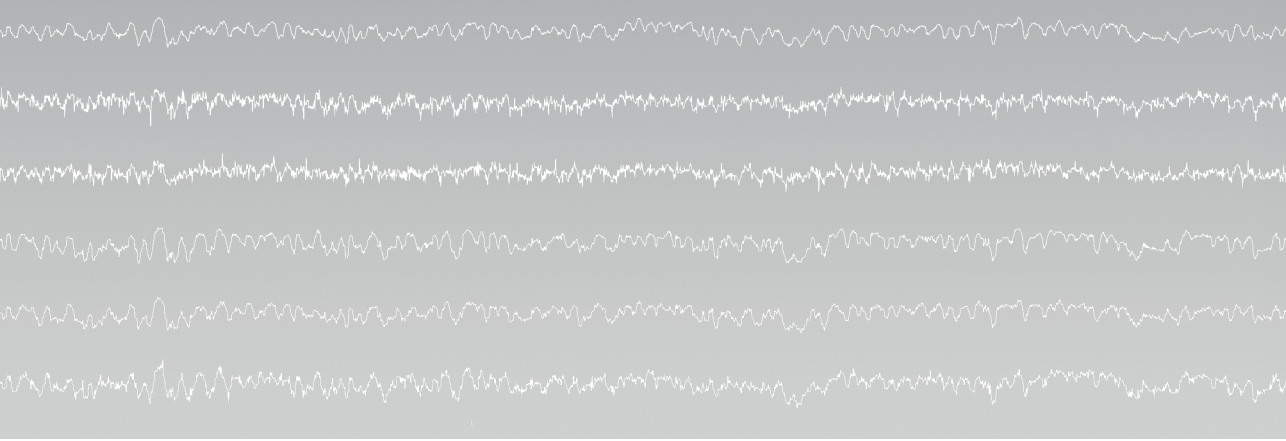
*These analyses, without correction for RASS score, were performed as sensitivity analysis in addition to the analyses corrected for RASS. The differences in relative delta power, functional connectivity, anterior-posterior ratio and MST-overlap measures were calculated by sedation-baseline and presented as mean ± standard deviation. Difference between effect of midazolam and effect of propofol were evaluated with linear regression model with sedative as fixed factor and baseline measure as covariate. AP-ratio = anterior-posterior ratio, MST = minimum spanning tree, PLI = phase lag index. *Significantly different (p<0.05).*

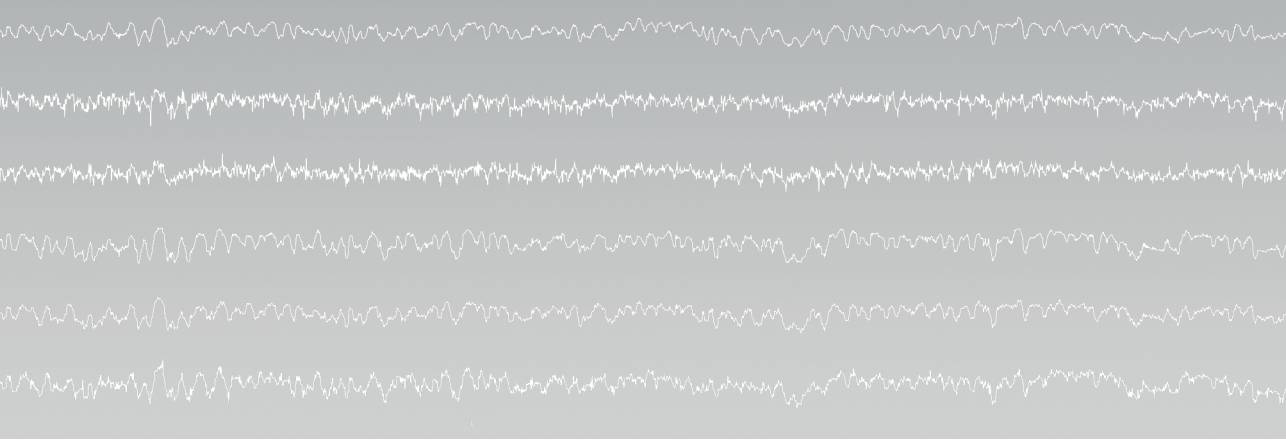


PART 2

VALIDATION OF AN EEG-BASED DELIRIUM MONITOR







5

RECOGNITION OF DELIRIUM IN POSTOPERATIVE ELDERLY PATIENTS: A MULTICENTER STUDY

T. Numan
M. van den Boogaard
A.M. Kamper
P.J.T. Rood
L.M. Peelen
A.J.C. Slooter
for the Dutch Delirium Detection Study Group

Journal of the American Geriatrics Society, epub ahead of print

ABSTRACT

OBJECTIVE: To evaluate to what extent delirium experts agree on the diagnosis of delirium when assessing independent exactly the same information. Secondly to evaluate the sensitivity of delirium screening tools in routine daily practice by clinical nurses.

DESIGN: Prospective observational longitudinal study.

SETTING: Multicenter study in three medical centers in the Netherlands.

PARTICIPANTS: Elderly postoperative patients (n=167).

MEASUREMENTS: Patients were daily (postoperative day 1-3) examined for delirium by a researcher using standardized cognitive assessment and interview including the delirium rating scale revised version-98 as global impression without any cut-off values, which was recorded on video. The videos and clinical information over the last 24 hours from the patients' record were evaluated by two delirium experts independently, who classified each assessment as delirious, possible delirious, or non-delirious. Interrater agreement between delirium experts was determined using weighted Cohen's kappa. When there was no consensus, a third expert was consulted. Final classification was based on median score and compared to the results of the confusion assessment method for intensive care unit and delirium observation screening scale that clinical nurses administered.

RESULTS: We included 424 postoperative assessments of 167 patients. The overall kappa was κ=0.61 (95% confidence interval 0.53-0.68). In 89 (21.0%) assessments there was no agreement between the experts, and a third delirium expert was needed for the final classification. Assessments performed by nurses detected 32% of the assessments which were diagnosed as (possible) delirious by the experts.

CONCLUSION: There was considerable disagreement in classification of delirium by experts who independently assessed exactly the same information, showing the difficulty of delirium diagnosis. Furthermore, the sensitivity of daily delirium screening by clinical nurses was poor. Future research should focus on the development of objective instruments to diagnose delirium.

INTRODUCTION

Delirium is characterized by an acute disturbance of attention and awareness, with additional changes in cognition that tend to fluctuate over time[1]. The incidence of delirium during hospital stay is high, especially in elderly postoperative and critically ill patients[2–4]. Consequences of delirium are longer intensive care unit (ICU) and hospital stay[5–7], increased health care costs[2,8,9] and deterioration of long-term cognitive function[4,7,8,10]. Therefore, delirium is a serious health care problem.

The reference standard for diagnosing delirium is the assessment by a delirium expert: geriatrician, psychiatrist, or neurologist, usually using the criteria of the Diagnostic and Statistical Manual of Mental Disorders (DSM)-5[1]. Delirium diagnosis may be based on one or a combination of clinical impression, cognitive testing, clinical notes, physical examination, and lab results, which was found to differ greatly for the reference standard in literature[11]. The consistency of the delirium diagnosis is however currently unclear, as it has never been studied to what extent different delirium experts agree on the delirium diagnosis when independently assessing exactly the same information.

In clinical practice, it is infeasible for delirium experts to examine all patients. Therefore, delirium screening tools have been developed, and routine delirium screening is recommended by several organizations amongst others the American Geriatrics Society[12] and the Society of Critical Care Medicine[13]. The feasibility and interrater reliability of delirium instruments has been investigated in several studies[14–18]. Importantly, almost all of these studies were performed in a research setting where dedicated researchers performed the assessments. Further, most of these studies were performed in centers where the screening tools were developed. The sensitivity of delirium screening in routine, clinical practice by clinical nurses in ICU patients was low (overall 47% using the confusion assessment method for the ICU (CAM-ICU))[19], as well as in non-ICU patients when using the CAM, was also low, 25%[20].

The aim of this study was to determine agreement in diagnosing delirium by delirium experts when independently confronted with exactly the same clinical information and to determine the sensitivity of delirium screening tools in routine daily practice by clinical nurses.

METHODS

DESIGN, SETTING AND STUDY POPULATION

This prospective, multicenter observational study was performed in three centers: University Medical Center Utrecht (UMCU), Radboud University Medical Center Nijmegen, and Isala Klinieken Zwolle, a nonacademic teaching hospital (clinicaltrials.gov identifier NCT02404181). The current study was performed in the context of the validation of an electroencephalogram (EEG)-based delirium monitor, which was approved by the medical ethical board of the UMCU (protocol number 13-643).

Patients scheduled for major surgery, (i.e., expected hospital stay of ≥2 days) aged 60 years and older and considered to be at risk for delirium were informed about the study and provided written informed consent. Patients were considered to be at risk of delirium if they were aged 75 and older, had a history of transient ischemic attack (TIA), stroke, alcohol abuse, depression, dementia, delirium, cognitive problems, or severe cardiac or respiratory disease[2]. Exclusion criteria were inability to speak Dutch or English, deafness, neurological surgery, and inability to perform EEG-based delirium monitor recording.

REFERENCE STANDARD DELIRIUM DIAGNOSIS

Patients were examined by trained researchers preoperatively (time point T-1) and the first three consecutive days after surgery (T1, T2 and T3). Patients were interviewed using the 13-severity items of the delirium rating scale revised edition '98 (DRS-R-98)[21,22] and the CAM-ICU[15]. The DRS-R-98 is a specialist instrument designed for psychiatrists or other professionals with substantial training in mental health assessment (see supplementary Table S5.1 for scoring). Items were rated between 0 and 3, yielding a total DRS-R-98 score ranging between 0 and 39, higher scores indicate more severe delirium symptoms. No predefined cut-off values were used to ensure unbiased delirium diagnosis by the experts. All interviews including cognitive assessment were recorded on video and took approximately 10-15 minutes. When patients refused to participate on a specific recording day, no video assessment was performed and no formal expert diagnosis was available for that day.

Delirium experts from academic and nonacademic hospitals in the Netherlands evaluated the videos. The experts were psychiatrists (n=17), geriatricians (n=15), neurologists (n=4), neuropsychologist (n=1), or nurse-scientist (n=1) with a median of 11.5 years (interquartile range; 1st-3rd quartile (IQR) 6.3-19.0) of experience and

saw an estimated median of 15 (IQR 10-25) patients with delirium monthly. All experts received written instruction about the evaluation procedure. The experts evaluated all video recordings combined with patients' medical and nursing record information of the last 24 hours, and filled in the DRS-R-98. Each patient was assessed by two delirium experts, who were blinded to each other's evaluations, with different combinations of delirium experts in all patients.

Two classifications had to be made: one based solely on the cognitive testing recorded on video, and one based on all information over the last 24 hours including the description in the medical and nursing files and the video. Both classifications were based on DSM-5 criteria. For each classification a likelihood of the patient having delirium on a numeric rating scale (NRS: 0=definite no delirium, 10=definite delirium), as well as a final diagnosis: 'no delirium', 'possible delirium', or 'delirium' had to be reported. When the patient was classified as having (possible) delirium, the motor subtype (hypoactive, hyperactive, or mixed) was reported.

Classifications of the two delirium experts were compared. If there was no consensus on either the video recording or all information in the last 24 hours including the video recording, a third expert was consulted. The third expert was blinded for the classification of the first two experts, but knew that he/she was consulted because of disagreement. Final classification and motor subtype were based on the median score of the three experts. The final NRS and DRS-R-98 scores of the two or three (when a third expert was consulted) delirium experts were averaged.

DELIRIUM SCREENING BY THE CLINICAL NURSE

In the participating hospitals, the CAM-ICU and delirium observation screening (DOS) scale were used in daily practice for delirium assessment two or three times a day, in the ICU and general wards, respectively. The DOS scale consists of 13 items (supplementary Table S5.2)[17,23]. Patients' record was reviewed for available documentation of the CAM-ICU and/or DOS-score for the last 24 hours before the video recording. In patients who were transferred from the ICU to the general ward, both scores could have been administered in the last 24 hours. The highest DOS score per day was used, when both positive and negative CAM-ICU scores were documented, the positive score was used. A combined clinical nurse score was calculated and defined positive when either one or both CAM-ICU and DOS scores were positive. Clinical nurses and physicians were blinded for the classification of the delirium experts.

OTHER DATA COLLECTION

The following baseline characteristics were recorded: age, sex, alcohol consumption (no alcohol, 1-14, or >14 alcohol units per week), transient ischemic attack (TIA) or stroke in the medical history, preexisting psychiatric disease, and previous cognitive problems defined by mini-mental state examination (MMSE) as assessed preoperatively. Furthermore, characteristics of the surgical procedure were documented, including type and duration of surgery.

STATISTICAL ANALYSIS

To assess the interrater agreement between two experts for all postoperative assessments, we used linear weighted Cohen's kappa statistic, separately for the classification based on cognitive testing on video only, and for the classification based on all information over the last 24 hours that included cognitive testing on video. To assess reliability, the intraclass correlation coefficient was calculated, based on a single measure two-way random model, between the ratings of the two delirium experts for the likelihood (NRS) and severity (DRS-R-98) scores.

Furthermore, variability in NRS and DRS-R-98 scores between the two delirium experts was evaluated for the motor subtypes by calculating the average and the absolute

Table 5.1. Baseline characteristics of the study population.

	All patients (n=167)
Age (years), mean (SD)	76.8 (6.4)
Male, n (%)	113 (67.7%)
Alcohol consumption*, n (%)	
None	64 (38.3)
1-14 per week	73 (43.7)
> 14 per week	26 (15.6)
Medical history, n (%)	
TIA or Stroke	47 (28.1)
Psychiatric disease	11 (6.6)
MMSE*, median (range)	28 (14-30)
Surgery type, n (%)	
Cardiothoracic or vascular	147 (88.6)
Orthopedic	12 (7.2)
Other	7 (4.2)
Duration of surgery (min.) *, median (range)	169 (20-401)

**Number of missings: alcohol consumption 4 (2.4%), MMSE 43 (25.7%), duration of surgery 4 (4.2%). Surgery type other: otorhinolaryngology (n=3), gastrointestinal or general surgery (n=3), urologic (n=1). MMSE = mini-mental state examination, n = number, SD = standard deviation.*

difference of the NRS and DRS-R-98 scores between the initial two experts. These scores were compared between the different motor subtypes using the Mann Whitney U test.

Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the CAM-ICU, DOS, and the combined clinical nurse score were calculated using the classification of delirium experts based on all information over the last 24 hours including cognitive testing as reference. For these calculations, 'possible delirium' and 'delirium' were combined and compared to 'no delirium'. To assess the effect of including 'possible delirium' in the delirium group, we performed a sensitivity analysis where 'possible delirium' was included in the group without delirium.

All postoperative assessments were included as independent observations. To assess the influence of repeated measures, we recalculated the diagnostic values of the CAM-ICU scored by clinical nurses based on T1 only. P-values <0.05 were assumed statistically significant. Statistical analyses were performed in IBM SPSS Statistics (Version 21).

RESULTS

Of 196 patients from whom informed consent was obtained, 29 (14.8%) were excluded for logistical reasons. The remaining study population consisted of 167 patients, in whom 424 postoperative delirium assessments were performed. Of the 501 potential postoperative assessments, 31 (6.2%) were excluded because the patient refused participation on that day and 46 (9.2%) because of logistical reasons, for example discharge. None of the patients were intubated at the time of cognitive testing.

POPULATION CHARACTERISTICS

Baseline characteristics of the study population are presented in Table 5.1. The vast majority of patients underwent cardiothoracic surgery (88.6%). The median preoperative MMSE score was 28 (range 14-30), 10 patients had an MMSE below 24. Assessments were performed in the ICU (n=71, 16.7%), medium care/coronary care unit (n=93, 21.9%), and general ward (n=260, 61.3%).

INTERRATER RELIABILITY OF THE DIAGNOSIS OF DELIRIUM

Of the 424 assessments, 344 (81.5%) were classified similarly by both experts into 'delirium', 'possible delirium', or 'no delirium' based on cognitive testing only (Table 5.2A), two observations were excluded due to an incomplete video. In 78 (18.5%)

assessments there was disagreement between the experts on the classification of video-recorded cognitive testing. This resulted in κ=0.53 (95% CI 0.44-0.62).

The extent of agreement between experts classifying delirium based on all information over the last 24 hours including cognitive testing, resulted in κ=0.61 (95% CI 0.53-0.68, Table 5.2B). In the 89 (21.0%) assessments with disagreement based on all information, a third expert was consulted to obtain final classification. Based on all information including cognitive testing recorded on video, 50 (11.8%) postoperative assessments were classified as delirious and 50 (11.8%) as possible delirious, which were combined as 100 (23.6%) positive delirious assessments. The other 324 (76.4%) assessments were classified as no delirium. This corresponded with 31 (18.6%) patients classified

Table 5.2. Interrater variability of the classification of delirium by delirium experts.

A		**Expert X**		
Classification: cognitive tests only (422 assessments*)		No delirium	Possible delirium	Delirium
Expert Y	No delirium	313 (74.2%)	-	-
	Possible delirium	45 (10.7%)	19 (4.5%)	-
	Delirium	9 (2.1%)	24 (5.7%)	12 (2.8%)

B		**Expert X**		
Classification: all information in last 24 hours (424 assessments)		No delirium	Possible delirium	Delirium
Expert Y	No delirium	286 (67.5%)	-	-
	Possible delirium	49 (11.6%)	16 (3.8%)	-
	Delirium	11 (2.6%)	29 (6.8%)	33 (7.8%)

*In each patient the two initial experts classified the patients as having 'no delirium', 'possible delirium' or 'delirium' based on (A) cognitive tests recorded on video (weighted kappa 0.53 (95% CI 0.44-0.62)) and based on (B) all information in the last 24 hours, including the cognitive testing recorded on video (weighted kappa 0.61 (95% CI 0.53-0.68)). *The videos of two postoperative assessments were incomplete and therefore not included in this comparison.*

as delirious and 30 (18.0%) patients classified as possible delirious in the first three postoperative days. Median DRS-R-98 score was 3.0 (IQR 2.0-4.0) for patients without delirium, 7.7 (IQR 6.3-12.7) for those with possible delirious, and 12.7 (IQR 9.6-17.5) for those with delirium. Further analyses were performed using the final delirium expert classification based on all information over the last 24 hours.

VARIABILITY OF THE LIKELIHOOD AND SEVERITY OF DELIRIUM

The intraclass correlation coefficient between the scores of the first two delirium experts was 0.73 (95% CI 0.68-0.77) for NRS and 0.76 (95% CI 0.72-0.80) for DRS-R-98 (Figure 5.1). The median absolute difference in NRS scores of both experts was 1 (IQR 0-2), the median absolute difference in DRS-R-98 scores was also 1 (IQR 1-3). Agreement between the first two experts on the individual DRS-R-98 items is shown in supplementary Table S5.3.

The NRS scores that the experts assigned for the likelihood of the patient being delirious did not differ between hypoactive (median 6.3 (IQR 4.4-7.2)), mixed (median 6.2 (IQR 4.6-9.0)), and hyperactive (median 7.0 (IQR 5.5-9.0), p=0.18) subtypes of delirium. The DRS-R-98 that the experts assigned for the severity of delirium did not differ over the three motor subtypes either (hypoactive: median 10.0 (IQR 7.8-13.3), mixed: median 9.7 (IQR 6.9-12.2), hyperactive: median 13.0 (IQR 6.5-20.3), p=0.33).

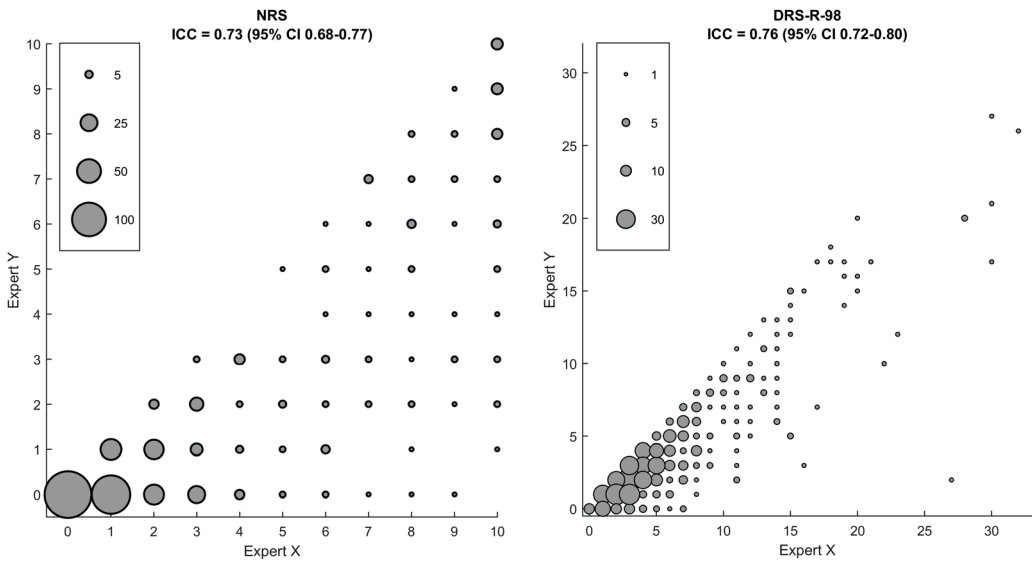


Figure 5.1. Scores of first two delirium experts of the likelihood (numeric rating scale, NRS, left panel) and severity of delirium (DRS-R-98, right panel) based on all information within the last 24 hours including cognitive testing as recorded on video.

Expert X represents the highest NRS/DRS-R-98 score of the two experts, Expert Y had the lowest NRS/DRS-R-98 score of the two experts. The size of the circles indicates the number of postoperative assessments with these scores. ICC = intraclass correlation coefficent.

There was more disagreement on the likelihood of the patient being delirious between the experts in hypoactive delirium than in the other types of delirium: the absolute difference of the NRS score between the initial two experts was significantly higher in the hypoactive subtype compared to mixed (p=0.01) and hyperactive subtypes (p=0.03). No differences in assigned DRS-R-98 scores were found between the delirium subtypes (p=0.45), see Table 5.3.

CLINICAL USE OF DELIRIUM SCREENING TOOLS

Of the 424 included assessments, in 117 (27.6%) only the CAM-ICU was performed, in 87 (20.5%) only the DOS was documented, in 14 (3.3%) both CAM-ICU and DOS were performed, and in 206 (48.6%) no clinical screening was documented. The combined score (i.e., either the CAM-ICU and/or DOS) was performed in 218 (51.4%) delirium assessments. In 32 of the 100 delirium assessments that were classified by the experts as (possible) delirium, the CAM-ICU or DOS was positive. Screening by clinical nurses was performed in 44.4% , 60.0% and 82.0% of the non-delirious, possible delirious and delirious assessments, respectively.

The CAM-ICU was performed in 44 assessments with (possible) delirium classification by the delirium experts, which was positive in 11 (25.0%) of these assessments. The DOS was documented in 37 (possible) delirium assessments and positive in 23 (62.2%) of the assessments (Table 5.4). Combining the two delirium screening tools resulted in a sensitivity of 43.2% (95% CI 31.6-55.2%) and specificity of 97.2% (95% CI 92.5-99.1%).

Table 5.3. Absolute difference in likelihood (NRS) and severity (DRS-R-98) of delirium between two initial experts.

Absolute difference between experts	Hypoactive (29 assessments)	Mixed (58 assessments)	Hyperactive (13 assessments)	p-value
Likelihood of delirium, (NRS)	3.0 (1.5-5.0)	2.0 (1.0-3.3)	1.0 (0.0-3.5)	0.01[a,b]
Severity of delirium, (DRS-R-98)	3.0 (1.0-8.0)	3.0 (1.8-5.3)	2.0 (0.5-5.0)	0.45

Motor subtype classification based on two (or majority vote of three) delirium experts based on all available information over the last 24 hours including cognitive testing recorded on video. Both 'possible delirium' and 'delirium' classified assessments were included. The median absolute difference (interquartile range) of the likelihood and severity of delirium were assessed with the Kruskal Wallis test, post-hoc Mann-Whitney U test showed; [a] significant difference between hypoactive delirium and mixed type, [b] significant difference between hypoactive and hyperactive delirium. DRS-R-98 = delirium rating scale revised in '98, averaged score between 0 and 39 for the severity of delirium; NRS = numeric rating scale, averaged score between 0 and 10 for the likelihood of a patient being delirious.

Table 5.4. Evaluation of clinical nurse delirium screening in clinical practice.

Delirium screening method	Sensitivity (% (95% CI))	Specificity (% (95% CI))	PPV (% (95% CI))	NPV (% (95% CI))
CAM-ICU (131 assessments, 33.6% (possible) delirious‡)	25.0 (13.7-40.6)	96.6 (89.5-99.1)	78.6 (48.8-99.1)	71.8 (62.6-79.5)
DOS (101 assessments, 36.6% (possible) delirious‡)	62.2 (44.8-77.1)	98.4 (90.4-99.9)	95.8 (90.4-99.9)	81.8 (71.0-89.4)
Combined CAM-ICU and/or DOS (218 assessments, 34.4% (possible) delirious‡)	43.2 (31.6-55.2)	97.2 (92.5-99.1)	88.9 (73.0-96.4)	76.9 (70.0-82.7)

‡Number of assessments classified as 'possible delirious' or 'delirious' by the final classification based on all information in the last 24 hours, including cognitive testing stored on video. In 14 assessments both CAM-ICU and DOS were documented, which are included in all three groups. DOS ≥3 was used as threshold for delirium detection. For combined CAM-ICU and/or DOS the highest score (either positive for the CAM-ICU or DOS ≥3) documented in the last 24 hours. CAM-ICU = confusion assessment method for intensive care unit, CI = confidence interval, DOS = delirium observation screening, NPV = negative predictive value, PPV = positive predictive value.

The diagnostic performance of the clinical nurses with screening improved slightly when 'possible delirium' was grouped as 'non-delirious' (sensitivity: 55.8%, specificity: 93.1%, PPV: 66.7%, and NPV: 89.6%).

Finally, the diagnostic performance of the CAM-ICU was calculated for T1 only, to exclude effects of repeated measurements. This resulted in a sensitivity of 21.9% (95% CI 9.9-40.4%), a specificity of 95.7% (95% CI 87.2-98.9%), PPV was 70.0% (95% CI 35.4-91.9%), and NPV 72.8% (95% CI 62.4-81.3%), which did not show major differences compared to the main analysis over all postoperative days.

DISCUSSION

In this study we found a considerable amount of disagreement in the classification of delirium by experts who independently assessed exactly the same information. The variability between experts in the likelihood of a patient being delirious was highest in the hypoactive motor subtype. Recognition of delirium by clinical nurses using screening appeared to be insufficient with not more than 32% of the postoperative cases identified.

RECOGNITION OF DELIRIUM IN CLINICAL PRACTICE

The CAM-ICU and DOS have been validated in research settings with good diagnostic properties[15,17,24,25], however not all subsequent studies were able to reproduce these high diagnostic values[14,19,26]. In this study we found a sensitivity of 25.0% and 62.2% for CAM-ICU and DOS, respectively, indicating under-detection of delirium in clinical practice. Cautious interpretation is needed, because only cases were included when the results of the delirium screening tool administered by clinical nurses was documented in the patient record, which was present in half of the study population. We did not investigate underlying reasons of this rather low number of documented screening tools. It is not clear how this might have influenced the diagnostic values. Another explanation of the low diagnostic performance is the included 'possible delirium' assessments, which might not fulfill all criteria and therefore result in false negative scores. Repeating the analysis with 'possible delirium' as negative assessment resulted in a slightly higher sensitivity of 55.8%, although still not sufficient for clinical practice.

STRENGTHS AND LIMITATIONS

Strengths of this study are the prospective, multicenter setting, in academic and nonacademic hospitals, and large sample size. The first assessment was performed before surgery, facilitating the classification of delirium by the reference standard. Cognitive testing was extensive and standardized. As we recorded these assessments, we had a unique opportunity to investigate to what extent there is agreement of the delirium diagnosis by different delirium experts truly independent of each other. The reference standard was robust, as it was based on all available information compatible with usual clinical practice, and not based on a cut-off score of the DRS-R-98. Moreover, consensus by two delirium experts was required, and in case of disagreement, a third expert was consulted. To account for the fluctuating nature of delirium, experts had excess to the medical and nursing files in addition to the videos with the interview including cognitive assessment.

Some potential limitations need to be addressed. First, the category 'possible delirium' may have been difficult to interpret and may have influenced the interrater agreement. Second, each postoperative assessment was analyzed as individual case and therefore the effect of repeated measures within patients was not taken into account. However, sensitivity analysis evaluating the diagnostic values of the CAM-ICU based on one measure of each patient (T1), showed similar results. Third, 31 (6.2%) formal assessments could not be performed because the patient refused participation, which may have led to

underdetection of delirium. However, we do not recall cases that showed signs of delirium based on clinical impression of the visiting researcher or the patient record. Finally, agreement between the experts and detection rate by clinical nurses might have been higher when the included patients had more severe delirium.

RECOMMENDATIONS FOR FUTURE STUDIES

There is a clear need for an easy-to-use, reliable method to detect delirium in clinical practice. Most of the current tools are based on observation, simple tests[27], or a combination of both approaches[15]. Several cognitive tests have been described and validated in the literature, such as the abbreviated cognitive test for delirium[28], Edinburgh delirium test box[29,30], or smartphone application[31]. These tools showed good performance, but interpretation by the performer is needed, and interaction with the patient is essential but not possible in all patients, for example due to depressed levels of consciousness. Therefore, an objective tool is needed which is applicable in every patient and unambiguous to interpret. Bipolar EEG may be a candidate for objective delirium detection[32]. One minute recording without artifacts was shown to be sufficient to distinguish delirious from non-delirious patients[32]. It should, however, be noted that this study was performed in patients with either definite delirium or definite no delirium, findings should be confirmed in an independent study population. It is not clear whether this approach could replace assessment of all aspects of delirium, or whether some features (e.g., hallucinations or distress) would still require interaction.

CONCLUSION

This study showed a considerable amount of disagreement in the classification of delirium by experts who independently assessed exactly the same information, indicating the difficulty of delirium diagnosis. Moreover, delirium recognition in clinical practice by clinical nurses was poor and a relative high number of delirious cases remained unrecognized despite the use of delirium screening tools. An objective and reliable screening tool may improve the detection of delirium.

ACKNOWLEDGEMENTS

Financial disclosure: ESICM Clinical Research Award, Technology Foundation STW Take-off grant 14066, EU H2020-PHC-2014-672974 grant. The authors develop an EEG-based monitor for delirium detection in routine clinical practice. Any (future) profits from EEG-based delirium monitoring will be used for future scientific research.

REFERENCES

1 American Psychiatric Association. Diagnostics and statistical manual of mental disorders. 5th edition. (2013).
2 Inouye, S. K., Westendorp, R. G. & Saczynski, J. S. Delirium in elderly people. Lancet 383, 911–922 (2014).
3 Koster, S., Hensens, A. G., Schuurmans, M. J., et al. Risk factors of delirium after cardiac surgery. Eur. J. Cardiovasc. Nurs. 10, 197–204 (2011).
4 Wolters, A. E., van Dijk, D., Pasma, W., et al. Long-term outcome of delirium during intensive care unit stay in survivors of critical illness: a prospective cohort study. Crit. Care 18, R125 (2014).
5 Gleason, L. J., Schmitt, E. M., Kosar, C. M., et al. Effect of delirium and other major complications on outcomes after elective surgery in older adults. JAMA Surg. 150, 1134 (2015).
6 Salluh, J. I. F., Wang, H., Schneider, E. B., et al. Outcome of delirium in critically ill patients: systematic review and meta-analysis. BMJ 350, h2538 (2015).
7 Klein Klouwenberg, P. M. C., Zaal, I. J., Spitoni, C., et al. The attributable mortality of delirium in critically ill patients: prospective cohort study. BMJ 349: g6652, 1–10 (2014).
8 Rizk, P., Morris, W., Oladeji, P., et al. Review of postoperative delirium in geriatric patients undergoing hip surgery. Geriatr. Orthop. Surg. Rehabil. 7, 100–105 (2016).
9 Leslie, D. L., Marcantonio, E. R., Zhang, Y., et al. One-year health care costs associated with delirium in the elderly population. Arch. Intern. Med. 168, 27–32 (2008).
10 Fong, T. G., Davis, D., Growdon, M. E., et al. The interface between delirium and dementia in elderly adults. Lancet. Neurol. 14, 823–832 (2015).
11 Neufeld, K. J., Nelliot, A., Inouye, S. K., et al. Delirium diagnosis methodology used in research: a survey-based study. Am. J. Geriatr. Psychiatry 22, 1513–1521 (2014).
12 American Geriatrics Society Expert Panel on Postoperative Delirium in Older Adults. Postoperative delirium in older adults: best practice statement from the American Geriatrics Society. J. Am. Coll. Surg. 220, 136–148.e1 (2015).
13 Barr, J., Fraser, G. L., Puntillo, K., et al. Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit. Crit. Care Med. 41, 263–306 (2013).
14 Van Velthuijsen, E. L., Zwakhalen, S. M. G., Warnier, R. M. J., et al. Psychometric properties and feasibility of instruments for the detection of delirium in older hospitalized patients: a systematic review. International Journal of Geriatric Psychiatry 31, 974–989 (2016).
15 Ely, E. W., Margolin, R., Francis, J., et al. Evaluation of delirium in critically ill patients: validation of the confusion assessment method for the intensive care unit (CAM-ICU). Crit. Care Med. 29, 1370–1379 (2001).
16 Scheffer, A. C., van Munster, B. C., Schuurmans, M. J., et al. Assessing severity of delirium by the Delirium Observation Screening Scale. Int. J. Geriatr. Psychiatry 26, 284–291 (2011).
17 Gemert van, L. A. & Schuurmans, M. J. The Neecham confusion scale and the delirium observation screening scale: capacity to discriminate and ease of use in clinical practice. BMC Nurs. 6, 3 (2007).
18 De, J. & Wand, A. P. F. Delirium Screening: A systematic review of delirium screening tools in hospitalized patients. Gerontologist 55, 1079–1099 (2015).
19 Eijk, M. M. van, Boogaard, M. van den, Marum, R. J. van, et al. Routine use of the confusion assessment method for the intensive care unit: a multicenter study. Am. J. Respir. Crit. Care Med. 184, 340–344 (2011).

20 Rice, K. L., Bennett, M., Gomez, M., et al. Nurses' recognition of delirium in the hospitalized older adult. Clin. Nurse Spec. 25, 299–311 (2011).
21 Trzepacz, P. T., Mittal, D., Torres, R., et al. Validation of the delirium rating scale-revised-98: comparison with the delirium rating scale and the cognitive test for delirium. J. Neuropsychiatry Clin. Neurosci. 13, 229–242 (2001).
22 De Rooij, S. E., van Munster, B. C., Korevaar, J. C., et al. Delirium subtype identification and the validation of the delirium rating scale-revised-98 (Dutch version) in hospitalized elderly patients. Int. J. Geriatr. Psychiatry 21, 876–882 (2006).
23 Koster, S., Hensens, A. G., Oosterveld, F. G. J., et al. The delirium observation screening scale recognizes delirium early after cardiac surgery. Eur. J. Cardiovasc. Nurs. 8, 309–314 (2009).
24 Ely, E., SK, Inouye., GR, Bernard., et al. Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU). JAMA 286, 2703–2710 (2001).
25 Neto, A. S., Nassar, A. P., Cardoso, S. O., et al. Delirium screening in critically ill patients. Crit. Care Med. 40, 1946–1951 (2012).
26 Eijk, M. M. J. van, Marum, R. J. van, Klijn, I. A. M., et al. Comparison of delirium assessment tools in a mixed intensive care unit. Crit. Care Med. 37, 1881–1885 (2009).
27 Adamis, D., Meagher, D., Murray, O., et al. Evaluating attention in delirium: a comparison of bedside tests of attention. Geriatr. Gerontol. Int. 16. 1028-1035 (2016).
28 Hart, R. P., Best, A. M., Sessler, C. N., et al. Abbreviated cognitive test for delirium. J. Psychosom. Res. 43, 417–423 (1997).
29 Brown, L. J. E., Fordyce, C., Zaghdani, H., et al. Detecting deficits of sustained visual attention in delirium. J. Neurol. Neurosurg. Psychiatry 82, 1334–1340 (2011).
30 Tieges, Z., McGrath, A., Hall, R. J., et al. Abnormal level of arousal as a predictor of delirium and inattention: An exploratory study. Am. J. Geriatr. Psychiatry 21, 1244–1253 (2013).
31 Tieges, Z., Stíobhairt, A., Scott, K., et al. Development of a smartphone application for the objective detection of attentional deficits in delirium. Int. Psychogeriatrics 27, 1251–1262 (2015).
32 Van der Kooi, A. W., Zaal, I. J., Klijn, F. A., et al. Delirium detection using EEG. Chest 147, 94–101 (2015).

SUPPLEMENTARY MATERIAL

Table S5.1. Delirium rating scale revised-98 (DRS-R-98).

Item	Score patient				
1. Sleep-wake cycle	0	1	2	3	
2. Perceptual disturbances	0	1	2	3	
3. Delusions	0	1	2	3	
4. Lability of affect	0	1	2	3	
5. Language	0	1	2	3	□ Check here if intubated, mute, etc.
6. Thought process	0	1	2	3	
7. Motor agitation	0	1	2	3	□ Check here if restrained *Type of restraints:*
8. Motor retardation	0	1	2	3	
9. Orientation	0	1	2	3	Date: Place: Person:
10. Attention	0	1	2	3	
11. Short-term memory	0	1	2	3	□ Record # of trials for registration of items: Check here if category cueing helped
12. Long-term memory	0	1	2	3	
13. Visuospatial ability	0	1	2	3	□ Check here if unable to use hands
Total score:					

Total score DRS-R-98: sum score of the 13 items.

Table S5.2. Delirium observation screening (DOS) scale.

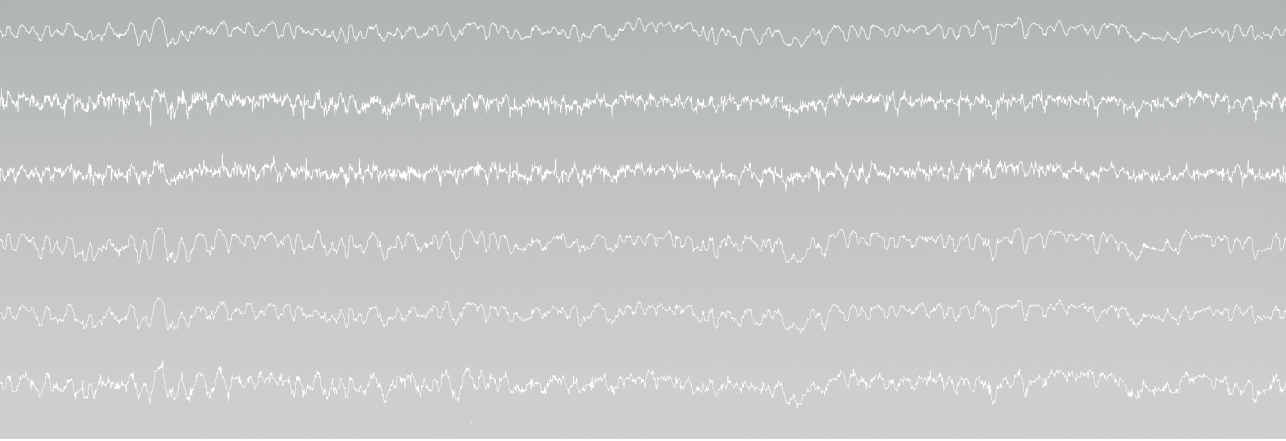
Item	Scoring	
	Never	Sometimes or always
1. Dozes during conversation or activities	0	1
2. Is easy distracted by stimuli from the environment	0	1
3. Maintains attention to conversation or action	1	0
4. Does not finish question or answer	0	1
5. Gives answers which do not fit the question	0	1
6. Reacts slowly to instructions	0	1
7. Thinks to be somewhere else	0	1
8. Knows which part of the day it is	1	0
9. Remembers recent event	1	0
10. Is picking, disorderly, restless	0	1
11. Pulls IV tubes, feeding tubes, catheters etc.	0	1
12. Is easy or sudden emotional (frightened, angry, irritated)	0	1
13. Sees persons/things as somebody/something else	0	1
Total score (sum):		

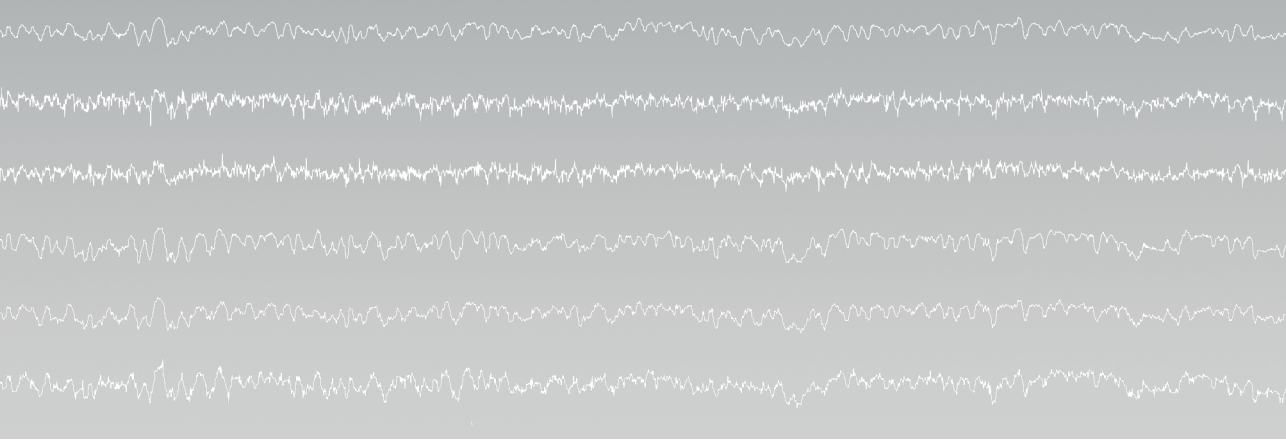
The items of the delirium observation screening score (DOS) are depicted. Each item is scored 0 or 1. The cut-off value for delirium is a sum score of 3 and higher.

Table S5.3. Agreement between the first two experts on each individual DRS-R-98 item.

DRS-R-98 item	Weighted kappa	Number of missings
1. Sleep-wake cycle	0.39 (95% CI 0.32-0.46)	6
2. Perceptual disturbances	0.65 (95% CI 0.55-0.75)	9
3. Delusions	0.45 (95% CI 0.28-0.62)	4
4. Lability of affect	0.26 (95% CI 0.15-0.37)	7
5. Language	0.37 (95% CI 0.27-0.48)	11
6. Thought process	0.39 (95% CI 0.27-0.51)	19
7. Motor agitation	0.42 (95% CI 0.29-0.55)	9
8. Motor retardation	0.49 (95% CI 0.37-0.61)	8
9. Orientation	0.63 (95% CI 0.55-0.72)	12
10. Attention	0.43 (95% CI 0.36-0.50)	10
11. Short-term memory	0.66 (95% CI 0.60-0.71)	18
12. Long-term memory	0.49 (95% CI 0.37-0.61)	17
13. Visuospatial ability	0.51 (95% CI 0.42-0.59)	38

The total number of included assessments was n=424. CI = confidence interval





6

DELIRIUM DETECTION BASED ON A BRIEF, SINGLE-CHANNEL QUANTITATIVE ELECTROENCEPHALOGRAPHY RECORDING IN OLDER POSTOPERATIVE PATIENTS: A PROSPECTIVE, MULTICENTER VALIDATION STUDY

T. Numan
M. van den Boogaard
A.M. Kamper
P.J.T. Rood
A. Wassenaar
M. Abawi
J.A.H.R. Claassen
M. Coesmans
P.L.J. Dautzenberg
T.A.D.F. Dhondt
S.B. Diraoui
P. Eikelenboom
M.H. Emmelot-Vonk
R.A. Faaij
W.A. van Gool
E.R. Groot
C.Hagestein- de Bruijn
J.G.F.M. Hovens
M. van der Jagt
A. de Jonghe
H.L. Koek
A.W. van der Kooi
M. Kromkamp
J. Lagro
A.F.G. Leentjens
G.J. Lefeber
F.S.S. Leijten
C. Leue
T. de Man
R.J. van Marum
R.C. van der Mast
B.C. van Munster
R.J. Osse
C.B. Portier
N. Rius Ottenheim
C.H. Röder
Y. Schoon
A. Tromp
J.J.B. van der Vlugt
A.M. Vondeling
H. Weinstein
J. Witlox
J.S. van Zanten
P.M. Zeman
L.M. Peelen
A.J.C. Slooter

Submitted

ABSTRACT

IMPORTANCE: Delirium due to encephalopathy is associated with poor outcome but is frequently unrecognized, which hampers treatment. Electroencephalography (EEG) is an objective tool to measure brain activity that shows slowing of background activity in delirium. However, conventional EEG recordings are unsuitable for routine clinical applications as they can be time consuming and require trained personnel for interpretation.

OBJECTIVE: To validate a brief recording with quantitative, single-channel EEG to diagnose delirium in an independent cohort of older postoperative patients.

DESIGN: Prospective, multicenter study.

SETTING: Two university hospitals and one nonacademic teaching hospital in the Netherlands.

PARTICIPANTS: Postoperative patients, aged 60 years or older (n=159).

EXPOSURES: Patients were seen preoperatively and on the first three postoperative days to perform a five minute EEG recording with two channels (Fp2-Pz and T8-Pz according to the international 10-20 system), followed by a standardized cognitive assessment recorded on video, including the delirium rating scale revised version-98. In addition, we recorded all information from the medical- and nursing files. Two, or in case of disagreement, three delirium experts classified each postoperative day based on the videos and chart review, Diagnostic and Statistical Manual of Mental Disorders (DSM–5) criteria, and majority vote, into 'no delirium' versus 'possible delirium', or 'delirium'. This classification was compared to the EEG recordings.

MAIN OUTCOMES AND MEASURES: The relative delta power (1-4 Hz), an estimate of slowing of EEG activity based, on 1-minute artifact-free EEG recording was compared with the final delirium classification by the experts, using a receiver operative characteristic curve (ROC).

RESULTS: The area under the ROC curve (AUROC) was 0.75 (95% confidence interval (CI) 0.69-0.82) for relative delta power (1-4 Hz) on derivation Fp2-Pz. Additional analysis showed that widened relative delta power (1-6 Hz) on Fp2-Pz had a significantly higher AUROC (0.78, 95% CI 0.72-0.84, p=0.014).

CONCLUSIONS AND RELEVANCE: Delirium can be detected in older postoperative patients based on a brief single-channel EEG recording. This approach enables objective detection of delirium, providing a continuous scale instead of a dichotomized outcome.

TRIAL REGISTRATION: Clinicaltrials.gov identifier NCT02404181.

INTRODUCTION

Delirium due to encephalopathy is a common disorder in hospitalized patients, particularly after surgery[1–4] and in the intensive care unit (ICU)[5,6]. Delirium is associated with prolonged hospitalization[7–9], long-term cognitive decline[4,5,8,10,11], increased mortality[8,12], institutionalization[13] and costs[12,14], and is therefore a serious health care problem. Early detection enables early treatment of delirium. However, delirium is often not recognized[12]. To improve detection of delirium, several screening tools have been developed. In a research setting with a limited number of researchers, these tools could reliably classify delirium[15–18]. However, these instruments appeared to be less sensitive in routine, daily practice[19]. The most frequently used screening tool in the ICU, the confusion assessment method (CAM)-ICU, had a sensitivity of 47% in a clinical setting when administered by numerous bedside ICU nurses[19]. The CAM assessed by clinical nurses in hospitalized older patients showed a sensitivity of 25%[20]. These findings raise the question whether an objective tool could be used to detect delirium.

Delirium is associated with slowing of electroencephalogram (EEG) background activity[21,22]. EEG-based monitoring and quantification could have great potential for detecting delirium in routine daily practice as it is objective, applicable in all patients despite language- or sensory barriers, and could quantify delirium severity on a continuous scale, whereas commonly used screening tools yield a dichotomized score. However, performing conventional EEG in all patients at risk for delirium is unsuitable for daily clinical practice, since it can be time consuming and requires trained personnel for interpretation. In our previous proof of concept study, we showed that one minute EEG recording with three electrodes was enough to distinguish cases with delirium from non-delirious controls after cardiac surgery using the relative delta power (i.e., the power of slow waves below 4 Hz)[23]. Based on these findings we developed an EEG-based delirium monitor.

The objective of the current study was to validate this EEG-based monitor, by assessing its discriminative performance in an independent cohort of older postoperative patients. We hypothesized that delirium can be distinguished from non-delirium based on the relative delta power.

METHODS

DESIGN, SETTING AND STUDY POPULATION

In this prospective, multicenter study, patients who were scheduled for surgery were included between June 2014 and July 2016, in University Medical Center (UMC) Utrecht, Radboud University Medical Center Nijmegen, and the nonacademic teaching hospital Isala in Zwolle. The ethical committee of UMC Utrecht approved the study protocol (number 13-634) and the protocol was registered at ClinicalTrials.gov (identifier: NCT02404181).

Eligible patients of 60 years or older, scheduled for major surgery, (i.e., an expected hospital stay of at least two postoperative days), were visited at the preoperative outpatient screening clinic or at the ward at hospital admission and asked for written informed consent. Patients at risk for delirium were invited to participate. Patients were considered to be at risk for delirium when: aged ≥ 75 years, medical history of a transient ischemic attack (TIA), stroke, depression, dementia, alcohol abuse, cognitive problems, severe cardiac or respiratory disease, or previous episode of delirium[12]. Exclusion criteria were admission for neurological surgery and inability to perform cognitive testing (e.g., language barrier or deafness).

STUDY PROCEDURE

A researcher visited the patients before surgery (T-1), and on the first three consecutive days after surgery (T1, T2, and T3). During each visit an EEG recording was performed followed by standardized cognitive assessment. In addition, all information from the medical and nursing files during the study period (T-1 to T3) was collected.

Index test - EEG recording

Ag/AgCl electrodes were applied using EEG paste for conductivity and a simple headband for positioning the electrodes at Fp2, T8, Pz, and the ground electrode at FPz according to the international 10-20 system. A MobiMini system (TMSi, Oldenzaal, the Netherlands) was used as amplifier. Data were transmitted by a Bluetooth wireless sender to a laptop. Bipolar recording was performed for Fp2-Pz and T8-Pz with a sample frequency of 512 Hz, re-referencing gave the third EEG derivation Fp2-T8. Patients were instructed to relax, and close their eyes for five minutes. The researcher ensured that the patients were kept awake with verbal commands.

The first 60 seconds of artifact-free data were selected by a researcher, blinded to the classification of the experts described below. Data were filtered in the following order: first, the 50 Hz component and its visible harmonics in the Fourier spectrum were removed using notch filters. Second, drift (i.e., frequencies <0.15 Hz) was removed using an infinite impulse response (IIR) filter. Phase distortion was avoided by processing the data in a forward and reversed arrangement. Third, using the obtained transfer function between the apparatus of our previous study[23] and the MobiMini system in the current study, we corrected for differences of the EEG signals between the systems using a combination of finite impulse response (FIR) and IIR filtering.

The selected EEG segments were assessed with different methods. First, the relative delta power was computed, as the ratio between the absolute delta power between 1 and 4 Hz and the total power (i.e., power between 1 and 30 Hz). For exploratory analyses, several other EEG parameters were calculated for the three deviations (Fp2-Pz, T8-Pz, and Fp2-T8): relative powers of widened ranges of delta (1-5 Hz and 1-6 Hz), theta (4-8 Hz), alpha (8-13 Hz), and beta (13-30 Hz) power, slow-fast ratio (power between 1-8 Hz/8-30 Hz), and peak frequency (between 4-13 Hz)[23].

Reference test - Delirium assessment

Directly after each EEG recording, the Dutch version of the delirium rating scale – revised in 1998 (DRS-R-98) concerning the last 24 hours and the CAM-ICU were administered by a researcher[24,15]. This structured examination took 10-15 minutes and was recorded on video. The level of consciousness was assessed with the Richmond agitation-sedation scale (RASS), ranging between -5 (i.e., unarousable) to +4 (i.e., combative), with 0 being alert and calm[25].

Diagnoses of delirium were made by delirium experts, geriatricians (n=17), neurologists (n=4), psychiatrists (n=15), a psychologist (n=1) and a nurse-scientist (n=1) with a median of 11.5 (interquartile range (IQR) 6.3-19.0) years of experience after board certification. In their practice, these experts diagnosed an estimated 15 (median, IQR 10-25) patients with delirium per month.

For each patient (at T-1, T1, T2 and T3), a unique couple of two delirium experts was assigned to classify, independently of each other, each day as either ‘no delirium’, ‘possible delirium’ or ‘delirium’, based on cognitive examination as recorded on video combined with chart review, using Diagnostic and Statistical Manual of mental

disorders (DSM–5) criteria[26]. In addition, the experts independently scored severity of delirium based on the DRS-R-98 (0- 39, higher values indicates more severe delirium), and the likelihood of a patient being delirious on a 0-10 numeric rating scale (NRS, 0 meaning 'not delirious' and 10 meaning 'definite delirious'). In case of disagreement in classification, a third expert was consulted who received all clinical information and was unaware of the classification of the first two experts. The majority vote of the three experts was used as final classification. The average of the DRS-R-98 and NRS values of the two, or in case of disagreement of three, delirium experts were used in the analyses. The experts were blinded for the EEG recordings.

Other data collection

For all included patients, we recorded age, sex, self-reported alcohol consumption per week, self-reported history of transient ischemic attack (TIA) or stroke, self-reported pre-existing psychiatric disease, as well as type and duration of surgery from the electronic patient record. Preoperative cognitive deficits were assessed with the mini-mental state examination (MMSE). For each visit, benzodiazepine-use in the preceding 24 hours was extracted from the patient record.

SAMPLE SIZE CALCULATION

The sample size was calculated based on a conservatively expected delirium incidence of 10% during the first three postoperative days[27]. The minimum acceptable specificity of a delirium monitor was set on 90%, based on a meta-analysis of previous studies of the CAM-ICU in a research setting[17,28]. With a precision of 5% and standard formula for diagnostic studies, we calculated the required minimum number of patients to be 154[29].

STATISTICAL ANALYSIS

Variables were tested for normality using the Kolmogorov-Smirnov test and described as mean (standard deviation (SD)) or median (IQR) where appropriate. EEG characteristics were calculated for each delirium classification (i.e., no delirium, possible delirium, and delirium), and compared using the Kruskal-Wallis test, due to skewed distribution. When statistically significant differences were found, a post-hoc Mann-Whitney U test was performed.

In the main analyses, we combined the possible delirium and the delirium groups, since both were considered to be clinically relevant for timely delirium detection, and computed a receiver operating characteristic (ROC) curve. Only complete cases (i.e., an artifact-free one minute EEG recording and final diagnosis by the delirium experts)

were included for diagnostic accuracy analysis. The area under the ROC curve (AUROC) was calculated based on different sensitivity and specificity points with accompanying 95% confidence intervals (CI). With exploratory analyses, the AUROC values of the above described EEG characteristics were calculated for the three derivations. Inverse parameter was calculated for relative theta, alpha and beta power (e.g., 1-alpha power), and peak frequency (e.g., 1/(peak frequency)) for comparison of the AUROC values. The Delong method was used to assess differences in AUROC of these EEG characteristics with the relative delta power (1-4 Hz)[30].

Spearman's rho correlation coefficient was calculated between the relative delta power at Fp2-Pz and the NRS (i.e., likelihood of delirium), the DRS-R-98 score (i.e., severity of delirium), score of attentional deficits (i.e., item-10 of the DRS-R-98 score, 0 meaning no-, and 3 indicating severe deficits) and the RASS score (i.e., level of consciousness). For the EEG algorithm that showed the highest discrimination in the aforementioned AUROC analyses, we also calculated these correlations.

We repeated the analyses in two a priori defined subgroups: patients with or without TIA or stroke in medical history, because this can be associated with focal increased slowing[31,32], and patients with or without benzodiazepine administration in the 24 hours prior to the EEG recording, as this is associated with increased beta activity (13-30 Hz)[33]. Additionally, we assessed the discrimination based on assessments of time point T1 only.

To correct for multiple testing, the Benjamini-Hochberg correction[34] based on the false discovery rate, was applied for 24 EEG characteristics (i.e., eight variables for three bipolar derivations). Analyses were performed in MATLAB (version 2015a) and SPSS (version 21, SPSS Inc., Chicago, IL). Data were reported according to the standards for reporting diagnostic accuracy studies[35].

RESULTS

STUDY POPULATION AND DELIRIUM ASSESSMENT

In total, 159 patients were included in whom 360 postoperative assessment were performed (Figure 6.1). In total, 8.6% (15/174) of patients and 8.6% (34/394) of recordings were not included because of technical difficulties. Table 6.1 shows patients characteristics. The EEG recordings were in all cases immediately followed by the video-recorded interview with a standardized cognitive assessment. In 77 (21.4%) of

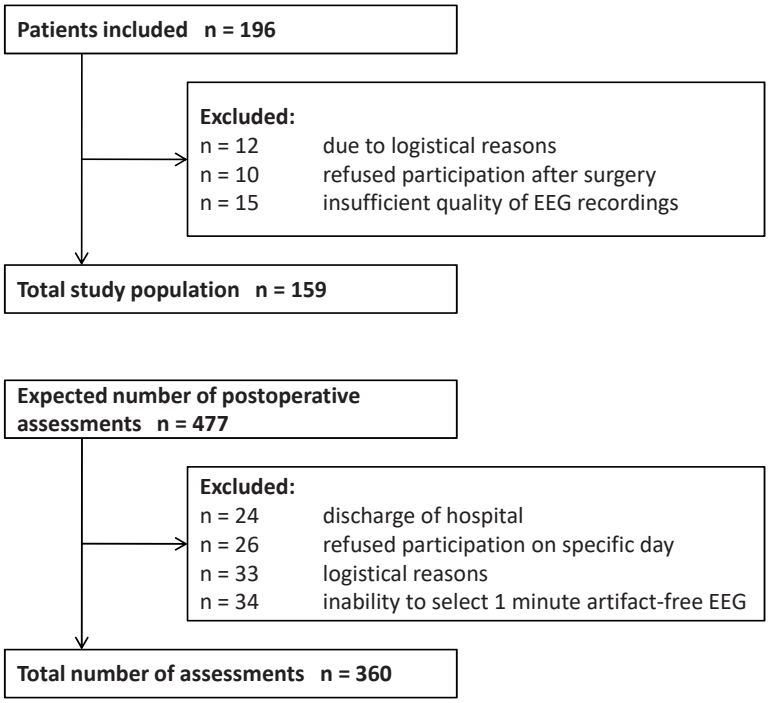


Figure 6.1. Flow-chart of included patients and included postoperative assessments.

Table 6.1. Patient characteristics.

	All patients (n=159)
Mean age (years)	76.9 (SD 6.2)
Male sex (%)	106 (66.7%)
Alcohol consumption*, n (%)	
None	60 (37.7%)
1-14 per week	70 (44.0%)
> 14 per week	24 (15.1%)
Medical history, n (%)	
TIA or Stroke	46 (28.9%)
Psychiatric disease	10 (6.3%)
Median MMSE (range)*	28 (18-30)
Surgery type, n (%)	
Cardiothoracic or vascular	139 (87.4%)
Orthopedic	12 (7.5%)
Other	8 (5.0%)
Median duration of surgery (min, range)*	171 (26-436)

*Values are presented as mean (SD), number (%), or median (range). *We had 5 missing values for alcohol consumption, 41 missing values for MMSE, and 4 missing values for duration of surgery. Psychiatric disease were depression (n=6), (early) dementia (n=3), and bipolar disease (n=1). Other types of surgery were gastrointestinal (n=4), otorhinolaryngologic (n=3), urologic (n=1), and general surgery (n=1). MMSE = mini-mental state examination, n = number, SD = standard deviation.*

assessment-days there was no agreement between the first two experts and a third delirium expert was consulted. In the final classification 43 postoperative assessment-days (11.9%) were labeled as 'delirium', 41 (11.4%) as 'possible delirium' and 276 (76.7%) as 'no delirium'. This corresponded with 29 (18.2%) patients who were classified as delirious at least once during follow-up, 26 (16.4%) patients with at most possible delirium, and 104 (65.4%) patients without any sign of delirium.

VALIDATION OF RELATIVE DELTA POWER IN FP2-PZ

The relative delta power (1-4 Hz) was significantly different over the three groups: no delirium median 0.44 (IQR 0.32-0.58), possible delirium median 0.64 (IQR 0.48-0.73), and delirium median 0.67 (IQR 0.55-0.74), p<0.001. The no delirium group had a lower relative delta power (1-4 Hz) than the possible delirium (p<0.001) and the delirium groups (p<0.001). Possible delirium and delirium did not differ in relative delta power (1-4 Hz) (p=0.187), and were grouped. Figure 6.2 shows the ROC curve based on distinction into no delirium vs. (possible) delirium, and the relative delta power on Fp2-Pz (AUROC 0.75, 95% CI 0.69-0.81).

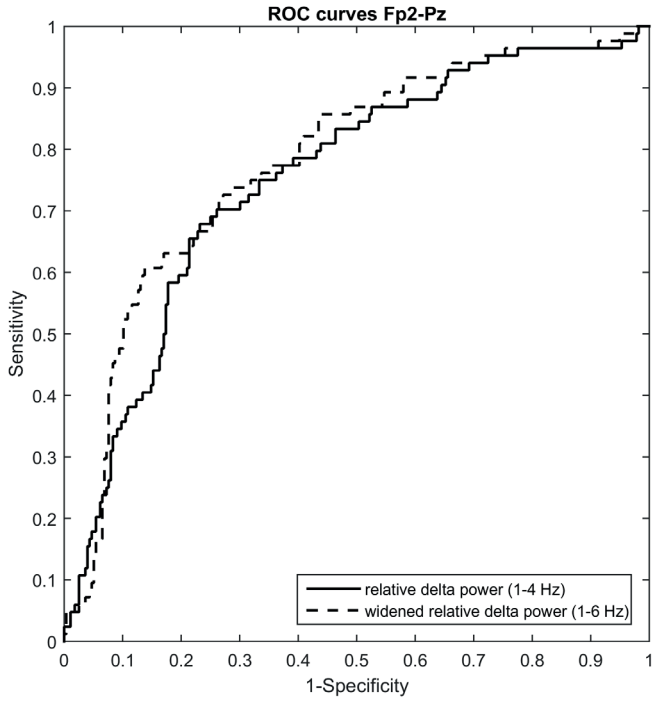


Figure 6.2. ROC curves of relative delta power (1-4 Hz) and widened relative delta power (1-6 Hz) on deviation Fp2-Pz. *Postoperative assessments were defined positive when classified as either possible delirium (n=41) or delirium (n=43). Relative delta power (1-4 Hz) AUROC=0.75 (95% CI 0.69-0.81), widened relative delta power (1-6 Hz) AUROC=0.78 (95% CI 0.72-0.84). AUROC = area under the receiver operating characteristic curve, CI = confidence interval.*

ADDITIONAL EEG ALGORITHMS AND DERIVATIONS

Of various EEG algorithms, the widened relative delta power (1-6 Hz) on Fp2-Pz showed a statistically significantly higher AUROC value (0.78, 95% CI 0.72-0.84) than the relative delta power (1-4 Hz) at the same derivation (p=0.014, Figure 6.2). See supplementary Table S6.1 for the results of all candidate EEG algorithms. Moreover, the differences compared to baseline (T-1) for relative delta power (1-4 Hz) and widened relative delta power (1-6 Hz) were calculated. This yielded similar AUROC values as in analyses on uncorrected EEG parameters (supplementary Table S6.2).

CORRELATION OF EEG PARAMETERS WITH CLINICAL SCORES

The likelihood of delirium (NRS), severity of delirium (DRS-R-98), attentional level (DRS-R-98 item 10), and level of consciousness (RASS) were all, although statistically significant, weakly correlated with both the relative delta power (1-4 Hz) and the widened relative delta power (1-6 Hz), see Table 6.2.

SUBGROUP ANALYSIS

When we stratified according to the presence or absence of a previous TIA or stroke, there were no major differences in AUROC values for the relative delta power (1-4 Hz) or the widened relative delta power (1-6 Hz) (supplementary Table S6.3-I). Patients who had used benzodiazepines in the preceding 24 hours had AUROC values of 0.77 (95% CI 0.68-0.87) for relative delta power (1-4 Hz) and 0.81 (95% CI 0.72-0.89) for widened relative delta power (1-6 Hz), whereas patients without benzodiazepines in the previous 24 hours had AUROC values of 0.74 (95% CI 0.66-0.82) and 0.76 (95% CI 0.68-0.84) for relative delta power (1-4 Hz) and widened relative delta power (1-6 Hz), respectively (Table S6.3-II).

Table 6.2. Correlation of (widened) relative delta power with the likelihood and severity of delirium, attention and consciousness level.

	Relative delta power(1-4 Hz) Fp2-Pz		Widened relative delta power (1-6 Hz) Fp2-Pz	
	Correlation coefficient (95% CI)	p-value	Correlation coefficient (95% CI)	p-value
Likelihood of delirium (NRS)	0.33 (0.23-0.42)	<0.001	0.38 (0.28-0.47)	<0.001
Severity of delirium (DRS-R-98)	0.39 (0.29-0.47)	<0.001	0.44 (0.34-0.53)	<0.001
Attention level (item 10 of DRS-R-98)	0.27 (0.17-0.37)	<0.001	0.31 (0.21-0.40)	<0.001
Level of consciousness (RASS)	-0.29 (-0.38--0.19)	<0.001	-0.30 (-0.40--0.19)	<0.001

CI = confidence interval, DRS-R-98 = delirium rating scale revised '98, NRS = numeric rating scale, RASS = Richmond agitation-sedation scale.

EFFECT OF REPEATED MEASUREMENTS

To assess the effect of patients being repeatedly assessed in our main analyses, the AUROC values were recalculated for the relative delta power (1-4 Hz) and widened relative delta power (1-6 Hz) for assessments on time point T1 only (supplementary Table S6.4 and Figure S6.1). There was a minor reduction in AUROC values when only T1 assessments were included compared to the main analysis. We did however not conduct a formal statistical comparison because of the different number of assessments in T1-only versus all postoperative assessment-days.

DISCUSSION

MAIN RESULTS

In this study, we used an independent cohort of older postoperative patients to validate our previous finding that delirium can be detected using relative delta power based on 1-minute of single-channel EEG. We found an AUROC value of 0.75 with the relative delta power (1-4 Hz), and a statistically significantly higher AUROC value (0.78) when the widened relative delta power (1-6 Hz) was used. With these algorithms, we found weak correlations with scores on the likelihood and severity of delirium, inattention, and level of consciousness.

METHODOLOGICAL CONSIDERATIONS

In the main analysis, 'possible delirium' and 'delirium' were combined, as in clinical practice doubtful cases of delirium may represent a prodromal phase in which treating the underlying cause may prevent progression to delirium. Moreover, 'possible delirium' and 'delirium' were associated with similar values on a variety of EEG parameters (Table S6.1), supporting grouping of both entities into one category.

The presented ROC curves were calculated without additional clinical information, since the EEG-based delirium monitor was developed with the aim to replace current screenings methods and also be applicable without the need of further clinical characteristics. Since delirium can be regarded as a condition with a spectrum of different severities, no cut-off values for EEG parameters were determined. Also for monitoring delirium over the course of days, a continuous representation seems to be more appropriate.

Based on previous studies[23,36], we investigated several EEG parameters. Other known EEG characteristics of delirium or encephalopathy, which were not included in the current study, are spectral variability and complexity[37], triphasic waves, and (inter)ictal epileptiform discharges[38–41]. The current algorithm could potentially be improved using some of these features.

STRENGTHS AND LIMITATIONS

Strengths of this investigation include the prospective nature of this multicenter study with a large sample size. We used a thorough diagnostic work-up. The final diagnosis was based on the judgment of two or three delirium experts, who evaluated all available information truly independently. Moreover, the availability of preoperative assessments facilitated the diagnosis of postoperative delirium. All experts had large experience in diagnosing delirium.

Several limitations of this study need to be addressed. Firstly, 8.6% of the EEG recordings had to be excluded for analysis due to technical difficulties. Secondly, the EEG recordings were performed once a day, which might have resulted in overlooking of periods of slowing of the EEG. Finally, we analyzed repeated measures of each patient as independent observations, which could have influenced the results. However, this is unlikely as restriction to T1 assessments only resulted in similar results as in analysis on all postoperative recordings.

CLINICAL AND RESEARCH IMPLICATIONS

As the algorithms that we used can be applied real-time, and a brief EEG registration with three electrodes is easy to perform, our findings suggest that this approach can be used to monitor delirium by bedside nurses as part of clinical routine. An objective delirium monitor with a continuous scale instead of a dichotomized outcome would be a major step forward for both clinical practice and research on delirium. Currently used tools to estimate delirium have subjective elements and disappointing sensitivity in routine daily practice[19,20], where numerous beside nurses assess delirium at different time points.

FUTURE RESEARCH

Before a delirium monitor can be used in clinical or research settings, several steps should be made. Firstly, the current study is a technical validation of a prototype monitor. Optimizing feasibility with input from nurses and patients is necessary for successful implementation. Secondly, diagnostic performance of the algorithm could be further optimized. This includes the use of different algorithms simultaneously and real-time automatic artifact-detection and -rejection. Thirdly, the optimal number of recordings per day should be established. Furthermore, it is not clear to what extent the current findings can be generalized to other patient populations, such as ICU patients in whom sedatives are applied, or dementia patients. Both, use of sedatives and dementia, result in slowing of the EEG[33,42], and may therefore interfere with the current algorithm.

CONCLUSION

Delirium can be detected in older postoperative patients based on a brief bipolar EEG recording. This approach enables objective detection of delirium, providing a continuous scale instead of a dichotomized outcome.

ACKNOWLEDGEMENTS

Individual funding grants: Dr. A.F.G Leentjens: Stichting Internationaal Parkinsonfonds (IPF), Michael J Fox Foundation for Parkinson research. Royalties: Reed Business, Van Gorcum, Springer. Honoraria: Elsevier Inc (as editor-in-chief of the J Psychosom Res). Dr. P.L.J. Dautzenberg is principle investigator is some ongoing trials with patients with Alzheimer's disease (MSD, Boehringer Ingelheim), Lewy body dementia (Axovant) and healthy elderly with a greater risk for Alzheimer's disease (Janssen). Dr. A.J.C. Slooter and A.W. van der Kooi have a patent pending: Method and system for determining a parameter which is indicative for whether a patient is delirious. Dr. F.S.S. Leijten hold shares in ProLira, a company developing a delirium monitor. Funding sources (grants or institutional or corporate support): Clinical Research Award ESICM, Technology Foundation STW (14066), European Union (H2020-PHC-12-2014-672974) and institutional support.

REFERENCES

1 Kalish, V. B., Gillham, J. E. & Unwin, B. K. Delirium in older persons: evaluation and management. Am. Fam. Physician 90, 150–158 (2014).
2 Marcantonio, E. R., Flacker, J. M., Wright, R. J., et al. Reducing delirium after hip fracture: a randomized trial. J. Am. Geriatr. Soc. 49, 516–522 (2001).
3 Neufeld, K. J., Leoutsakos, J. S., Sieber, F. E., et al. Evaluation of two delirium screening tools for detecting post-operative delirium in the elderly. Br. J. Anaesth. 111, 612–618 (2013).
4 Saczynski, J. S., Marcantonio, E. R., Quach, L., et al. Cognitive trajectories after postoperative delirium. N. Engl. J. Med. 367, 30–39 (2012).
5 Wolters, A. E., van Dijk, D., Pasma, W., et al. Long-term outcome of delirium during intensive care unit stay in survivors of critical illness: a prospective cohort study. Crit. Care 18, R125 (2014).
6 Ely, E. W., Shintani, A., Truman, B., et al. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. JAMA 291, 1753–1762 (2004).
7 Gleason, L. J., Schmitt, E. M., Kosar, C. M., et al. Effect of delirium and other major complications on outcomes after elective surgery in older adults. JAMA Surg. 150, 1134 (2015).
8 Salluh, J. I. F., Wang, H., Schneider, E. B., et al. Outcome of delirium in critically ill patients: systematic review and meta-analysis. BMJ 350, h2538 (2015).
9 Klein Klouwenberg, P. M. C., Zaal, I. J., Spitoni, C., et al. The attributable mortality of delirium in critically ill patients : prospective cohort study. BMJ 349: g6652, 1–10 (2014).
10 Pandharipande, P. P., Girard, T. D., Jackson, J. C., et al. Long-term cognitive impairment after critical illness. N. Engl. J. Med. 369, 1306–1316 (2013).
11 Fong, T. G., Davis, D., Growdon, M. E., et al. The interface between delirium and dementia in elderly adults. Lancet. Neurol. 14, 823–832 (2015).
12 Inouye, S. K., Westendorp, R. G. & Saczynski, J. S. Delirium in elderly people. Lancet 383, 911–922 (2014).
13 Khan, B. A., Zawahiri, M., Campbell, N. L., et al. Delirium in hospitalized patients: Implications of current evidence on clinical practice and future avenues for research: a systematic evidence review. J. Hosp. Med. 7, 580–589 (2012).
14 Leslie, D. L., Marcantonio, E. R., Zhang, Y., et al. One-year health care costs associated with delirium in the elderly population. Arch. Intern. Med. 168, 27–32 (2008).
15 Ely, E. W., Margolin, R., Francis, J., et al. Evaluation of delirium in critically ill patients: validation of the confusion assessment method for the intensive care unit (CAM-ICU). Crit. Care Med. 29, 1370–1379 (2001).
16 Koster, S., Hensens, A. G., Oosterveld, F. G. J., et al. The delirium observation screening scale recognizes delirium early after cardiac surgery. Eur. J. Cardiovasc. Nurs. 8, 309–314 (2009).
17 Shi, Q., Warren, L., Saposnik, G., et al. Confusion assessment method: a systematic review and meta-analysis of diagnostic accuracy. Neuropsychiatr. Dis. Treat. 9, 1359–1370 (2013).
18 Van Velthuijsen, E. L., Zwakhalen, S. M. G., Warnier, R. M. J., et al. Psychometric properties and feasibility of instruments for the detection of delirium in older hospitalized patients: a systematic review. International Journal of Geriatric Psychiatry 31, 974–989 (2016).
19 Eijk, M. M. van, Boogaard, M. van den, Marum, R. J. van, et al. Routine use of the confusion assessment method for the intensive care unit: a multicenter study. Am. J. Respir. Crit. Care Med. 184, 340–344 (2011).
20 Rice, K. L., Bennett, M., Gomez, M., et al. Nurses' recognition of delirium in the hospitalized older adult. Clin. Nurse Spec. 25, 299–311 (2011).
21 Engel, G. L. & Romano, J. Delirium, a syndrome of cerebral insufficiency. J. Chronic Dis. 9, 260–277 (1959).

22 Engel, G. L. & Romano, J. Delirium, a syndrome of cerebral insufficiency. 1959. J. Neuropsychiatry Clin. Neurosci. 16, 526–538 (2004).
23 Van der Kooi, A. W., Zaal, I. J., Klijn, F. A., et al. Delirium detection using EEG. Chest 147, 94–101 (2015).
24 De Rooij, S. E., van Munster, B. C., Korevaar, J. C., et al. Delirium subtype identification and the validation of the delirium rating scale-revised-98 (Dutch version) in hospitalized elderly patients. Int. J. Geriatr. Psychiatry 21, 876–882 (2006).
25 Sessler, C. N., Gosnell, M. S., Grap, M. J., et al. The Richmond agitation–sedation scale. Am. J. Respir. Crit. Care Med. 166, 1338–1344 (2002).
26 American Psychiatric Association. Diagnostics and statistical manual of mental disorders. 5th edition. (2013).
27 Sauër, A. C., Slooter, A. J. C., Veldhuijzen, D. S., et al. Intraoperative dexamethasone and delirium after cardiac surgery. Anesth. Analg. 119, 1046–1052 (2014).
28 Gusmao-Flores, D., Martins, J. C. S., Amorin, D., et al. Tools for diagnosing delirium in the critically ill: is calibration needed for the less sedated patient? Intensive Care Med. 40, 137–138 (2014).
29 Hajian-Tilaki, K. Sample size estimation in diagnostic test studies of biomedical informatics. Journal of Biomedical Informatics 48, 193–204 (2014).
30 DeLong, E. R., DeLong, D. M. & Clarke-Pearson, D. L. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametetric approach. Biometrics 44, 837–845 (1988).
31 Rabiller, G., He, J.-W., Nishijima, Y., et al. Perturbation of brain oscillations after ischemic stroke: a potential biomarker for post-stroke function and therapy. Int. J. Mol. Sci. 16, 25605–25640 (2015).
32 Moyanova, S. G. & Dijkhuizen, R. M. Present status and future challenges of electroencephalography- and magnetic resonance imaging-based monitoring in preclinical models of focal cerebral ischemia. Brain Res. Bull. 102, 22–36 (2014).
33 De Visser, S. J., Van der Post, J., Pieters, M. S. M., et al. Biomarkers for the effects of antipsychotic drugs in healthy volunteers. Br. J. Clin. Pharmacol. 51, 119–132 (2001).
34 Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. 57, 289–300 (1995).
35 Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. BMJ 326, 41–44 (2003).
36 Van der Kooi, A. W., Leijten, F. S. S., van der Wekken, R. J., et al. What are the opportunities for EEG-based monitoring of delirium in the ICU? J. Neuropsychiatry Clin. Neurosci. 24, 472–477 (2012).
37 Van der Kooi, A. W., Slooter, A. J. C., van het Klooster, M. A., et al. EEG in delirium: increased spectral variability and decreased complexity. Clin. Neurophysiol. 125, 2137–2139 (2014).
38 Neufeld, K. J. & Thomas, C. Delirium: definition, epidemiology, and diagnosis. J. Clin. Neurophysiol. 30, 438–442 (2013).
39 Kaplan, P. W. & Sutter, R. Seeing more clearly through the fog of encephalopathy. J. Clin. Neurophysiol. 30, 431–434 (2013).
40 Sutter, R., Stevens, R. D. & Kaplan, P. W. Clinical and imaging correlates of EEG patterns in hospitalized patients with encephalopathy. J. Neurol. 260, 1087–1098 (2013).
41 Naeije, G., Bachir, I., Gaspard, N., et al. Epileptic activities are common in older people with delirium. Geriatr. Gerontol. Int. 14, 447–451 (2014).
42 Moretti, D. V., Babiloni, C., Binetti, G., et al. Individual analysis of EEG frequency and band power in mild Alzheimer's disease. Clin. Neurophysiol. 115, 299–308 (2004).

SUPPLEMENTARY MATERIAL

Table S6.1. Median values with interquartile ranges on various EEG parameters with accompanying AUROC values according to the presence of (possible) delirium.

	EEG parameter	No delirium (n=276)	Possible delirium (n=41)	Delirium (n=43)	Comparison between groups p-value	AUROC (95% CI)	Comparison AUROC values Delong Method p-value
F_{p2}-P_z	**Relative delta power (1-4 Hz)**	0.44 (0.32-0.58)	0.64 (0.48-0.73)	0.67 (0.55-0.74)	<0.001[a,b]	0.75 (0.69-0.81)	Reference
	1-Relative theta power (4-8 Hz)	0.78 (0.67-0.86)	0.78 (0.69-0.85)	0.81 (0.73-0.87)	0.292	0.55 (0.48-0.62)	<0.001*
	1-Relative alpha power (8-13 Hz)	0.84 (0.74-0.90)	0.92 (0.84-0.95)	0.93 (0.89-0.96)	<0.001[a, b]	0.75 (0.69-0.81)	0.800
	1-Relative beta power (13-30 Hz)	0.91 (0.87-0.95)	0.95 (0.92-0.97)	0.96 (0.93-0.98)	<0.001[a,b]	0.74 (0.67-0.80)	0.617
	Widened delta power (1-5 Hz)	0.49 (0.36-0.64)	0.71 (0.54-0.77)	0.75 (0.61-0.81)	<0.001[a,b]	0.77 (0.71-0.83)	0.006*
	Widened delta power (1-6 Hz)	0.55 (0.41-0.70)	0.78 (0.60-0.82)	0.82 (0.66-0.86)	<0.001[a,b]	0.78 (0.72-0.84)	0.014*
	Slow-fast ratio	2.72 (1.58-4.87)	6.87 (3.39-9.58)	8.20 (4.58-13.65)	<0.001[a,b]	0.76 (0.71-0.82)	0.624
	1/Peak frequency	0.13 (0.12-0.16)	0.16 (0.14-0.22)	0.19 (0.14-0.24)	<0.001[a,b]	0.72 (0.66-0.79)	0.350

Table S6.1 (continued). Median values with interquartile ranges on various EEG parameters with accompanying AUROC values according to the presence of (possible) delirium.

	EEG parameter	No delirium (n=252)	Possible delirium (n=37)	Delirium (n=38)	Comparison between groups p-value	AUROC (95% CI)	Comparison AUROC values Delong Method p-value
T8-Pz	**Relative delta power (1-4 Hz)**	0.37 (0.25-0.48)	0.55 (0.44-0.66)	0.64 (0.53-0.73)	<0.001[a,b]	0.79 (0.72-0.85)	0.19
	1-Relative theta power (4-8 Hz)	0.74 (0.64-0.82)	0.73 (0.65-0.82)	0.80 (0.67-0.84)	0.239	0.55 (0.48-0.63)	<0.001*
	1-Relative alpha power (8-13 Hz)	0.78 (0.67-0.86)	0.89 (0.83-0.94)	0.91 (0.87-0.95)	<0.001[a,b]	0.78 (0.72-0.84)	0.418
	1-Relative beta power (13-30 Hz)	0.92 (0.89-0.95)	0.96 (0.94-0.98)	0.97 (0.95-0.98)	<0.001[a,b]	0.80 (0.74-0.86)	0.228
	Widened delta power (1-5 Hz)	0.41 (0.30-0.53)	0.62 (0.53-0.71)	0.70 (0.60-0.81)	<0.001[a,b,c]	0.80 (0.73-0.86)	0.099
	Widened delta power (1-6 Hz)	0.47 (0.35-0.62)	0.72 (0.61-0.79)	0.77 (0.65-0.87)	<0.001[a,b]	0.80 (0.74-0.86)	0.099
	Slow-fast ratio	2.13 (1.11-4.20)	5.67 (3.54-9.52)	7.70 (4.27-13.45)	<0.001[a,b]	0.80 (0.73-0.86)	0.192
	1/Peak frequency	0.13 (0.12-0.16)	0.15 (0.13-0.19)	0.19 (0.13-0.24)	<0.001[a,b]	0.69 (0.63-0.76)	0.089

Table S6.1 (continued). Median values with interquartile ranges on various EEG parameters with accompanying AUROC values according to the presence of (possible) delirium.

	EEG parameter	No delirium (n=252)	Possible delirium (n=37)	Delirium (n=38)	Comparison between groups p-value	AUROC (95% CI)	Comparison AUROC values Delong Method p-value
Fp2-T8	**Relative delta power (1-4 Hz)**	0.51 (0.37-0.65)	0.63 (0.42-0.74)	0.64 (0.51-0.72)	<0.001[a,b]	0.65 (0.58-0.72)	<0.001*
	1-Relative theta power (4-8 Hz)	0.82 (0.74-0.87)	0.79 (0.74-0.86)	0.81 (0.74-0.86)	0.653	0.47 (0.40-0.54)	<0.001*
	1-Relative alpha power (8-13 Hz)	0.88 (0.81-0.93)	0.92 (0.86-0.95)	0.93 (0.88-0.95)	<0.001[a,b]	0.65 (0.58-0.72)	<0.001*
	1-Relative beta power (13-30 Hz)	0.87 (0.80-0.93)	0.93 (0.87-0.95)	0.91 (0.84-0.95)	<0.001[a,b]	0.66 (0.58-0.73)	<0.001*
	Widened delta power (1-5 Hz)	0.55 (0.42-0.70)	0.74 (0.59-0.84)	0.73 (0.60-0.83)	<0.001[a,b]	0.66 (0.60-0.73)	<0.001*
	Widened delta power (1-6 Hz)	0.60 (0.46-0.74)	0.74 (0.59-0.84)	0.73 (0.60-0.83)	<0.001[a,b]	0.67 (0.61-0.74)	<0.001*
	Slow-fast ratio	2.69 (1.38-4.99)	4.77 (2.30-7.89)	4.81 (2.74-9.29)	<0.001[a,b]	0.67 (0.61-0.74)	<0.001*
	1/Peak frequency	0.14 (0.12-0.20)	0.19 (0.14-0.23)	0.21 (0.14-0.22)	<0.001[a,b]	0.67 (0.61-0.73)	0.007*

*Post-hoc analyses were performed when significant after Benjamini-Hochberg correction for 24 explorative EEG characteristics. AUROC is calculated between no delirium cases and (possible) delirious cases. [a] significantly different between no delirium and possible delirium, [b] significantly different between no delirium and delirium, [c] significant different between possible delirium and delirium, n = number of postoperative assessments, CI = confidence interval, EEG = electro-encephalography, AUROC = area under receiver operating characteristic curve. *Significantly different from reference using Delong method p-values after Benjamini-Hochberg correction.*

Table S6.2. Absolute difference in relative delta power (1-4 Hz) and widened relative delta power (1-6 Hz) compared to the patient's baseline values.

Measure	No delirium (n=219)	Possible delirium (n=36)	Delirium (n=36)	Comparison between groups (p-value)	AUROC (95% CI)	Comparison ROC curves with Delong Method p-value
Relative delta power (1-4 Hz)	0.42 (0.29-0.56)	0.64 (0.47-0.73)	0.68 (0.58-0.76)	<0.001[a,b]	0.79 (0.73-0.85)	Reference
Δ Relative delta power (1-4 Hz) compared to T-1	0.03 (-0.13-0.15)	0.24 (0.03-0.37)	0.29 (0.09-0.41)	<0.001[a,b]	0.75 (0.69-0.82)	0.245
Widened relative delta power (1-6 Hz)	0.52 (0.38-0.66)	0.78 (0.60-0.82)	0.82 (0.68-0.87)	<0.001[a,b]	0.82 (0.76-0.88)	Reference
Δ Widened relative delta power (1-6 Hz) compared to T-1	0.06 (-0.10-0.20)	0.28 (0.06-0.40)	0.34 (0.13-0.46)	<0.001[a,b]	0.77 (0.71-0.84)	0.149

Post-hoc analyses were performed when significant between overall comparison. [a] significant difference between possible delirium and no delirium, [b] significant difference between delirium and no delirium, AUROC = area under the receiver operating characteristic curve, CI = confidence interval.

Table S6.3-I. Subgroup analysis with or without TIA or stroke in medical history.

	No TIA/Stroke in medical history	TIA/Stroke in medical history
Number of positive assessments	56	28
Number of negative assessments	208	68
AUROC relative delta power (1-4 Hz) Fp2-Pz	0.75 (95% CI 0.68-0.82)	0.75 (95% CI 0.64-0.86)
AUROC widened relative delta power (1-6 Hz) Fp2-Pz	0.78 (95% CI 0.71-0.85)	0.77 (95% CI 0.67-0.87)

AUROC = area under the receiver operating characteristic curve, CI = confidence interval, TIA = transient ischemic attack.

Table S6.3-II. Subgroup analysis benzodiazepine versus no benzodiazepine administration in last 24 hours.

	No benzodiazepines in <24h	Benzodiazepines <24h
Number of positive assessments	50	34
Number of negative assessments	210	66
AUROC relative delta power (1-4 Hz) Fp2-Pz	0.74 (95% CI 0.66-0.82)	0.77 (95% CI 0.68-0.87)
AUROC widened relative delta power (1-6 Hz) Fp2-Pz	0.76 (95% CI 0.68-0.84)	0.81 (95% CI 0.72-0.89)

AUROC = area under the receiver operating characteristic curve, CI = confidence interval.

Table S6.4. Effect of repeated measures.

	All postoperative assessments	T1 assessments
# positive assessments	84	43
# negative assessments	276	96
AUROC relative delta power (1-4 Hz) Fp2-Pz	0.75 (95% CI 0.69-0.81)	0.72 (95% CI 0.62-0.81)
AUROC widened relative delta power (1-6 Hz) Fp2-Pz	0.78 (95% CI 0.72-0.84)	0.75 (95% CI 0.66-0.84)

AUROC = area under the receiver operating characteristic curve, CI = confidence interval.

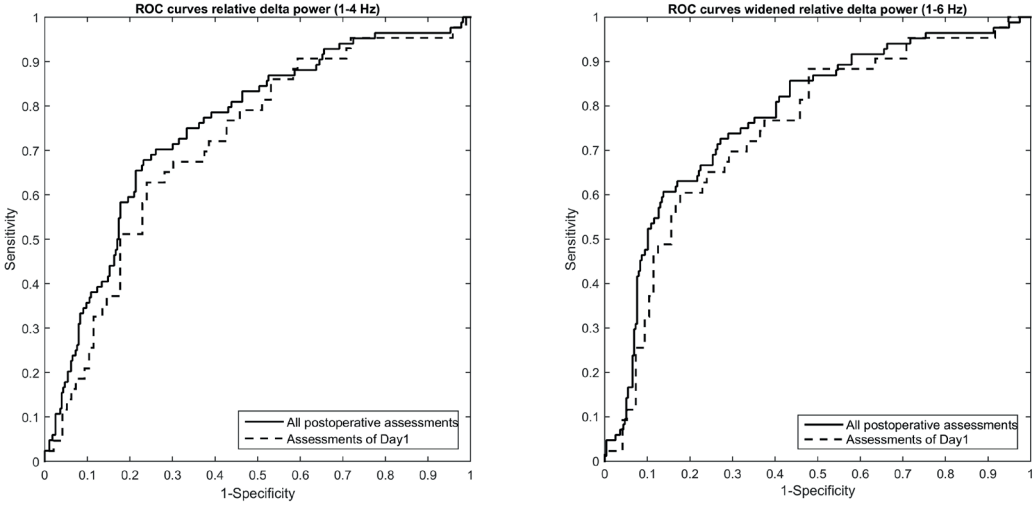
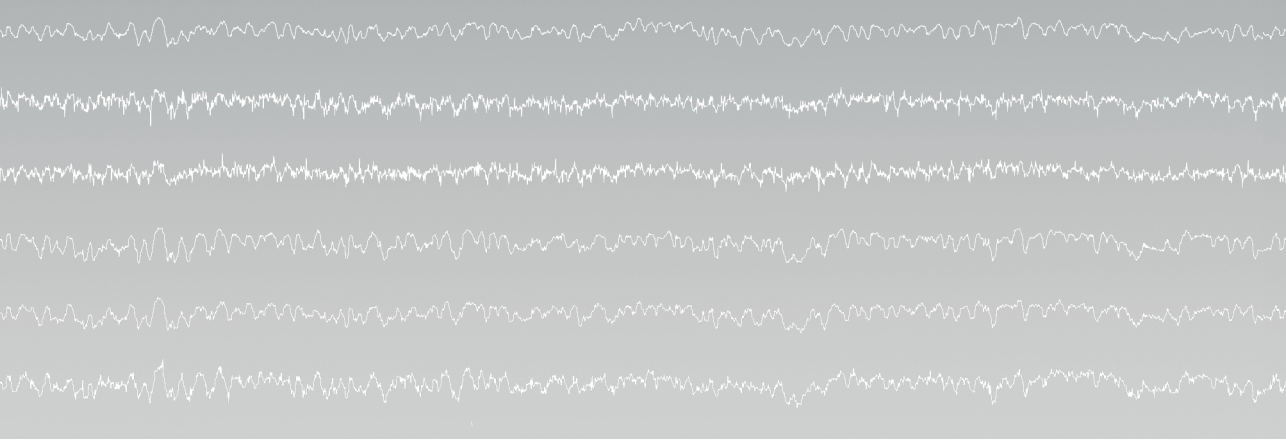
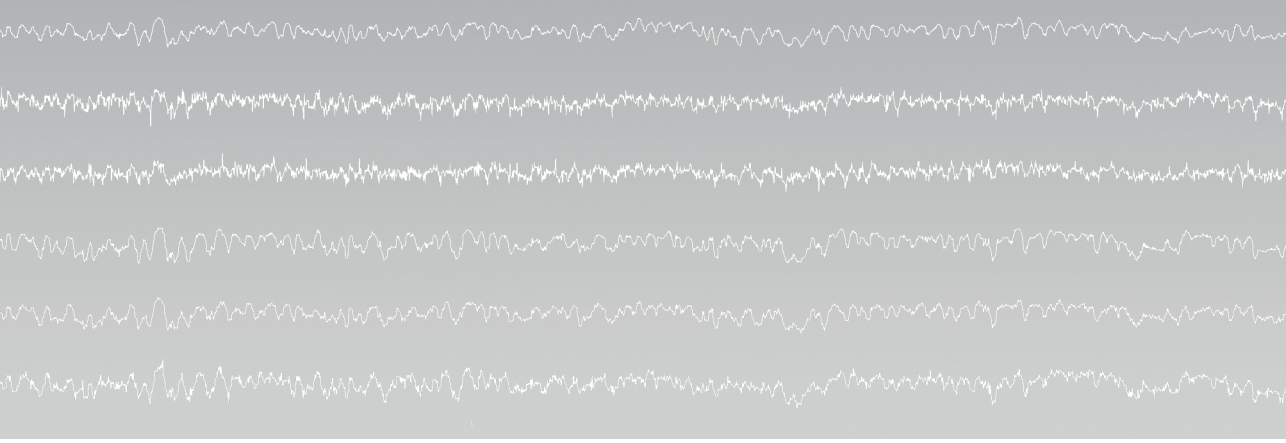


Figure S6.1. Receiver operating characteristic (ROC) curves of (left) relative delta power (1-4 Hz) and (right) widened relative delta power (1-6 Hz) related to final delirium classification for all postoperative assessment-days and T1 (day 1) only.





7

PREDICTING DELIRIUM WITH A SINGLE-CHANNEL EEG RECORDING

T. Numan
M. van den Boogaard
A.M. Kamper
P.J.T. Rood
L.M. Peelen
A.J.C. Slooter
for the Dutch Delirium Detection Study Group

In preparation

ABSTRACT

BACKGROUND: Early detection of delirium enables early interventions and improves outcome of the patients. The aim of the current study was to assess whether one-minute, single-channel quantitative electroencephalography (EEG) can predict delirium the next day.

METHODS: Older postoperative patients were daily visited (first to third day after surgery) for a bipolar EEG recording and a video-taped, standardized cognitive assessment. Delirium was diagnosed by delirium experts, who evaluated the video recordings and accompanying patient record. Relative delta power (1-4 Hz) and widened relative delta power (1-6 Hz) were calculated based on one-minute EEG. EEG parameters of the current day were compared between delirious and non-delirious patients on the next day. Logistic regression analysis including age and duration of surgery (model 1), EEG parameters only (model 2), and combination of model 1 and model 2 (model 3) were performed to assess the predictive ability of delirium the next day.

RESULTS: Significantly higher relative delta power was found in patients who were delirious the next day (n=11) compared to non-delirious patients (n=153), (1-4 Hz; p=0.035, 1-6 Hz; p=0.033). The logistic regression analyses showed area under the receiving operating characteristic curve (AUROC) values of model 1; 0.56 (95% confidence interval (CI) 0.55-0.84), model 2; 0.69 (95% CI 0.55-0.84), model 3; 0.70 (95% CI 0.56-0.84). Model 3 resulted in a significant higher AUROC compared to model 1 (p=0.036).

CONCLUSIONS: This is the first study showing the ability to predict delirium the next day based on a simple EEG recording.

INTRODUCTION

Delirium is a serious health care problem occurring frequently in elderly[1–3] and critically ill patients[4–6]. Severe outcomes of delirious patients concern prolonged intensive care unit (ICU) and hospital stay, higher mortality, and long-term cognitive problems[1–6]. Early recognition or prediction of delirium is essential for early interventions, particularly early treatment of causes of delirium, which can improve outcome[7–9].

Currently available screening methods, such as the confusion assessment method (CAM) and for ICU patients (CAM-ICU), were shown to have high sensitivity and specificity in research setting[10,11]. However, in daily clinical practice these tools were less sensitive, resulting in missed cases of delirium[12,13]. Therefore, an objective and simple method for delirium detection is needed, to enable interventions for early treatment of delirium. Quantitative electroencephalography (EEG) with three electrodes and one minute recording was shown to be able to recognize delirium in postoperative older patients[14]. Patients with delirium were found to have increased relative delta power (i.e., frequencies between 1 and 4 Hz) compared to non-delirious patients as defined by delirium experts according to the Diagnostic and Statistical Manual fifth edition (DSM-5) criteria[15]. Detection prior to the occurrence of clinical symptoms of delirium would be optimal, as timely prediction of delirium may facilitate an early window for prevention or early interventions. A previous study showed that in ICU patients with sepsis, delirium in the following days could be predicted based on EEG recordings[16]. Therefore, the aim of the current study was to assess whether increased relative delta power recorded with EEG could predict the diagnosis of delirium the next day.

METHODS

DESIGN, SETTING AND POPULATION

Data was collected as part of the prospective, multicenter observational study entitled ‘Validation of an EEG-based delirium monitor’ (ClinicalTrials.gov identifier: NCT02404181), approved by the medical ethical committee of the University Medical Center Utrecht (UMC Utrecht). In short, patients scheduled for major surgery in UMC Utrecht, Radboud University Medical Center Nijmegen or the non-academic teaching hospital Isala in Zwolle, were included after written informed consent.

Criteria for inclusion were an age of 60 years or older, being scheduled for major surgery, an expected postoperative stay of at least 2 days, and at risk for delirium. Delirium risk was defined as having at least one of the following criteria: aged ≥ 75 years, medical history of a transient ischemic attack (TIA), stroke, depression, dementia, alcohol abuse, cognitive problems, severe cardiac or respiratory disease, or a previous episode of delirium[1]. Exclusion criteria were neurological surgery, inability to perform cognitive tests due to language barrier, or inability to perform a recording with the EEG-based delirium monitor.

DATA COLLECTION

EEG recording and processing

Patients were visited preoperatively (T-1) and postoperatively on all three consecutive days after surgery (T1, T2, and T3) or until discharge, whichever came first. The patients were not visited on the day of their surgery, because of the possible remaining effect of anesthetics during surgery on the EEG. The visiting researcher performed a 7-minute EEG recording with Ag/AgCl electrodes positioned on Fp2, T8, Pz, and ground electrode at FPz according to the international 10-20 system. The EEG recording with a sample frequency of 512 Hz was performed using a MobiMini-system (TMSi, Oldenzaal, the Netherlands). Patients were instructed to relax and have their eyes open during the first two minutes, thereafter they were instructed to keep their eyes closed.

Of each recording, the first 60 seconds without major artifacts were selected by a researcher of the eyes closed segment. Data was filtered with a notch filter to remove 50 Hz components and its harmonics, followed by an infinite impulse response (IIR) to remove the drift in the signal (i.e., frequencies below 0.15 Hz), and finally it was filtered with a combination of IIR and a finite impulse response filter to correct for differences between the current used system and the system used in the proof of concept study[14].

Based on the Fp2-Pz selected segment, the relative delta power (i.e., the amount of frequencies between 1 and 4 Hz) and widened relative delta power (1-6 Hz) were calculated, as the ratio between the absolute power (i.e., the recorded power between 1-4 Hz and 1-6 Hz, respectively) divided by the total power between 1 and 30 Hz. The relative powers were calculated to eliminate differences in signal strength between subjects. Relative delta power (1-4 Hz) was found to distinguish delirium from non-delirium in the proof of concept study[14]. The widened relative delta power (1-6 Hz) has previously been shown as currently best parameter to distinguish delirium from non-delirium.

Delirium diagnosis

After the EEG recording, a standardized cognitive test was performed by the visiting researcher and recorded on video. This cognitive test was based on the severity items of the Dutch version of the delirium rating scale – revised in 1998 (DRS-R-98)[17], which consists of 13 items, and the confusion assessment method for ICU (CAM-ICU)[11,17,18]. Delirium experts (n=38: n=17 geriatricians, n=4 neurologists, n=15 psychiatrists, n=1 psychologist, and n=1 nurse-scientist; years of experience: median 11.5 (interquartile range (IQR) 6.3-19.0); number of delirious patients per month in regular clinical care: median 15 (IQR 10-25)) assessed all the four videos in different couples with accompanying description of the patient record (T-1, T1, T2 and T3) of ten to fifteen patients per expert. Patients were randomly assigned to two experts who diagnosed patients as delirious, possible delirious or non-delirious independently of each other. The two diagnoses were compared, in case of disagreement a third expert was consulted. This third expert did not receive the diagnosis of the first two experts, but did know that there was no agreement in diagnosis. Final classification was based on the median score of the two or three experts. Possible delirium and delirium diagnosis were combined into the delirium group, since both entities may be clinically relevant to detect, and compared to the non-delirious cases. In the main study, 196 patients were included. For the current study, we selected the patients who were non-delirious on day X and had a delirium diagnosis on the next day (day X+1).

The following patient characteristics were collected preoperatively: age, sex, alcohol consumption per week, history of transient ischemic attack (TIA) or stroke, pre-existing psychiatric disease, mini-mental state examination (MMSE) to assess preoperative cognitive deficits, type of surgery, and duration of the surgical procedure.

STATISTICAL ANALYSIS

Continuous patient characteristics and EEG variables were tested for a normal distribution and reported as mean with standard deviation (SD) when normally distributed, or median (interquartile range (IQR)) in case of a skewed distribution, unless otherwise stated. Categorical variables were described as numbers (%).

We compared the EEG parameters recorded on the current day (day X) with the delirium diagnoses (i.e., delirium or non-delirium), on the next day (day X+1), using Mann Whitney U test, because of the deviation from a normal distribution for the EEG parameters.

To assess whether the relative delta power (1-4 Hz) of a non-delirious day, could predict delirium the next day, logistic regression analyses was performed. First, model 1 was based on age and duration of surgery only, as model for currently available, easily assessable, clinical information. Due to the limited number of patients with delirium, only two clinical parameters were included in this model. Moreover, other important preoperative risk factors for delirium, such as cognitive impairment, dementia and alcohol misuse[1], were not sufficiently collected in the current study and therefore not included. Secondly, a logistic regression analysis based on either relative delta power (1-4 Hz) or widened relative delta power (1-6 Hz) was performed (model 2). Finally, to assess additional information of EEG parameters to clinical variables, we performed logistic regression analysis including both, EEG parameters, age and duration of surgery (model 3). We repeated these analyses for the widened relative delta power (1-6 Hz), since this variable was found to best distinguish delirium from non-delirium in the validation study of the EEG-based delirium monitor.

For all models we assessed predictive performance in terms of discrimination and calibration. Based on the predicted probabilities a receiver operating characteristic (ROC) curve was created and the accompanying area under the ROC curve (AUROC) with 95% confidence interval (CI) was calculated. The ROC curves were compared with the ROC curve of model 1 (clinical parameters) as reference using the Delong method[19]. To assess the calibration of each model the Hosmer-Lemeshow statistic was calculated based on deciles.

Finally, to assess whether including the repeated measures in the same individuals as independent observations have influenced the results, we repeated the analyses for all three models with only one measure per patient as sensitivity analysis: the EEG recordings of day 1 and delirium diagnosis of day 2 after surgery. Statistical significance was defined as $p<0.05$. Analyses were performed in MATLAB (Version 2015a, The MathWorks, Inc.) and IBM SPSS Statistics (Version 21, SPSS Inc., Chicago, IL).

RESULTS

POPULATION AND DELIRIUM ASSESSMENTS

Of 196 included patients in the main analysis, of whom 25 patients were excluded for analysis due to logistical reasons (e.g., surgery was canceled, refusal of further

Table 7.1. Characteristics of all patients.

	Never delirious during follow-up (n=99)	Ever delirious during follow-up (n=11)	p-value
Mean age ± SD (years)	76.3 ± 6.5	77.9 ± 7.9	0.45
Male sex, n (%)	70 (70.7)	7 (63.6)	0.63
Alcohol consumption*, n (%)			1.00
None	36 (36.4)	4 (36.4)	
1-14 per week	42 (42.4)	5 (45.5)	
> 14 per week	19 (19.2)	2 (18.2)	
Medical history, n (%)			
TIA or Stroke	26 (26.3)	2 (18.2)	0.73
Psychiatric disease	5 (5.1)	1 (9.1)	0.48
Median MMSE (min-max)*	29 (18-30)	29 (26-30)	0.73
Surgery type, n (%)			0.56
Cardiothoracic or vascular	87 (87.9)	9 (91.8)	
Orthopedic	8 (8.1)	1 (9.1)	
Other	4 (4.0)	1 (9.1)	
Median duration of surgery (minutes, min-max)	159 (107-213)	181 (120-227)	0.36

*Values are presented as mean ± SD, frequency (%), or median (minimum-maximum). The variables between patients who were never delirious and ever delirious during follow-up were compared; t-test for mean age, Mann Whitney U test for MMSE and duration of surgery, chi-square for sex, and Fisher exact test for the remaining categorical variables. Surgery type other was otorhinolaryngologic (n=2), gastrointestinal (n=1), and general surgery (n=1) for never delirious group and otorhinolaryngologic (n=1) for ever delirious group. *Missing values; alcohol consumption 2 missings and MMSE 22 missing in the never delirious group, 1 MMSE missing in the delirious group. MMSE = mini-mental state examination, n = number, SD = standard deviation.*

Table 7.2. Characteristics of all transitions from non-delirium to delirium or non-delirium the next day.

	Non-delirious the next day (n=153)	Delirious the next day (n=11)	p-value
Mean age ± SD (years)	76.6 ± 6.5	77.9 ± 7.9	0.54
Male sex, n (%)	111 (72.5)	7 (63.6)	0.53
Alcohol consumption*, n (%)			1.00
None	55 (35.9)	4 (36.4)	
1-14 per week	65 (42.5)	5 (45.5)	
> 14 per week	31 (20.3)	2 (18.2)	
Medical history, n (%)			
TIA or Stroke	39 (25.5)	2 (18.2)	0.73
Psychiatric disease	9 (5.9)	1 (9.1)	0.51
Median MMSE (min-max)*	29 (18-30)	29 (26-30)	0.73
Surgery type, n (%)			0.38
Cardiothoracic or vascular	135 (88.2)	9 (81.8)	
Orthopedic	13 (8.5)	1 (9.1)	
Other	5 (3.3)	1 (9.1)	
Median duration of surgery (minutes, min-max)	159 (107-208)	181 (120-227)	0.29

*Values are presented as mean ± SD, frequency (%), or median (minimum-maximum). Variables were compared between non-delirious or delirious the next day; t-test for age, Mann Whitney U test for MMSE and duration of surgery, chi-square for sex, and Fisher exact test for the remaining categorical variables. Surgery type other was otorhinolaryngologic (n=2), gastrointestinal (n=1), and general surgery (n=1) for non-delirious the next day group and otorhinolaryngologic (n=1) for delirious the next day group. *Missing values; alcohol consumption 2 missings and MMSE 31 missing in the non-delirious the next day group, 1 MMSE missing in the delirious the next day group. MMSE = mini-mental state examination, n = number, SD = standard deviation.*

participation, insufficient quality of EEG recording). In addition, 20 patients were excluded since they only completed one postoperative assessment and three patients were excluded because no assessment was performed on the second day after surgery (T2), therefore, it was not possible to compare current day diagnosis with preceding day EEG recording. Finally, twelve patients were excluded because of insufficient EEG recordings within the transition of interest. The study population therefore consisted of 110 patients, 77 (70%) male, and mean age 76.5 ± 6.6 years, with 164 accompanying transitions, from which 11 (6.7%) were non-delirious to delirious the next day. In Table 7.1 the patients' characteristics are presented. The characteristics of the transitions were shown in Table 7.2. No differences were found between patients or between the characteristics of the transitions. Most delirium cases included in this analysis were diagnosed on day T2 (n=8) versus day T3 (n=3).

EEG PARAMETERS

The EEG parameters (relative delta power (1-4 Hz) and widened relative delta power (1-6 Hz)) recorded on day X were compared between the patients who were classified as delirious the next day (n=11) and patients who were classified as non-delirious (n=153), see Figure 7.1. Relative delta power (1-4 Hz) and widened relative delta power (1-6 Hz) were significantly increased in patients who were diagnosed as delirious the next day (day X+1), compared to patients who were not delirious on the next day (p=0.035 and p=0.033, respectively).

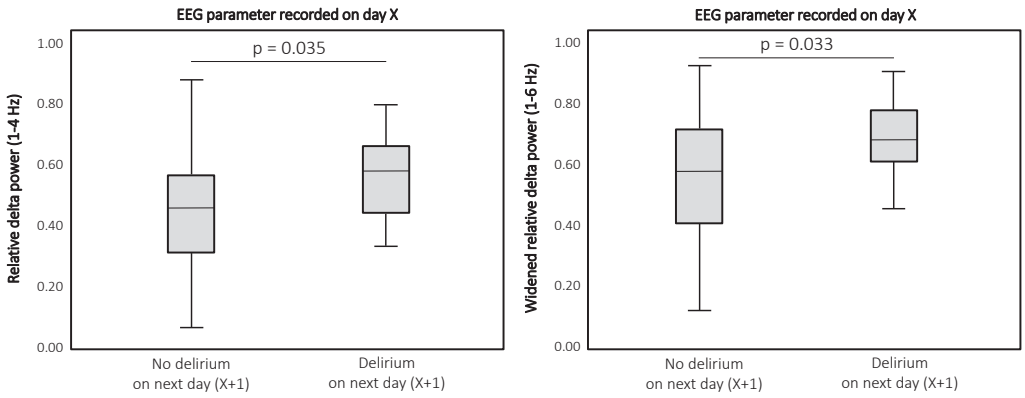


Figure 7.1. Relative delta power (1-4 Hz, left) and widened relative delta power (1-6 Hz, right) recorded on day X in patients diagnosed as no delirium (n=153) or delirium (n=11) on the next day (day X+1).
P-values are presented in the Figures.

PREDICTION OF DELIRIUM THE NEXT DAY

First, the predictive ability of easily assessable, clinical variables (i.e., age and duration of surgery) was evaluated (model 1), Table 7.3. The corresponding AUROC curve was presented in Figure 7.2, AUROC value was 0.56 (95% CI 0.42-0.70). Secondly, model 2 based on EEG parameters only was assessed. Model 2 showed an odds ratio of 1.48 (95% CI 1.01-2.15) for relative delta power (1-4 Hz) and 1.52 (95% CI 1.03-2.24) for the widened relative delta power (1-6 Hz). Thus, both EEG parameters were associated with the diagnosis of delirium the next day. The accompanying AUROC values were 0.69 (95% CI 0.56-0.83) and 0.69 (95% CI 0.55-0.84) for relative delta power (1-4 Hz) and widened relative delta power (1-6 Hz), see Figure 7.2.

Model 3, including the EEG parameter and clinical parameters, showed and odds ratio of 1.46 (95% CI 1.00-2.15) for the relative delta power (1-4 Hz), which was not significantly associated (p=0.051) with the delirium diagnosis the next day. The widened relative delta power (1-6 Hz) in model 3 was significantly associated with the diagnosis the next day (OR 1.50 (95% CI 1.01-2.24), p=0.046). The accompanying AUROC was 0.70 (95% CI 0.56-0.84), see Figure 7.2.

Table 7.3. Logistic regression models for prediction of delirium diagnosis the next day.

Variables	OR (95% CI)	p-value	Variables	OR (95% CI)	p-value
Model 1 (clinical parameters only)			**Model 1** (clinical parameters only)		
Age	1.05 (0.95-1.16)	0.379	Age	1.05 (0.95-1.16)	0.379
Duration of surgery	1.04 (0.96-1.13)	0.342	Duration of surgery	1.04 (0.96-1.13)	0.342
Model 2 (EEG parameter only)			**Model 2** (EEG parameter only)		
Relative delta power (1-4 Hz)	1.48 (1.01-2.15)	0.041*	Widened relative delta power (1-6 Hz)	1.52 (1.03-2.24)	0.037*
Model 3 (EEG parameter and clinical parameters)			**Model 3** (EEG parameter and clinical parameters)		
Relative delta power (1-4 Hz)	1.46 (1.00-2.15)	0.051	Widened relative delta power (1-6 Hz)	1.50 (1.01-2.24)	0.046*
Age	1.04 (0.94-1.16)	0.451	Age	1.05 (0.94-1.17)	0.426
Duration of surgery	1.04 (0.95-1.13)	0.440	Duration of surgery	1.03 (0.94-1.13)	0.501

*Model 2 and model 3 were separately presented for relative delta power (1-4 Hz, left) and widened relative delta power (1-6 Hz, right). Model 1 was presented in both columns. EEG parameters were included in the models with steps of 0.10 (for example the odds on having delirium the next day is 1.48 higher when a patients has a relative delta power of 0.60 compared to a relative delta power of 0.50), duration of surgery was included per 10 minutes, and age was included per year. Hosmer and Lemeshow test was not significant for all models (p>0.05). *Significant, p<0.05. CI = confidence interval, OR = odds ratio.*

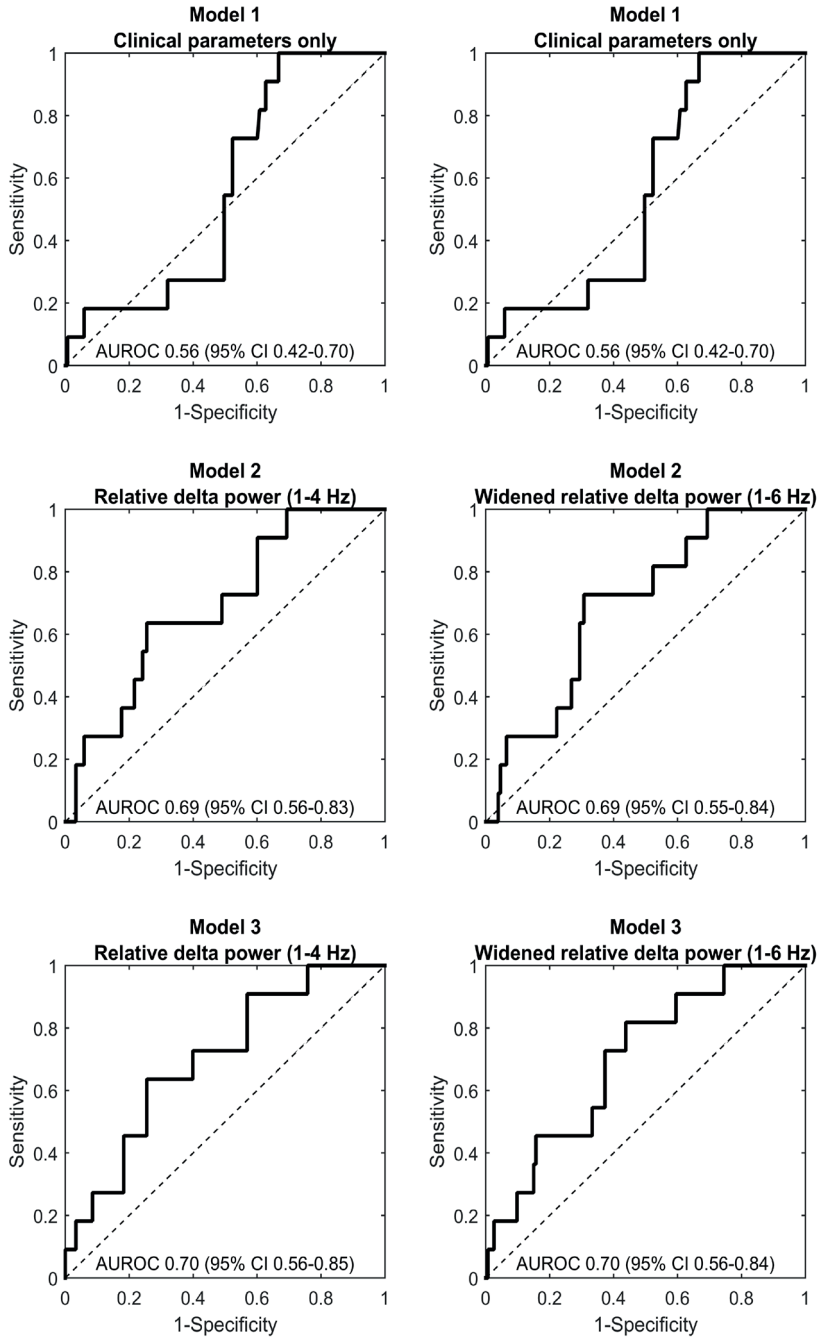


Figure 7.2. ROC curves of the three models predicting delirium the next day.
Model 1 on upper panels; clinical parameters only (age and duration of surgery), note that ROC curve of model 1 is presented in both columns. Model 2 on middle panels; EEG parameters only. Model 3 on the lower panels; combination of model 1 and 2: clinical parameters and EEG parameters. The AUROC values Model 3 with relative delta power (1-4 Hz, left panels) and widened relative delta power (1-6 Hz, right panels) were significantly higher compared to model 1 according to Delong method (p=0.041 and p=0.036, respectively). AUROC = area under the receiver operating characteristic curve, CI = confidence interval.

Using Delong method to assess the improvement of the ROC curves compared to model 1, as currently available clinical information, we found significant improvement in model 3, for both relative delta power (1-4 Hz, p=0.041) and widened relative delta power (1-6 Hz, p=0.036). The ROC curves of model 2 based on EEG parameters only, were not significantly higher than model 1.

As sensitivity analysis for the effect of repeated measures, we repeated the logistic regression analyses. Results were based on 80 independent cases with a non-delirious classification the next day and 8 delirium cases the next day. Data are presented in supplementary Table S7.1 and Figure S7.1. Model 2 and model 3, based on 1 transition day only, showed significant association between both EEG parameters and the diagnosis of delirium the next day. Moreover, the resulting AUROC values were significantly higher compared to model 1.

DISCUSSION

In this multicenter study we found that one minute of a single-channel quantitative EEG recording could predict delirium the next day. In addition, EEG parameters significantly improved the prediction of next day delirium compared to clinical characteristics only.

Delirium can be diagnosed based on the Diagnostic and Statistical Manual of mental disorders (DSM)-5[15] by an expert. However, in daily clinical practice delirium screening is performed by clinical nurses, based on one of the several available screening tools. Although diagnosis is a dichotomous outcome variable, there is a broad range of severity in delirium, for example based on the DRS-R-98[17,20], and difference in motor subtype (i.e., hyperactive, hypoactive, and mixed type)[21,22]. Therefore, delirium may be seen as a spectrum[23], with the possibility to detect delirium in an early stage, possible prior to the fulfilling of all symptoms according to DSM-5 criteria or the definitions by screening tools. However, this is not possible with the current screening tools. Since EEG changes, especially increase in delta and theta activity, are related to diagnosis of delirium[14,24,25], the onset of delirium prior to full syndrome, may also show slowing of the EEG. The results of our study showed an increase of relative delta power and widened relative delta power one day before the diagnosis of delirium by experts, which suggest an opportunity to detect delirium before the onset of the full syndrome enabling early treatment or prevention steps for these patients. However, the discriminative ability of the EEG parameters only to predict delirium the next day was moderate.

To assess the ability to predict delirium the next day based on clinical parameters only, we included age and duration of surgery, as these variables were previously described as risk factors for delirium[1,26–29]. In the current study, other preoperative risks factors for delirium, such as cognitive impairment and alcohol misuse[1], were not sufficiently collected and therefore not included in the model. Moreover, based on the relative low number of delirium cases only age and duration of surgery were included in the clinical parameter model. The low AUROC value of the clinical model (model 1: AUROC of 0.56) indicates that age and duration of surgery are not reliable predictors for delirium in this study population. This was not expected based on previous results, were age is one of the important variables predicting delirium[1]. However, these clinical parameters predict delirium after surgery instead of a day-to-day prediction. The clinical model could have benefitted when precipitating risk factors for delirium, such as administration of medication, infection, or immobilization, were included[27], however, these were not collected. Model 3 based on clinical parameters and the widened relative delta power (1-6 Hz), showed significantly better discriminative values compared to the model of clinical parameters only. Model 2, based on EEG parameters only did not had a significant higher AUROC, although the AUROC was moderate (0.69) compared to the poor AUROC of model 1 (0.56). This suggests that the EEG findings describe different aspects of the risk on delirium, because there was no difference between model 2 (i.e., only EEG parameters) and model 3 (i.e., EEG parameters and clinical parameters) when comparing the AUROC values.

Currently available prediction models for delirium are based on patient characteristics, such as age, hearing impairment or previous delirium, and daily clinical parameters, such as administered benzodiazepines and blood levels[30–32], which are known to be risk factors for delirium[1]. In the current study, we measured the brain activity before the delirium diagnosis, which may be the ongoing process of the encephalopathy. Therefore, the suggested single-channel EEG recording may improve day-to-day delirium prediction in clinical practice in combination with already existing risk prediction models for delirium, however, this should be confirmed in a study collecting all clinical parameters.

The strengths of the current study include high quality of delirium diagnosis by two or three experts based on comprehensive cognitive assessment and multicenter nature of the study. Moreover, this is the first study using EEG in day-to-day prediction of delirium in postoperative elderly patients. This study also has several limitations which need to be addressed. First, the EEG recording and cognitive assessment were performed once a day

due to logistical reasons. This may have resulted in missed cases due to the fluctuating course of delirium, however, providing the information of the patient record should have reduced the missed number of delirium cases. Secondly, the number of positive cases was relatively small (n=11) and included both possible delirium and delirium diagnosis. This limited the number of variables which could be included in the logistic regression models, and therefore could have negatively affected the results. Finally, repeated measures were included as independent observations. The effect of this clustering of the data was assessed by a sensitivity analysis, repeating the logistic regression analysis based on one day only (i.e., EEG recording of T1, first day after surgery and delirium diagnosis on T2, second day after surgery). The resulting AUROC values were higher compared to the main analysis, although not statistically compared. This suggests that including two transitions of one patient in the main analysis as independent observations did not overestimate the predictive ability of the EEG parameters.

Single-channel EEG was shown to be promising in diagnosing delirium[14]. Moreover, the results of the current study suggest that a single-channel EEG recording could also predict delirium in postoperative older patients. However, the predictive ability should be studied in a larger study population, to increase the number of delirious cases. Moreover, the EEG-recording could start during surgery or directly after surgery for an earlier prediction. For clinical purposes, it would be interesting to repeat this study in geriatric ward patients or ICU patients, because of the high incidence of delirium. Moreover, future studies should evaluate the effects of used sedatives and cognitive problems such as dementia on the EEG parameters, since these also result in slowing of the EEG[33,34]. This may hamper prediction of delirium based on a single-channel EEG. Finally, although logistical difficult, recording continuous EEG together with multiple delirium assessments per day, would give more insight in the relation between course of delirium (i.e., from the prodromal phase toward the recovery of delirium) and EEG fluctuations.

CONCLUSION

This is the first study showing that it is possible to predict delirium the next day based on a one channel EEG recording in older postoperative patients. Prediction of delirium one day ahead based on the one channel EEG recording in combination with already available prediction models, should enable early intervention and may improve clinical outcome of these patients.

ACKNOWLEDGEMENTS

This work was supported by Clinical Research Award ESICM, Technology Foundation STW (14066), European Union (H2020-PHC-12-2014-672974) and institutional support.

REFERENCES

1 Inouye, S. K., Westendorp, R. G. & Saczynski, J. S. Delirium in elderly people. Lancet 383, 911–922 (2014).
2 Kalish, V. B., Gillham, J. E. & Unwin, B. K. Delirium in Older persons: evaluation and management. Am. Fam. Physician 90, 150–158 (2014).
3 Siddiqi, N., House, A. O. & Holmes, J. D. Occurrence and outcome of delirium in medical in-patients: a systematic literature review. Age Ageing 35, 350–364 (2006).
4 Wolters, A. E., van Dijk, D., Pasma, W., et al. Long-term outcome of delirium during intensive care unit stay in survivors of critical illness: a prospective cohort study. Crit. Care 18, R125 (2014).
5 Ely, E. W., Shintani, A., Truman, B., et al. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. JAMA 291, 1753–1762 (2004).
6 Salluh, J. I. F., Wang, H., Schneider, E. B., et al. Outcome of delirium in critically ill patients: systematic review and meta-analysis. BMJ 350, h2538 (2015).
7 Milisen, K., Foreman, M. D., Abraham, I. L., et al. A nurse-led interdisciplinary intervention program for delirium in elderly hip-fracture patients. J. Am. Geriatr. Soc. 49, 523–532 (2001).
8 O'Hanlon, S., O'Regan, N., Maclullich, A. M. J., et al. Improving delirium care through early intervention: from bench to bedside to boardroom. J Neurol Neurosurg Psychiatry 85, 207–213 (2014).
9 Heymann, A., Radtke, F., Schiemann, A., et al. Delayed treatment of delirium increases mortality rate in intensive care unit patients. J. Int. Med. Res. 38, 1584–1595 (2010).
10 Inouye, S. K., Van Dyck, C. H., Alessi, C. A., et al. Clarifying confusion: the confusion assessment method: a new method for detection of delirium. Ann. Intern. Med. 113, 941–948 (1990).
11 Ely, E. W., Margolin, R., Francis, J., et al. Evaluation of delirium in critically ill patients: validation of the confusion assessment method for the intensive care unit (CAM-ICU). Crit. Care Med. 29, 1370–1379 (2001).
12 Eijk, M. M. van, Boogaard, M. van den, Marum, R. J. van, et al. Routine use of the confusion assessment method for the intensive care unit: a multicenter study. Am. J. Respir. Crit. Care Med. 184, 340–344 (2011).
13 Rice, K. L., Bennett, M., Gomez, M., et al. Nurses' recognition of delirium in the hospitalized older adult. Clin. Nurse Spec. 25, 299–311 (2011).
14 Van der Kooi, A. W., Zaal, I. J., Klijn, F. A., et al. Delirium detection using EEG. Chest 147, 94–101 (2015).
15 American Psychiatric Association. Diagnostics and statistical manual of mental disorders. 5th edition. (2013).

16 Azabou, E., Magalhaes, E., Braconnier, A., et al. Early standard electroencephalogram abnormalities predict mortality in septic intensive care unit patients. PLoS One 10, 1–18 (2015).
17 Trzepacz, P. T., Mittal, D., Torres, R., et al. Validation of the delirium rating scale-revised-98: comparison with the delirium rating scale and the cognitive test for delirium. J. Neuropsychiatry Clin. Neurosci. 13, 229–242 (2001).
18 De Rooij, S. E., van Munster, B. C., Korevaar, J. C., et al. Delirium subtype identification and the validation of the delirium rating scale-revised-98 (Dutch version) in hospitalized elderly patients. Int. J. Geriatr. Psychiatry 21, 876–882 (2006).
19 DeLong, E. R., DeLong, D. M. & Clarke-Pearson, D. L. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametetric approach. Biometrics 44, 837–845 (1988).
20 Boettger, S., Nuñez, D. G., Meyer, R., et al. Subsyndromal delirium in the intensive care setting: Phenomenological characteristics and discrimination of subsyndromal delirium versus no and full-syndromal delirium. Palliat. Support. Care 6, 1–11 (2017).
21 Meagher, D. Motor subtypes of delirium: Past, present and future. Int. Rev. Psychiatry 21, 59–73 (2009).
22 Grover, S., Sharma, A., Aggarwal, M., et al. Comparison of symptoms of delirium across various motoric subtypes. Psychiatry Clin. Neurosci. 68, 283–291 (2014).
23 Ouimet, S., Riker, R., Bergeon, N., et al. Subsyndromal delirium in the ICU: evidence for a disease spectrum. Intensive Care Med. 33, 1007–1013 (2007).
24 Kooi, A. W. van der, Leijten, F. S. S., Wekken, R. J. van der, et al. What are the opportunities for EEG-based monitoring of delirium in the ICU? J. Neuropsychiatry Clin. Neurosci. 24, 472–477 (2012).
25 Engel, G. L. & Romano, J. Delirium, a syndrome of cerebral insufficiency. 1959. J. Neuropsychiatry Clin. Neurosci. 16, 526–538 (2004).
26 Afonso, A., Scurlock, C., Reich, D., et al. Predictive model for postoperative delirium in cardiac surgical patients. Semin. Cardiothorac. Vasc. Anesth. 14, 212–217 (2010).
27 Zaal, I. J. & Slooter, A. J. C. Delirium in critically ill patients: Epidemiology, pathophysiology, diagnosis and management. Drugs 72, 1457–1471 (2012).
28 Hatano, Y., Narumoto, J., Shibata, K., et al. White-matter hyperintensities predict delirium after cardiac surgery. Am. J. Geriatr. Psychiatry 21, 938–945 (2013).
29 Raats, J. W., Steunenberg, S. L., de Lange, D. C., et al. Risk factors of post-operative delirium after elective vascular surgery in the elderly: A systematic review. Int. J. Surg. 35, 1–6 (2016).
30 Boogaard, M. V. D., Pickkers, P., Slooter, A. J. C., et al. Development and validation of PRE-DELIRIC (PREdiction of DELIRium in ICu patients) delirium prediction model for intensive care patients: observational multicentre study. BMJ 344, e420 (2012).
31 De Wit, H. A J. M., Winkens, B., Mestres Gonzalvo, C., et al. The development of an automated ward independent delirium risk prediction model. Int. J. Clin. Pharm. 38, 915–923 (2016).
32 Kim, M. Y., Park, U. J., Kim, H. T., et al. DELirium Prediction Based on Hospital Information (Delphi) in General Surgery Patients. Medicine (Baltimore). 95, e3072 (2016).
33 De Visser, S. J., Van der Post, J., Pieters, M. S. M., et al. Biomarkers for the effects of antipsychotic drugs in healthy volunteers. Br. J. Clin. Pharmacol. 51, 119–132 (2001).
34 Moretti, D. V., Babiloni, C., Binetti, G., et al. Individual analysis of EEG frequency and band power in mild Alzheimer's disease. Clin. Neurophysiol. 115, 299–308 (2004).

SUPPLEMENTARY MATERIAL

Table S7.1. Repeated logistic regression models for prediction of delirium diagnosis the next day, T1 to T2 transition only.

Variables	OR (95% CI)	p-value	Variables	OR (95% CI)	p-value
Model 1 (clinical parameters only)			**Model 1** (clinical parameters only)		
Age	1.00 (0.89-1.13)	0.984	Age	1.00 (0.89-1.13)	0.984
Duration of surgery	1.05 (0.95-1.15)	0.369	Duration of surgery	1.05 (0.95-1.15)	0.369
Model 2 (EEG parameter only)			**Model 2** (EEG parameter only)		
Relative delta power (1-4 Hz)	1.84 (1.10-3.09)	0.021*	Widened relative delta power (1-6 Hz)	1.97 (1.07-3.01)	0.027*
Model 3 (EEG parameter and clinical parameters)			**Model 3** (EEG parameter and clinical parameters)		
Relative delta power (1-4 Hz)	1.81 (1.08-3.05)	0.025*	Widened relative delta power (1-6 Hz)	1.76 (1.04-2.98)	0.035*
Age	0.99 (0.86-1.13)	0.834	Age	0.99 (0.87-1.14)	0.898
Duration of surgery	1.02 (0.92-1.14)	0.696	Duration of surgery	1.02 (0.91-1.13)	0.760

*Model 2 and model 3 were separately presented for relative delta power (1-4 Hz, left) and widened relative delta power (1-6 Hz, right). Model 1 was presented in both columns. EEG parameters were included in the models with steps of 0.10, duration of surgery was included per 10 minutes, and age was included per year. Hosmer and Lemeshow test was not significant for all models (p>0.05). *Significant, p<0.05. CI = confidence interval, OR = odds ratio.*

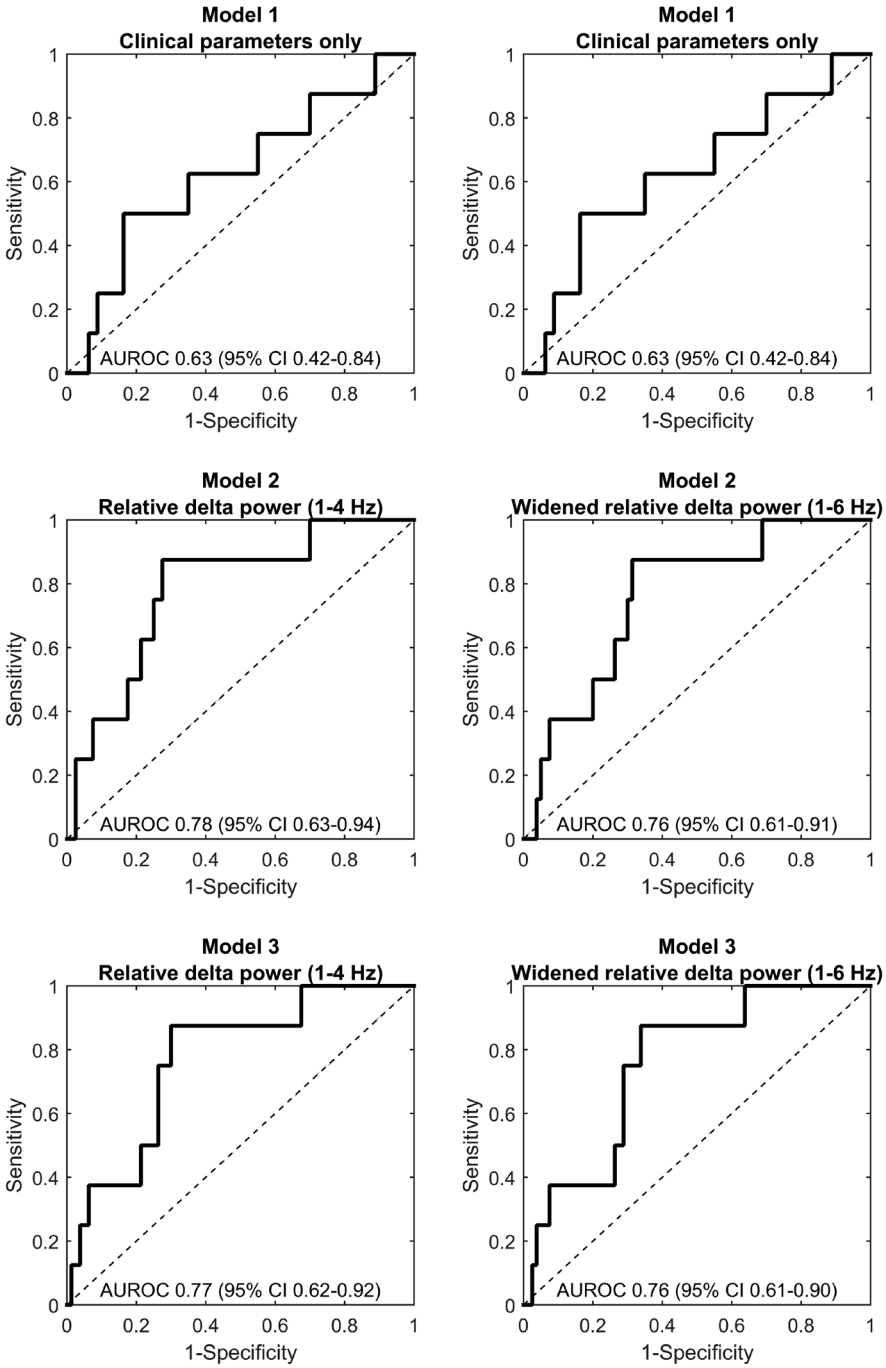
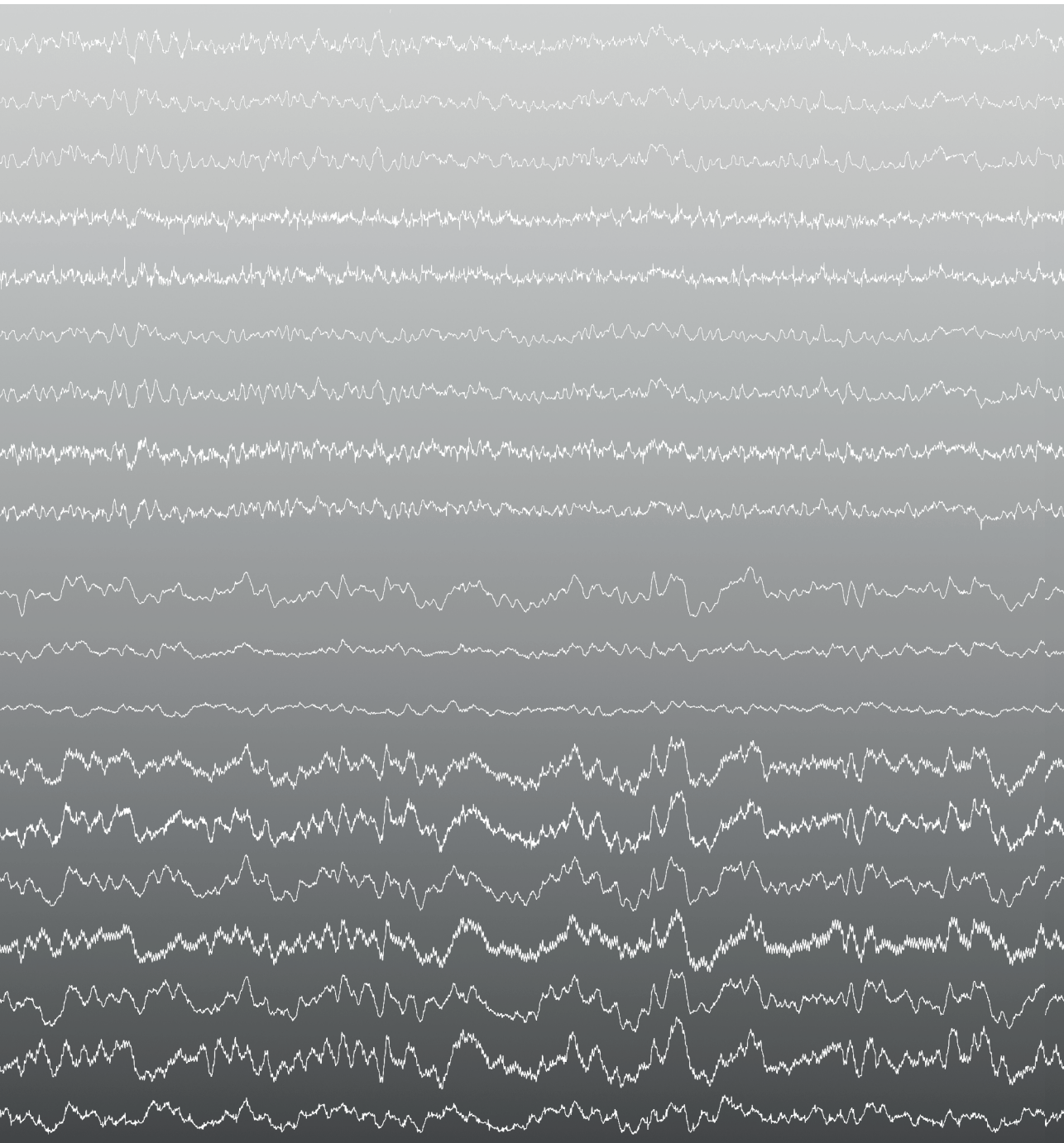


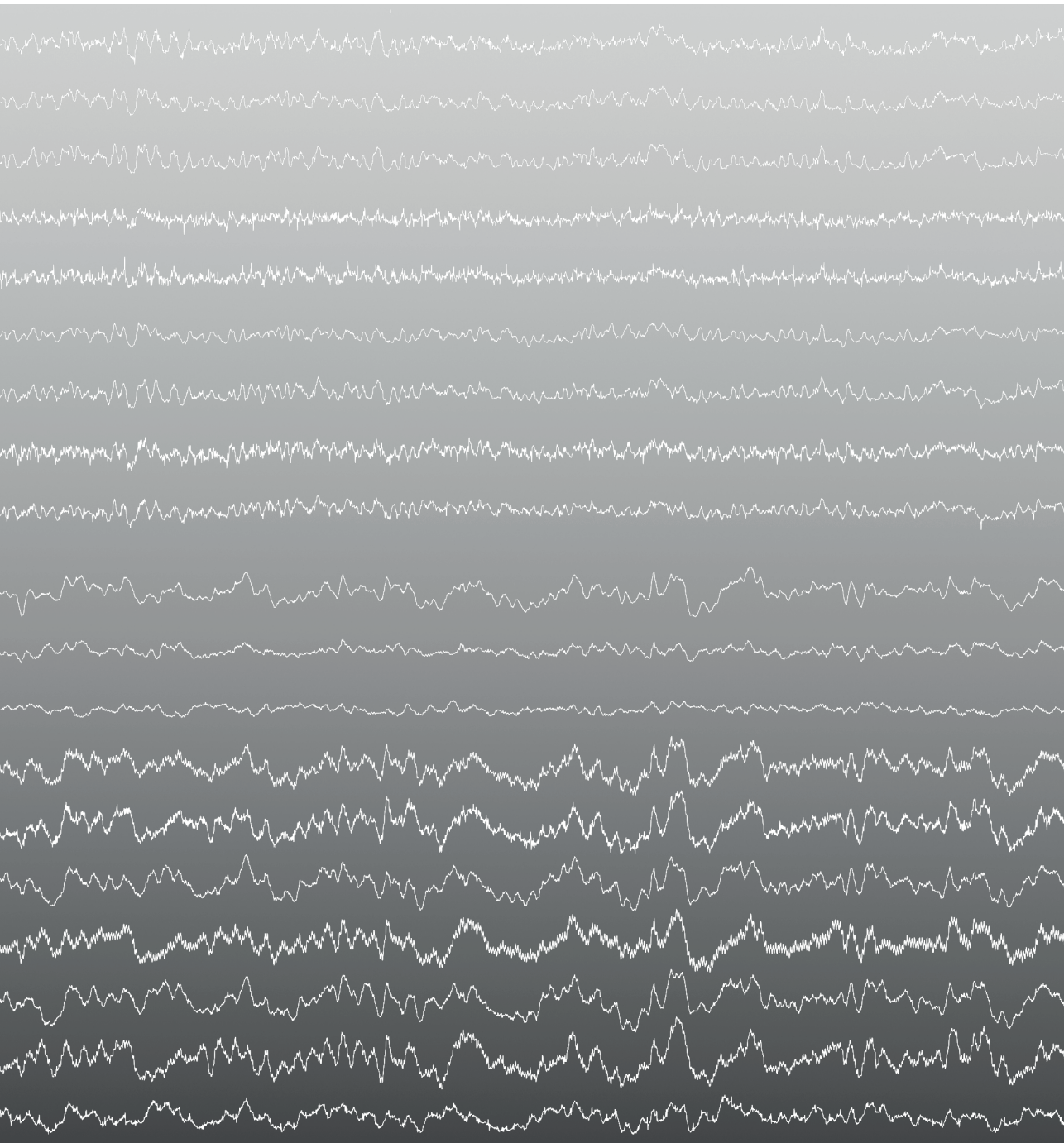
Figure S7.1. The ROC curves with accompanying AUROC values for the prediction models 1, 2, and 3 without repeated measures.

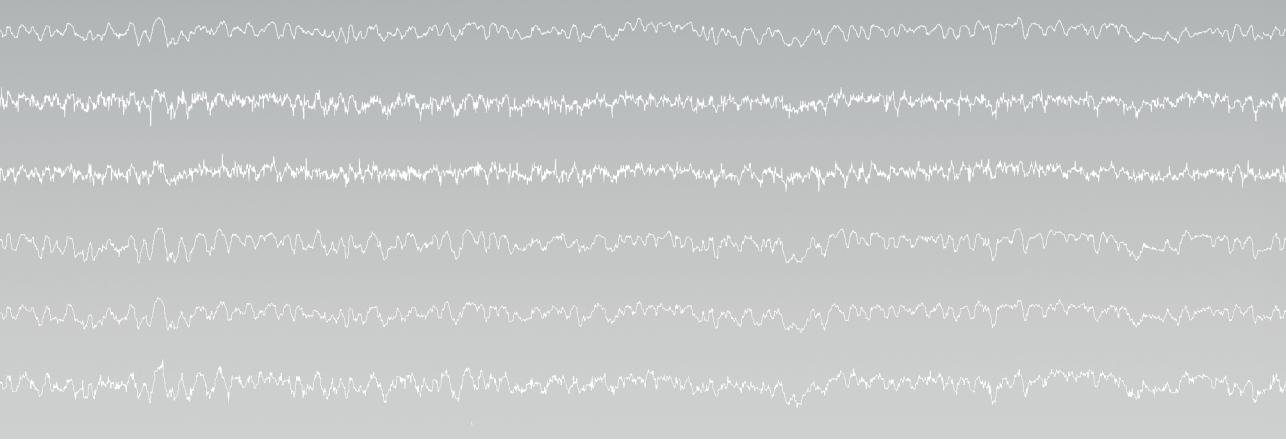
Model 1; clinical parameters only (age and duration of surgery), note that ROC curve of model 1 is presented twice. Model 2; EEG parameters only (left, relative delta power (1-4 Hz), right, widened relative delta power (1-6 Hz)). Model 3; model 1 and model 2 combined (left, relative delta power (1-4 Hz), right, widened relative delta power (1-6 Hz)). AUROC values were compared to model 1 using Delong method: Model 2 relative delta power, p=0.005 and widened relative delta power, p=0.014. Model 3 relative delta power, p=0.006 and widened relative delta power, p=0.011.



PART 3

SUMMARY AND GENERAL DISCUSSION





8

SUMMARY AND GENERAL DISCUSSION

SUMMARY

Delirium is a serious health care problem, which is characterized by an acute disturbance of cognition and attention which fluctuates during the day. The pathophysiology of delirium is not fully understood and the diagnosis of delirium can be difficult, which leads to under detection of delirium. In the intensive care unit (ICU), the incidence of delirium is high, especially the hypoactive subtype when the patient is lethargic and withdrawn. In clinical practice, it can be difficult to distinguish hypoactive delirium from the effects of sedation.

The first aim of this thesis was to improve our knowledge of reduced levels of consciousness during hypoactive delirium and light levels of sedation, by studying interactions between brain areas, using functional network analysis. The second aim was to improve delirium detection in clinical, routine practice.

PART I: RESTING STATE EEG DURING HYPOACTIVE DELIRIUM AND LIGHT LEVELS OF SEDATION

The brain can be regarded as a complex network. In the last two decades, an increasing number of papers have been published, studying the brain as functional network, using different modalities such as functional magnetic resonance imaging (fMRI), magneto-encephalography (MEG), and electroencephalography (EEG). The main advantage of EEG is that it can be used at the bedside. A major limitation is the low spatial resolution compared to fMRI or MEG, and therefore it is difficult to specify the anatomical location of the recorded functional network. A broad range of functional connectivity measures has been developed in recent years, enabling a more thorough interpretation of functional network observations. However, the comparison between studies with different modalities or different functional network measures remains difficult. **Chapter 2** describes the pipeline of functional connectivity and network analysis with EEG or MEG, from recording settings towards assessment of the functional networks. For comparison between studies on functional brain networks, it is essential to know the influence of methodological differences.

In ICU patients disturbed consciousness is common, mainly due to administration of sedatives or hypoactive delirium. Although they can be clinically similar, the cause of reduced consciousness is different, and therefore the underlying mechanism might be different as well. Changes in functional connectivity, directed connectivity, and

the organization of the network were described previously for the levels of reduced consciousness, however not compared to hypoactive delirium which is a similar state of reduced consciousness. Therefore, the functional brain network during hypoactive delirium and recovery from anesthesia is studied in **chapter 3**, and the effects of midazolam and propofol individually in **chapter 4**. Hypoactive delirium and recovery from anesthesia after cardiac surgery were found to be associated with reduced functional connectivity and loss of the normal back-to-front pattern of information flow. Therefore, functional connectivity and directed connectivity might be related to reduced levels of consciousness. Differences between hypoactive delirium and recovery from anesthesia were found in the organization of the functional network, or the topology. Hypoactive delirium showed a less integrated, and therefore less efficient network, whereas the recovery from anesthesia patients showed a more integrated organization of their functional network. This difference could be explained by the difference in state of the patients: the group recovering from anesthesia is regaining consciousness, whereas hypoactive patients are not in a transition towards a fully awake state.

In **chapter 4** the effects of light sedation due to midazolam and propofol in patients undergoing an endoscopic procedure are described. Both sedatives resulted in slowing of the EEG, reduced functional connectivity and loss of the back-to-front pattern in the alpha frequency band. No differences in functional network topology based on the minimum spanning tree (MST) were found compared to the baseline, the awake state. The MST is the backbone structure of the network including the strongest connections without forming loops in the network. A significantly greater reduction of alpha band functional connectivity and disruption of the directed delta band connectivity was found with propofol compared to midazolam. This suggests that midazolam and propofol could have a distinct neurophysiological signature for light levels of sedation.

PART II: VALIDATION OF AN EEG-BASED DELIRIUM MONITOR

The second aim of this thesis was to improve delirium detection in clinical practice, because delirium detection in the clinical practice is insufficient and delirium cases can be missed. First, the current clinical situation related to delirium detection was evaluated. In **chapter 5** we assessed the agreement between delirium experts, who based their conclusion whether a patient was delirious or not on exactly the same clinical information. Postoperative patients were visited by a researcher performing standardized cognitive assessments, which were recorded on video. Delirium experts

evaluated these videos with additional patient record information. A considerable amount of disagreement (21%) was found between the classifications of the different experts. This indicates the difficulty of the diagnosis of delirium, even by experts. In case of disagreement, a third expert was consulted for the final classification.

Using this final classification of the delirium experts, the sensitivity of the delirium screening tools for routine daily practice by clinical nurses was evaluated. In about 50% of the cases either the confusion assessment method of intensive care unit (CAM-ICU) or the delirium observation screening (DOS) scale was applied and documented by the nurses. The combined sensitivity was 43%, which is considered to be insufficient for delirium detection in clinical practice. This indicates the need for an objective tool to detect delirium in daily practice.

Second, a prototype electroencephalography (EEG)-based delirium monitor was developed and validated in **chapter 6**. In 159 older postoperative patients, a single-channel EEG recording was performed followed by a standardized cognitive assessment, which was recorded on video. The videos were evaluated by two, or in case of disagreement, three delirium experts. Several EEG-based measures were calculated and compared with the final classification using a receiver operating characteristic (ROC) curve. The relative delta power (1-4 Hz) resulted in an area under the ROC curve (AUROC) of 0.75 (95% confidence interval (CI) 0.69-0.82). The AUROC of the widened relative delta power (1-6 Hz) was significantly higher, 0.78 (95% CI 0.72-0.84). These results suggest that delirium can be detected in postoperative patients, using a simple single-channel EEG recording.

Finally, in **chapter 7**, the predictive ability of the EEG-based delirium monitor for observed delirium the next day was assessed. Both, relative delta power (1-4 Hz) and widened relative delta power (1-6 Hz) were significantly higher in non-delirious patients on the day of recording who developed delirium the next day compared to non-delirious patients who remained non-delirious the day after the recording. Prediction models including clinical parameters and relative delta power (1-4 Hz) resulted in an AUROC of 0.70 (95% CI 0.56-0.85). Therefore, this study suggests that it is possible to predict delirium based on a single-channel EEG recording one day before delirium experts diagnosed delirium. However, this should be confirmed in a larger study population with a higher number of delirious cases.

GENERAL DISCUSSION

The aim of this thesis was twofold: to obtain a better understanding of the neurophysiological mechanisms underlying delirium and light levels of sedation, and to improve delirium detection for routine clinical practice using EEG. The methodology, results and implications of the previous chapters will be described here in a broader perspective.

OPTIMAL FUNCTIONAL CONNECTIVITY AND NETWORK ANALYSES

It is essential to reliable measure brain activity and use reliable methods for research, which encompasses several aspects such as the instruction to the patients and selection of correct methods. Resting state is defined as the state in which a subject is awake but not performing any mental or physical task. Instructing subjects to be in this resting state can be difficult. These instructions varied between different studies and the interpretation and experience of the instructions can differ between subjects (**chapter 2**). Moreover, it is even more difficult to instruct patients, for example with reduced level of consciousness (**chapter 4**) or during delirium with reduced attention (**chapter 3**). One could even argue that during reduced levels of consciousness subjects cannot have a formal resting state, because they are not fully awake. However, spontaneous brain activity can be recorded using one of the imaging modalities and compared across different levels of consciousness, which can be compared to the resting state.

A large number of connectivity measures is available (**chapter 2**), all with their own strengths and limitations. Based on the imaging modality, a suitable connectivity measure should be selected. For EEG analysis, it is important to consider the effects of volume conduction and field spread. When multiple EEG channels record the activity of a single underlying source, these time series will be highly similar. For example, coherence, a frequently used measure of functional connectivity (**chapter 2**), is likely to be affected by volume conduction[1], and therefore may lead to erroneous results of functional connectivity. Phase-based approaches which exclude zero-lag phase differences, such as the phase lag index (PLI) are less sensitive to these spurious interactions[1,2].

When the optimal functional connectivity measure is selected, an adjacency matrix can be calculated. This matrix contains the connection strength between all nodes (i.e., in case of EEG all electrodes). Based on this adjacency matrix, the functional network can

be computed. In this thesis the MST was used to evaluate network topology. Several MST measures can be used to quantify the MST structure. Hypoactive delirium patients showed a more path-like MST (i.e., less integrated) whereas patients recovering from anesthesia had a more star-like MST (i.e., more integrated). Light levels of sedation, due to either midazolam or propofol, had no apparent effect on MST topology. The selection of frequency band of interest might have affected this result. The EEG time series were analyzed in four different frequency bands, delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), and beta (13-20 Hz). Moreover, six MST parameters were calculated to describe the different characteristics of the networks. Increasing the number of tests, will also increase the risk of false positive results (i.e., type I error). It is therefore important to critically select the number of tests applied to the data. To reduce the number of tests, we tried to select the frequency band of interest using MST-overlap, to find differences in topography between the groups[3]. A reference MST was created based on the control subjects for each frequency band. The percentage of overlap of a patient's MST with the reference MST was calculated. It was assumed that if there was a significant difference in MST overlap between the groups, this was the result of different underlying topology, and therefore MST measures of this frequency band were further investigated.

In **chapter 3** the MST overlap showed significant differences between the hypoactive delirium, recovery from anesthesia, and non-delirious controls for the delta, alpha, and beta frequency bands. MST measures were calculated and compared between the groups in these frequency bands, showing different topologies. In **chapter 4** however, significant difference in MST overlap were found in the beta frequency band between baseline and propofol sedation. Unexpectedly, the MST measures were not significantly different between baseline and propofol sedation. This may suggest that propofol alters some of the connections of the MST, however the topology (i.e., the global organization) is not affected. Another explanation is that the MST-overlap measure is more sensitive to differences in the MST networks than the individual MST measures. The values of MST-overlap were relatively low, which may implicate high variability of the MST structures of each patient, based on either relative high noise levels or a more dynamic character of the MST structures. It remains unclear if this selection procedure using MST-overlap of the frequency bands of interest for MST analyses is an optimal method. However, it is also unclear if calculation of all MST measures in the frequency bands with sufficient correction for multiple testing, would have led to different results.

IS DELIRIUM A DISCONNECTION SYNDROME?

A previous EEG functional connectivity study in delirium patients suggested that delirium is a disconnection syndrome[4], which is confirmed by the results of **chapter 3**. However, also recovery from anesthesia resulted in a similar reduction of functional connectivity. Moreover, other diseases in which cognition is affected, such as schizophrenia and Alzheimer's disease, are also described as disconnection syndromes[5,6], and therefore disconnection does not seem to be specific for delirium.

The direction of the information flow was also assessed in **chapter 3** and **chapter 4**. Using the directed phase transfer entropy, significant loss of the back-to-front pattern was found during hypoactive delirium as well as during recovery from anesthesia and light levels of sedation due to midazolam or propofol. This loss of directionality was also found in previous studies, assessing the effect of propofol in healthy subjects, and in patients with dementia with Lewy bodies[7–9]. Therefore, loss of directionality seems not to be a specific characteristic for delirium either, but may rather be a characteristic of loss of attention and awareness.

One previous study investigated the functional brain network using fMRI, both during delirium and after recovery of an episode of delirium[10]. The effort of performing fMRI recordings in delirious patients is admirable, because of the difficulty to instruct these patients. Therefore, this is a highly interesting dataset enabling the opportunity to evaluate the functional brain connectivity according to anatomical locations. Choi et al. (2012) used Pearson's correlation as measure of functional connectivity, and found a positive correlation between the time series of the dorsolateral prefrontal cortex and the posterior cingulate cortex during delirium episodes, whereas non-delirious control subjects had a negative correlation between these brain regions[10]. Although these results are not directly comparable to the EEG studies in this thesis, the disturbance of the back-to-front pattern is suggested to point in the same direction: a disturbance of the communication between the frontal and posterior parts of the cortex. Frontoparietal communication is assumed to be essential for attention, thus disruption of this connectivity may lead to inattention, which is a central feature of delirium. However, this seems not to be exclusively related to delirium, since reduced levels of consciousness due to sedation showed a similar loss of back-to-front pattern.

Moreover, Choi et al. also performed an fMRI recording after resolution of the delirious episode, meaning that the symptoms were markedly reduced[10]. However, the disruption of the negative correlation in activity between dorsolateral prefrontal cortex and the posterior cingulate cortex found during the delirious episode was not fully recovered by then. The different transition phases of anesthesia (i.e., induction, maintenance, and emergence) were also found to have unique functional connectivity characteristics, especially based on the directed connectivity[7,8]. Therefore, it would be interesting to study the entire course of delirium using EEG, to evaluate when the functional brain network is fully restored in relation to the symptoms, and to study whether remaining EEG abnormalities predict long-term cognitive impairment.

The differences between hypoactive delirium, recovery from anesthesia, and non-delirious control subjects were found with the MST measures. Betweenness centrality was most specific to distinguish hypoactive delirium from recovery from anesthesia. Choi et al. showed increased connectivity between precuneus and posterior cingulate cortex in the delirious patients, which was suggested as increased integration of the functional network[10]. The functional connectivity between several subcortical brain regions was reduced[10]. How the fMRI results of Choi et al. relate to the MST measures based on EEG recordings is unclear. It would be interesting to perform resting state network analysis using fMRI and EEG simultaneously, enabling interpretation of both results related to anatomical location of the disturbances.

Taken together, we can conclude that delirium can be regarded as a disconnection syndrome, and that more specific characteristics of the functional network such as betweenness centrality, can be used to distinguish delirium from related conditions.

RECOGNITION OF DELIRIUM IN CURRENT CLINICAL PRACTICE

The case description of the delirious patient in **chapter 1** Box 1.1, Mrs.R., is an example of an older frail patient who had major surgery and develops delirium. The first morning after surgery, the patient was drowsy and slept through the day. Although disturbance of the sleep-wake cycles is one of the symptoms of delirium, delirium may go unnoticed in these patients. Mrs. R. was diagnosed with delirium when she was agitated and removed the i.v. tube, which alarmed the nurse.

The hypoactive motor subtype is if often missed by clinicians as a delirium[11]. To improve delirium detection in clinical practice several screening tools have been developed.

Currently, the CAM for general ward patients[12], and the CAM-ICU for the intensive care[13] are the most frequently used delirium screening tools[14,15]. In a research setting these tools can identify delirium cases correctly, however, in clinical practice the detection rate appears to be disappointing. In **chapter 5** the detection rate of delirium by bedside nurses was also low, only 32 out of the 100 (possible) delirium cases were detected and documented. This emphasizes the need of improved delirium detection for clinical practice.

GOLD STANDARD VERSUS REFERENCE STANDARD IN DELIRIUM CLASSIFICATION

For delirium research, etiological studies, clinical trials, and validation studies, it is essential to have a common agreement on the diagnosis of delirium in the study population. The gold standard for delirium diagnosis in most studies is an evaluation by a delirium expert[15]. The term gold standard originates from economy field, signifying a monetary standard. A gold standard has a sensitivity and specificity of 100% by definition, which is not feasible in clinical practice or a medical research setting[16]. For delirium diagnosis, interpretation is needed by the expert, which is subjective and implies the impossibility of having no false positives or false negatives. Moreover, the information available for delirium diagnosis varies substantially in different studies: clinical impression, cognitive testing, clinical notes, physical examination, or a combination of all[17].

Furthermore, in validation studies of delirium tools it is essential to have a small time window between gold standard evaluation and the new proposed delirium screening tool, because of the fluctuating character of delirium. Unfortunately, not all studies on delirium have a reliable gold standard diagnosis, for example due to unclear time delay between gold standard and the new test[14]. Moreover, the DSM criteria have been changed over the last decades, also resulting in slightly different results[18].

In the second part of this thesis, 38 delirium experts evaluated patients on delirium based on extensive and validated cognitive tests recorded on video as approximation of the gold standard diagnosis of delirium. The participating delirium experts diagnosed approximately 15 delirious patients on a monthly basis and they had numerous years of clinical experience. Therefore, they were assumed to be able to diagnose patients correctly, having a robust delirium diagnosis. Two delirium experts assessed each video and accompanying patient record information of the last 24 hours. No consensus was reached in 21% of the cases based on exactly the same video recording and patient

record information. Moreover, the delirium experts were asked to give a likelihood of a patient being delirious. Comparing these likelihoods of the delirium experts, the highest variability was found in the hypoactive motor subtype.

Together, these results indicate that it is difficult to classify delirium, even for delirium experts with broad experience. Moreover, this shows that what is currently considered the gold standard for delirium diagnosis, is rather a reference standard being the best method currently available. In future delirium research, it is essential that the studies use a reliable method for delirium diagnosis and have this well-defined in the method section. Despite imperfect diagnoses by delirium experts, they should be considered as the best reference standard available at this time.

DELIRIUM AS A DICHOTOMOUS OR A CONTINUOUS SYNDROME

Chapter 6 describes the results of the validation of the EEG-based delirium monitor according to the 'dichotomous reference standard'. The delirium experts classified patients as 'non-delirious', 'possible delirious', or 'delirious'. The 'possible delirious' and 'delirious' classified patients were combined because these were all cases who should be detected in clinical practice, allowing possible interventions. Moreover, a dichotomous comparison calculating the sensitivity, specificity, and accompanying ROC curve enables comparison with previously validated delirium tools, such as the CAM-ICU[13]. It is unclear how giving the experts the opportunity to classify the patients as 'possible delirious' has influenced the results.

Delirium is classified according to the DSM criteria, resulting in either delirium or non-delirium, there is nothing in between. However, it is more plausible that delirium is a spectrum or a continuum instead of a dichotomous diagnosis[19]. In clinical practice, the physician should critically follow the patients who do not fulfill all criteria for possible deterioration towards delirium and non-pharmacological interventions can be started. The results of **chapter 7** showed that there is increased EEG delta activity in patients who developed delirium the day after the EEG recording, compared to those who do not. This indicates a mild encephalopathy, without delirious symptoms or not fulfilling all criteria for delirium diagnosis. In the literature, the term prodromal delirium or subsyndromal delirium is used to classify patients who are not fulfilling all criteria. Subsyndromal delirium has also been associated with impaired outcome and therefore relevant to recognize[19].

Delirium can range in severity, according to the number and severity of the symptoms and the duration of the delirious episode. This information may be relevant for the trajectory of delirium and long-term consequences[20], and therefore for possible interventions. If a patient can be defined as 'subsyndromal delirium' the physician could start non-medical interventions trying to prevent the transition towards the full syndrome. On the other hand, when the severity of the symptoms decreases, this may suggest recovering of the delirious episode. Therefore, delirium should be regarded and treated as a continuous spectrum. The EEG-based delirium monitor gives a continuous result and seems more compatible with delirium as continuous spectrum. However, the relative delta power (1-4 Hz) was only weakly correlated with the severity of delirium (**chapter 6**). Future research should evaluated which EEG features can be used best to monitor the delirious episode as a continuous measure. Not only the frequency related measures should be considered, also more complex measures, such as the functional connectivity and MST measures, since delirium and reduced levels of consciousness without delirium showed similar results on the frequency measures. Using multiple EEG measures instead of one could improve detection, but also a stronger correlation with severity of delirium, which is desirable for clinical practice.

SEDATION-INDUCED DELIRIUM

In the ICU sedatives are frequently administered to comfort patients[21]. In patients who are mechanically ventilated, daily interruption of sedatives enables neurological evaluation and was shown to reduce the number of days of mechanical ventilation, and the duration of ICU and hospital stay[22]. During the interruption of sedation, patients can also be evaluated on delirium. However, it is not clear how long it takes to fully regain consciousness after interruption of the sedatives. In a prospective study, the long-term outcome between rapidly reversible, sedation-related delirium and persistent delirium was evaluated by assessing delirium before and after the cessation of sedatives[23]. This study showed that rapidly reversible, sedation-related delirium (i.e., delirium before cessation of sedatives, but not thereafter) does not have the same poor long-term prognosis as persistent delirium. Delirium was assessed using the CAM-ICU by researchers. The authors suggested that the timing of delirium assessment in relation to sedative interruption should be taken into account, and that delirium screening should preferably be performed in those who are awake[23].

In **chapter 3** we assessed the differences between hypoactive delirium and recovery from anesthesia. This recovery of anesthesia may be assumed similar as the rapidly reversible, sedation-related delirium as defined this previous study. Although we did not perform any delirium assessment in these patients during the EEG recording, these patients were visited the two following days for delirium evaluation and were excluded when found delirious. The full-EEG recording showed a similar power spectrum as was found in the hypoactive delirium patients. Increasing the number of electrodes in the EEG-based delirium monitor, may facilitate the opportunity to distinguish sedation effects from delirium using MST measures. However, it is unclear what the minimum number of EEG channels should be.

A different study assessed patients for delirium directly after surgery at the post-anesthesia care unit (PACU) and during the following days, according to the DSM criteria[24]. The patients who were only delirious at the PACU and not during the following days had no long-term cognitive problems at three months after surgery, which was similar to the non-delirious patients. Moreover, the patients who were diagnosed as delirious on the consecutive days after surgery, did have an increased risk of postoperative cognitive decline[24]. Therefore, this study indicates that there is a significant difference in long-term outcome between sedation-delirium (i.e., directly after surgery) and delirium on the consecutive days after surgery. Hence, it is important to distinguish these two groups in clinical practice in relation to long-term prognosis.

IMPROVED DELIRIUM DETECTION USING AN EEG-BASED DELIRIUM MONITOR

One of the aims of this thesis was to improve delirium detection using an EEG-based delirium monitor. **Chapter 6** describes the results of the validation study of a brief, single-channel EEG recording in older postoperative patients. Based on the dichotomous classification by the delirium experts, either non-delirious or (possible) delirious, and several cut-off points of the relative delta power, an ROC curve was created and the calculated area under the ROC (AUROC) was 0.75 (95% CI 0.69-0.82). A slightly higher AUROC value (0.78 (95% CI 0.72-0.84)) was found for the widened relative delta power (1-6 Hz). These results are significantly better than chance (AUROC=0.50), although deviate largely from the perfect discriminating test (AUROC=1.00). In the proof-of-concept study AUROC values of 0.99 were found using the relative delta power to discriminate between definitely delirious patients from definitely non-delirious patients[25]. Several explanations can be given for the lower AUROC values found in the validation study,

chapter 6. First, the used reference standard is not perfect. Although a third expert was consulted in case of disagreement between the first two delirium experts, this indicates that the diagnosis of delirium is not straightforward. In the proof-of-concept study, only the definite delirious patients and definite non-delirious patients were included[25], excluding possible doubtful cases which will occur in clinical practice. Moreover, the dichotomization of the classification of the experts, may have influenced the results negatively, including the possible delirious cases into the delirium group. However, the possible delirium patients had similar increased relative delta power as the delirium cases, supporting the merge of the possible delirium cases with the delirium cases.

A second explanation for the lower AUROC values compared to the proof-of-concept study is that patients were only visited once during the day, whereas delirium has a fluctuating character. If patients did not show delirium symptoms during the visit, but for example only during the evening, this was not recorded on video. We tried to prevent under-detection of these cases by including the information of the patient record. However, if a patient showed symptoms in the evening but this was not noted by the nurse or physician or not documented in the patient record, this episode would have still be missed in the current study. This would lead to false positive cases if the EEG would be affected.

Finally, the lower AUROC values can be explained by the possible difference in the course of the delirium symptoms compared to the course of the underlying encephalopathy, which is recorded by the EEG. In Figure 8.1, a hypothetical relation between time and the development of the encephalopathy is shown. Delirium diagnosed by experts according to the DSM-5 criteria is feasible when the underlying encephalopathy is sufficient to lead to delirious symptoms. The EEG measures may be more accurate and enable the opportunity to detect the ongoing encephalopathy before the symptoms arise. Moreover, in **chapter 7** the EEG variables of the current day, in non-delirious patients, were studied in relation to the development of delirium the day after the recording. Patients who were diagnosed as delirious on the day after the recording had significant higher relative delta power than patients who remained non-delirious. Relating these results to the diagnostic comparison (i.e., the results of the EEG recording vs. the diagnosis on the same day), indicates that patients who were classified as non-delirious, could have already higher relative delta power, related to delirium the next day. This also leads to false positive results regarding the ROC curves in **chapter 6**.

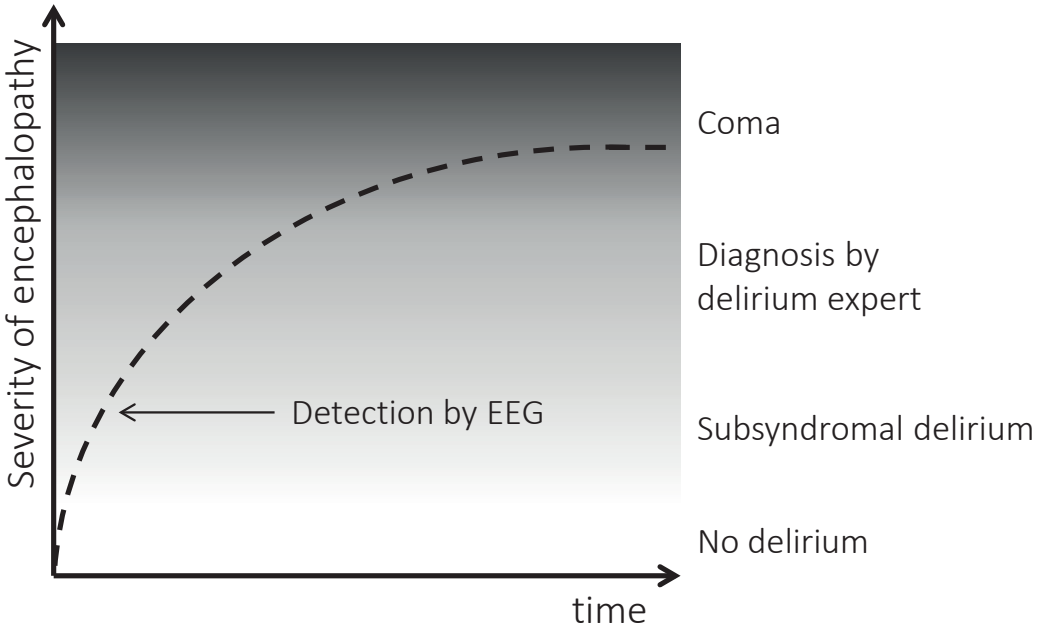


Figure 8.1. Hypothetical opportunity of EEG based delirium detection prior delirium diagnosis by experts.
Severity of the encephalopathy describes the underlying abnormalities that can result in delirium or coma. Delirium detection with EEG may be feasible in an earlier stage compared to the expert classification.

Based on these considerations, the results of **chapter 6** may be valuable and reliable for use in clinical practice. The cases that are false positive in **chapter 6**, but true positive cases in **chapter 7** should be considered correct, since the EEG predicts delirium before the delirium experts considered the diagnosis. This should improve the observed AUROC values of **chapter 6**.

Mrs. R., the case description in **chapter 1** Box 1.1, may have benefited when the EEG-based delirium monitor was used to record the relative delta power several times a day or continuously during the day. The ongoing encephalopathy could have been detected the first morning after surgery, instead of during the evening. However, whether this would have led to an improved outcome depends on the treatment options started by the physician. A previous study showed that increased adherence to the available delirium screening protocol resulted in reduced in-hospital mortality and reduced number of days on the mechanical ventilator[26]. This is remarkable, since delirium detection itself does not cure the problem. However, the development of delirium is by definition caused by underlying medical conditions, and detection of new onset delirium should be followed by the search for these underlying problems. Early treatment of the underlying problem could improve the outcome of the patients.

FUTURE PERSPECTIVES

Delirium is proposed to result from the breakdown, or disconnection, of the functional brain network[27]. Patients with higher functional connectivity are hypothesized to be less vulnerable for stressors such as anesthesia and surgery, and will develop delirium less often. Preoperative evaluation of the functional brain network may therefore be able to predict which patients will suffer from a delirious episode after surgery and possibly long-term postoperative cognitive decline. It would be valuable to detect surgical patients with lower functional connectivity, and improve this to reduce the risk of postoperative delirium. Behavioral or pharmacological interventions may recover or improve the functional connectivity of the brain network. Neurostimulation, with transcranial magnetic stimulation (TMS) or transcranial direct-current stimulation (tDCS), is thought to improve functional connectivity[27]. With tDCS a constant low electrical current is applied by electrodes to brain areas. Using TMS, a magnetic field is generated above the head of a patient, resulting in activation of the neurons in the cortex. Using tDCS or TMS is suggested to improve the functional connectivity, making patients less vulnerable for major stressors[27].

In this thesis the functional brain network during hypoactive delirium and the effects of sedatives were studied on a static level, the information across four epochs of eight seconds was averaged to gain stable results. However, it might also be of interest to study the dynamic components of the functional network, especially in relation to the increased variability of the EEG signal found during delirium[28]. In the different transition phases of general anesthesia, different dynamic subgraphs of the backbone structure were found in a previous study[29]. These dynamic subgraphs were obtained with a moving window over the time series instead of averaging the information over the entire epoch. This suggests that both the static (i.e., the averaged result over the epochs) and dynamic networks (i.e., the fluctuation within the epochs) are related to levels of consciousness.

In this thesis we tried to find the best EEG measures to distinguish delirium from non-delirium as classified by delirium experts. However, the long-term consequences of delirium may be more important than the detection of delirium itself. It would be of interest to relate the EEG measures recorded during the first three days after surgery with long-term outcome of these patients. It is possible that EEG features can make a better distinction between patients who will and will not suffer from cognitive decline 1 year after surgery, compared to the delirium diagnosis by experts during the three postoperative days. For example, a patient who does not fulfill all DSM-5 criteria for

delirium, but does suffer from surgery-related brain damage such as micro-emboli after heart surgery[30], may have long-term cognitive problems. Moreover, it was suggested that in some diseases a long-term follow-up variable could help in case of an imperfect gold standard[16]. It would be of great interest to detect these patients who will suffer from long-term cognitive decline. Firstly, this would allow us to inform the patients about the possible cognitive decline. Secondly, future treatment could be started based on the EEG measures, instead of delirium diagnosis based on symptoms only. Finally, pharmacological and non-pharmacological interventions could be improved based on EEG measures which relate to long-term outcomes. In the ICU, antipsychotics such as haloperidol are frequently used to reduce the symptoms of delirium. However, sufficient evidence of its efficacy and safety is lacking[21], and it does not treat the underlying cause. Using a reliable method to detect delirium or possible long-term outcome is essential in research estimating the efficacy of interventions.

Other effects influencing the EEG pattern should be evaluated more extensively and included in the EEG detection algorithm. In this thesis the effects of recovery from anesthesia in relation to hypoactive delirium were studied. However, the observed difference in network topology is not applicable in the single-channel EEG-based delirium monitor. Thus, other options should be explored. EEG in delirious patients was shown to have increased variability over time and reduced complexity[28]. This increased variability is probably the result of the intermittent rhythmic delta activity[31]. During the transition from consciousness to sedative induced unconsciousness, the EEG pattern becomes more regular, but also less complex[32]. Despite the similar reduced complexity of the EEG signal in delirium and anesthesia, the variability related to the intermittent rhythmic delta activity in delirium may be useful to make a distinction and should be evaluated in comparable populations.

Dementia is an important risk factor for delirium[33] and the incidence of dementia in elderly patients is significant. The delirium monitor should be able to distinguish delirium from dementia, as well as delirium superimposed on dementia from dementia. Using resting state EEG recording only showed minor differences between patient with dementia and patients with delirium superimposed on dementia[31]. Adding a prolonged activation procedure (i.e., the patients were instructed to keep their eyes open), improved the discriminating power using EEG characteristics[31]. The diagnostic value of single-channel EEG based delirium monitor should be further studied in dementia patients with and without delirium.

CONCLUSION

This thesis describes the opportunities to detect delirium based on a simple quantitative EEG recording. Early prediction of delirium prior to a clinical diagnosis seems feasible. As this allows early detection of underlying causes of delirium, the EEG-based delirium monitor is a useful tool for clinical practice. Sophisticated measures of functional network analysis were found to distinguish the effects of sedatives from delirium, however multiple EEG channels are needed. Improved algorithms should be developed to detect delirium distinct from sedatives and other states influencing the EEG recording.

REFERENCES

1 Numan, T., Stam, C. J., Slooter, A. J. C., et al. Being conscious of methodological pitfalls in functional brain network analysis. Anesthesiology 123, 484–485 (2015).

2 Stam, C. J., Nolte, G. & Daffertshofer, A. Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. Hum. Brain Mapp. 28, 1178–1193 (2007).

3 Van Dellen, E., de Waal, H., van der Flier, W. M., et al. Loss of EEG network efficiency is related to cognitive impairment in dementia with Lewy bodies. Mov. Disord. 30, 1785–1793 (2015).

4 Van Dellen, E., van der Kooi, A. W., Numan, T., et al. Decreased functional connectivity and disturbed directionality of information flow in the electroencephalography of intensive care unit patients with delirium after cardiac surgery. Anesthesiology 121, 328–335 (2014).

5 Stam, C. J., de Haan, W., Daffertshofer, A., et al. Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease. Brain 132, 213–224 (2008).

6 Friston, K. J. Schizophrenia and the disconnection hypothesis. Acta Psychiatr Scand 99, 68–79 (1999).

7 Lee, H., Mashour, G. A., Noh, G.-J., et al. Reconfiguration of network hub structure after propofol-induced unconsciousness. Anesthesiology 119, 1347–1359 (2013).

8 Lee, U., Ku, S., Noh, G., et al. Disruption of frontal–parietal communication by ketamine, propofol, and sevoflurane. Anesthesiology 118, 1264–1275 (2013).

9 Dauwan, M., van Dellen, E., van Boxtel, L., et al. EEG-directed connectivity from posterior brain regions is decreased in dementia with Lewy bodies: a comparison with Alzheimer's disease and controls. Neurobiol. Aging 41, 122–129 (2016).

10 Choi, S.-H., Lee, H., Chung, T.-S., et al. Neural network functional connectivity during and after an episode of delirium. Am. J. Psychiatry 169, 498–507 (2012).

11 Inouye, S. K., Foreman, M. D., Mion, L. C., et al. Nurses' recognition of delirium and its symptoms: comparison of nurse and researcher ratings. Arch. Intern. Med. 161, 2467–2473 (2001).

12 Inouye, S. K., Van Dyck, C. H., Alessi, C. A., et al. Clarifying confusion: the confusion assessment method: a new method for detection of delirium. Ann. Intern. Med. 113, 941–948 (1990).

13 Ely, E. W., Margolin, R., Francis, J., et al. Evaluation of delirium in critically ill patients: validation of the confusion assessment method for the intensive care unit (CAM-ICU). Crit. Care Med. 29, 1370–1379 (2001).

14 De, J. & Wand, A. P. F. Delirium Screening: A systematic review of delirium screening tools in hospitalized patients. Gerontologist 55, 1079–1099 (2015).

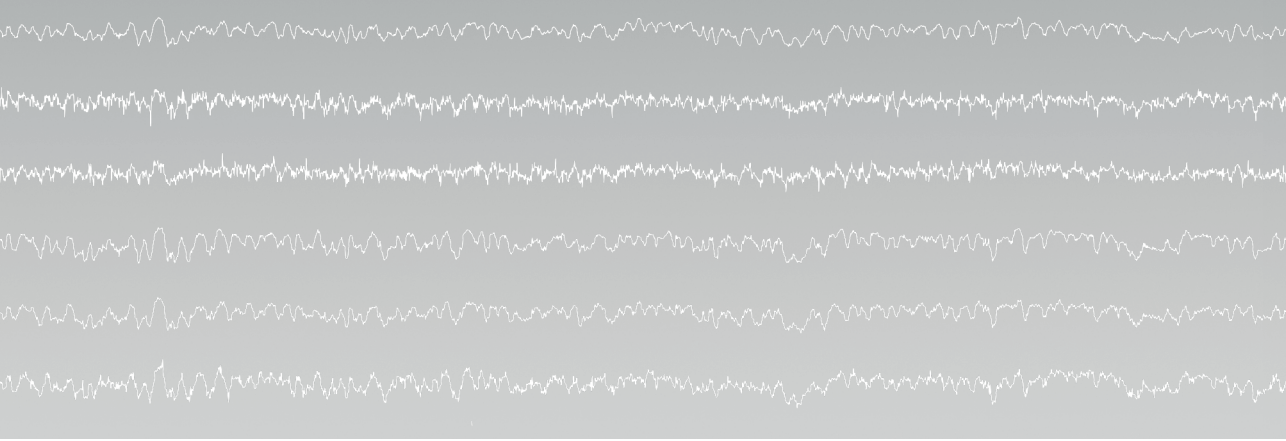
15 Neto, A. S., Nassar, A. P., Cardoso, S. O., et al. Delirium screening in critically ill patients. Crit. Care Med. 40, 1946–1951 (2012).

16 Gold, R., Reichman, M., Greenberg, E., et al. Developing a new reference standard. Acad. Radiol. 17, 1079–1082 (2010).

17 Neufeld, K. J., Nelliot, A., Inouye, S. K., et al. Delirium diagnosis methodology used in research: a survey-based study. Am. J. Geriatr. Psychiatry 22, 1513–1521 (2014).

18 Meagher, D. J., Morandi, A., Inouye, S. K., et al. Concordance between DSM-IV and DSM-5 criteria for delirium diagnosis in a pooled database of 768 prospectively evaluated patients using the delirium rating scale-revised-98. BMC Med. 12, 164 (2014).

19 Ouimet, S., Riker, R., Bergeon, N., et al. Subsyndromal delirium in the ICU: evidence for a disease spectrum. Intensive Care Med. 33, 1007–1013 (2007).
20 Jackson, T. A., Wilson, D., Richardson, S., et al. Predicting outcome in older hospital patients with delirium: a systematic literature review. Int. J. Geriatr. Psychiatry 31, 396–403 (2016).
21 Barr, J., Fraser, G. L., Puntillo, K., et al. Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit. Crit. Care Med. 41, 263–306 (2013).
22 Girard, T. D., Kress, J. P., Fuchs, B. D., et al. Efficacy and safety of a paired sedation and ventilator weaning protocol for mechanically ventilated patients in intensive care (awakening and breathing controlled trial): a randomised controlled trial. Lancet 371, 126–134 (2008).
23 Patel, S. B., Poston, J. T., Pohlman, A., et al. Rapidly reversible, sedation-related delirium versus persistent delirium in the intensive care unit. Am. J. Respir. Crit. Care Med. 189, 658–665 (2014).
24 Franck, M., Nerlich, K., Neuner, B., et al. No convincing association between post-operative delirium and post-operative cognitive dysfunction: a secondary analysis. Acta Anaesthesiol. Scand. 60, 1404–1414 (2016).
25 Van der Kooi, A. W., Zaal, I. J., Klijn, F. A., et al. Delirium detection using EEG. Chest 147, 94–101 (2015).
26 Luetz, A., Weiss, B., Boettcher, S., et al. Routine delirium monitoring is independently associated with a reduction of hospital mortality in critically ill surgical patients: a prospective, observational cohort study. J. Crit. Care 35, 168–173 (2016).
27 Shafi, M. M., Santarnecchi, E., Fong, T. G., et al. Advancing the neurophysiological understanding of delirium. J. Am. Geriatr. Soc. 65, 1114-1118 (2017).
28 Van der Kooi, A. W., Slooter, A. J. C., van het Klooster, M. A., et al. EEG in delirium: increased spectral variability and decreased complexity. Clin. Neurophysiol. 125, 2137–2139 (2014).
29 Shin, J., Mashour, G. A, Ku, S., et al. Subgraph 'backbone' analysis of dynamic brain networks during consciousness and anesthesia. PLoS One 8, e70899 (2013).
30 Goto, T. & Maekawa, K. Cerebral dysfunction after coronary artery bypass surgery. J. Anesth. 28, 242–248 (2014).
31 Thomas, C., Hestermann, U., Walther, S., et al. Prolonged activation EEG differentiates dementia with and without delirium in frail elderly patients. J. Neurol. Neurosurg. Psychiatry 79, 119–125 (2008).
32 Bai, Y., Liang, Z., Li, X., et al. Permutation Lempel – Ziv complexity measure of electroencephalogram in GABAergic anaesthetics. Physiol. Meas. 36, 2483 (2015).
33 Inouye, S. K., Westendorp, R. G. & Saczynski, J. S. Delirium in elderly people. Lancet 383, 911–922 (2014).



A

NEDERLANSE SAMENVATTING

(DUTCH SUMMARY)

SAMENVATTING

Het brein kan worden gezien als een complex netwerk, waarbij verschillende hersengebieden met elkaar communiceren door middel van elektrische signalen tussen de vele hersencellen (neuronen). De globale activiteit van de hersenen kan bijvoorbeeld worden gemeten met een elektro-encefalogram (EEG). Met behulp van netwerkanalyse, is het mogelijk om te onderzoeken hoe het hersennetwerk functioneert in gezonde personen en hoe dit anders is in patiënten met een verstoring van het brein, zoals bij delirium.

Delirium, ook wel delier genoemd, is een plotselinge verstoring van cognitie en aandacht en komt vaak voor bij oudere en ernstig zieke patiënten. Patiënten met een delier zijn in de war, gedesoriënteerd in plaats en tijd en kunnen de aandacht niet goed vasthouden. Er worden drie vormen van een delier beschreven, namelijk: hyperactief, hypoactief en de gemengde variant. Een patiënt met een hyperactief delier is onrustig en geagiteerd. Een patiënt met een hypoactief delier, ook wel stil delier genoemd, is juist slaperig en inactief. Een patiënt die afwisselend hyperactieve en hypoactieve periodes laat zien, heeft de gemengde variant. De stille variant komt veel voor op de intensive care (IC) en laat zich moeilijk onderscheiden van de effecten van slaapmiddelen. Veel patiënten op de IC krijgen slaapmiddelen om hun herstel en comfort te bevorderen. Juist het onderscheid tussen patiënten met een delier en patiënten die slaperig zijn vanwege het slaapmiddel is belangrijk, omdat de effecten van het slaapmiddel vanzelf uitwerken, terwijl de patiënt met een delier behandeld moet worden. Het is noodzakelijk om het delier te behandelen, omdat het een negatieve invloed heeft op het herstel van de patiënt.

In het eerste deel van dit proefschrift zijn met behulp van netwerkanalyses verschillende oorzaken van een verlaagd bewustzijn met elkaar vergeleken. In het tweede deel van dit proefschrift is gekeken naar de herkenning van delier in de klinische praktijk en is onderzocht of we dit kunnen verbeteren met behulp van EEG.

DEEL I: RESTING STATE EEG TIJDENS STIL DELIER EN MILDE SEDATIE

Het doel van het eerste deel van dit proefschrift was het verbeteren van de kennis van verlaagd bewustzijn tijdens een delier met behulp van netwerkanalyse van het brein. Er werd eerst onderzocht wat belangrijke keuzes zijn voor het uitvoeren van netwerkanalyse. In de afgelopen 20 jaar is het aantal onderzoeken naar het brein als functioneel netwerk enorm toegenomen. Het functionele hersennetwerk kan met

verschillende modaliteiten worden gemeten, onder andere met functionele MRI (fMRI), magneto-encefalografie (MEG) en EEG. Een voordeel van fMRI en MEG is dat ze met grote nauwkeurigheid de locatie van de hersenactiviteit kunnen vaststellen. Het grote voordeel aan EEG is dat het gebruikt kan worden aan het bed van de patiënt, wat met fMRI en MEG niet mogelijk is.

Er zijn veel verschillende maten ontwikkeld om de functionele connectiviteit, de interactie tussen verschillende hersengebieden, te berekenen. Alle maten hebben hun eigen voor- en nadelen. Het grote probleem is dat de resultaten van de verschillende maten en verschillende modaliteiten niet goed met elkaar vergeleken kunnen worden. Voor het vergelijken van onderzoeken naar functionele hersennetwerken is het van groot belang om te weten wat de invloed is van de verschillen in methodologische keuzes die gemaakt zijn. Daarom wordt in **hoofdstuk 2** beschreven welke keuzes er vooraf en tijdens de analyses gemaakt moeten worden en hoe dat invloed kan hebben op het resultaat.

In **hoofdstuk 3** zijn EEG netwerkanalyses gedaan op twee groepen patiënten met een verlaagd bewustzijn, namelijk patiënten met een stil delier en patiënten die wakker worden na de operatie. In beide patiëntengroepen vonden we afgenomen functionele connectiviteit en verlies van de gerichte connectiviteit van posterieur naar anterieur (van achter naar voren). Deze afnames zouden kunnen samenhangen met het verlaagde bewustzijn. Het verschil in het hersennetwerk van stil delier ten opzichte van verlaagd bewustzijn door slaapmiddelen, werd gevonden in de organisatie van het functionele netwerk. Het netwerk gemeten bij een stil delier werd gekarakteriseerd door een minder sterk geïntegreerd netwerk. Hierdoor kost het meer moeite om informatie van de ene kant van het netwerk naar de andere kant van het netwerk te versturen. Het netwerk is dus minder efficiënt geworden. De patiënten met een verlaagd bewustzijn door slaapmiddelen lieten juist een sterker geïntegreerd netwerk zien. Het verschil zou verklaard kunnen worden door de fase van het verlaagde bewustzijn van de patiënten. De patiënten met verlaagd bewustzijn door slaapmiddelen zaten in de fase waarin ze wakker werden, terwijl patiënten met een stil delier mogelijk nog niet aan het herstellen waren.

In **hoofdstuk 4** hebben we het verschil tussen twee soorten slaapmiddelen onderzocht, namelijk midazolam en propofol. Patiënten die een roesje kregen (een lichte vorm van sedatie) tijdens een endoscopische ingreep, werden geïncludeerd. Met behulp van EEG-metingen voor en tijdens het roesje konden de twee situaties met elkaar worden

vergeleken. Ook in deze studie zagen we een afname van de functionele connectiviteit en verlies van de gerichte connectiviteit bij zowel midazolam als propofol, ten opzichte van de wakkere meting voor het roesje. Echter, hier werden geen verschillen gevonden in de organisatie van het netwerk tussen de meting tijdens het roesje en de wakkere meting voor het roesje. Bij het vergelijken van de effecten van midazolam en propofol werd een grotere afname van de functionele connectiviteit in de alfa frequentieband en een grotere verstoring van de gerichte connectiviteit in de delta frequentieband gevonden bij propofol ten opzichte van midazolam. Dit verschil suggereert dat midazolam en propofol mogelijk een ander werkingsmechanisme hebben op het functionele hersennetwerk tijdens een roesje.

DEEL II: VALIDATIE VAN EEN DELIER MONITOR OP BASIS VAN EEG

Het tweede doel van dit proefschrift was het verbeteren van de herkenning van delier in de klinische praktijk, omdat delierherkenning op dit moment ondermaats is. Als eerste werd de huidige klinische situatie geëvalueerd. In **hoofdstuk 5** werd onderzocht in hoeverre de diagnose delier overeenkwam tussen verschillende delier experts. Een onderzoeker bezocht oudere patiënten (60 jaar en ouder) de eerste drie dagen na een grote operatie en nam een cognitieve screeningtest af. Dit werd vastgelegd op video. De experts kregen deze video's te zien en kregen informatie van de afgelopen 24 uur uit de klinische status van de patiënten. De experts beoordeelden de patiënten als het hebben van geen delier, mogelijk delier of delier. Bij het vergelijken van de beoordelingen van de experts werd in 21% geen overeenstemming gevonden. Dit benadrukt hoe lastig de diagnose delier kan zijn. Indien er geen overeenstemming tussen twee experts was, werd een derde expert gevraagd voor een uiteindelijke classificatie. Die uiteindelijke classificatie werd gebruikt om te testen hoe goed het delier in de klinische praktijk wordt herkend. Er zijn verschillende screeningmethodes ontwikkeld voor verpleegkundigen en artsen. In de deelnemende ziekenhuizen werd de *confusion assessment method for intensive care unit* (CAM-ICU) gebruikt voor IC-patiënten en de delirium observatie screening (DOS) schaal voor patiënten op andere verpleegafdelingen. De verpleegkundigen zouden deze screeningmethodes 2 of 3 keer per dag uit moeten voeren en documenteren in de status van de patiënt. Uit het huidige onderzoek bleek dat slechts in ongeveer de helft van de gevallen een CAM-ICU of DOS-score was ingevuld. Van de helft die wel was gedocumenteerd, bleek de sensitiviteit 43% te zijn. Hieruit blijkt dat in de klinische praktijk delier onvoldoende wordt herkend. Er is dus een betere en objectieve methode nodig om delier in de dagelijkse praktijk goed te kunnen herkennen. Daarom werd een delier monitor op basis van EEG ontwikkeld.

Hoofdstuk 6 beschrijft de validatiestudie van dit op EEG-gebaseerde prototype. Bij 159 patiënten van 60 jaar en ouder werd de eerste drie dagen na hun operatie een enkelkanaals EEG gemaakt, gevolgd door een cognitieve screeningtest die op video werd vastgelegd. De video's werden beoordeeld door twee of drie delier experts. Een minuut storingsvrij EEG-segment werd geselecteerd waarop meerdere EEG-maten werden berekend. Op basis van verschillende afkapwaardes van deze EEG maten werd gekeken hoe goed het EEG onderscheid kan maken tussen delier en geen delier op basis van de conclusies van de experts. Op basis van de relatieve delta power, de langzame frequenties tussen 1 en 4 Hertz (Hz), verkregen we een *receiver operating characteristic* (ROC) curve met een oppervlakte van 0.75 (95% betrouwbaarheidsinterval 0.69-0.82). De brede relatieve delta power (1-6 Hz) had een iets hogere oppervlakte onder de ROC curve, namelijk 0.78 (95% betrouwbaarheidsinterval 0.72-0.84). Het onderscheidend vermogen, vooral van de brede relatieve delta power, suggereert dat we een delier kunnen aantonen bij postoperatieve patiënten op basis van een simpele enkelkanaals EEG-meting.

In **hoofdstuk 7** werd tot slot onderzocht of het ook mogelijk is om het optreden van een delier een dag eerder al te voorspellen met behulp van het enkelkanaals EEG. Zowel de relatieve delta power als de brede relatieve delta power waren significant hoger in patiënten die de volgende dag een delier ontwikkelden. Daarnaast werd een voorspelmodel gemaakt op basis van klinische parameters en de relatieve delta power. Dit voorspelmodel bleek in staat te zijn om delier te voorspellen, voordat de delier expert de diagnose kan stellen. Echter moeten deze resultaten bevestigd worden in een grotere onderzoekspopulatie, met een hoger aantal delirante patiënten.

Het laatste deel van dit proefschrift bevat een Engelse samenvatting en een algemene discussie over de gebruikte methodes, resultaten en implicaties van dit proefschrift. Concluderend beschrijft dit proefschrift de mogelijkheid om met een EEG meting delier te kunnen herkennen en te onderscheiden van de effecten van slaapmedicatie. Op basis van een volledig EEG met meerdere kanalen bleek het mogelijk om de functionele hersennetwerken van een stil delier te onderscheiden van de effecten van slaapmiddelen. Daarnaast laat dit proefschrift zien dat het mogelijk is om delier aan tonen met behulp van een simpele, kwantitatieve, enkelkanaals EEG-meting. Voordat delier betrouwbaar op de intensive care gemeten kan worden, moet het huidige algoritme verbeterd worden, zodat er beter onderscheid gemaakt kan worden tussen delier en andere vormen van verlaagd bewustzijn op basis van het enkelkanaals EEG.