

**PHYSIOLOGICAL INFORMATION IN NEAR-INFRARED  
SPECTROSCOPY: FROM ADULTS TO INFANTS**

Naser Hakimi

# **PHYSIOLOGICAL INFORMATION IN NEAR-INFRARED SPECTROSCOPY: FROM ADULTS TO INFANTS**

## **FYSIOLOGISCHE INFORMATIE IN NEAR-INFRARED SPECTROSCOPY: VAN VOLWASSENEN TOT ZUIGELINGEN**

(met een samenvatting in het Nederlands)

### **Proefschrift**

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

woensdag 28 augustus 2024 des middags 4.15 uur

door

**Naser Hakimi**

geboren op 17 juni 1993  
te Nahavand, Iran

**Promotor:**

Dr. J. Dudink

**Co-promotoren:**

Dr. J.M. Horschig

Dr. T. Alderliesten

**Beoordelingscommissie:**

Prof. dr. M.N. Bekker

Prof. dr. D. van Dijk

Dr. F. Groenendaal

Prof. dr. G. Naulaers

Prof. dr. A.A. Salah (voorzitter)

# TABLE OF CONTENTS

Chapter 1	General introduction	p. 1
-----------	----------------------	------

## **Part I: physiological information in adult NIRS**

Chapter 2	Signal Quality Index: a novel algorithm for quantitative assessment of functional near infrared spectroscopy signal quality	p. 25
Chapter 3	fNIRS signal quality estimation by means of a machine learning algorithm trained on morphological and temporal features	p. 53
Chapter 4	Estimation of respiratory rate from functional near-infrared spectroscopy (fNIRS): a new perspective on respiratory interference	p. 69
Chapter 5	Estimation of respiratory rate during biking with a single sensor functional near-infrared spectroscopy (fNIRS) system	p. 89

## **Part II: physiological information in neonatal NIRS**

Chapter 6	Neuroprotection of the perinatal brain by early information of cerebral oxygenation and perfusion patterns	p. 109
Chapter 7	Heart rate extraction from neonatal near-infrared spectroscopy signals	p. 131
Chapter 8	Respiratory rate extraction from neonatal near-Infrared spectroscopy signals	p. 157
Chapter 9	Near-infrared spectroscopy for neonatal sleep classification	p. 179
Chapter 10	Discussion and conclusion	p. 201
Chapter 11	Nederlandse samenvatting (summary in Dutch)	p. 215
Chapter 12	List of abbreviations List of coauthors List of publications Curriculum vitae Acknowledgements	p. 221





# CHAPTER 1

General Introduction



# 1. Introduction to Near-Infrared Spectroscopy (NIRS)

## 1.1. Definition of NIRS

Near-infrared spectroscopy (NIRS) is an optical monitoring technique that measures cerebral oxygenation noninvasively. This technology uses the distinctive characteristics of near-infrared light, specifically within the 700 to 1000 nanometer wavelength range, to noninvasively assess the oxygenation levels of living tissues, with a particular focus on brain tissue. The underlying principle of this approach is that biological tissues exhibit a degree of transparency to light in this wavelength spectrum. This transparency allows the near-infrared light to penetrate the skull, enabling the measurement of the relative amounts of oxygenated hemoglobin (O<sub>2</sub>Hb) and deoxygenated hemoglobin (HHb) in the tissue. This capability enables NIRS to determine regional oxygen saturation, thereby providing insights into the oxygenation status of brain tissue [1,2].

NIRS has emerged as a valuable instrument in both perioperative and intensive care environments. In these settings, the continuous monitoring of cerebral oxygenation can be used to steer treatment [3]. NIRS offers a direct bedside insight into the balance of oxygen supply and demand within the brain and other tissues, although it does not differentiate between venous, capillary, and arterial blood [4]. This is particularly beneficial in the neonatal intensive care setting, where the noninvasive application of NIRS probes can yield immediate readings of end-organ oxygenation status, providing an indirect assessment of organ blood flow [5,6]. The simplicity and ease of application of this technology make it an appealing option for continuous bedside monitoring [7].

## 1.2. Brief history, basic principles, and development of the technology

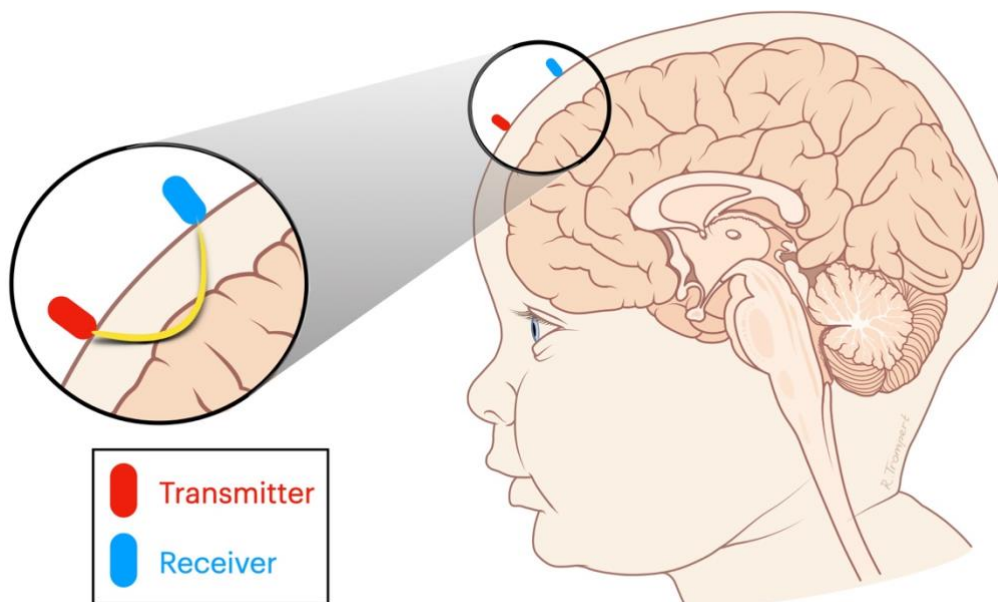
Transillumination, the passage of light through the body, has been explored since the early 1800s. Richard Bright, in 1831, first described transillumination of the head as a diagnostic technique for identifying hydrocephalus [8]. However, the history of NIRS technology dates back to the pioneering work of Frans Jöbsis in the 1970s [9]. Jöbsis' groundbreaking discovery that near-infrared light could penetrate biological tissues and that the differential absorption of this light by HbO<sub>2</sub> and Hb could be measured non-invasively, paved the way for NIRS' application in monitoring cerebral oxygenation. The pediatric application of NIRS for monitoring cerebral oxygenation in sick preterm infants began in 1986 [10], marking a significant expansion of NIRS possibilities. A crucial milestone came in 1993 with the introduction of INVOS 3100® as the first FDA-approved commercial cerebral oximetry device [11]. Since then, NIRS has undergone significant evolution, from early devices that could only measure two chromophores using two wavelengths to modern systems that utilize multiple wavelengths [12–15].

NIRS is grounded in the Beer-Lambert law (BLL), which establishes a connection between light absorbance and tissue properties [16]. When near-infrared light is transmitted into brain tissue, certain chromophores, primarily oxygenated and deoxygenated hemoglobins, absorb a portion of this light. The BLL expresses light absorbances  $\epsilon \cdot c \cdot l$ , where  $\epsilon$  is the molar

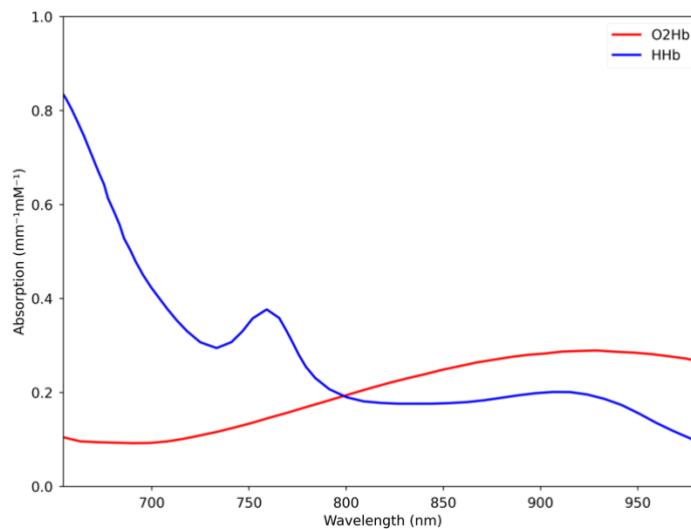


extinction coefficient,  $l$  is the optical path length, and  $c$  is the concentration of the chromophores.

However, in brain tissue, near-infrared light not only gets absorbed but also scatters. It follows a “banana-shaped” trajectory before reaching the receiver (see Figure 1). To account for this, the modified version of the BLL, known as modified BLL (MBLL) [17], includes additional terms:  $G$  compensating for the scatter; differential path length factor (DPF) accounting for the increased distance the light travels. Thus, the MBLL formula becomes  $OD = \epsilon \cdot c \cdot DPF \cdot l + G$  [17]. Here,  $OD = -\log_{10} \frac{I}{I_0}$  represents the optical density (OD) of the tissue, providing a measure of light attenuation. In addition,  $I$  and  $I_0$  denote the intensity of light received by the receiver and transmitted through the tissue, respectively. Assuming that changes in scattering are minor compared to changes in absorption, the  $G$  is treated as time-invariant and is then neglected when assessing changes in OD. Finally, by analyzing changes in OD at two different wavelengths, with one above and one below the isosbestic point of 800 nm—where oxygenated hemoglobin and deoxygenated hemoglobin have equal absorbance (see Figure 2)—the concentration changes in O<sub>2</sub>Hb and HHb are computed.



**Figure 1:** Schematic depicting the operational principle of NIRS, where light penetrates the brain tissue through a transmitter (or optode) and is detected by a receiver (or optode) along a banana shaped pathway. The image was adapted with permission of J. Dudink.



**Figure 2:** Absorption by hemoglobin chromophores with respect to the near-infrared light wavelength (nm). HHb: deoxygenated hemoglobin. O2Hb: oxygenated hemoglobin.

Over the decades, advancements in NIRS technology have been driven by the need to account for light scatter more accurately. The integration of complex algorithms to correct this phenomenon has enabled NIRS to provide precise measurements of chromophore concentration in the brain [18,19]. This capability has made NIRS an invaluable tool in both medical and research contexts, from monitoring neonates in intensive care units to evaluating cognitive functions during tasks [20].

Today, a range of NIRS devices is approved for clinical use, each with their specific configurations of emitters, wavelengths, and algorithms, providing clinicians with a spectrum of oxygenation values. This variety allows NIRS to be adapted to different clinical needs and patient populations. Research has demonstrated that different oximeters produce varying oxygenation values [21], but mathematical equations can translate data between devices by using a validation with a liquid phantom [22], ensuring consistent application of treatment guidelines. Despite the differences among devices, the fundamental goal of NIRS remains unchanged: to provide real-time evaluations of tissue oxygenation [23].

## 2. NIRS in Neonatology

### 2.1. Importance and an overview of neonatal physiological monitoring

During the neonatal period, especially for infants who are extremely vulnerable—such as those born preterm (before 37 weeks of gestation) or term but critically ill—managing the transition from fetal to postnatal life poses significant challenges. These infants face a heightened risk of adverse developmental outcomes and mortality due to potential complications including hemodynamic instability, hypothermia, apnea, bradycardia, and poor perfusion [24]. In this context, the deployment of non-invasive sensor-based technologies in the neonatal intensive care unit (NICU) becomes crucial. These technologies continuously monitor the vital signs of

these fragile infants and provide caregivers with immediate alerts for even the slightest indications of distress or emerging complications [25][26][27]. Additionally, they are essential for documenting, assessing, and managing the impact of medical interventions and therapies on these high-risk infants [28,29]. Furthermore, the evolution of sensor-based technologies tailored for NICU monitoring equipment has partly resulted from the diverse challenges encountered by neonatal intensive care staff, e.g., fragile skin, too many wires, and continuous monitoring [30–32], in dealing with conditions ranging from apnea and hypoglycemia to sepsis and seizures [33–35].

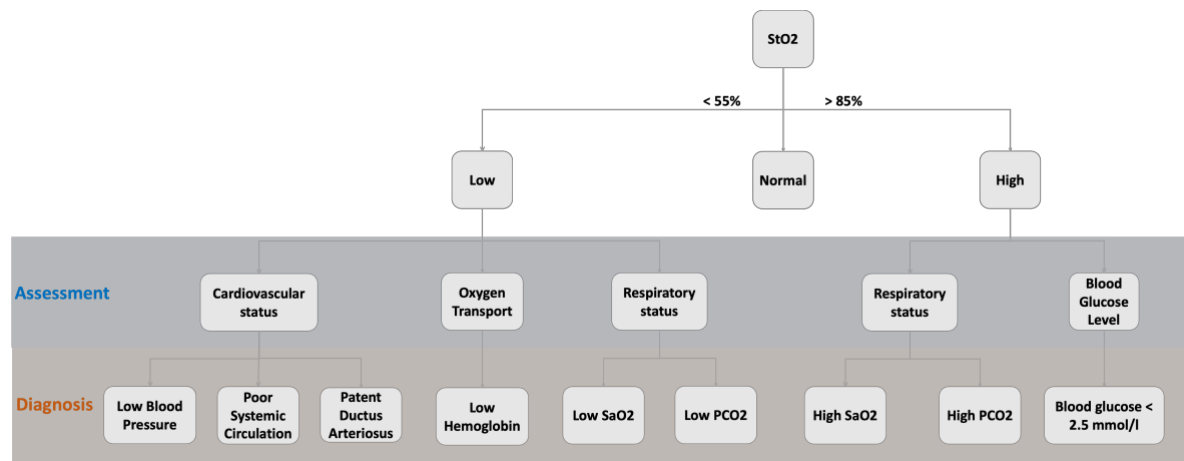
To monitor the physiological status of neonates, various non-invasive sensors are used to measure different vital parameters, such as temperature, respiration, heart rate (HR), blood pressure, and pulse oximetry or peripheral oxygen saturation (SpO<sub>2</sub>) [36,37]. These parameters can indicate the efficiency of their respiratory and circulatory systems as well as signs of stress and/or pain [38]. HR, derived from the electrocardiogram (ECG), can be monitored using electrodes on the skin surface, while pulse oximetry can be monitored using optical sensors in photoplethysmography (PPG), which also determine pulse rate [39,40]. Respiration can be monitored using methods such as electrical strain gauges or thoracic impedance plethysmography, which are more reliable in minimizing motion artifacts [39].

Physiological monitoring serves as a responsive guard in neonatal care. Its advancements significantly influence clinical management, leading to improved quality of care. This, in turn, enhances survival rates and promotes better long-term health and neurodevelopmental outcomes for the youngest and most vulnerable patients.

## 2.2. Applications of NIRS in neonates

Monitoring the balance between oxygen delivery and consumption is also important in infants admitted to the NICU because it provides key insights into the adequacy of tissue perfusion and oxygenation [41]. This balance is crucial not just on a global scale but also at the regional level, where early changes in blood flow, oxygen consumption, and extraction rates can signal variations in the health status of these vulnerable patients. To detect these early changes, additional monitoring systems such as laser Doppler flowmetry and NIRS are indispensable [42–44]. They enable the assessment of peripheral perfusion and microcirculation, which are often evaluated through clinical parameters and blood gas analysis [24,45].

NIRS stands out for its ability to continuously estimate regional blood flow and tissue oxygen saturation (StO<sub>2</sub>, TSI, rSO<sub>2</sub>, or TOI) in various organs, such as the brain, kidney, liver, and intestine, as well as the mixed venous StO<sub>2</sub> in the superior vena cava [28,41]. This non-invasive method is essential for assessing the balance between oxygen delivery and consumption in neonatal tissues [23,46,47], providing vital real-time information for comprehensive care and direct bedside interventions in NICUs, particularly for extremely and very preterm newborns (<32 gestational weeks) [43,48]. The SafeBoosC Phase II trial illustrates the potential of close monitoring of rStO<sub>2</sub> and adherence to specific treatment guidelines to reduce the risks of hypo- and hyperoxia in neonates [47]. Figure 3 shows the normal and abnormal NIRS StO<sub>2</sub> ranges of brain oxygenation index in neonates, offering a framework for evaluating and diagnosing abnormal values as described in the SafeBoosC Phase II study.



**Figure 3:** Normal and abnormal NIRS StO<sub>2</sub> ranges of brain oxygenation index in neonates: evaluation and diagnosis of abnormal values based on the SafeBoosC Phase II study [47].

One of the specific applications of NIRS in neonatal studies is the assessment of cerebral autoregulation in preterm and term infants [49–52]. Cerebral autoregulation involves the brain’s ability to maintain a constant blood flow despite changes in perfusion pressure. NIRS allows clinicians to evaluate how well the neonate’s brain is regulating its blood flow and oxygenation, which is particularly important in the management of newborns with hypoxic-ischemic encephalopathy (HIE), a condition that can lead to significant neurological damage [53]. NIRS can also be used to monitor the cerebral hemodynamic response to various stimuli, such as changes in arterial carbon dioxide, blood pressure, and oxygenation, as well as pharmacological agents and blood transfusions [54].

In the field of pediatric congenital heart disease, NIRS is increasingly used to monitor tissue oxygenation during and after cardiac surgeries [55]. This application of NIRS helps in safeguarding the brain and other vital organs during surgeries that carry a risk of impaired oxygen delivery.

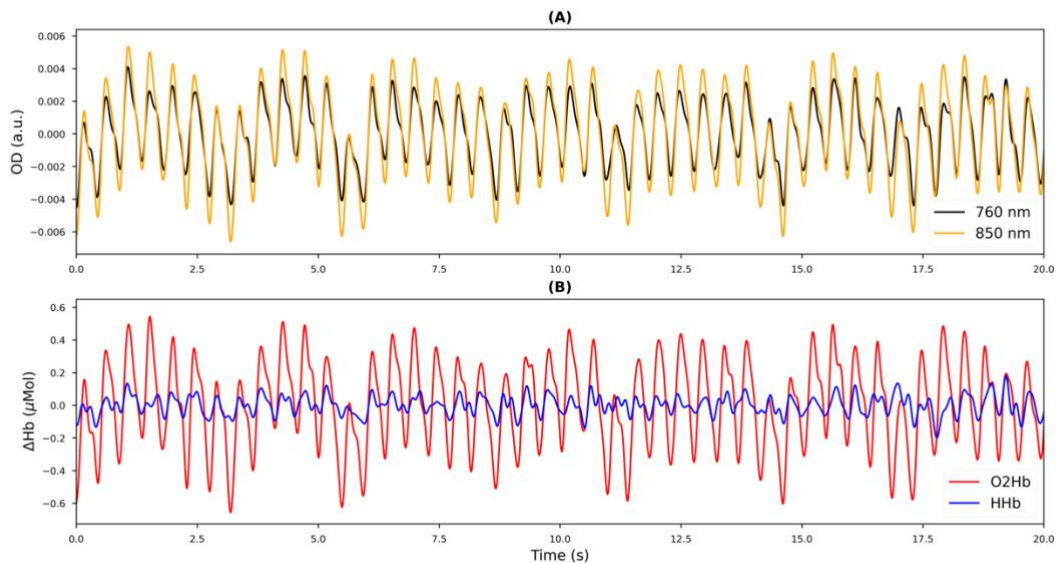
NIRS is also valuable during the critical transition from fetal to neonatal life, particularly in the immediate postnatal period [56]. Reference ranges for healthy term and late preterm infants have been established, aiding in the management of oxygen supply during this vital transition. Studies have shown that by using NIRS to guide medical support, alongside standard pulse oximetry, it is possible to reduce the burden of cerebral hypoxia, thus potentially improving neonatal outcomes [47,57,58].

## 3. Understanding NIRS Data

### 3.1. Types of physiological information obtainable through NIRS

NIRS captures a spectrum of physiological information associated with the hemodynamics and oxygenation of both the cerebral and extracerebral regions [59,60]. In clinical contexts, NIRS is primarily recognized for its role in measuring the cerebral oxygenation index and neuronal activity through monitoring changes in O<sub>2</sub>Hb and HHb concentrations in the brain’s cortex. Upon closer scrutiny, alterations in NIRS signals can be categorized into two compartments (regions): cerebral and extracerebral [61]. Cerebral changes stem from factors such as neuronal activity, systemic variations (e.g., changes in the cardiorespiratory system, blood

pressure, and autonomic nervous system activity), and vascular activity linked to vasomotion. In contrast, the extracerebral compartment undergoes only changes due to systemic, vascular, and muscular activities. Consequently, NIRS encompasses systemic physiological changes in both compartments, including changes in heartbeats, breathing, and blood pressure. Figure 4 shows a sample of raw NIRS signals, i.e., ODs and concentration signals (O<sub>2</sub>Hb and HHb), recorded from a neonate in NICU with a high sampling rate NIRS system (Artinis Medical Systems B.V., Elst, the Netherlands) at 100 Hz. Notably, pulsatile fluctuations attributed to systematic artifacts of heartbeats are distinctly evident across all raw signals.



**Figure 4:** A sample of raw NIRS signals, depicting (A) optical densities (ODs) at wavelengths of 760 nm and 850 nm, along with (B) concentration signals (O<sub>2</sub>Hb and HHb), obtained from a neonate in the NICU, with a high sampling rate NIRS system (Artinis Medical Systems B.V., Elst, the Netherlands) sampled at 100 Hz.

### 3.2. Common artifacts in NIRS data and their impact on data quality

The non-neuronal or systemic physiological information from cerebral and extracerebral compartments in NIRS signals is considered as artifacts in studies where NIRS is used to measure oxygenation or neural activity [62,63]. These artifacts can distort NIRS signals and conceal the true hemodynamic response to a stimulus or task. Therefore, applying suitable methods to detect and remove these artifacts is essential for measuring oxygenation or neural activity with NIRS. Various techniques have been proposed in the literature to address different types of artifacts in NIRS data, such as incorporating short separation channels in the general linear model paradigm, utilizing wavelet filtering, and applying principal component analysis [64–68].

NIRS measurements often face the challenge of motion artifacts, particularly in studies involving infants. These artifacts arise due to relative movements between optodes and the scalp, commonly caused by head or facial motion [69]. Such motion disrupts optical contact, leading to noticeable artifacts in NIRS signals that can obscure subtle hemodynamic changes. Overcoming the difficulty of extracting the genuine physiological NIRS signal from motion-contaminated data is a complex task [70]. Various post-measurement techniques exist to address motion artifacts, with some requiring additional movement measurements (e.g.,



adaptive filtering) and others not necessitating them (e.g., principal component analysis, Kalman filtering, wavelet-based filtering, and spline interpolation) [66,71–73]. Nevertheless, these correction methods are not as effective as preventing motion artifacts from occurring in the first place. Enhancing and stabilizing the coupling between optodes and the scalp can significantly enhance the quality of NIRS data.

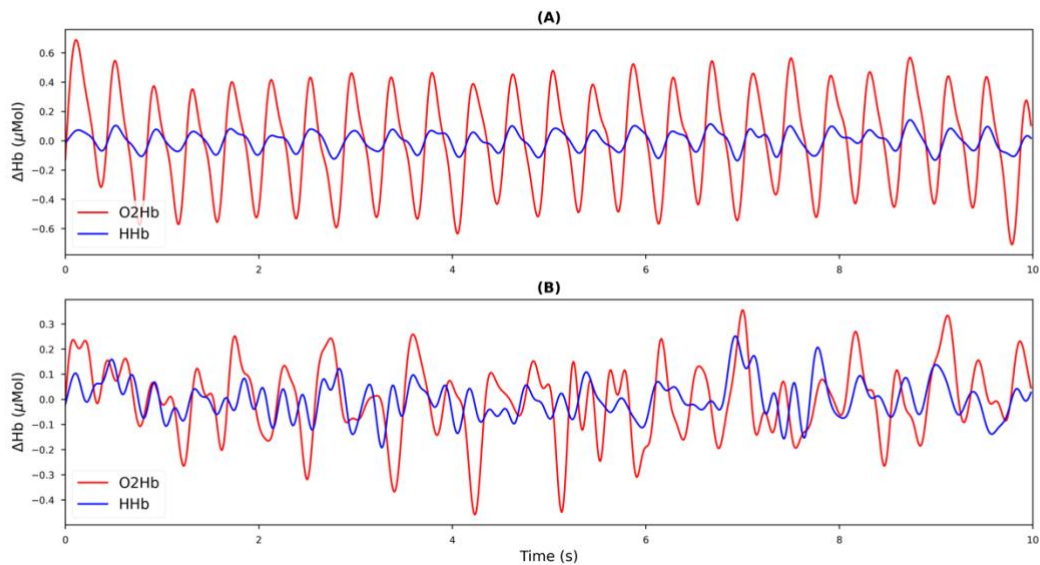
## 4. Signal Quality in NIRS

### 4.1. Importance of signal quality assessment in NIRS

Assessing signal quality in NIRS is paramount for reliable data collection. NIRS signal quality directly impacts measurements related to hemodynamic changes and cerebral oxygenation levels. Factors like motion artifacts or poorly placed optodes can compromise signal quality, obscuring subtle physiological variations and affecting study validity [74]. In both clinical and research settings, it is essential to meticulously evaluate signal quality to ensure data is precise and reliable for drawing meaningful conclusions. This involves optimizing the placement of optodes on the scalp for optimal contact, both before and during recording. In neonatal healthcare, assessing the quality of NIRS signals is vital. Accurate monitoring of oxygenation levels in the neonatal brain helps in early detection of irregularities, allowing for timely interventions that can improve clinical outcomes. High-quality signals also enhance the reliability of continuous monitoring systems and reduce false alarms.

### 4.2. Current methods for assessing signal quality using heartbeat information

A high-quality NIRS signal is distinguished by a robust cardiac component, serving as a key indicator of optimal optode-scalp coupling and forming the basis for evaluating NIRS signal quality [75]. This is attributed to the near-infrared light's journey through both extracerebral and cerebral, where intrinsic and extrinsic factors influence light absorption and scattering. Intrinsic factors involve hemodynamics from systemic artifacts in cerebral and extracerebral compartments, alongside functional brain hemodynamics [59]. The heartbeat systemic artifacts signify sufficient light reaching the brain and predominant absorption and scattering from intrinsic factors. Conversely, compromised optode-scalp coupling, diminishing NIRS signal quality, results from extrinsic factors such as optode looseness, scalp and skull thickness, skin properties, and hair density/color in hair-present cases. Hence, evaluating signal quality involves measuring the intensity of heartbeat artifacts within the raw NIRS signals recorded at a high sampling rate. For instance, Figure 5 displays concentration signals (i.e., O<sub>2</sub>Hb and HHb) sampled at 100 Hz from two recordings, indicating stronger heartbeat artifacts (as quasi-periodic pulses) in one recording (Figure 5(A)), indicative of superior signal quality, compared to the other recording (Figure 5(B)).



**Figure 5:** An example of concentration signals (O2Hb and HHb) from two recordings, measured with a high sampling rate NIRS system (Artinis Medical Systems B.V., Elst, the Netherlands) sampled at 100 Hz. (A) Illustrates pronounced heartbeat artifacts, indicative of potentially superior signal quality. Conversely, (B) displays reduced heartbeat artifacts, suggesting inferior signal quality compared to (A).

Two algorithms, the Scalp Coupling Index (SCI) [75] and PHOEBE (placing headgear optodes efficiently before experimentation) [74], have been proposed to assess scalp-optode coupling strength and the presence of a clear heartbeat in NIRS signals. The SCI, introduced by Pollonini et al. in 2013 [75], measures signal quality by filtering and normalizing photodetected (intensity) signals for each NIRS channel and computing the cross-correlation at a time lag of 0. The resulting zero-lag cross-correlation (i.e., SCI) serves as a quantitative measure of the channel's signal-to-noise ratio (SNR). However, motion artifacts can introduce signal anomalies in the SCI, potentially falsely elevating it. In 2016, Pollonini et al. proposed spectral power analysis to enhance robustness against artifacts, establishing a threshold at 0.1 on the peak value of power [74]. This method objectively distinguishes clean from noisy optical channels, introducing an objective measure of SNR and real-time software for optimal scalp coupling adjustment, aiming to streamline pre-acquisition preparation and maximize SNR for functional hemodynamic mapping.

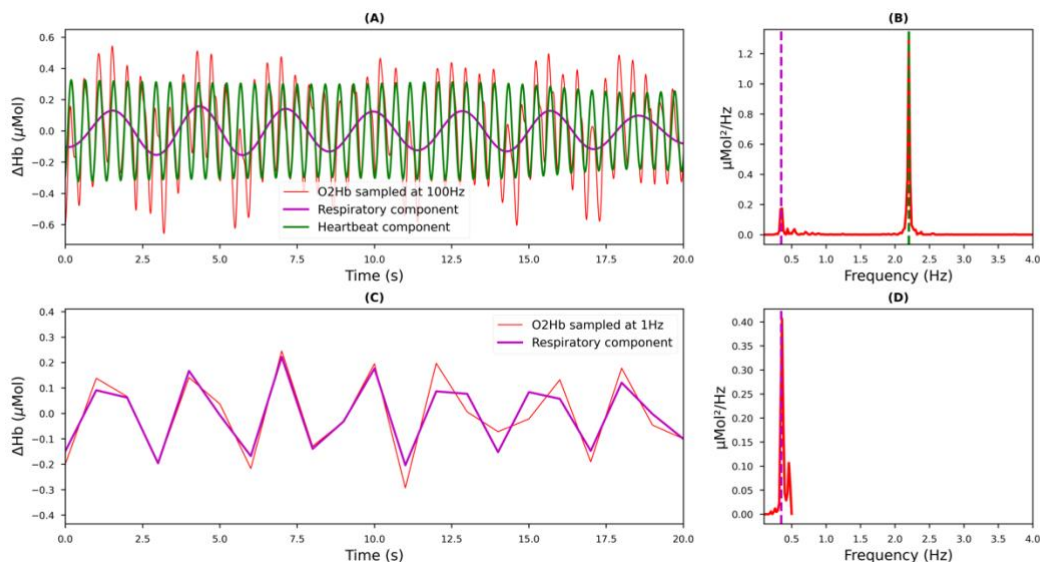
The current signal quality algorithms [74,75] are suboptimally tailored for application to clinical data. Clinical data, in contrast to data recorded in controlled settings from normal subjects, exhibits a greater degree of variability and is less regulated [76]. Consequently, it is imperative for the algorithm to possess the capability to quantify or scale the signal quality in such diverse clinical scenarios. The existing algorithms, however, are primarily designed to make a binary distinction between signals of good and bad quality. This binary approach fails to account for the nuanced levels of signal quality that may be present in signals categorized as good. In challenging situations where external factors severely limit the amount of light reaching the brain, the best achievable signal quality will not necessarily exhibit the characteristics of a high-quality signal obtained under less challenging circumstances. Recognizing and reporting these distinctions in the quality of collected data is essential for subsequent analyses and interpretations of the clinical NIRS data.

## 5. Vital Signs Measurement with NIRS

### 5.1. HR measurement

The raw NIRS signals, encompassing OD signals and concentrations of O<sub>2</sub>Hb and HHb, are often overlooked due to contamination by systemic artifacts [77]. Interestingly, these artifacts present an untapped potential for extracting additional physiological information. Among these artifacts, cardiac pulsation stands out as a prominent source of physiological information within NIRS signals, facilitating the extraction of HR [78].

In the time-domain of raw NIRS signals recorded at a high sampling rate, the pulsatile fluctuations reveal the cardiac physiological information (see Figure 6(A)). This cardiac information, when analyzed in the frequency-domain, manifests as a prominent peak within the subject's HR range (observed in Figure 6(B) around 2.2 Hz), representing the heartbeat as a quasi-periodic component. As a result, HR can be calculated from either the time or frequency-domain of the raw NIRS signals recorded at a high sampling rate. However, if the NIRS system lacks access to raw NIRS signals or if the signals are recorded at a low sampling rate, such as 1 Hz, which is often the case in clinical settings, then the extraction of HR from NIRS is not feasible. This limitation also applies when evaluating the quality of NIRS signals, as the intensity of heartbeats is determined (refer to section 4.2). For instance, Figure 6(C) displays a sample O<sub>2</sub>Hb signal recorded from a neonate at 1 Hz, lacking of HR information evident both in the time-domain and the frequency-domain spectrum of the signal (see Figure 6(D)).



**Figure 6:** Comparison of NIRS signal characteristics between O<sub>2</sub>Hb signals recorded with a NIRS system (Artinis Medical Systems B.V., Elst, the Netherlands) sampled at 100 Hz (top plots) and 1 Hz (bottom plots) from the same subject. (A) Time-domain representation of the O<sub>2</sub>Hb signal (depicted in red) sampled at 100 Hz, showcasing its associated heartbeat (in green) and respiratory (in purple) components derived through band-pass filtering. (B) Frequency-domain spectrum of the O<sub>2</sub>Hb signal in (A), highlighting a distinct peak (represented by a dashed green line) within the subject's HR range (approximately 2.2 Hz), along with a peak (dashed purple line) corresponding to the subject's respiratory rate (RR) range (around 0.3 Hz). (C) Time-domain illustration of the O<sub>2</sub>Hb signal (in red) sampled at 1 Hz, featuring its respiratory component (depicted in purple) isolated via band-pass filtering. (D) Frequency-domain spectrum of the O<sub>2</sub>Hb signal in (C), indicating a peak (dashed purple line) within the subject's RR range.

Numerous studies have affirmed the reliability of HR extracted from NIRS, comparing favorably with standard reference HR derived from ECG and PPG in healthy adults [79–82]. Scholkmann et al. [83] introduced the automatic multiscale-based peak detection (AMPD) algorithm for peak detection in noisy periodic and quasi-periodic signals, and Holper et al. [79] applied this algorithm to extract HR from NIRS data of healthy adults, demonstrating its utility in assessing physical stress. Additionally, Hakimi and Setarehdan [80] highlighted the superior performance of the AMPD algorithm in HR extraction from NIRS signals compared to existing algorithms for PPG. In the same study, they showcased the practicality of NIRS-derived HR in evaluating mental stress in healthy adults.

Continuing the exploration of NIRS applications in different populations, the extraction of HR from infant NIRS presents additional challenges compared to adult NIRS, including a lower SNR, increased motion artifacts, and a wider physiological HR range [76]. Consequently, a system with a high sampling rate becomes imperative to effectively capture the nuances of infant HR. In response to this demand, Perdue et al. [76] introduced an algorithm tailored for extracting HR from NIRS signals recorded in a controlled setting from healthy infants, approximately 7 months old, with a sampling rate of 10 Hz. Extending this research into a longitudinal study, Perdue et al. [84] delved into the applicability of the extracted HR in infants aged 3–12 months with familial risk for autism spectrum disorder. The study aimed to observe HR developmental trajectories in response to speech stimuli, demonstrating the potential of NIRS-derived HR to offer valuable insights into the developmental aspects of infants at risk for autism spectrum disorder.

The use of NIRS for HR extraction in infants presents significant advantages over traditional ECG and PPG methods. While all three techniques are noninvasive, NIRS, which does not require adhesive electrodes or tight-fitting sensors, could potentially reduce the discomfort and stress associated with the routine placement and removal of sensors on the delicate skin of newborns [80,81]. In neonates, NIRS signals on the forehead have been reported to be less susceptible to motion artifacts compared to peripheral PPG signals, given that hand movements are more common than head movements [71]. Furthermore, under stress conditions in adults, NIRS-derived HR has been shown to provide more information than PPG and ECG [71,72].

However, there is a noticeable absence of studies evaluating the performance of the existing algorithms in extracting HR from NIRS data recorded in clinical settings, particularly in neonatal patients in the NICU. Noninvasive monitoring of vital functions is of utmost importance in very small and vulnerable (preterm) neonates with extremely sensitive skin. In this patient group, the goal is to obtain information of body functions with minimal wiring and burden on the patients. A neonatal NIRS system with a robust HR extraction algorithm validated in clinical settings would enable a complementary view of the autonomic nervous system, systemic circulation, and cerebral hemodynamics within a single device.

## 5.2. Respiratory rate (RR) measurement

While the majority of studies traditionally treat the systemic physiological information in NIRS as a source of artifacts, there is a subset of investigations highlighting the potential benefits of such interference in improving the precision of cerebral activity analysis. Notably, Svinkunaite et al. [85] demonstrated that incorporating cardiac and respiratory features extracted from the NIRS spectrum can enhance the accuracy of mental workload classification when combined with NIRS temporal analysis. Similarly, Hakimi et al. [82] showcased the synergistic

advantages of integrating NIRS temporal analysis with extracted HR for stress assessment, resulting in a notable 10% improvement in classification accuracy.

Moving beyond considerations of HR, respiratory patterns emerge as another noticeable systemic physiological information in the NIRS spectrum [86]. For instance, in Figure 6(A), a sample O<sub>2</sub>Hb signal (depicted in red) recorded from a neonate at 100 Hz reveals its corresponding respiratory component, visualized as a sinusoidal pattern. This respiratory component is also apparent in the frequency-domain spectrum of the signal, represented by a peak around 0.3 Hz in Figure 6(B). Unlike HR extraction and signal quality assessment, RR can be derived from NIRS signals with low sampling rates depending on the subject's respiratory frequency range, as illustrated in Figures 6(C) and 5(D), which exhibit a sample O<sub>2</sub>Hb signal sampled at 1 Hz in both time and frequency domains, respectively, from the same neonate. However, it's noteworthy that NIRS systems with such low sampling rates need to provide raw NIRS signals.

Therefore, despite the broad applications of NIRS, there is a conspicuous lack of research exploring the feasibility of estimating RR from NIRS. Such an estimation could provide additional information that could be effectively incorporated into cerebral activity analysis in various contexts. The extraction of RR from NIRS is particularly beneficial for neonates admitted to NICUs, where respiratory distress is a prevalent cause of morbidity [82,83]. Furthermore, RR serves as a vital sign that is crucial for predicting early clinical deterioration in children [84]. The NIRS system, equipped with extracted RR, would allow for the simultaneous assessment of respiration and cerebral perfusion. This could provide a valuable tool for identifying abnormal respiratory events such as tachypnea and apnea, and for analyzing their impact on cerebral hemodynamics.

## 6. Sleep assessment: a possibility of integrating HR and RR with NIRS

The importance of sleep for the brain development of newborns is increasingly recognized [87–89]. It is thought that sleep disruptions during the neonatal period can have developmental implications [90]. Continuous monitoring of neonatal sleep in the NICU could be beneficial in two ways. One, neonatal sleep, which is thought to mirror normal brain maturation, could act as a predictive marker for future outcomes such as the severity of illness or the maturation state of an infant [91]. Two, if sleep is monitored in real-time, elective care could be tailored to sleep stages to promote optimal brain development [92,93].

The use of NIRS to extract HR and RR not only could simplify the collection of diverse physiological data with a single sensor but also could facilitate multimodal feature extraction for assessing sleep stages—an area still largely unexplored in infants. For instance, the Sleep Well Baby algorithm developed by Sentner et al. [94] incorporates HR and RR in its data analysis, significantly enhancing sleep stage accuracy in neonates. The multimodal approach with NIRS could provide a less intrusive and simpler alternative to traditional polysomnography [95], especially beneficial for preterm infants, by using just a single head-mounted sensor.

## Overall aim of the thesis

The overall aim of this thesis is to develop and refine algorithms for continuously evaluating signal quality and extracting RR and HR from NIRS data, with the ultimate goal of improving monitoring techniques for neonates in the NICU. This aim is pursued through a two-part investigation, starting with adult subjects in a controlled environment before transitioning to the more complex clinical setting of the NICU.

The rationale for beginning the studies with healthy adults in a controlled environment before moving to the NICU setting is twofold: First of all, to establishing baseline algorithms and techniques. Working with healthy adult subjects allows us to develop, test, and refine algorithms for signal quality quantification and physiological measurements (RR and HR) without the added complexity of the neonatal physiology and the NICU environment. Healthy adults can offer more stable and predictable physiological signals, fewer motion artifacts due to less movement, and a controlled environment minimizes external variables. This phase ensures that the methodologies are robust and reliable. Second, once the algorithms have proven effective in adults, they can be adapted to the specific challenges presented by neonates, such as their smaller size, more delicate physiology, and the presence of disease or preterm birth complications. The knowledge gained from adult studies provides a foundation for addressing these challenges. For example, algorithms that were devised to handle motion artifacts in adults can be adapted to the more sensitive context of neonates, who may experience different types and frequencies of movement.

## Outline of the thesis

Part I explores the potential of continuously evaluating signal quality and extracting the RR from NIRS in adult subjects. Part II discusses the transition from adult to neonatal NIRS data, leveraging the acquired knowledge to derive HR and RR from NIRS, and conduct a sleep assessment study.

### **Part I: physiological information in adult NIRS**

**Chapter two** describes an algorithm for quantifying signal quality in adult NIRS data.

**Chapter three** introduces a machine learning-based algorithm for quantifying signal quality in adult NIRS data.

**Chapter four** shows the feasibility of deriving RR from adult NIRS data during resting state.

**Chapter five** describes a devised algorithm for extracting RR from adult NIRS data tainted with motion artifacts.

### **Part II: physiological information in neonatal NIRS**

In **Chapter six** a literature review is provided examining early detection of abnormal brain oxygenation and advancements in NIRS technology.

**Chapter seven** presents an algorithm for deriving HR from NIRS data collected from hospitalized (pre)term neonates.

**Chapter eight** presents an algorithm for deriving RR from NIRS data collected from hospitalized (pre)term neonates.

**Chapter nine** introduces a novel NIRS-based method for sleep assessment in hospitalized (pre)term neonates by extracting various physiological features.

## References

1. Ferrari, M.; Quaresima, V. A Brief Review on the History of Human Functional Near-Infrared Spectroscopy (fNIRS) Development and Fields of Application. *Neuroimage* **2012**, *63*, 921–935, doi:10.1016/J.NEUROIMAGE.2012.03.049.
2. Scholkmann, F.; Kleiser, S.; Metz, A.J.; Zimmermann, R.; Mata Pavia, J.; Wolf, U.; Wolf, M. A Review on Continuous Wave Functional Near-Infrared Spectroscopy and Imaging Instrumentation and Methodology. *Neuroimage* **2014**, *85 Pt 1*, 6–27, doi:10.1016/J.NEUROIMAGE.2013.05.004.
3. Yu, Y.; Zhang, K.; Zhang, L.; Zong, H.; Meng, L.; Han, R. Cerebral Near-infrared Spectroscopy (NIRS) for Perioperative Monitoring of Brain Oxygenation in Children and Adults. *Cochrane Database Syst Rev* **2018**, *2018*, doi:10.1002/14651858.CD010947.PUB2.
4. Stepan, J.; Hogue, C.W. Cerebral and Tissue Oximetry. *Best Pract Res Clin Anaesthesiol* **2014**, *28*, 429, doi:10.1016/J.BPA.2014.09.002.
5. Petrova, A.; Mehta, R. Near-Infrared Spectroscopy in the Detection of Regional Tissue Oxygenation during Hypoxic Events in Preterm Infants Undergoing Critical Care. *Pediatr Crit Care Med* **2006**, *7*, 449–454, doi:10.1097/01.PCC.0000235248.70482.14.
6. Kaufman, J.; Almodovar, M.C.; Zuk, J.; Friesen, R.H. Correlation of Abdominal Site Near-Infrared Spectroscopy with Gastric Tonometry in Infants Following Surgery for Congenital Heart Disease. *Pediatr Crit Care Med* **2008**, *9*, 62–68, doi:10.1097/01.PCC.0000298640.47574.DA.
7. Zhong, W.; Ji, Z.; Sun, C. A Review of Monitoring Methods for Cerebral Blood Oxygen Saturation. *Healthcare* **2021**, *9*, doi:10.3390/HEALTHCARE9091104.
8. Hintz, S.R. Near-Infrared Spectroscopy: Neonatal and Perinatal Applications. *Neoreviews* **2001**, *2*, e22–e28, doi:10.1542/NEO.2-1-E22.
9. Vliet, F.F.J. Discovery of the Near-Infrared Window into the Body and the Early Development of near-Infrared Spectroscopy. <https://doi.org/10.1117/1.429952> **1999**, *4*, 392–396, doi:10.1117/1.429952.
10. Wyatt, J.S.; Delpy, D.T.; Cope, M.; Wray, S.; Reynolds, E.O.R. QUANTIFICATION OF CEREBRAL OXYGENATION AND HAEMODYNAMICS IN SICK NEWBORN INFANTS BY NEAR INFRARED SPECTROPHOTOMETRY. *The Lancet* **1986**, *328*, 1063–1066, doi:10.1016/S0140-6736(86)90467-8.
11. Murkin, J.M.; Arango, M. Near-Infrared Spectroscopy as an Index of Brain and Tissue Oxygenation. *Br J Anaesth* **2009**, *103*, i3–i13, doi:10.1093/bja/aep299.
12. Jones, S.; Chiesa, S.T.; Chaturvedi, N.; Hughes, A.D. Recent Developments in Near-Infrared Spectroscopy (NIRS) for the Assessment of Local Skeletal Muscle Microvascular Function and Capacity to Utilise Oxygen. *Artery Res* **2016**, *16*, 25, doi:10.1016/J.ARTRES.2016.09.001.
13. Hamaoka, T.; McCully, K.K. Review of Early Development of Near-Infrared Spectroscopy and Recent Advancement of Studies on Muscle Oxygenation and Oxidative Metabolism. *The Journal of Physiological Sciences* **2019**, *69:6* **2019**, *69*, 799–811, doi:10.1007/S12576-019-00697-2.



14. Gomez, A.; Sainbhi, A.S.; Froese, L.; Batson, C.; Alizadeh, A.; Mendelson, A.A.; Zeiler, F.A. Near Infrared Spectroscopy for High-Temporal Resolution Cerebral Physiome Characterization in TBI: A Narrative Review of Techniques, Applications, and Future Directions. *Front Pharmacol* **2021**, *12*, doi:10.3389/FPHAR.2021.719501/FULL.
15. Zhou, X.; Xia, Y.; Xia, Y.; Uchitel, J.; Collins-Jones, L.; Yang, S.; Yang, S.; Loureiro, R.; Cooper, R.J.; Zhao, H.; et al. Review of Recent Advances in Frequency-Domain near-Infrared Spectroscopy Technologies [Invited]. *Biomedical Optics Express*, Vol. 14, Issue 7, pp. 3234–3258 **2023**, *14*, 3234–3258, doi:10.1364/BOE.484044.
16. Elwell, C.E.; Cope, M.; Edwards, A.D.; Wyatt, J.S.; Delpy, D.T.; Reynolds, E.O.R. Quantification of Adult Cerebral Hemodynamics by Near-Infrared Spectroscopy. <https://doi.org/10.1152/jappl.1994.77.6.2753> **1994**, *77*, 2753–2760, doi:10.1152/JAPPL.1994.77.6.2753.
17. Delpy, D.T.; Cope, M.; Van Der Zee, P.; Arridge, S.; Wray, S.; Wyatt, J. Estimation of Optical Pathlength through Tissue from Direct Time of Flight Measurement. *Phys Med Biol* **1988**, *33*, 1433, doi:10.1088/0031-9155/33/12/008.
18. Kovacsova, Z.; Bale, G.; Bale, G.; Bale, G.; Mitra, S.; Lange, F.; Tachtsidis, I. Absolute Quantification of Cerebral Tissue Oxygen Saturation with Multidistance Broadband NIRS in Newborn Brain. *Biomedical Optics Express*, Vol. 12, Issue 2, pp. 907–925 **2021**, *12*, 907–925, doi:10.1364/BOE.412088.
19. Pinti, P.; Tachtsidis, I.; Burgess, P.W.; Hamilton, A.F. de C. Non-Invasive Optical Imaging of Brain Function with FNIRS: Current Status and Way Forward. In: *Grafman, Jeremy, (ed.) Encyclopedia of the human brain, 2nd edition: reference collection in reference module in Neuroscience and Biobehavioral Psychology, 2023. Elsevier (2023)* **2023**, doi:10.1016/B978-0-12-820480-1.00028-0.
20. Peng, C.; Hou, X. Applications of Functional Near-Infrared Spectroscopy (FNIRS) in Neonates. *Neurosci Res* **2021**, *170*, 18–23, doi:10.1016/J.NEURES.2020.11.003.
21. Dix, L.M.L.; Van Bel, F.; Baerts, W.; Lemmers, P.M.A. Comparing Near-Infrared Spectroscopy Devices and Their Sensors for Monitoring Regional Cerebral Oxygen Saturation in the Neonate. *Pediatric Research* **2013**, *74*:5 **2013**, *74*, 557–563, doi:10.1038/pr.2013.133.
22. Greisen, G.; Wolf, M.; Kleiser, S.; Nasser, N.; Andresen, B. Comparison of Tissue Oximeters on a Liquid Phantom with Adjustable Optical Properties. *Biomedical Optics Express*, Vol. 7, Issue 8, pp. 2973–2992 **2016**, *7*, 2973–2992, doi:10.1364/BOE.7.002973.
23. Scheeren, T.W.L.; Schober, P.; Schwarte, L.A. Monitoring Tissue Oxygenation by near Infrared Spectroscopy (NIRS): Background and Current Applications. *J Clin Monit Comput* **2012**, *26*, 279, doi:10.1007/S10877-012-9348-Y.
24. Vrancken, S.L.; van Heijst, A.F.; de Boode, W.P. Neonatal Hemodynamics: From Developmental Physiology to Comprehensive Monitoring. *Front Pediatr* **2018**, *6*, 1, doi:10.3389/FPED.2018.00087.
25. Zhu, Z.; Liu, T.; Li, G.; Li, T.; Inoue, Y. Wearable Sensor Systems for Infants. *Sensors (Basel)* **2015**, *15*, 3721, doi:10.3390/S150203721.
26. Tang, C.; Lei, F.; Liu, J.; Gong, F. Infection Prevention and Early Warning in Neonatal Intensive Care Unit Based on Physiological Sensor Monitoring. *Front Bioeng Biotechnol* **2023**, *11*, doi:10.3389/FBIOE.2023.1241287.

27. Ryan, R.; Santesso, N.; Lowe, D.; Hill, S.; Grimshaw, J.; Prictor, M.; Kaufman, C.; Cowie, G.; Taylor, M. Interventions to Improve Safe and Effective Medicines Use by Consumers: An Overview of Systematic Reviews. *Cochrane Database Syst Rev* **2014**, *2014*, 7768, doi:10.1002/14651858.CD007768.PUB3.
28. Garvey, A.A.; Dempsey, E.M. Applications of near Infrared Spectroscopy in the Neonate. *Curr Opin Pediatr* **2018**, *30*, 209–215, doi:10.1097/MOP.0000000000000599.
29. Green, M.S.; Sehgal, S.; Tariq, R. Near-Infrared Spectroscopy: The New Must Have Tool in the Intensive Care Unit? *Semin Cardiothorac Vasc Anesth* **2016**, *20*, 213–224, doi:10.1177/1089253216644346/ASSET/IMAGES/LARGE/10.1177\_1089253216644346-FIG3.JPEG.
30. Anton, O.; Dore, H.; Rendon-Morales, E.; Aviles-Espinosa, R.; Seddon, P.; Wertheim, D.; Fernandez, R.; Rabe, H. Non-Invasive Sensor Methods Used in Monitoring Newborn Babies after Birth, a Clinical Perspective. *Maternal Health, Neonatology and Perinatology* **2022** *8:1* **2022**, *8*, 1–11, doi:10.1186/S40748-022-00144-Y.
31. Davoudi, A.; Shickel, B.; Tighe, P.J.; Bihorac, A.; Rashidi, P. Potentials and Challenges of Pervasive Sensing in the Intensive Care Unit. *Front Digit Health* **2022**, *4*, 773387, doi:10.3389/FDGTH.2022.773387/BIBTEX.
32. Tang, C.; Lei, F.; Liu, J.; Gong, F. Infection Prevention and Early Warning in Neonatal Intensive Care Unit Based on Physiological Sensor Monitoring. *Front Bioeng Biotechnol* **2023**, *11*, 1241287, doi:10.3389/FBIOE.2023.1241287/BIBTEX.
33. Trollmann, R. Neuromonitoring in Neonatal-Onset Epileptic Encephalopathies. *Front Neurol* **2021**, *12*, 623625, doi:10.3389/FNEUR.2021.623625.
34. Mukai, T.; Takahashi, N. Neonatal Hypoglycemia. *Cerebral Palsy: Perspective and Clinical Relation to Perinatal Complications/Events in Japan* **2023**, 271–276, doi:10.1007/978-981-19-2217-6\_30.
35. Variane, G.F.T.; Pietrobono, R.F.R.; Noh, C.Y.; Van Meurs, K.P.; Chock, V.Y. Newer Indications for Neuromonitoring in Critically Ill Neonates. *Front Pediatr* **2023**, *11*, 1111347, doi:10.3389/FPED.2023.1111347/BIBTEX.
36. Kumar, N.; Akangire, G.; Sullivan, B.; Fairchild, K.; Sampath, V. Continuous Vital Sign Analysis for Predicting and Preventing Neonatal Diseases in the Twenty-First Century: Big Data to the Forefront. *Pediatr Res* **2020**, *87*, 210, doi:10.1038/S41390-019-0527-0.
37. Villarroel, M.; Chaichulee, S.; Jorge, J.; Davis, S.; Green, G.; Arteta, C.; Zisserman, A.; McCormick, K.; Watkinson, P.; Tarassenko, L. Non-Contact Physiological Monitoring of Preterm Infants in the Neonatal Intensive Care Unit. *NPJ Digit Med* **2019**, *2*, doi:10.1038/S41746-019-0199-5.
38. Reuter, S.; Moser, C.; Baack, M. Respiratory Distress in the Newborn. *Pediatr Rev* **2014**, *35*, 417, doi:10.1542/PIR.35-10-417.
39. Murković, I.; Steinberg, M.D.; Murković, B. Sensors in Neonatal Monitoring: Current Practice and Future Trends. *Technology and Health Care* **2003**, *11*, 399–412, doi:10.3233/THC-2003-11602.
40. Mishra, S.; Khouqeer, G.A.; Aamna, B.; Alodhayb, A.; Ali Ibrahim, S.J.; Hooda, M.; Jayaswal, G. A Review: Recent Advancements in Sensor Technology for Non-Invasive

Neonatal Health Monitoring. *Biosens Bioelectron X* **2023**, *14*, 100332, doi:10.1016/J.BIOSX.2023.100332.

41. Sood, B.G.; McLaughlin, K.; Cortez, J. Near-Infrared Spectroscopy: Applications in Neonates. *Semin Fetal Neonatal Med* **2015**, *20*, 164–172, doi:10.1016/J.SINY.2015.03.008.
42. Maitoza, L.A.; Neeman, E.; Funaro, M.; Pierce, R.W. Relevance of Microvascular Flow Assessments in Critically Ill Neonates and Children: A Systematic Review. *Pediatr Crit Care Med* **2020**, *21*, 373, doi:10.1097/PCC.0000000000002201.
43. Dix, L.M.L.; van Bel, F.; Lemmers, P.M.A. Monitoring Cerebral Oxygenation in Neonates: An Update. *Front Pediatr* **2017**, *5*, 1, doi:10.3389/FPED.2017.00046.
44. Weindling, A.M. Peripheral Oxygenation and Management in the Perinatal Period. *Semin Fetal Neonatal Med* **2010**, *15*, 208–215, doi:10.1016/J.SINY.2010.03.005.
45. Arteaga, G.M.; Crow, S. End Organ Perfusion and Pediatric Microcirculation Assessment. *Front Pediatr* **2023**, *11*, 1123405, doi:10.3389/FPED.2023.1123405/BIBTEX.
46. Korček, P.; Straňák, Z.; Širc, J.; Naulaers, G. The Role of Near-Infrared Spectroscopy Monitoring in Preterm Infants. *Journal of Perinatology 2017 37:10* **2017**, *37*, 1070–1077, doi:10.1038/jp.2017.60.
47. Pellicer, A.; Greisen, G.; Benders, M.; Claris, O.; Dempsey, E.; Fumagalli, M.; Gluud, C.; Hagmann, C.; Hellström-Westas, L.; Hyttel-Sorensen, S.; et al. The SafeBoosC Phase II Randomised Clinical Trial: A Treatment Guideline for Targeted Near-Infrared-Derived Cerebral Tissue Oxygenation versus Standard Treatment in Extremely Preterm Infants. *Neonatology* **2013**, *104*, 171–178, doi:10.1159/000351346.
48. Mitra, S.; Bale, G.; Meek, J.; Tachtsidis, I.; Robertson, N.J. Cerebral Near Infrared Spectroscopy Monitoring in Term Infants With Hypoxic Ischemic Encephalopathy—A Systematic Review. *Front Neurol* **2020**, *11*, 1–17, doi:10.3389/FNEUR.2020.00393.
49. Pavlek, L.R.; Mueller, C.; Jebbia, M.R.; Kiehl, M.J.; Fathi, O. Near-Infrared Spectroscopy in Extremely Preterm Infants. *Front Pediatr* **2021**, *8*, 624113, doi:10.3389/FPED.2020.624113/BIBTEX.
50. Kooi, E.M.W.; Verhagen, E.A.; Elting, J.W.J.; Czosnyka, M.; Austin, T.; Wong, F.Y.; Aries, M.J.H. Measuring Cerebrovascular Autoregulation in Preterm Infants Using Near-Infrared Spectroscopy: An Overview of the Literature. *Expert Rev Neurother* **2017**, *17*, 801–818, doi:10.1080/14737175.2017.1346472.
51. Caicedo, A.; De Smet, D.; Vanderhaegen, J.; Naulaers, G.; Wolf, M.; Lemmers, P.; Van Bel, F.; Ameye, L.; Van Huffel, S. Impaired Cerebral Autoregulation Using Near-Infrared Spectroscopy and Its Relation to Clinical Outcomes in Premature Infants. *Adv Exp Med Biol* **2011**, *701*, 233–239, doi:10.1007/978-1-4419-7756-4\_31/COVER.
52. Thewissen, L.; Caicedo, A.; Lemmers, P.; Van Bel, F. V.; Van Huffel, S. V.; Naulaers, G. Measuring Near-Infrared Spectroscopy Derived Cerebral Autoregulation in Neonates: From Research Tool toward Bedside Multimodal Monitoring. *Front Pediatr* **2018**, *6*, 341349, doi:10.3389/FPED.2018.00117/BIBTEX.
53. Allen, K.A.; Brandon, D.H. Hypoxic Ischemic Encephalopathy: Pathophysiology and Experimental Treatments. *Newborn Infant Nurs Rev* **2011**, *11*, 125, doi:10.1053/J.NAINR.2011.07.004.

54. Ranger, M.; Johnston, C.C.; Limperopoulos, C.; Rennick, J.E.; Du Plessis, A.J. Cerebral Near-Infrared Spectroscopy as a Measure of Nociceptive Evoked Activity in Critically Ill Infants. *Pain Research & Management: The Journal of the Canadian Pain Society* **2011**, *16*, 331, doi:10.1155/2011/891548.
55. Zaleski, K.L.; Kussman, B.D. Near-Infrared Spectroscopy in Pediatric Congenital Heart Disease. *J Cardiothorac Vasc Anesth* **2020**, *34*, 489–500, doi:10.1053/J.JVCA.2019.08.048.
56. Bruckner, M.; Pichler, G.; Urlesberger, B. NIRS in the Fetal to Neonatal Transition and Immediate Postnatal Period. *Semin Fetal Neonatal Med* **2020**, *25*, doi:10.1016/J.SINY.2020.101079.
57. Greisen, G.; Hansen, M.L.; Rasmussen, M.I.S.; Vestager, M.; Hyttel-Sørensen, S.; Hahn, G.H. Cerebral Oximetry in Preterm Infants—To Use or Not to Use, That Is the Question. *Front Pediatr* **2021**, *9*, 747660, doi:10.3389/FPED.2021.747660.
58. Hyttel-Sorensen, S.; Greisen, G.; Als-Nielsen, B.; Gluud, C. Cerebral Near-infrared Spectroscopy Monitoring for Prevention of Brain Injury in Very Preterm Infants. *Cochrane Database Syst Rev* **2017**, *2017*, doi:10.1002/14651858.CD011506.PUB2.
59. Tachtsidis, I.; Scholkmann, F. False Positives and False Negatives in Functional Near-Infrared Spectroscopy: Issues, Challenges, and the Way Forward. <https://doi.org/10.1117/1.NPh.3.3.031405> **2016**, *3*, 031405, doi:10.1117/1.NPH.3.3.031405.
60. Caldwell, M.; Scholkmann, F.; Wolf, U.; Wolf, M.; Elwell, C.; Tachtsidis, I. Modelling Confounding Effects from Extracerebral Contamination and Systemic Factors on Functional Near-Infrared Spectroscopy. *Neuroimage* **2016**, *143*, 91–105, doi:10.1016/J.NEUROIMAGE.2016.08.058.
61. Systemic Physiology Augmented Functional Near-Infrared Spectroscopy: A Powerful Approach to Study the Embodied Human Brain Available online: [https://www.spiedigitallibrary.org/journals/neurophotonics/volume-9/issue-3/030801/Systemic-physiology-augmented-functional-near-infrared-spectroscopy--a-powerful/10.1117/1.NPh.9.3.030801.full#\\_=\\_](https://www.spiedigitallibrary.org/journals/neurophotonics/volume-9/issue-3/030801/Systemic-physiology-augmented-functional-near-infrared-spectroscopy--a-powerful/10.1117/1.NPh.9.3.030801.full#_=_) (accessed on 19 January 2024).
62. Yücel, M.A.; Selb, J.; Aasted, C.M.; Lin, P.-Y.; Borsook, D.; Becerra, L.; Boas, D.A. Mayer Waves Reduce the Accuracy of Estimated Hemodynamic Response Functions in Functional Near-Infrared Spectroscopy. *Biomed Opt Express* **2016**, *7*, 3078, doi:10.1364/BOE.7.003078.
63. Reducing False Discoveries in Resting-State Functional Connectivity Using Short Channel Correction: An FNIRS Study Available online: <https://www.spiedigitallibrary.org/journals/neurophotonics/volume-9/issue-1/015001/Reducing-false-discoveries-in-resting-state-functional-connectivity-using-short/10.1117/1.NPh.9.1.015001.full> (accessed on 19 January 2024).
64. Virtanen, J.; Noponen, T.E.J.; Meriläinen, P. Comparison of Principal and Independent Component Analysis in Removing Extracerebral Interference from Near-Infrared Spectroscopy Signals. <https://doi.org/10.1117/1.3253323> **2009**, *14*, 054032, doi:10.1117/1.3253323.
65. Quantitative Comparison of Correction Techniques for Removing Systemic Physiological Signal in Functional Near-Infrared Spectroscopy Studies Available online: <https://www.spiedigitallibrary.org/journals/neurophotonics/volume-7/issue->

3/035009/Quantitative-comparison-of-correction-techniques-for-removing-systemic-physiological-signal/10.1117/1.NPh.7.3.035009.full (accessed on 19 January 2024).

66. Molavi, B.; Dumont, G.A. Wavelet-Based Motion Artifact Removal for Functional near-Infrared Spectroscopy. *Physiol Meas* **2012**, *33*, 259, doi:10.1088/0967-3334/33/2/259.
67. von Lüthmann, A.; Ortega-Martinez, A.; Boas, D.A.; Yücel, M.A. Using the General Linear Model to Improve Performance in FNIRS Single Trial Analysis and Classification: A Perspective. *Front Hum Neurosci* **2020**, *14*, 514061, doi:10.3389/FNHUM.2020.00030/BIBTEX.
68. Gagnon, L.; Yücel, M.A.; Boas, D.A.; Cooper, R.J. Further Improvement in Reducing Superficial Contamination in NIRS Using Double Short Separation Measurements. *Neuroimage* **2014**, *85*, 127–135, doi:10.1016/J.NEUROIMAGE.2013.01.073.
69. Sweeney, K.T.; Ayaz, H.; Ward, T.E.; Izzetoglu, M.; McLoone, S.F.; Onaral, B. A Methodology for Validating Artifact Removal Techniques for FNIRS. *Annu Int Conf IEEE Eng Med Biol Soc* **2011**, *2011*, 4943–4946, doi:10.1109/IEMBS.2011.6091225.
70. Cooper, R.J.; Caffini, M.; Dubb, J.; Fang, Q.; Custo, A.; Tsuzuki, D.; Fischl, B.; Wells, W.; Dan, I.; Boas, D.A. Validating Atlas-Guided DOT: A Comparison of Diffuse Optical Tomography Informed by Atlas and Subject-Specific Anatomies. *Neuroimage* **2012**, *62*, 1999–2006, doi:10.1016/J.NEUROIMAGE.2012.05.031.
71. Izzetoglu, M.; Chitrapu, P.; Bunce, S.; Onaral, B. Motion Artifact Cancellation in NIR Spectroscopy Using Discrete Kalman Filtering. *Biomed Eng Online* **2010**, *9*, doi:10.1186/1475-925X-9-16.
72. Robertson, F.C.; Douglas, T.S.; Meintjes, E.M. Motion Artifact Removal for Functional near Infrared Spectroscopy: A Comparison of Methods. *IEEE Trans Biomed Eng* **2010**, *57*, 1377–1387, doi:10.1109/TBME.2009.2038667.
73. Zhang, Q.; Brown, E.N.; Strangman, G.E. Adaptive Filtering for Global Interference Cancellation and Real-Time Recovery of Evoked Brain Activity: A Monte Carlo Simulation Study. *J Biomed Opt* **2007**, *12*, 044014, doi:10.1117/1.2754714.
74. Pollonini, L.; Bortfeld, H.; Oghalai, J.S. PHOEBE: A Method for Real Time Mapping of Optodes-Scalp Coupling in Functional near-Infrared Spectroscopy. *Biomedical Optics Express, Vol. 7, Issue 12, pp. 5104-5119* **2016**, *7*, 5104–5119, doi:10.1364/BOE.7.005104.
75. Pollonini, L.; Olds, C.; Abaya, H.; Bortfeld, H.; Beauchamp, M.S.; Oghalai, J.S. Auditory Cortex Activation to Natural Speech and Simulated Cochlear Implant Speech Measured with Functional Near-Infrared Spectroscopy. *Hear Res* **2014**, *309*, 84–93, doi:10.1016/J.HEARES.2013.11.007.
76. Perdue, K.L.; Westerlund, A.; McCormick, S.A.; Nelson, C.A.; Iii, C.A.N. Extraction of Heart Rate from Functional Near-Infrared Spectroscopy in Infants. <https://doi.org/10.1117/1.JBO.19.6.067010> **2014**, *19*, 067010, doi:10.1117/1.JBO.19.6.067010.
77. Izzetoglu, M.; Bunce, S.; Izzetoglu, K.; Onaral, B.; Pourrezaei, A. Functional Brain Imaging Using Near-Infrared Technology. *IEEE Engineering in Medicine and Biology Magazine* **2007**, *26*, 38–46, doi:10.1109/MEMB.2007.384094.

78. Tong, Y.; Lindsey, K.P.; Frederick, B.D. Partitioning of Physiological Noise Signals in the Brain with Concurrent Near-Infrared Spectroscopy and fMRI. *J Cereb Blood Flow Metab* **2011**, *31*, 2352–2362, doi:10.1038/JCBFM.2011.100.
79. Holper, L.; Seifritz, E.; Scholkmann, F. Short-Term Pulse Rate Variability Is Better Characterized by Functional near-Infrared Spectroscopy than by Photoplethysmography. *J Biomed Opt* **2016**, *21*, 091308, doi:10.1117/1.JBO.21.9.091308.
80. Hakimi, N.; Kamaledin Setarehdan, S. Stress Assessment by Means of Heart Rate Derived from Functional Near-Infrared Spectroscopy. <https://doi.org/10.1117/1.JBO.23.11.115001> **2018**, *23*, 115001, doi:10.1117/1.JBO.23.11.115001.
81. Trajkovic, I.; Scholkmann, F.; Wolf, M. Estimating and Validating the Interbeat Intervals of the Heart Using Near-Infrared Spectroscopy on the Human Forehead. <https://doi.org/10.1117/1.3606560> **2011**, *16*, 087002, doi:10.1117/1.3606560.
82. Hakimi, N.; Jodeiri, A.; Mirbagheri, M.; Setarehdan, S.K. Proposing a Convolutional Neural Network for Stress Assessment by Means of Derived Heart Rate from Functional near Infrared Spectroscopy. *Comput Biol Med* **2020**, *121*, 103810, doi:10.1016/J.COMPBIOMED.2020.103810.
83. Scholkmann, F.; Boss, J.; Wolf, M. An Efficient Algorithm for Automatic Peak Detection in Noisy Periodic and Quasi-Periodic Signals. *Algorithms* **2012**, *Vol. 5*, Pages 588–603 **2012**, *5*, 588–603, doi:10.3390/A5040588.
84. Perdue, K.L.; Edwards, L.A.; Tager-Flusberg, H.; Nelson, C.A. Differing Developmental Trajectories in Heart Rate Responses to Speech Stimuli in Infants at High and Low Risk for Autism Spectrum Disorder. *J Autism Dev Disord* **2017**, *47*, 2434–2442, doi:10.1007/S10803-017-3167-4/FIGURES/4.
85. Svinkunaite, L.; Horschig, J.M.; Floor-Westerdijk, M.J. Employing Cardiac and Respiratory Features Extracted from FNIRS Signals for Mental Workload Classification. <https://doi.org/10.1117/12.2587155> **2021**, *11638*, 53–61, doi:10.1117/12.2587155.
86. Izzetoglu, M.; Holtzer, R. Effects of Processing Methods on FNIRS Signals Assessed during Active Walking Tasks in Older Adults. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2020**, *28*, 699–709, doi:10.1109/TNSRE.2020.2970407.
87. Yang, F.N.; Xie, W.; Wang, Z. Effects of Sleep Duration on Neurocognitive Development in Early Adolescents in the USA: A Propensity Score Matched, Longitudinal, Observational Study. *Lancet Child Adolesc Health* **2022**, *6*, 705–712, doi:10.1016/S2352-4642(22)00188-2.
88. Petty, J.; van den Hoogen, A. Brain Development, Promoting Sleep and Well-Being in the Context of Neonatal Developmental Care. In *Neonatal Nursing: A Global Perspective*; Springer International Publishing: Cham, 2022; pp. 135–149.
89. Dereymaeker, A.; Pillay, K.; Vervisch, J.; Van Huffel, S.; Naulaers, G.; Jansen, K.; De Vos, M. An Automated Quiet Sleep Detection Approach in Preterm Infants as a Gateway to Assess Brain Maturation. *Int J Neural Syst* **2017**, *27*, doi:10.1142/S012906571750023X.
90. Besedovsky, L.; Lange, T.; Born, J. Sleep and Immune Function. *Pflugers Arch* **2012**, *463*, 121–137, doi:10.1007/S00424-011-1044-0/FIGURES/4.

91. Scher, M.S.; Loparo, K.A. Neonatal EEG/Sleep State Analyses: A Complex Phenotype of Developmental Neural Plasticity. *Dev Neurosci* **2009**, *31*, 259–275, doi:10.1159/000216537.
92. Allen, K.A. Promoting and Protecting Infant Sleep. *Adv Neonatal Care* **2012**, *12*, 288, doi:10.1097/ANC.0B013E3182653899.
93. Colombo, G.; De Bon, G. Strategies to Protect Sleep. *J Matern Fetal Neonatal Med* **2011**, *24 Suppl 1*, 30–31, doi:10.3109/14767058.2011.607571.
94. Sentner, T.; Wang, X.; De Groot, E.R.; Van Schaijk, L.; Tataranno, M.L.; Vijlbrief, D.C.; Benders, M.J.N.L.; Bartels, R.; Dudink, J. The Sleep Well Baby Project: An Automated Real-Time Sleep-Wake State Prediction Algorithm in Preterm Infants. *Sleep* **2022**, *45*, doi:10.1093/SLEEP/ZSAC143.
95. Werth, J.; Atallah, L.; Andriessen, P.; Long, X.; Zwartkruis-Pelgrim, E.; Aarts, R.M. Unobtrusive Sleep State Measurements in Preterm Infants – A Review. *Sleep Med Rev* **2017**, *32*, 109–122, doi:10.1016/J.SMRV.2016.03.005.



# Part I

Physiological Information  
in Adult NIRS







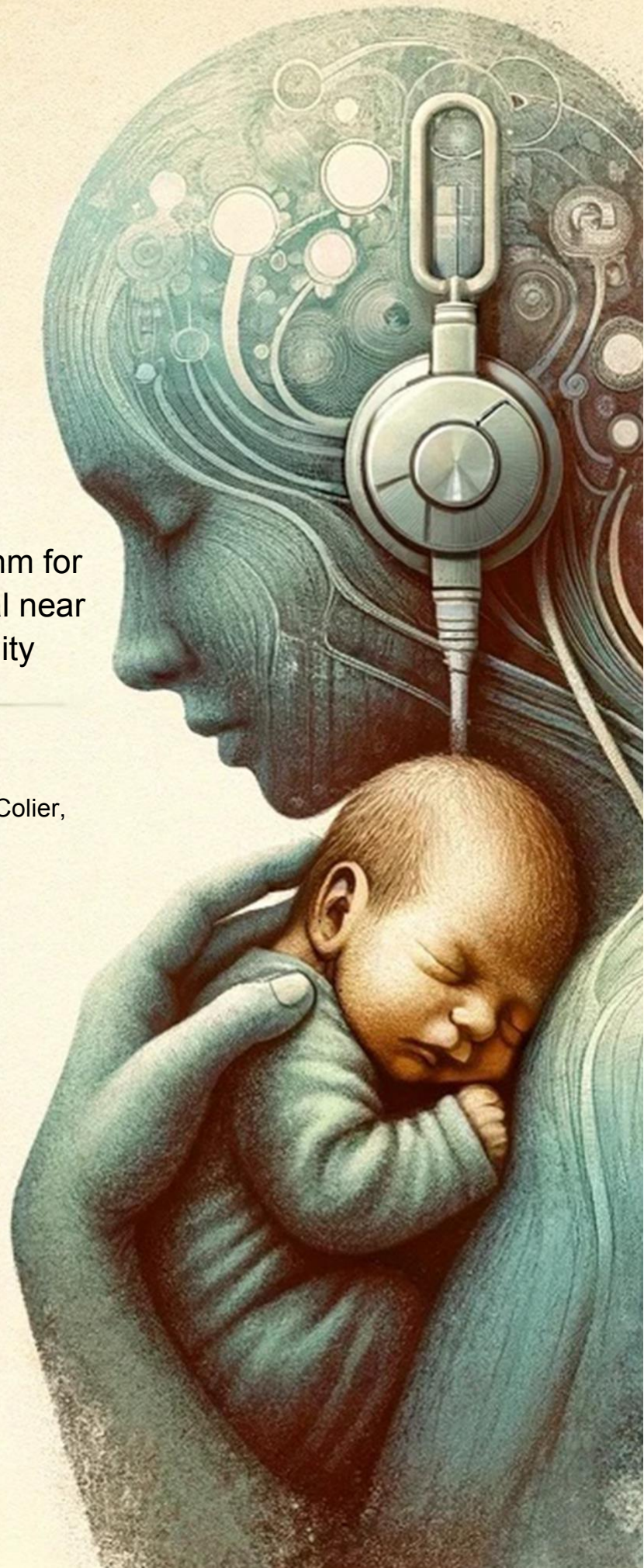
## CHAPTER 2

Signal Quality Index: a novel algorithm for quantitative assessment of functional near infrared spectroscopy signal quality

M. Sofia Sappia\*, Naser Hakimi\*, Willy N.J.M. Colier,  
Jörn M. Horschig

\* shared first author

doi: <https://doi.org/10.1364/BOE.409317>



## Abstract

We propose the Signal Quality Index (SQI) algorithm as a novel tool for quantitatively assessing the functional near infrared spectroscopy (fNIRS) signal quality in a numeric scale from 1 (very low quality) to 5 (very high quality). The algorithm comprises two preprocessing steps followed by three consecutive rating stages. The results on a dataset annotated by independent fNIRS experts showed SQI performed significantly better ( $p < 0.05$ ) than PHOEBE (Placing Headgear Optodes Efficiently Before Experimentation) and SCI (Scalp Coupling Index), two existing algorithms, in both quantitatively rating and binary classifying the fNIRS signal quality. Employment of the proposed algorithm to estimate the signal quality before processing the fNIRS signals increases certainty in the interpretations.



# 1. Introduction

Functional near-infrared spectroscopy (fNIRS) neuroimaging technique makes it possible to non-invasively investigate brain activity in both experimental and clinical settings [1-12]. In order to guarantee a reliable estimation of the functional hemodynamics in the brain cortex, the initial task of a researcher is to collect signals with a good quality. A poor signal quality may lead to wrong interpretations of the collected data and to consequent findings of false positives and false negatives in the analysis. False positives and false negatives in fNIRS functional experiments are changes in O<sub>2</sub>Hb and HHb concentrations over a certain brain area that can be caused by task related systemic activity and/or extracerebral hemodynamics, but could be mistakenly interpreted as increased neural activity (false positives) in that brain area, or could mask the neuronally induced hemodynamic response (false negatives) [13].

To the best of our knowledge, there are no standardized criteria for quantitatively assessing fNIRS signal quality unlike, for example, in electroencephalography (EEG), where the impedance of the electrodes with the human scalp can be assessed to infer good scalp contact. These impedances are usually measured and reported, providing researchers with a reference metric for conducting proper set-up and contributing to a standardized assessment of EEG signal quality over the entire research community [14]. In the case of fNIRS, however, it is currently up to the researcher's expertise and subjective judgment to deem a signal good enough, which makes it difficult to have reliable and reproducible studies. In addition, researchers new to the field of fNIRS are often faced with the issue that they lack the experience to judge the quality of the signals, leading to recordings of poor or inconsistent quality data. Therefore, it is necessary to have an objective measure that quantifies fNIRS signal quality independently of the researcher's experience and subjective judgment.

A high quality fNIRS signal is characterized by the presence of a strong cardiac component, which is the main indicator of a good optode-scalp coupling and can be employed for the assessment of fNIRS signal quality. The reason behind this is that in fNIRS, emitted near infrared light travels through superficial and cerebral layers. When the light passes through these layers, intrinsic and extrinsic factors affect the absorption and scattering of the transmitted light. The intrinsic factors are hemodynamics caused by systemic artefacts in the cerebral and extracerebral compartments, as well as the functional brain hemodynamics in the cerebral compartment [10, 15]. The heartbeat is one of such systemic artefacts, present in both compartments [8, 13]. Hence, its presence in fNIRS signals indicates that enough light has reached the brain and that most of the absorption and scattering are caused by intrinsic factors. This is not the case when extrinsic factors excessively limit the amount of light reaching the brain, causing a decrease in the optode-scalp coupling and compromising fNIRS signal quality. Examples of extrinsic factors are looseness of the optodes, scalp and skull thickness, skin properties, and hair density and color in the cases where hair is present [16].

With the purpose of assessing the strength of the scalp-optode coupling and the presence of a clear heartbeat in fNIRS signals, two algorithms have been proposed for distinguishing between good and bad quality fNIRS signals: SCI (Scalp Coupling Index, [17]) and PHOEBE (Placing Headgear Optodes Efficiently Before Experimentation, [18]). SCI and PHOEBE, though, are not designed to discern low and high quality signals from medium quality signals. Instead, they are designed to binary discriminate between good and bad quality signals. Such a distinction is sharp and does not consider the different levels of signal quality that a good

quality signal may have. For challenging cases in which the amount of light reaching the brain is strongly limited due to unavoidable extrinsic factors, the best quality signal that can be achieved will not have the characteristics of a high quality signal achieved for a less challenging case. Being aware of these differences in the signal quality of the collected data and reporting them is important for subsequent analysis and interpretations. While for photoplethysmography (PPG), an optical modality sharing the same principles as fNIRS, several signal quality measures have been proposed [19, 20], for fNIRS no efforts have been conducted so far to rate fNIRS signal quality in more than two levels. Here, we propose an algorithm for quantitatively rating the fNIRS signal quality based on the strength of the optodes coupling with the scalp, providing the fNIRS research community with an objective estimate of fNIRS signal quality and eliminating the researchers' subjective bias. In this study, the proposed Signal Quality Index (SQI) algorithm is compared with the SCI and PHOEBE algorithms in quantitative rating as well as binary classification of the fNIRS signal quality.

## 2. Materials and methods

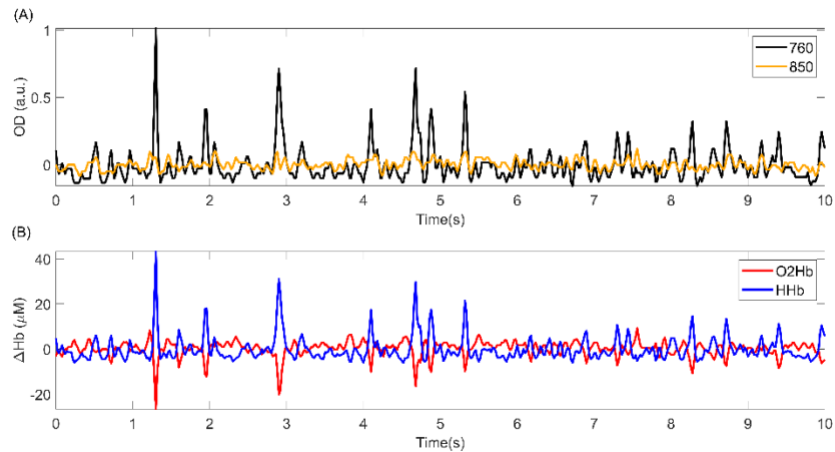
### 2.1. Data collection and annotation

We assembled two sets of data from previously recorded, unpublished data: a training dataset, to develop and fit the parameters of the algorithm; and a validation dataset, to provide an unbiased evaluation of the performance of the algorithm on an unseen set of data.

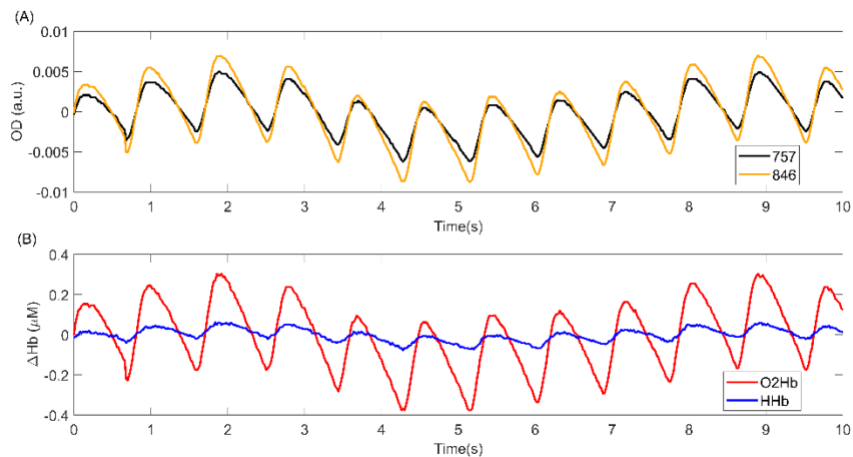
#### 2.1.1. Training dataset

The training dataset used in this study consisted of fNIRS recordings of adult healthy volunteers collected with the following devices using OxySoft software (Artinis Medical Systems B.V., Elst, The Netherlands): OctaMon, Brite 23, Brite 24, and OxyMon (all by Artinis Medical Systems B.V., Elst, The Netherlands). The transmitters for all of these devices emit light at two different wavelengths in the near infrared spectrum. This dataset consisted of 158 10-second signal segments of optical densities and HHb and O<sub>2</sub>Hb concentration changes, which were sampled at 50 Hz. The dataset contains two optical density signals per signal segment, corresponding to the two wavelengths used by each device. The data was collected from 14 participants. The signal segments were sampled from arbitrary channels of the dataset. Source detector distances ranged from 3 to 3.5 cm, and the brain areas measured were prefrontal cortex, temporal cortex and parietal cortex.

Seven independent fNIRS experts (annotators A-G) working at Artinis Medical Systems B.V. rated all 158 signal segments on the presence of motion artefacts and overall signal quality. Presence of motion artefacts was assessed as a "Yes/No" question. Signal quality was rated on an ordinal scale ranging from 1 (very low quality) to 5 (very high quality). We chose these 5 levels of quality because it is the scale most often used by fNIRS experts working at Artinis Medical Systems B.V. (the annotators) for discriminating between different levels of fNIRS signal quality. For each signal segment, both the changes in optical densities and the corresponding changes in O<sub>2</sub>Hb and HHb concentrations were shown to the annotator. The signal segments were presented to the annotators in randomized order. Figure 1 and Figure 2 show, respectively, an example of a signal segment rated as having very low and very high signal quality by all annotators.

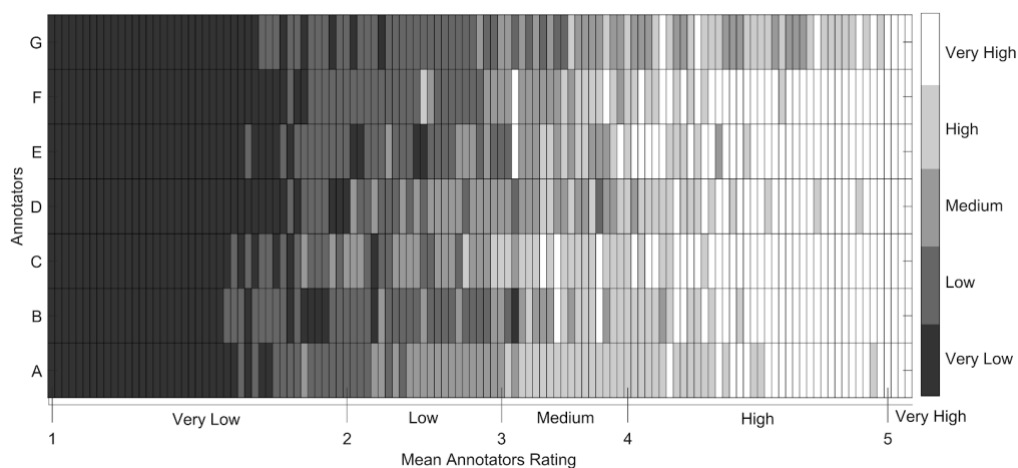


**Figure 1:** A sample signal segment having a very low quality. This signal segment was rated as having very low quality by all annotators in the training phase. (A) Orange and black curves represent the detrended optical density signals for the wavelengths 850 and 760 nm, respectively. (B) Red and blue curves represent the detrended changes in O2Hb and HHb concentrations, respectively.



**Figure 2:** A sample signal segment having a very high quality. This signal segment was rated as having very high quality by all annotators in the training phase. (A) Orange and black curves represent the detrended optical density signals for the wavelengths 846 and 757 nm, respectively. (B) Red and blue curves represent the detrended changes in O2Hb and HHb concentrations, respectively.

Signal segments containing motion artefacts according to at least half of the annotators were excluded, leaving a total of 123 signal segments. The resulting dataset includes a similar number of signal segments for each of the five different signal quality levels. Figure 3 shows the ratings given by the annotators for each of the 123 signal segments. Generally, very low (1) and very high (5) rated signal segments received similar ratings across annotators, while there was less agreement across the annotators ratings for signal segments that were rated from low (2) to high quality (4). We used the mean annotators rating as the reference rating for the quality of the fNIRS signal segments in the training phase (see section 2.3.1).

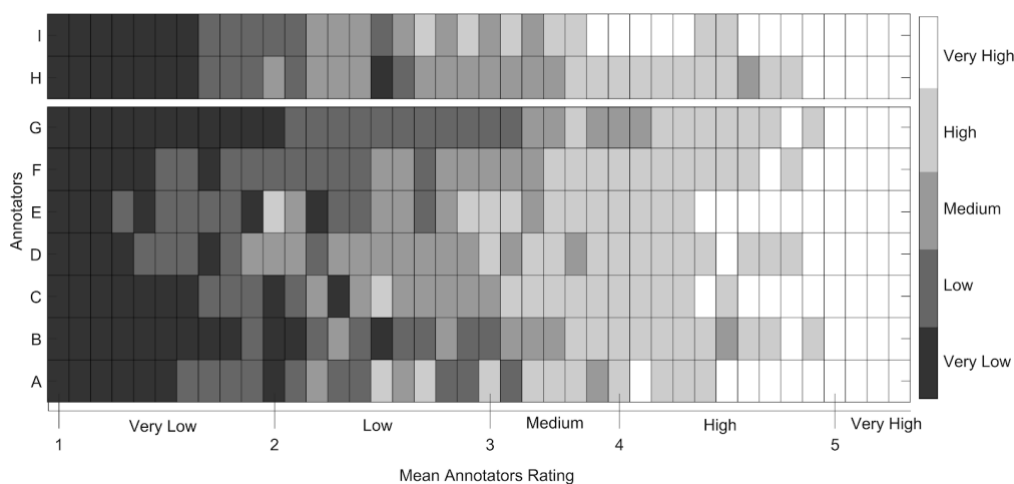


**Figure 3:** The ratings given by the annotators for each of the signal segments included in the training dataset. The x-axis represents the mean annotators rating for each of the signal segments from very low to very high signal quality. The y-axis shows the ratings of the seven annotators A-G. The colorbar on the right represents the ordinal rating scale considered by the annotators: 1 (Very Low), 2 (Low), 3 (Medium), 4 (High), and 5 (Very High). The signal segments were sorted in ascending order based on their mean annotators rating.

### 2.1.2. Validation dataset

For the validation phase, a new set of data was assembled, which comprised fNIRS recordings of adult healthy volunteers collected with the same devices as the training dataset, using OxySoft software (Artinis Medical Systems B.V., Elst, The Netherlands): OctaMon, Brite 23, Brite 24, and OxyMon (all by Artinis Medical Systems B.V., Elst, The Netherlands). This dataset consisted of 40 10-second signal segments of optical densities, which were sampled at 50 Hz (as the data in the training dataset). The data was collected from 4 participants, 3 of which were also included in the training dataset. The signal segments were sampled from arbitrary channels of the dataset. Source detector distances ranged from 3 to 3.5 cm, and the brain areas measured were prefrontal cortex, temporal cortex and parietal cortex.

The validation dataset was rated by the same annotators as in the training dataset, and two additional fNIRS experts working at Artinis Medical Systems B.V., each with more than 10 years of experience in the field of fNIRS. The validation set was free of motion artefacts, and hence the annotators were only asked to rate the signal quality on an ordinal scale ranging from 1 (very low quality) to 5 (very high quality). For the rating, the signal segments were presented to the annotators in the same manner as was done for the training dataset. Figure 4 shows the ratings given by the annotators for each of the 40 signal segments. Although there was less agreement between the annotators ratings for the signal segments that were annotated as having low (2) and medium (3) quality; very low (1), high (4), and very high (5) rated signal segments received similar ratings across annotators. The mean of the annotators ratings was considered as the reference rating for the quality of the fNIRS signal segments in the validation phase (see section 2.3.2).



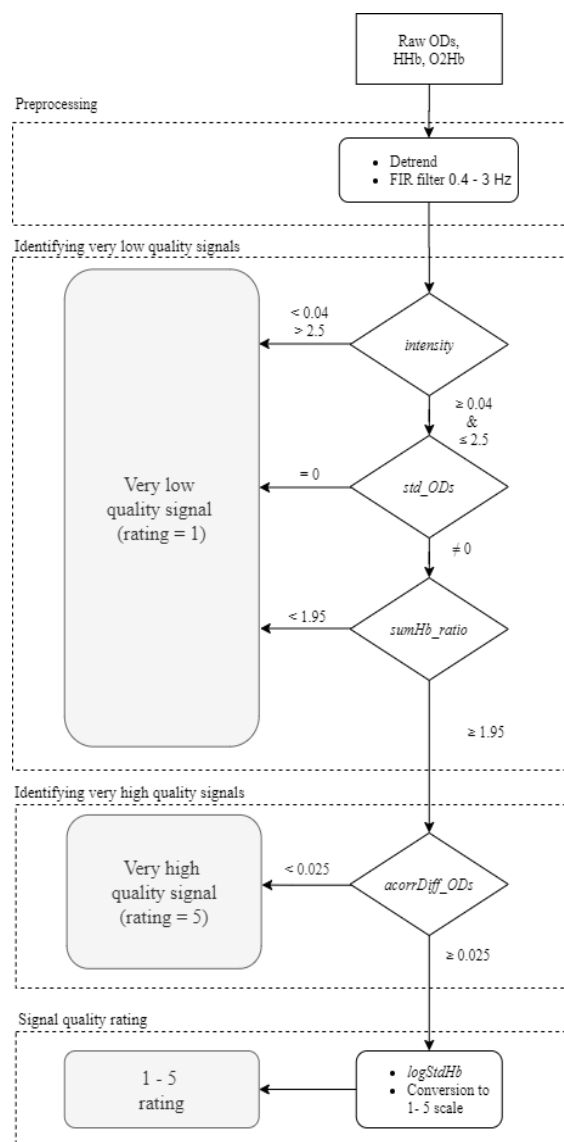
**Figure 4:** The ratings given by the annotators for each of the signal segments included in the validation dataset. The x-axis represents the mean annotators rating for each of the signal segments from very low to very high signal quality. The y-axis shows the ratings of the nine annotators: the seven annotators that rated the training dataset (A-G) and the two additional annotators I-H. The colorbar on the right represents the ordinal rating scale considered by the annotators: 1 (Very Low), 2 (Low), 3 (Medium), 4 (High), and 5 (Very High). The signal segments were sorted in ascending order based on their mean annotators ratings.

## 2.2. SQI algorithm workflow

The here developed Signal Quality Index (SQI) algorithm comprises two preprocessing steps (see section 2.2.1) followed by three consecutive rating stages: 1) identifying very low quality signals, 2) identifying very high quality signals and 3) signal quality rating. Rating stage one and two hereby serve as heuristics to identify signal segments with clear characteristics of very low quality signals or very high quality signals, respectively, early on. In the third stage, a signal segment could still be rated as having very low or very high quality, but could also receive a rating in between.

The workflow of the SQI algorithm per signal segment is as follows (see Figure 5). After signal preprocessing (see section 2.2.1), the signal segment enters rating stage one (for more details, see Materials and Methods, section Rating stage one: identifying very low quality signals), where it is identified as a very low quality signal or otherwise enters rating stage two. If the signal segment is identified as a very low quality signal in rating stage one, it is rated with the lowest rating (i.e. 1) and does not go through the subsequent stages of the algorithm. Similarly, in rating stage two (for more details, see Materials and Methods, section Rating stage two: identifying very high quality signals) the signal segment is identified as a very high quality signal or not. If the signal segment is identified as a very high quality signal, it is rated with the highest rating (i.e. 5) and does not enter the signal quality rating stage. If instead the signal segment is not identified as a very high quality signal in rating stage two, then it enters the signal quality rating stage. In this third stage (for more details, see Materials and Methods, section Rating stage three: signal quality rating), a rating between 1 (very low quality) and 5 (very high quality) is assigned to the signal segment.





**Figure 5:** Flowchart of the rating process implemented by the SQI algorithm. The SQI algorithm comprises two preprocessing steps followed by three consecutive rating stages: 1) identifying very low quality signals, 2) identifying very high quality signals, and 3) signal quality rating. A 10-second signal segment is input into the algorithm in the form of raw optical densities and HHb and O2Hb concentration changes. After signal preprocessing (see section 2.2.1 for details), the signal segment enters rating stage one, in which it is evaluated by three consecutive features (see Materials and Methods, section Rating stage one: identifying very low quality signals): absolute light intensity (abbreviated as *intensity*) assesses whether the intensity of the measured light is within a certain range, standard deviation of the ODs per wavelength (*std\_ODs*) detects whether one or both optical density signal segments are flat, and ratio of oxygenated and deoxygenated hemoglobin summation (*sumHb\_ratio*) assesses the difference in amplitude between O2Hb and HHb. If any of these features exceeds the thresholds for the signal segment, it is identified as a very low quality signal, obtaining the lowest rating (i.e. 1), and does not go through the subsequent stages of the algorithm. Otherwise, it enters rating stage two. In this stage, the feature standard deviation of the difference between optical density signals autocorrelation (*acorrDiff\_ODs*) evaluates whether the two optical densities in the signal segment have similar amplitude and shape. If the value obtained for this feature is lower than the threshold, the signal segment is identified as a very high quality signal, it is rated with the highest rating (i.e. 5), and does not enter the signal quality rating stage. Otherwise, it enters the signal quality rating stage. In this stage, the signal segment is evaluated by the feature logarithm of the ratio between the standard deviation of O2Hb and HHb (*logStdHb*). This feature characterizes the relative presence of the heartbeat in O2Hb and HHb signals. The raw value obtained for this feature is then converted to assign a rating between 1 (very low quality) and 5 (very high quality).

## 2.2.1. Preprocessing

First, the received light intensities digitized with 16-bit resolution were converted to optical densities (OD). The optical densities were then converted into oxygenated hemoglobin (O2Hb) and deoxygenated hemoglobin (HHb) changes in concentration by applying the modified Beer-Lambert law [21]. In the following, we processed the OD values as well as O2Hb and HHb, as each of them was required for different features. Two preprocessing steps were applied to ODs, O2Hb and HHb. Firstly, HHb, O2Hb and OD signal segments were detrended by subtracting the least-squares fit of a straight line to the data. Secondly, these signal segments were band-pass filtered using a zero-phase forward Hamming-windowed sinc FIR filter of order 208 with cutoff frequencies (-6 dB) of 0.4 Hz and 3 Hz. The transition width of the filter was of 0.8 Hz, with a passband between 0.8-2.6 Hz, and a stopband (-53 dB) between 3.4 - 25 Hz. The filter order was estimated based on the normalized transition width for the Hamming window [22]. This band-pass filter was implemented using `ft_preproc_bandpassfilter` function in the FieldTrip toolbox [23], commit `62c9a0d` on master branch. In order to avoid edge artefacts, prior to filtering, the signal segments were zero-padded with a sample length corresponding to two seconds of data both in the beginning and end of the signal segment.

## 2.2.2. SQI features

### 2.2.2.1. Rating stage one: identifying very low quality signals

We used three different features for assessing whether a signal segment is identified as a very low quality signal in this stage:

- *Absolute light intensity* (abbreviated as *intensity*): this feature represents the intensity of the measured light, which should be within a certain range in order to guarantee a linear response from the system and a sufficient signal-to-noise ratio. Bright ambient light, low pigmentation in skin and hair, or lack of hair can saturate the detector, while too little or no light could indicate that there is too much light absorption, e.g. by hair between the scalp and the optodes. We use a thresholding of the raw light intensity, as has previously been used in [16, 24]. Lower (too much light) and upper thresholds (too little light) were respectively set to 0.04 OD and 2.5 OD. If the signal segment exceeded these thresholds, it was identified as a very low quality signal.
- *Standard deviation of the ODs per wavelength* (*std\_ODs*): this feature detects whether one or both optical density signal segments are flat, and thus have a standard deviation of zero, indicating a poor coupling with the scalp. This feature was calculated as the standard deviation of the optical density over the whole 10-second length of the signal segment. If the signal segment obtained a value of zero for either wavelength, it was identified as a very low quality signal. The feature *std\_ODs* was constrained to be equal to zero rather than being evaluated by a higher threshold, because empirical evaluation showed that values different than zero were not discriminative.
- *Ratio of oxygenated and deoxygenated hemoglobin summation* (*sumHb\_ratio*): high quality fNIRS signals are characterized by a clear heartbeat for O2Hb signals, with a smaller and attenuated one for HHb signals (see Figure 2). This difference makes the sum of the absolute values of the O2Hb signals greater than that of the HHb signals. In contrast, for very low quality fNIRS signals, the opposite happens. Feature *sumHb\_ratio*, considers this difference by computing the ratio between the sum over samples of the absolute value of O2Hb and the sum over samples of the absolute

value of HHb for the 10-second signal segment (Eq. (1), where  $|x|$  represents the absolute value of  $x$ ). Low values for this feature indicate very low signal quality due to the prominence of HHb signal amplitude over that of O2Hb. The threshold for this feature was empirically derived from the training dataset and set to 1.95. If the signal segment obtained a value lower than this threshold, it was identified as a very low quality signal.

$$sumHb\_ratio = \sum |O_2Hb| / |HHb| \quad (1)$$

### 2.2.2.2. Rating stage two: identifying very high quality signals

In this stage, we used one feature for assessing whether a signal segment is identified as a very high quality signal:

- *Standard deviation of the difference between optical density signals autocorrelation (acorrDiff\_ODs)*: for very high quality signals, the autocorrelation for the two optical densities of a signal segment have similar amplitude and shape. Hence, the standard deviation of the difference between them is low. This feature was calculated by Eq. (2), where  $OD1acorr$  and  $OD2acorr$  are the autocorrelations of the optical densities measured for each wavelength, and  $std()$  is the standard deviation.

$$acorrDiff\_ODs = std(OD1acorr - OD2acorr) \quad (2)$$

A threshold of 0.025 was used, as empirically derived from the training dataset. If the signal segment obtained a value below this threshold, it was identified as a very high quality signal.

### 2.2.2.3. Rating stage three: signal quality rating

In this stage, we used one feature for rating the signal segment between 1 (very low quality) and 5 (very high quality).

- *Logarithm of the ratio between the standard deviation of O2Hb and HHb (logStdHb)*: the natural logarithm of the ratio of the standard deviation of the 10-second O2Hb and HHb time series was computed by Eq. (3). Its output value was used to quantitatively rate the quality of the signal segment in an adimensional magnitude. High quality fNIRS signals are characterized by a strong clear heartbeat as the strongest signal component. For high quality signals, the magnitude of the heartbeat is higher in the O2Hb signal than in HHb. Therefore, the standard deviation of O2Hb is higher than the standard deviation of HHb for high quality fNIRS signals. Consequently, the higher the signal quality, the higher the ratio between them, which adopts an exponential trend. This trend is made linear by means of the natural logarithm implemented in this feature.

$$logStdHb = \ln \left( std \left( \frac{O_2Hb}{HHb} \right) \right) \quad (3)$$

## 2.3. Performance assessment

We assessed the performance of the SQI algorithm in two different phases: training phase and validation phase. All computation and analysis in this study were done using MATLAB R2019b (MathWorks, Natick, Massachusetts), on an ASUS workstation with Intel Core-i7-8565U @ 1.99 GHz CPU and 16 GB RAM.

### 2.3.1. Training phase

First, we assessed the correctness of each of the features of the SQI algorithm on the training dataset. Next, we assessed the performance of the SQI algorithm with respect to the mean annotators rating.

#### 2.3.1.1. Features correctness assessment

In order to assess and compare the correctness of the different features included in rating stage one (identifying very low quality signals) and rating stage two (identifying very high quality signals) of the algorithm on the training dataset, we computed the following quantitative metrics: precision in correct identifications, number of signal segments identified as very low quality signals (for features in rating stage one) and number of signal segments identified as very high quality signals (for the feature in rating stage two), correct and incorrect decisions. For rating stage one of the algorithm, the identified signal segments for which the mean annotators ratings were below or equal to 2 (low quality) were classified as correct decisions, and otherwise classified as incorrect decisions. Similarly, for rating stage two of the algorithm, the identified signal segments for which the mean annotators ratings were above or equal to 4 (high quality) were classified as correct decisions, and otherwise classified as incorrect decisions.

For evaluating the correctness of the rating feature *logStdHb* in the third stage of the algorithm, we used Pearson's correlation coefficient between the raw feature values and the mean annotators ratings for all 123 signal segments in the training dataset. To test for the significance of the correlation, we considered a significance level  $\alpha=5\%$  and computed the p-value using a Student's t distribution as implemented in Matlab's *corr* function. We applied the Benjamini-Hochberg method [25] to correct for multiple comparisons.

#### 2.3.1.2. Algorithm performance assessment

We assessed the performance of the SQI algorithm in 1) quantitative rating the signal quality, 2) binary classification of the signal quality (good or bad signal quality). We compared the performance of the SQI algorithm with the performance of two existing algorithms used for binary assessing fNIRS signal quality: SCI [17] and PHOEBE [18].

SCI quantifies the similarity of the cardiac component in both optical densities to determine the strength of the coupling between the scalp and the optodes. The algorithm computes the zero-lag cross-correlation between both optical density signal segments as a quantitative measure of the signal-to-noise ratio of the signal segment [17]. According to the original implementation of the SCI algorithm, signal segments with a zero-lag cross-correlation value above 0.75 are considered high quality signals. The PHOEBE algorithm evaluates the similarity of both optical density signal segments by means of the SCI metric, as well as the spectral power of their cross-correlation to determine the strength of the cardiac component. According to the original implementation of the PHOEBE algorithm, signal segments with a spectral value above 0.1 are considered high quality signals. Both SCI and PHOEBE were computed as implemented by the main author of the original papers [26].

##### 2.3.1.2.1. Quantitative rating performance

A regression analysis was performed to obtain the linear models that convert the raw values of the rating feature in the signal quality rating stage of SQI to the 1 - 5 scale that the annotators

used for rating. The linear model was built by computing a least squares regression between the mean annotators ratings and the raw values of the rating feature *logStdHb* on the training data to obtain the slope and intercept. This regression was carried out only on the signal segments that were not identified in rating stage one or two, and that hence entered rating stage three in the SQI algorithm workflow. The slope and intercept were then used to convert the raw value of the rating feature *logStdHb* to a rating on the 1-5 scale. In order to compare the performance of SCI and PHOEBE with the proposed algorithm, the raw values of the rating features used in SCI and PHOEBE were also converted into a 1-5 scale in the same manner. In this conversion, we considered the zero-lag cross-correlation between the two optical density signals and the peak spectral power of the cross-correlation as the rating features used in SCI and PHOEBE, respectively (see Materials and Methods, section Algorithm performance assessment). The ratings obtained for each of the algorithms in a 1-5 scale will be hereon referred to as estimated ratings.

Moreover, Bland Altman limits of agreement (BLA) [27] were computed between the mean annotators ratings and the estimated ratings introduced in this section for each of the evaluated algorithms SQI, SCI, and PHOEBE. Bland Altman limits of agreement were computed by calculating 1.96 times the standard deviation of the error, representing the range in which 95% of the differences between the mean annotators ratings and estimated ratings fell. Three quantitative measures were calculated to further compare the performance of the evaluated algorithms: mean of error (ME), standard deviation of error (StdE), and coefficient of determination ( $r^2$ ) assessed by Pearson's correlation coefficient. To test for the significance of the correlation, we considered a significance level  $\alpha=5\%$  and computed the p-value using a Student's t distribution as implemented in Matlab's *corr* function. In addition to the three quantitative measures, the computation time using our reference implementation in MATLAB R2019b (MathWorks, Natick, Massachusetts) for a single signal segment was measured five times and then averaged to compare the time required for each of the algorithms to obtain the estimated rating.

#### 2.3.1.2.2. Binary classification performance

To conduct a binary classification for the proposed SQI algorithm, signal segments identified in rating stage one (identifying very low quality signals) were classified as "bad quality" signals; while those identified in rating stage two (identifying very high quality signals) were classified as "good quality" signals. For the signal segments that entered rating stage three, a threshold for the raw value of the feature *logStdHb* was empirically derived on the training dataset. This threshold value was set to 1.478. Signal segments with a value equal or higher than this threshold were classified as "good quality" signals. They were otherwise classified as "bad quality" signals. For SCI and PHOEBE, we used the thresholds proposed in the original papers (see [17, 18]): a zero-lag cross-correlation value of 0.75 and a peak spectral power of 0.1. For both algorithms, signal segments with values equal or higher than the considered thresholds were classified as "good quality" signals. They were otherwise classified as "bad quality" signals.

To assess the binary classification performance of the algorithms, the mean annotators ratings were binarized by thresholding at a value of 3.5. Signal segments rated lower than this threshold were labeled as "bad quality" signals, while those rated equal to or higher than this threshold were labeled as "good quality" signals. Accuracy, sensitivity, specificity, precision, and F1-Score were computed as performance measures in the binary classification of the signal quality. We applied McNemar's binomial test [28] with a significance level  $\alpha=5\%$  to

assess whether the classification accuracies obtained for the different algorithms were significantly different from each other. We applied the Benjamini-Hochberg method [25] to correct for multiple comparisons.

### 2.3.2. Validation phase

In the validation phase, the performance of the SQI algorithm was assessed in the validation dataset in the same manner as performance was assessed in the training dataset. For the quantitative rating, this assessment included computing the estimated ratings of the SCI and PHOEBE algorithms for the validation dataset to allow a comparison with their performance. For this, the feature values of SCI and PHOEBE were converted to a continuous scale from 1 (very low quality) to 5 (very high quality) using the linear models fitted in the training phase (see Materials and Methods, section Quantitative rating performance). For the binary classification, we proceeded as explained in Materials and Methods, section Binary classification performance.

## 3. Result

### 3.1. Training phase

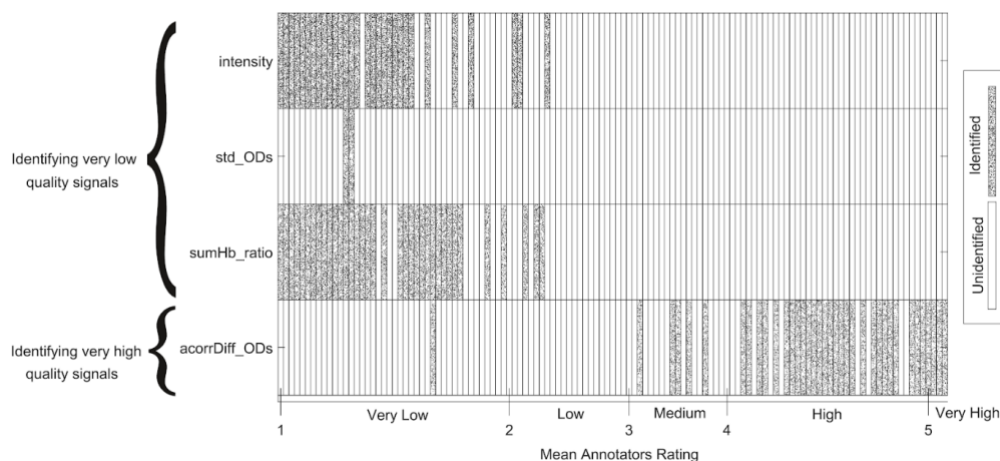
#### 3.1.1. Features correctness assessment

The here proposed fNIRS signal quality algorithm, SQI, assesses the quality of fNIRS signal segments in three consecutive rating stages: 1) identifying very low quality signals, 2) identifying very high quality signals and 3) signal quality rating (see section 2.2 for more details). Here, we assess the correctness of these features.

In rating stage one (identifying very low quality signals), the algorithm identifies very low quality signal segments based on three features: the absolute light intensity (abbreviated as *intensity*), the standard deviation of the ODs per wavelength (*std\_ODs*), and the ratio of oxygenated and deoxygenated hemoglobin summation (*sumHb\_ratio*). For each of these features, we applied an empirically derived threshold to identify signal segments having very low quality. Figure 6 shows the identification of each signal segment by these features. Each signal segment identified as a very low quality signal had a mean annotators rating indicating a low or very low signal quality. This means none of these features incorrectly identified medium, high, or very high quality signal segments as very low quality signals. Although the feature *std\_ODs* correctly identified signal segments with very low quality, its identification completely coincided with the identification of the features *intensity* and *sumHb\_ratio*. Moreover, the number of identified signal segments is lower than that of the other two features in this stage. Nonetheless, we decided to include this feature in the algorithm because of two reasons. Firstly, it is an intuitive feature that identifies signal segments having at least one optical density signal as a flatline, which is a clear sign of a very low quality signal. Secondly, it potentially avoids misclassification of signal segments in subsequent stages of the algorithm (see Appendix 1 for more details).

In rating stage two (identifying very high quality signals), the algorithm identifies very high quality signal segments based on one feature: the standard deviation of the difference between optical density signals autocorrelation (*acorrDiff\_ODs*). We applied an empirically derived threshold to identify signal segments with clear characteristics of very high quality

signals. Figure 6 shows the output of the feature *acorrDiff\_ODs* and its threshold. Note that we show all 123 signal segments, although the signal segments identified as very low quality signals in rating stage one would not enter rating stage two in the SQL algorithm. The figure shows that most of the signal segments identified as very high quality signals had mean annotators ratings of 4 (high quality) or higher. However, some medium quality signal segments as well as one very low quality signal segment were incorrectly identified as very high quality signals in this stage. As can be seen in Figure 6, this very low quality signal segment was identified as a very low quality signal by feature *sumHb\_ratio* in rating stage one. This means that in the SQL algorithm workflow, it would not enter rating stage two and would hence not be incorrectly classified by the algorithm.



**Figure 6:** Signal segments identified during rating stage one (identifying very low quality signals) and two (identifying very high quality signals) of the SQL algorithm. The y-axis shows the three features included in rating stage one at the top, and the feature included in rating stage two at the bottom. The features in rating stage one are: the absolute light intensity (abbreviated as *intensity*), the standard deviation of the ODs per wavelength (*std\_ODs*), and the ratio of oxygenated and deoxygenated hemoglobin summation (*sumHb\_ratio*). The feature in rating stage two is the standard deviation of the difference between optical density signals autocorrelation (*acorrDiff\_ODs*). The x-axis represents the mean annotators rating for each of the 123 signal segments. The signal segments were sorted in ascending order based on their mean annotators ratings. Cells in the plot filled with a speckle pattern represent the signal segments that were identified as either very low or very high quality signals in the respective rating stage by the feature. Cells left blank represent the signal segments that were not identified by the features in rating stage one and two.

Quantitative metrics evaluating the correctness of the features in rating stage one and two are reported in Table 1 (for more details, see Materials and Methods, section Features correctness assessment). Regarding the features included in rating stage one, the features *intensity* and *sumHb\_ratio* identified a greater number of signal segments as very low quality signals than the feature *std\_ODs*. Both features achieved a high precision in the identification (both above 94%). The feature *std\_ODs* identified less signal segments as having very low quality than the other two features, though with 100% precision. In rating stage two, feature *acorrDiff\_ODs* identified 37 signal segments as very high quality signals with a precision of 83.78%.

In rating stage three (signal quality rating), using the feature *logStdHb* (the logarithm of the ratio between the standard deviation of O<sub>2</sub>Hb and HHb), the SQL algorithm rates the signal quality for those signal segments which have not been identified in rating stage one and two. Figure 7 shows a scatter plot of this feature for all 123 signal segments, revealing a significant positive linear relationship ( $r^2 = 0.79$ ,  $p\text{-value} < 0.05$ ) between the raw values of the feature *logStdHb* and their mean annotators ratings. Signal segments that were identified by features

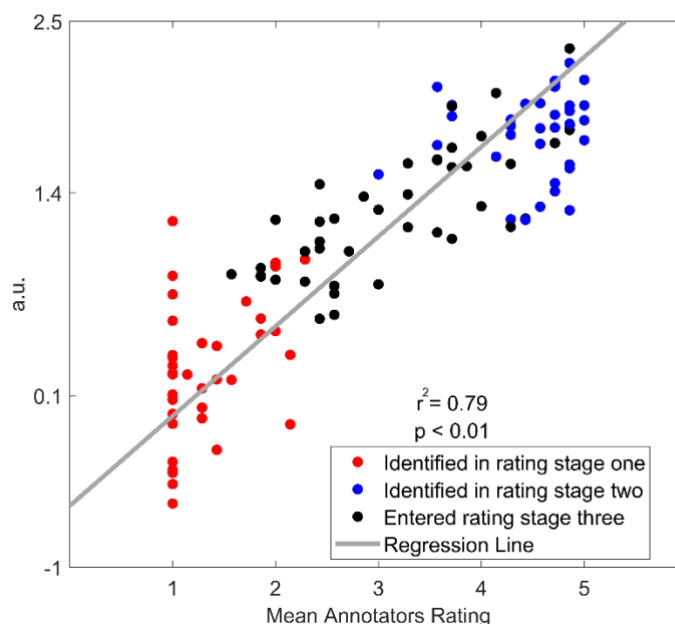


in rating stage one and two are shown in red (very low quality) and blue (very high quality), respectively.

**Table 1:** Quantitative metrics for correctness assessment of the proposed features identifying very low (*intensity*, *std\_ODs*, and *sumHb\_ratio*) and very high (*acorrDiff\_ODs*) quality signal segments in rating stage one and two of the SQI algorithm, respectively.

Feature	Precision (%)	Number of identified signal segments <sup>a</sup>	Number of correct decisions	Number of incorrect decisions
intensity	96.67	30	29	1
std_ODs	100	2	2	0
sumHb_ratio	94.44	36	34	2
acorrDiff_ODs	83.78	37	31	6

<sup>a</sup> Number of signal segments identified as very low quality signals (for features in rating stage one) and number of signal segments identified as very high quality signals (for the feature in rating stage two).



**Figure 7:** Scatter plot of the feature  $\logStdHb$  in rating stage three (signal quality rating) of the SQI algorithm. Raw values (y-axis) obtained for the feature  $\logStdHb$  (logarithm of the ratio between the standard deviation of O2Hb and HHb) with respect to the mean annotators ratings (x-axis) for each of the signal segments are depicted. Each dot represents one of the 123 signal segments included in the training dataset. Red dots represent signal segments that were identified as very low quality signals in rating stage one. Blue dots represent signal segments that were identified as very high quality signals in rating stage two. Black dots represent signal segments that were not identified as very low or very high quality signals in rating stage one and two and consequently entered rating stage three. The gray line represents the regression line obtained between the raw values for the feature  $\logStdHb$  and the mean annotators ratings.

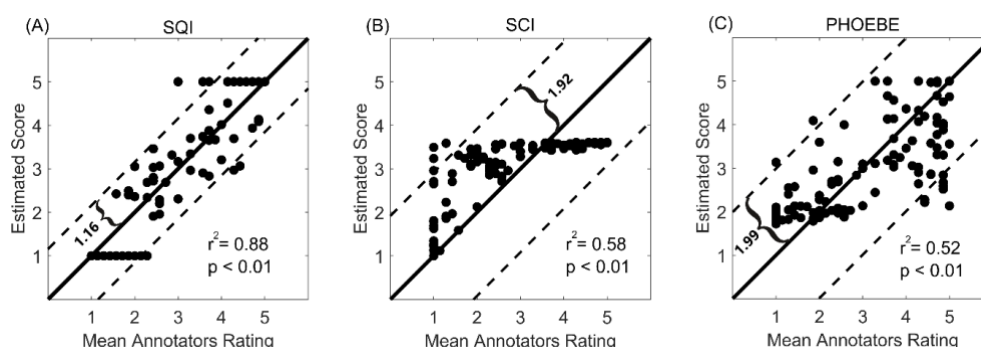
### 3.1.2. Algorithm performance assessment

#### 3.1.2.1. Quantitative rating performance

The raw values of the rating features were converted to a continuous scale from 1 to 5 (for more details, see Materials and Methods, section Quantitative rating performance) in order to

evaluate the performance of SQI, SCI and PHOEBE algorithms in quantitatively rating fNIRS signal quality. We refer to these as estimated ratings.

Figure 8 illustrates the scatter plots comparing the estimated ratings with the mean annotators ratings for SQI, SCI and PHOEBE algorithms. The proposed algorithm exhibits a better fit to the mean annotators ratings than PHOEBE and SCI, although a significant positive correlation can be observed for the three algorithms ( $p < 0.05$ ). While SQI explains 88% ( $p$ -value  $< 0.05$ ) of the variance, PHOEBE and SCI explain 52% and 58% ( $p$ -value  $< 0.05$ ) of the variance, respectively. Data points in the scatter plots for the PHOEBE and SCI algorithms are more sparsely distributed around the  $y=x$  line than for SQI. This is reflected by the higher Bland Altman limits of agreement obtained for PHOEBE (BLA=1.99) and SCI (BLA=1.92) with respect to SQI (BLA=1.16). Quantitative measures showing the similarity between the mean annotators ratings and estimated ratings for each of the algorithms are reported in Table 2. SQI had a lower standard deviation of error than SCI and PHOEBE. The mean of error of all considered algorithms was below 0.07. The computation time took the longest (53.11 ms) for the SQI algorithm, whereas for SCI and PHOEBE this time was approximately 20% and 60% of the SQI computation time, respectively.



**Figure 8:** Scatter plots for the estimated rating after the regression to a 1-5 continuous scale. The scatter plots show the estimated ratings (y axis) for each of the signal segments with respect to the mean annotators ratings (x axis). Each dot represents one of the 123 signal segments included in the training dataset. The full and dashed lines are, respectively, the  $y=x$  line and the Bland Altman limits of agreement. From left to right, the figure shows the scatter plots for: SQI (A), SCI (B) and PHOEBE (C).

**Table 2:** Quantitative measures for comparing the performance of the considered algorithms in quantitatively rating the fNIRS signal quality on the training dataset.

Method	ME <sup>a</sup>	StdE <sup>b</sup>	$r^2$	p-value of correlation <sup>c</sup>	Computation Time <sup>d</sup> (ms)
SCI	0.06	0.98	0.58	<0.01	10.77
PHOEBE	-0.004	1.01	0.52	<0.01	31.42
SQI	0.04	0.59	0.88	<0.01	53.11

<sup>a</sup> Mean of error.

<sup>b</sup> Standard deviation of error.

<sup>c</sup> P-values were corrected for multiple comparisons applying the Benjamini-Hochberg method.

<sup>d</sup> The computation time was calculated for a single signal segment using our reference implementation in MATLAB R2019b (MathWorks, Natick, Massachusetts) without considering the time to compute the linear regression.

### 3.1.2.2. Binary classification performance

We compared the binary classification performance of the three considered algorithms in classifying the signal segments into “bad quality” and “good quality” signals. To assign these

binary labels to the training dataset, the mean annotators ratings were thresholded using a threshold value of 3.5. The binary classification of fNIRS signal quality performed by the considered algorithms was conducted as explained in Materials and Methods, section Binary classification performance.

The binary classification performance results on the training dataset are reported in Table 3. The SQI, PHOEBE and SCI algorithms performed with an accuracy equal or higher than 65%, which is above chance level (50%). All three algorithms showed a high specificity and precision (above 92%), with SCI obtaining the highest for both measures. SQI showed the highest sensitivity (sensitivity = 92.86%) of all three algorithms, with PHOEBE and SCI showing a high (sensitivity = 78.57%) and low (sensitivity = 38.57%) sensitivity, respectively. The results of the statistical comparison of the binary performance of all algorithms against each other (see Table 4) exhibit that all three algorithms performed significantly differently from each other.

**Table 3:** Performance measures for comparing the performance of the considered algorithms in binary classification of the fNIRS signal quality on the training dataset.

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)
SCI	65.04	38.57	100	100	55.67
PHOEBE	84.55	78.57	92.45	93.22	85.27
SQI	92.68	92.86	92.45	94.20	93.53

**Table 4:** Z-scores and p-values calculated by applying McNemar's binomial test to compare the classification accuracies between each pair of algorithms with respect to the binarized mean annotators ratings on the training dataset. P-values were corrected for multiple comparisons applying the Benjamini-Hochberg method.

	SCI	PHOEBE	SQI
SCI	-	$z(123) = -4.15,$ $p < 0.01$	$z(123) = -5.21,$ $p < 0.01$
PHOEBE	$z(123) = -4.15,$ $p < 0.01$	-	$z(123) = -1.82,$ $p < 0.05$
SQI	$z(123) = -5.21,$ $p < 0.01$	$z(123) = -1.82,$ $p < 0.05$	-

## 3.2. Validation phase

We further assessed the performance of the SQI algorithm on the validation dataset in i) quantitative rating fNIRS signal quality and ii) binary classification of fNIRS signal quality. In both cases, we compared its performance with the performance of SCI and PHOEBE algorithms.

### 3.2.1. Quantitative rating performance

For all three algorithms, the raw values of the rating features were converted to the continuous scale from 1 to 5 (estimated rating) by using the linear models fitted on the training dataset. Quantitative measures showing the similarity between the mean annotators ratings and the estimated ratings for each of the algorithms are reported in Table 5. These results are

consistent with the results obtained for the training dataset. While SQI explained 85% (p-value < 0.05) of the variance, PHOEBE and SCI only explained 43% and 65% (p-value < 0.05), respectively. In addition, SQI showed a lower standard deviation of error than both SCI and PHOEBE algorithms. The mean of error (ME) for SQI was lower than 0.1 and was the lowest of all three algorithms.

**Table 5:** Quantitative measures for comparing the performance of the considered algorithms in quantitatively rating the fNIRS signal quality on the validation dataset.

Method	ME <sup>a</sup>	StdE <sup>b</sup>	r <sup>2</sup>	p-value of correlation <sup>c</sup>
SCI	-0.15	0.80	0.65	<0.01
PHOEBE	0.44	1.01	0.43	<0.01
SQI	-0.09	0.60	0.85	<0.01

<sup>a</sup> Mean of error.

<sup>b</sup> Standard deviation of error.

<sup>c</sup> P-values were corrected for multiple comparisons applying the Benjamini-Hochberg method.

### 3.2.2. Binary classification performance

We compared the binary classification performance of the three considered algorithms on the validation dataset in classifying the signal segments into “bad quality” and “good quality” signals. We binarized the output of the three considered algorithms as well as the mean annotators ratings as detailed in Materials and Methods, section Binary classification performance. The threshold used for the *logStdHb* feature in rating stage three of the SQI algorithm was the same as empirically derived in the training phase. The thresholds used for PHOEBE and SCI were those suggested in the original papers (see [17, 18]). The binary classification performance results are reported in Table 6. All considered algorithms performed with an accuracy equal or higher than 65%, which is above chance level (50%). However, the accuracy for the SQI algorithm (accuracy = 95%) was 30% higher than SCI (accuracy = 65%) and 20% higher than PHOEBE (accuracy = 75%). Although both PHOEBE and SCI obtained a 100% for both specificity and precision, they obtained low sensitivity values: 39.13% SCI and 56.52% PHOEBE. This shows that, despite correctly classifying all good quality signals, they did not perform well at classifying bad quality signals. This is reflected in the lower accuracy and F1-score obtained for both algorithms compared to SQI. Conversely, the SQI algorithm showed high values for the three metrics (specificity = 94.12%, precision = 95.65%, sensitivity = 95.65%).

**Table 6:** Performance measures for comparing the performance of the considered algorithms in binary classification of the fNIRS signal quality on the validation dataset.

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)
SCI	65	39.13	100	100	56.25
PHOEBE	75	56.52	100	100	72.22
SQI	95	95.65	94.12	95.65	95.65

We conducted a binomial test to determine whether the differences in performance of the considered algorithms were significant. The results for this test are reported in Table 7. The SQI algorithm performed significantly differently ( $p$ -value  $< 0.05$ ) from SCI and PHOEBE at binary classifying fNIRS signals quality, while SCI and PHOEBE did not perform significantly differently from each other. Furthermore, since 3 out of the total of 4 participants included in the validation dataset were also common to the training dataset (see section 2.1.2), in Appendix 2 we analyzed whether they introduced any bias in the classification.

**Table 7:** Z-scores and p-values calculated by applying McNemar’s binomial test to compare the classification accuracies between each pair of algorithms with respect to the binarized mean annotators ratings on the validation dataset. P-values were corrected for multiple comparisons applying the Benjamini-Hochberg method.

	SCI	PHOEBE	SQI
SCI	-	$z(40) = -1.53,$ $p > 0.05$	$z(40) = -2.76,$ $p < 0.01$
PHOEBE	$z(40) = -1.53,$ $p > 0.05$	-	$z(40) = -1.99,$ $p < 0.05$
SQI	$z(40) = -2.76,$ $p < 0.01$	$z(40) = -1.99,$ $p < 0.05$	-

## 4. Discussion

In this study, we developed a novel algorithm, Signal Quality Index (SQI), capable of quantitatively assessing the quality of fNIRS signals in a numeric scale from 1 (very low quality) to 5 (very high quality). To the best of our knowledge, the SQI algorithm is the first algorithm to quantitatively rate fNIRS signal quality in more than two quality levels. Current existing algorithms, like SCI [17] and PHOEBE [18], offer only a sharp distinction between two quality levels of good and bad signal quality. We found that, despite higher computation time, the SQI algorithm showed a significantly better performance ( $p < 0.05$ ) in quantitatively rating the fNIRS signal quality than PHOEBE and SCI (see Tables 2 and 5). SQI numerically showed a greater positive correlation ( $r^2 = 0.85$ ,  $p < 0.05$ ) between the estimated ratings and the mean annotators ratings than PHOEBE ( $r^2 = 0.43$ ,  $p < 0.05$ ) and SCI ( $r^2 = 0.65$ ,  $p < 0.05$ ). Furthermore, we found that SQI showed a significantly better performance in binary classification of the fNIRS signal quality ( $p < 0.05$ ) than PHOEBE and SCI, in terms of accuracy, sensitivity, and F1-score (see Tables 3 and 6).

The novelty of the here developed SQI algorithm is that, rather than using a single metric for directly rating the signal quality as SCI does, or a combination of two metrics as done by the PHOEBE algorithm, the SQI algorithm is composed of three different rating stages: identifying very low quality signals (rating stage one), identifying very high quality signals (rating stage two), signal quality rating (rating stage three). Each of these stages includes one or more features, which were selected in order to translate current heuristics in visual assessment of fNIRS signal quality into a numeric rating. The performance of the features in rating stage one and two on the training dataset (see Table 1) showed that they are well suited for identifying very low and very high quality signals with high precision ( $>80\%$ ). The performance results of the feature *logStdHb* in rating stage three demonstrated that it is well suitable for fNIRS signal

quality rating. This claim is supported by the significant positive linear relationship ( $r^2 = 0.79$ ,  $p\text{-value} < 0.05$ ) observed between the raw values of this feature and the mean annotators ratings (See Figure 7). Furthermore, including rating stage one and two in the algorithm improves its performance, since the features in these stages identify very low and very high quality signal segments that would otherwise be incorrectly rated by the feature *logStdHb* in rating stage three. This shows the efficacy of including three rating stages in the SQI algorithm workflow, outperforming the state-of-the-art algorithms SCI and PHOEBE.

In this study, the SCI and PHOEBE algorithms used for comparison were implemented as done by their authors (see [26]). We used the thresholds proposed in the original papers [17, 18] for binary classification. Though beyond the scope of this study, these thresholds could be optimized on the training dataset to improve their performance. Further studies could be conducted to evaluate the performance sensitivity of the SQI, SCI, and PHOEBE algorithms to different thresholds.

Binary approaches for signal quality assessment, such as SCI and PHOEBE, provide a sharp distinction between “bad” and “good” quality signals. Conversely, three or more levels of signal quality rating allow for the discrimination of “bad” and “good” quality signals into different levels of quality. In some experiments, recording “good” quality signals is very hard to achieve, e.g. when the signals are recorded from subjects or brain areas with a high density of dark hair. In such cases, although the signal might show a weak pulsatile heartbeat component because a great part of the transmitted light is absorbed by hair, the quality of such signal segments might still be considered sufficient for certain types of analysis. In a binary approach, these signal segments would either be classified as “bad” or “good” quality signals, depending on the sensitivity of the classifier. In a three or more levels signal quality approach, however, an intermediate quality rating for such signal segments allows the researcher to decide whether the achieved level of quality is sufficient for the analysis. Furthermore, the five-level scale provided in the SQI algorithm makes it possible to decrease the resolution of rating to either a three-level scale or a binary one, which is not possible for the other way around. Therefore, depending on the type of analysis that is to be conducted, the researcher can convert this five-level scale to fewer levels.

Although the results we obtained for the SQI algorithm are promising for both offline and online applications, further improvements are possible. We did not include a stage identifying and rating signal segments affected by motion, since it was beyond the scope of this study. Therefore, in the current work, the analysis of the algorithm performance was limited to data free of motion artifacts. However, motion artifacts are an important source of noise in fNIRS signals. These artifacts, often caused by head and body movements, are unavoidable especially in challenging subject groups like infants and in rehabilitation research and sport science, which often deal with subjects in motion [29-32]. Subject movement causes motion between the optodes and the scalp, leading to rapid shifts in the optical coupling. These rapid shifts take the form of transient spikes in fNIRS signals, and have a scale and frequency composition that are distinct from the background fNIRS signal [32].

The features included in the SQI algorithm use heuristics about meaningful physiological components and characteristics of fNIRS signals, therefore the presence of motion artifacts might affect the performance of the algorithm. We included a section in Appendix 3 where we compared the performance of the considered algorithms on a motion-corrupted dataset. We found that none of the algorithms performed well in quantitatively rating fNIRS signal quality. Therefore, to improve the performance of the SQI algorithm in the presence of motion, future

research should consider the effects of motion on fNIRS signal quality and include a stage for motion artefact recognition [29, 33].

In addition, both the training and validation of the SQI algorithm were conducted in a set of data collected from healthy young adults. Because functional and systemic-related hemodynamics depend on age [34, 35], further studies could be conducted to assess the SQI algorithm in a more extensive dataset, comprising data from infants and elderly people as well as from subjects with cardiac abnormalities.

In this study, we introduced an algorithm, SQI, which is capable of quantitatively rating and binary classifying fNIRS signal quality and outperforms current state-of-the-art algorithms. It can be used as an offline tool for identifying and rejecting channels and trials with a poor signal quality, as well as for online assessment of fNIRS signal quality during set-up and data acquisition. The SQI algorithm was designed to run over 10-second sliding windows. When applying the algorithm to either offline or online data, the user can compute the SQI values for every 10 seconds of data. This allows for an estimation of the varying signal quality of long recordings over time, making it possible to drop trials with poor signal quality rather than the channels themselves. We believe that the widespread use of this algorithm for reporting signal quality ratings for each of the measured channels would ensure that experiments are carried out in proper conditions for collecting actual hemodynamic information and reducing artefacts. Therefore, whether a novice or an experienced researcher conducts an experiment would no longer be an issue for guaranteeing the collection of good quality fNIRS data. The source code for the SQI algorithm is available at <https://github.com/Artinis-Medical-Systems-B-V/SignalQualityIndex>. The algorithm will be implemented in Artinis' fNIRS software OxySoft 3.3.

## 5. Conclusion

In this study we have developed an algorithm, signal quality index (SQI), to quantitatively assess NIRS signal quality in a numeric scale from 1 (very low quality) to 5 (very high quality). The results demonstrate the adequacy of the proposed algorithm for both binary and quantitatively rating the NIRS signal quality. The SQI algorithm performed better than SCI and PHOEBE, existing algorithms in the literature. The promising results obtained in this study suggest that the SQI algorithm could be exploited in different applications: from offline use after recording the signals to online use during recording or optodes setup.

## Appendix 1

Although the feature *std\_ODs* correctly identified signal segments with very low quality, its identification completely coincided with the identification of the features *intensity* and *sumHb\_ratio*. Moreover, the number of identified signal segments is very low compared with the number of signal segments identified by the other two features in this stage. Nonetheless, we decided to include this feature in the algorithm because of two reasons. Firstly, it is an intuitive feature that identifies signal segments having at least one optical density (OD) signal as a flatline, which is a clear sign of a very low quality signal. Secondly, it potentially avoids misclassification of signal segments in subsequent stages of the algorithm.



Misclassification of such signal segments could arise in cases where the amplitude of the flatline OD signal is within the permitted range for feature *intensity* in rating stage one. The signal segment would not be rated by this feature as a very low quality signal and would then be evaluated by feature *sumHb\_ratio*. This feature could fail at identifying the signal segment as a very low quality signal because of using concentration changes in O2Hb ( $\Delta[O_2Hb]$ ) and HHb ( $\Delta[HHb]$ ) signals, obtained from OD signals by means of the Modified Beer-Lambert Law (Eq. (4)). For illustrative purposes, we replaced in Eq. (4) the values for the molar extinction coefficients of each chromophore for two sample wavelengths used by the devices (1 =850 nm, 2=760 nm), the source detector distance ( $d = 3$  cm), and the differential pathlength factor (DPF = 6). This is simplified in Eq. (5). The resulting HHb and O2Hb signals are a linear combination of both OD signals, as shown in Eq. (6) and (7), respectively.

$$\begin{bmatrix} \Delta[HHb] \\ \Delta[O_2Hb] \end{bmatrix} = (d)^{-1} \begin{bmatrix} \varepsilon_{HHb,\lambda_1} & \varepsilon_{O_2Hb,\lambda_1} \\ \varepsilon_{HHb,\lambda_2} & \varepsilon_{O_2Hb,\lambda_2} \end{bmatrix}^{-1} \begin{bmatrix} \Delta OD(\Delta t, \lambda_1)/DPF(\lambda_1) \\ \Delta OD(\Delta t, \lambda_2)/DPF(\lambda_2) \end{bmatrix} \quad (4)$$

$$\begin{bmatrix} \Delta[HHb] \\ \Delta[O_2Hb] \end{bmatrix} = \begin{bmatrix} -23.2 & 44 \\ 63.6 & -29.9 \end{bmatrix} \begin{bmatrix} \Delta OD(\Delta t, \lambda_1) \\ \Delta OD(\Delta t, \lambda_2) \end{bmatrix} \quad (5)$$

$$\Delta[HHb] = -23.2 \times \Delta OD(\Delta t, \lambda_1) + 44 \times \Delta OD(\Delta t, \lambda_2) \quad (6)$$

$$\Delta[O_2Hb] = 63.6 \times \Delta OD(\Delta t, \lambda_1) - 29.9 \times \Delta OD(\Delta t, \lambda_2) \quad (7)$$

Feature *sumHb\_ratio* of rating stage one considers a scale ratio between HHb and O2Hb signals. If one of the OD signals is a flat line of constant amplitude, i.e. it has a zero standard deviation, the resulting O2Hb and HHb signals will be a factor of the OD signal having a non-zero standard deviation plus a constant value introduced by the OD signal with a zero standard deviation (see Eq. (6)). This constant value is removed from the O2Hb and HHb signals in the preprocessing stage. Depending on which of the OD signals is a flat line, the ratio between the filtered O2Hb and HHb signals will be either greater or lower than one. In the case in which this ratio is lower than one, the signal segment will be correctly identified as a very low quality signal by feature *sumHb\_ratio*. In the opposite case, however, it would enter subsequent stages of the algorithm, which could result in its misclassification. Therefore, considering that the computation time of the standard deviation is low, including the feature *std\_ODs* is necessary to guarantee a better performance of the SQI algorithm.

In the present study, the two signal segments having a standard deviation of zero were correctly identified as very low quality signals by feature *intensity*, because they had null amplitudes. They were also correctly identified by the feature *sumHb\_ratio*, because they are examples of the case where the ratio between O2Hb and HHb signals is lower than one. However, as explained above, this is not necessarily always the case.

## Appendix 2

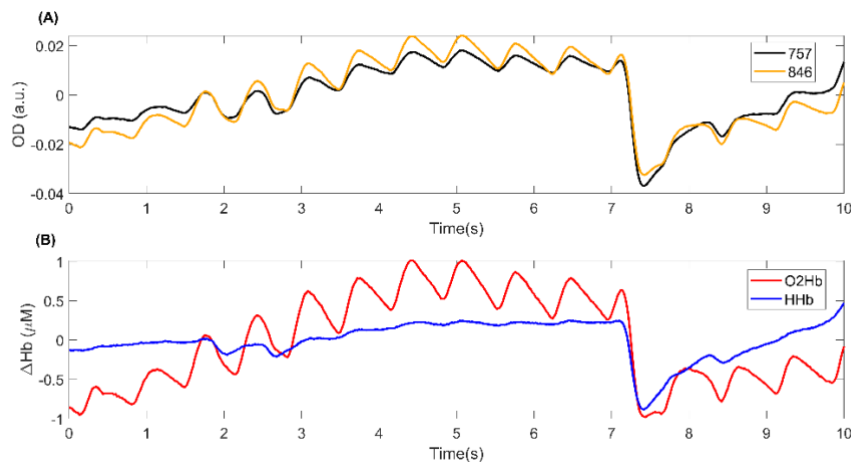
We checked whether the three participants common to the training and validation datasets introduced any bias in the classification. In order to do so, we split the validation dataset into two sets: one with only the common participants (set A) and the other with only the additional participant not present in the training dataset (set B). We compared the binary classification performance of the algorithms for each of these sets (see Table 8). We found that the performances of the algorithms were consistent for both sets of data and thus conclude that there was no bias in the classification.

**Table 8:** Accuracy for comparing the performance of the considered algorithms in binary classification of the fNIRS signal quality on two sets of the validation dataset: one with the participants common to both training and validation datasets (set A) and the other with only the additional participant not present in the training dataset (set B).

Method	Accuracy (%) for set A	Accuracy (%) for set B
SCI	63.3	70
PHOEBE	76	70
SQI	96.7	90

## Appendix 3

We further assessed the performance of the SQI algorithm on a validation dataset contaminated with motion artifacts, in both quantitatively rating and binary classifying fNIRS signal quality. We compared the performance of the SQI algorithm with the performance of SCI and PHOEBE. The dataset used comprised 35 10-second signal segments, which had been excluded in the training phase as signal segments containing motion (see section 2.1.1). Figure 9 shows one of the signal segments rated as having motion artifacts by all annotators in the training phase.



**Figure 9:** A sample signal segment contaminated with motion artifacts. This signal segment was rated as containing motion artifacts by all annotators in the training phase. (A) Orange and black curves represent the detrended optical density signals for the wavelengths 846 and 757 nm, respectively. (B) Red and blue curves represent the detrended changes in O2Hb and HHb concentrations, respectively.

### A3.1. Quantitative rating performance

For each of the three algorithms, estimated ratings in the 1-5 scale were obtained for the motion corrupted validation dataset by applying the linear models fitted on the training dataset (see Materials and Methods, section Quantitative rating performance). Quantitative measures showing the similarity between the mean annotators ratings and the estimated ratings are reported in Table 9. While PHOEBE explained 31% (p-value < 0.01) of the variance, SQI and SCI explained 20% and 11% (p-value < 0.05) of the variance, respectively. SCI and PHOEBE had a lower standard deviation of error than SQI. The mean of error of all considered algorithms was below 0.9, and was higher for SCI than for SQI and PHOEBE.

**Table 9:** Quantitative measures for comparing the performance of the considered algorithms in quantitatively rating the fNIRS signal quality on the motion-corrupted validation dataset.

Method	ME <sup>a</sup>	StdE <sup>b</sup>	r <sup>2</sup>	p-value of correlation <sup>c</sup>
SCI	0.87	1.06	0.11	<0.05
PHOEBE	-0.18	0.91	0.31	<0.01
SQI	0.03	1.39	0.20	<0.01

<sup>a</sup> Mean of error.

<sup>b</sup> Standard deviation of error.

<sup>c</sup> P-values were corrected for multiple comparisons applying the Benjamini-Hochberg method.

### A3.2. Binary classification performance

For the three considered algorithms, we compared the binary classification performance on the motion corrupted validation dataset as explained in Materials and Methods, section Binary classification performance. The binary classification performance results are reported in Table 10. Both SQI and PHOEBE algorithms performed with an accuracy higher than chance level (50%), while SCI performance was below chance level. Although SCI showed the highest specificity and precision (100%), its sensitivity and F1-score were very low, and were the lowest of the three algorithms.

**Table 10:** Performance measures for comparing the performance of the considered algorithms in binary classification of the fNIRS signal quality on the motion-corrupted validation dataset.

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)
SCI	28.57	7.41	100	100	13.79
PHOEBE	80	92.59	37.50	83.33	87.72
SQI	65.71	74.07	37.50	80	76.92

The results of the statistical comparison of the binary performance of all algorithms against each other are reported in Table 11. These results show that SQI and PHOEBE did not perform significantly differently from each other (p-value > 0.05). However, both algorithms performed significantly differently (p-value < 0.01) from SCI at binary classifying fNIRS signals quality.

**Table 11:** Z-scores and p-values calculated by applying McNemar's binomial test to compare the classification accuracies between each pair of algorithms with respect to the binarized mean annotators ratings on the motion-corrupted validation dataset. P-values were corrected for multiple comparisons applying the Benjamini-Hochberg method.

	SCI	PHOEBE	SQI
SCI	-	$z(35) = -2.94,$ $p < 0.01$	$z(35) = -2.33,$ $p < 0.01$
PHOEBE	$z(35) = -2.94,$ $p < 0.01$	-	$z(35) = -1.47,$ $p > 0.05$
SQI	$z(35) = -2.33,$ $p < 0.01$	$z(35) = -1.47,$ $p > 0.05$	-

## References

1. Chance, B.; Zhuang, Z.; UnAh, C.; Alter, C.; Lipton, L. Cognition-Activated Low-Frequency Modulation of Light Absorption in Human Brain. *Proceedings of the National Academy of Sciences of the United States of America* **1993**, *90*, 3770–3774, doi:10.1073/pnas.90.8.3770.
2. Hakimi, N.; Setarehdan, S.K. Stress Assessment by Means of Heart Rate Derived from Functional Near-Infrared Spectroscopy. *Journal of Biomedical Optics* **2018**, *23*, 1, doi:10.1117/1.JBO.23.11.115001.
3. Villringer, A.; Planck, J.; Hock, C.; Schleinkofer, L.; Dirnagl, U. Near Infrared Spectroscopy (NIRS): A New Tool to Study Hemodynamic Changes during Activation of Brain Function in Human Adults. *Neuroscience Letters* **1993**, *154*, 101–104, doi:10.1016/0304-3940(93)90181-J.
4. Quaresima; Ferrari A Mini-Review on Functional Near-Infrared Spectroscopy (FNIRS): Where Do We Stand, and Where Should We Go? *Photonics* **2019**, *6*, 87, doi:10.3390/photonics6030087.
5. Hakimi, N.; Jodeiri, A.; Mirbagheri, M.; Setarehdan, S.K. Proposing a Convolutional Neural Network for Stress Assessment by Means of Derived Heart Rate from Functional near Infrared Spectroscopy. *Computers in Biology and Medicine* **2020**, *121*, 103810, doi:10.1016/j.combiomed.2020.103810.
6. Boas, D.A.; Gaudette, T.; Strangman, G.; Cheng, X.; Marota, J.J.A.; Mandeville, J.B. The Accuracy of near Infrared Spectroscopy and Imaging during Focal Changes in Cerebral Hemodynamics. *NeuroImage* **2001**, *13*, 76–90, doi:10.1006/nimg.2000.0674.
7. Colier, W.N.J.M.; Quaresima, V.; Oeseburg, B.; Ferrari, M. Human Motor Cortex Oxygenation Changes Induced by Cyclic Coupled Movements of Hand and Foot. *Experimental Brain Research* **1999**, *129*, 457–461, doi:10.1007/s002210050913.
8. Scholkmann, F.; Kleiser, S.; Metz, A.J.; Zimmermann, R.; Mata Pavia, J.; Wolf, U.; Wolf, M. A Review on Continuous Wave Functional Near-Infrared Spectroscopy and Imaging Instrumentation and Methodology. *NeuroImage* **2014**, *85*, 6–27.
9. Hoshi, Y.; Tamura, M. Detection of Dynamic Changes in Cerebral Oxygenation Coupled to Neuronal Function during Mental Work in Man. *Neuroscience Letters* **1993**, *150*, 5–8, doi:10.1016/0304-3940(93)90094-2.
10. Kato, T.; Kamei, A.; Takashima, S.; Ozaki, T. Human Visual Cortical Function during Photic Stimulation Monitoring by Means of Near-Infrared Spectroscopy. *Journal of Cerebral Blood Flow and Metabolism* **1993**, *13*, 516–520, doi:10.1038/jcbfm.1993.66.
11. Pinti, P.; Tachtsidis, I.; Hamilton, A.; Hirsch, J.; Aichelburg, C.; Gilbert, S.; Burgess, P.W. The Present and Future Use of Functional Near-Infrared Spectroscopy (FNIRS) for Cognitive Neuroscience. *Annals of the New York Academy of Sciences* **2018**, *1464*, 5, doi:10.1111/nyas.13948.
12. Ehlis, A.C.; Schneider, S.; Dresler, T.; Fallgatter, A.J. Application of Functional Near-Infrared Spectroscopy in Psychiatry. *NeuroImage* **2014**, *85*, 478–488.

13. Tachtsidis, I.; Scholkmann, F. False Positives and False Negatives in Functional Near-Infrared Spectroscopy: Issues, Challenges, and the Way Forward. *Neurophotonics* **2016**, *3*, 031405, doi:10.1117/1.NPh.3.3.031405.
14. Huigen, E.; Peper, A.; Grimbergen, C.A. Investigation into the Origin of the Noise of Surface Electrodes. *Medical and Biological Engineering and Computing* **2002**, *40*, 332–338, doi:10.1007/BF02344216.
15. Caldwell, M.; Scholkmann, F.; Wolf, U.; Wolf, M.; Elwell, C.; Tachtsidis, I. Modelling Confounding Effects from Extracerebral Contamination and Systemic Factors on Functional Near-Infrared Spectroscopy. *NeuroImage* **2016**, *143*, 91–105, doi:10.1016/j.neuroimage.2016.08.058.
16. Orihuela-Espina, F.; Leff, D.R.; James, D.R.C.; Darzi, A.W.; Yang, G.Z. Quality Control and Assurance in Functional near Infrared Spectroscopy (FNIRS) Experimentation. *Physics in Medicine and Biology* **2010**, *55*, 3701–3724, doi:10.1088/0031-9155/55/13/009.
17. Pollonini, L.; Olds, C.; Abaya, H.; Bortfeld, H.; Beauchamp, M.S.; Oghalai, J.S. Auditory Cortex Activation to Natural Speech and Simulated Cochlear Implant Speech Measured with Functional Near-Infrared Spectroscopy. *Hearing Research* **2014**, *309*, 84–93, doi:10.1016/j.heares.2013.11.007.
18. Pollonini, L.; Bortfeld, H.; Oghalai, J.S. PHOEBE: A Method for Real Time Mapping of Optodes-Scalp Coupling in Functional near-Infrared Spectroscopy. *Biomedical Optics Express* **2016**, *7*, 5104, doi:10.1364/boe.7.005104.
19. Elgendi, M. Optimal Signal Quality Index for Photoplethysmogram Signals. *Bioengineering* **2016**, *3*, doi:10.3390/bioengineering3040021.
20. Sukor, J.A.; Redmond, S.J.; Lovell, N.H. Signal Quality Measures for Pulse Oximetry through Waveform Morphology Analysis. *Physiological Measurement* **2011**, *32*, 369–384, doi:10.1088/0967-3334/32/3/008.
21. Delpy, D.T.; Cope, M.; Van Der Zee, P.; Arridge, S.; Wray, S.; Wyatt, J. Estimation of Optical Pathlength through Tissue from Direct Time of Flight Measurement. *Physics in Medicine and Biology* **1988**, *33*, 1433–1442, doi:10.1088/0031-9155/33/12/008.
22. Ifeachor, E.C.; Jervis, B.W. *Digital Signal Processing: A Practical Approach*; Prentice Hall, Upper Saddle River, NJ, 2002;
23. Oostenveld, R.; Fries, P.; Maris, E.; Schoffelen, J.M. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience* **2011**, *2011*, doi:10.1155/2011/156869.
24. Huppert, T.J.; Diamond, S.G.; Franceschini, M.A.; Boas, D.A. HomER: A Review of Time-Series Analysis Methods for near-Infrared Spectroscopy of the Brain. *Applied Optics* **2009**, *48*, doi:10.1364/AO.48.00D280.
25. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **1995**, *57*, 289–300, doi:10.1111/j.2517-6161.1995.tb02031.x.
26. Pollonini, L.; Perry, J. PHOEBE (2020), Github Repository Available online: <https://bitbucket.org/lpollonini/phoebe/wiki/Home> (accessed on 1 September 2020).



27. Altman, D.G.; Bland, J.M. Measurement in Medicine: The Analysis of Method Comparison Studies †; 1983; Vol. 32;.
28. Dietterich, T.G. Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. *Neural Computation* **1998**, *10*, 1895–1923, doi:10.1162/089976698300017197.
29. Brigadoi, S.; Ceccherini, L.; Cutini, S.; Scarpa, F.; Scatturin, P.; Selb, J.; Gagnon, L.; Boas, D.A.; Cooper, R.J. Motion Artifacts in Functional Near-Infrared Spectroscopy: A Comparison of Motion Correction Techniques Applied to Real Cognitive Data. *NeuroImage* **2014**, *85*, 181–191, doi:10.1016/j.neuroimage.2013.04.082.
30. Pfeifer, M.D.; Scholkmann, F.; Labruyère, R. Signal Processing in Functional Near-Infrared Spectroscopy (fNIRS): Methodological Differences Lead to Different Statistical Results. *Frontiers in Human Neuroscience* **2018**, *11*, 641, doi:10.3389/fnhum.2017.00641.
31. Scholkmann, F.; Spichtig, S.; Muehlemann, T.; Wolf, M. How to Detect and Reduce Movement Artifacts in Near-Infrared Imaging Using Moving Standard Deviation and Spline Interpolation. *Physiological Measurement* **2010**, *31*, 649–662, doi:10.1088/0967-3334/31/5/004.
32. Cooper, R.J.; Selb, J.; Gagnon, L.; Phillip, D.; Schytz, H.W.; Iversen, H.K.; Ashina, M.; Boas, D.A. A Systematic Comparison of Motion Artifact Correction Techniques for Functional Near-Infrared Spectroscopy. *Frontiers in Neuroscience* **2012**, *6*, 147, doi:10.3389/fnins.2012.00147.
33. Scholkmann, F.; Spichtig, S.; Muehlemann, T.; Wolf, M. How to Detect and Reduce Movement Artifacts in Near-Infrared Imaging Using Moving Standard Deviation and Spline Interpolation. *Physiological Measurement* **2010**, *31*, 649–662, doi:10.1088/0967-3334/31/5/004.
34. Peng, T.; Ainslie, P.N.; Cotter, J.D.; Murrell, C.; Thomas, K.; Williams, M.J.A.; George, K.; Shave, R.; Rowley, A.B.; Payne, S.J. The Effects of Age on the Spontaneous Low-Frequency Oscillations in Cerebral and Systemic Cardiovascular Dynamics. *Physiological Measurement* **2008**, *29*, 1055–1069, doi:10.1088/0967-3334/29/9/005.
35. Safonova, L.P.; Michalos, A.; Wolf, U.; Wolf, M.; Hueber, D.M.; Choi, J.H.; Gupta, R.; Polzonetti, C.; Mantulin, W.W.; Gratton, E. Age-Related Changes in Cerebral Hemodynamics Assessed by near-Infrared Spectroscopy. *Archives of Gerontology and Geriatrics* **2004**, *39*, 207–225, doi:10.1016/j.archger.2004.03.007.

## CHAPTER 3

fNIRS signal quality estimation by means  
of a machine learning algorithm trained on  
morphological and temporal features

M. Sofia Sappia\*, Naser Hakimi\*, Liucija Svinkunaite,  
Thomas Alderliesten, Jörn M. Horschig, Willy N.J.M. Colier

\* shared first author

doi: <https://doi.org/10.1117/12.2587188>





## Abstract

Cortical hemodynamics are reliably estimated using functional near infrared spectroscopy (fNIRS) when the recorded signals have sufficient quality. A poor signal quality leads to false findings in the analysis. It is therefore important to assess the fNIRS signal quality prior to signal analysis. In this study, we developed a machine learning-based algorithm that quantitatively rates the fNIRS signal quality in a numeric scale from 1 (very low quality) to 5 (very high quality). The algorithm consists of three main stages: 1) preprocessing, 2) feature extraction, where 14 morphological and temporal features are extracted, 3) regression, where the signal quality is estimated using a Gaussian SVM regressor. The performance of the proposed algorithm was validated on a manually annotated fNIRS dataset and was compared with the performance of the state-of-the-art algorithms SCI (Scalp Coupling Index), PHOEBE (Placing Headgear Optodes Efficiently Before Experimentation), and SQI (Signal Quality Index). The results showed our proposed algorithm performed significantly better ( $p$ -value  $< 0.05$ ) than SCI, PHOEBE, and SQI in both quantitatively rating ( $r^2 = 0.76$ ) and binary rating (accuracy = 90%) fNIRS signal quality. The use of this algorithm for fNIRS signal quality assessment can help increase the reliability of statistical analyses and reduce misinterpretations.

# 1. Introduction

Functional near-infrared spectroscopy (fNIRS) is a non-invasive neuroimaging technique that monitors changes in brain hemodynamics (blood perfusion) and oxygenation [1–7]. fNIRS is commonly used to investigate brain activity not only in experimental conditions, but also in clinical settings, in which monitoring cerebral oxygenation can contribute to the identification of deficits in cerebral oxygenation [8]. However, brain monitoring can be highly compromised if low quality signals are collected, leading to misinterpretations and false findings in the analysis [9]. Therefore, it is essential to assess fNIRS signal quality prior to signal analysis.

To assess fNIRS signal quality, the presence of a strong cardiac component is often used as the main indicator of a good optode-scalp coupling [10–12]. A good optode-scalp coupling guarantees that near infrared light travels through both cerebral and extracerebral compartments. For reaching cerebral compartments, the distance between near-infrared transmitter and receiver has to be sufficient (~3 cm). Part of the light is absorbed by the hemoglobin chromophores in the tissue, while another part of the light is backscattered and measured by the receivers. The received light (fNIRS signals) contains functional and systemic hemodynamic information from both cerebral and extracerebral compartments. Among the systemic hemodynamic information, the most prominent component is the heartbeat, which manifests itself as a pulsatile component in the fNIRS signals. Therefore, heartbeat presence in the fNIRS signals indicates that optodes are well coupled to the scalp and that enough light has reached the brain and backscattered to the receiver. The more visible this pulsatile component is in the signal, the better the optode-scalp coupling and, hence, the signal quality. Extrinsic factors, such as loose connection between the optodes and the skin, hair or scalp properties can diminish this coupling [13], limiting the light pathway and thus reducing the quality of the fNIRS signals.

Three algorithms based on morphological characteristics of good quality signals have previously been proposed for fNIRS signal quality assessment: SCI (Scalp Coupling Index [10]), PHOEBE (Placing Headgear Optodes Efficiently Before Experimentation [11]), and SQI (Signal Quality Index [12]). SCI and PHOEBE were designed to binarily discriminate between good and bad quality signals. However, a binary classification is not always desirable as it provides only two quality levels and may not accurately reflect the different levels of a good quality signal. In challenging cases where light absorption in the brain is hindered due to extrinsic factors, very high quality signals are difficult to achieve. In these cases, signals having a quality level on the edge of the binary decision boundary can still contain functional information and be sufficient for certain types of analyses. A quantitative signal quality rating, therefore, provides more insight about different quality levels, which is beneficial for subsequent analyses and interpretations.

In a previous study [12], we proposed the signal quality index (SQI) algorithm to address this issue by providing a quantitative rating of 5 levels. We showed that SQI performed significantly better than SCI and PHOEBE in both binary classifying and quantitatively rating fNIRS signal quality. This algorithm was based on fNIRS experts' heuristics used in the visual quality assessment of fNIRS signal quality. We hypothesize that a better performance could be achieved by using a machine learning approach and incorporating more features. This approach allows us to leverage both human expertise and the mathematical capabilities of machine learning algorithms.

In this study we introduce a novel algorithm for fNIRS signal quality assessment, based on machine learning: Machine Learning Based Signal Quality Index (MLSQI). MLSQI is capable of assessing the fNIRS signal quality of 4-second signal segments, providing a quantitative rating on a scale from 1 (very low signal quality) to 5 (very high signal quality). MLSQI considers different morphological and temporal features in the fNIRS signals to determine the strength of the optodes coupling with the scalp, providing the fNIRS research community with an objective estimate of fNIRS signal quality and eliminating the researchers' subjective bias. We compare the performance of our proposed MLSQI algorithm with that of the currently existing algorithms, SCI, PHOEBE and SQI. We show that MLSQI outperforms these algorithms in quantitatively rating as well as in binary classifying the fNIRS signal quality.

## 2. Materials and methods

### 2.1. Data collection and annotation

We assembled two sets of data from previous recordings: a training dataset, to develop and fit the parameters of the algorithm; a validation dataset, to provide an unbiased evaluation of the performance of the algorithm on an unseen set of data.

#### 2.1.1. Training dataset

The training dataset consisted of 4-second signal segments extracted from the motion-free training dataset used in Sappia, Hakimi et al. [12]. These data consisted of fNIRS recordings of healthy adult volunteers collected with the following devices: OctaMon, Brite 23, Brite 24, and OxyMon (all by Artinis Medical Systems B.V., Elst, The Netherlands). Originally, the dataset comprised 123 10-second signal segments of optical densities measured at two different wavelengths and concentration changes in oxygenated hemoglobin (O<sub>2</sub>Hb) and deoxygenated hemoglobin (HHb), sampled at 50 Hz. The dataset was annotated by seven independent fNIRS experts working at Artinis Medical Systems B.V., who rated all 123 10-second signal segments on an ordinal scale ranging from 1 (very low quality) to 5 (very high quality). The 4-second signal segments used in this study were obtained by taking the initial and last 4 seconds of each 10-second signal segment, resulting in a total of 246 4-second signal segments. For each of these 4-second signal segments, the annotators ratings of the corresponding 10-second signal segment were used.

#### 2.1.2. Validation dataset

The validation dataset consisted of 4-second signal segments extracted from the validation dataset used in Sappia, Hakimi et al. [12], as well as from a new set of data assembled from previously recorded, unpublished data. The 4-second signal segments corresponding to the validation dataset used in Sappia, Hakimi et al. [12] were extracted in the same way as for the training dataset. The resulting validation dataset used in this study consisted of fNIRS recordings of healthy adult volunteers collected with the following devices: OctaMon, Brite 23, Brite 24, and OxyMon (all by Artinis Medical Systems B.V., Elst, The Netherlands). It comprised 123 4-second signal segments of optical densities measured at two different wavelengths, as well as HHb and O<sub>2</sub>Hb concentration changes. All signals were sampled at 50 Hz. Nine independent fNIRS experts working at Artinis Medical Systems B.V. rated the

signal quality of the signal segments on an ordinal scale ranging from 1 (very low quality) to 5 (very high quality).

## 2.2. Algorithm description

The MLSQI algorithm consists of three stages. In the first stage (preprocessing), the segmented light intensities are converted to optical densities (ODs) and concentration changes in O2Hb and HHb, which is followed by two preprocessing steps. In the second stage (feature extraction), 14 morphological and temporal features are extracted from the preprocessed ODs, O2Hb, and HHb signal segments. In the third stage (regression), an SVM regressor model is trained based on the extracted 14 features, giving as output a quantitative signal quality rating.

### 2.2.1. Preprocessing

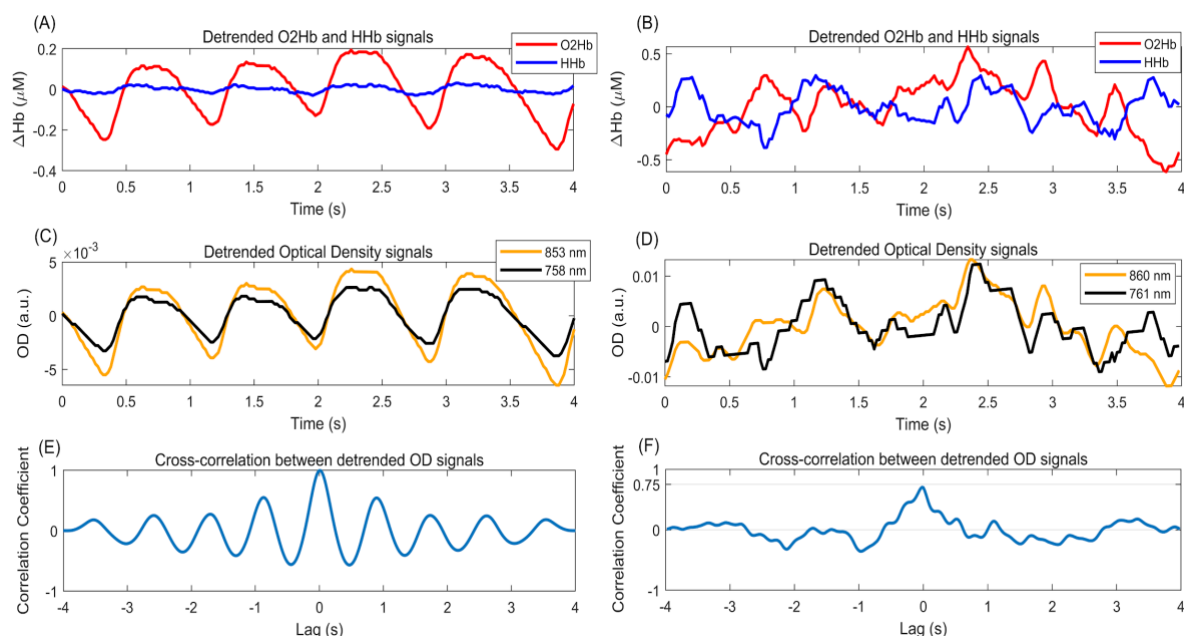
First, the received light intensities digitized with 16-bit resolution and segmented in 4-second windows were converted to optical densities (OD). We then converted the optical densities into O2Hb and HHb signals by applying the modified Beer-Lambert law [14]. Next, we applied two preprocessing steps to OD, O2Hb, and HHb signal segments. In the first step, we detrended the OD, O2Hb, and HHb signal segments by subtracting the least-squares fit of a straight line to the data. In the second step, we applied a band-pass filter on these signal segments using a zero-phase forward Hamming-windowed sinc FIR filter of order 208 with cutoff frequencies ( $-6$  dB) of 0.4 Hz and 3 Hz. The transition width of the filter was of 0.8 Hz, with a passband between 0.8 - 2.6 Hz, and a stopband ( $-53$  dB) between 3.4 - 25 Hz. The filter order was estimated based on the normalized transition width for the Hamming window [15]. This band-pass filter was implemented using the `ft_preproc_bandpassfilter` function in the FieldTrip toolbox [16], commit 62c9a0d on master branch. To avoid edge artefacts, prior to filtering, the signal segments were zero-padded with a sample length corresponding to two seconds of data both in the beginning and end of the signal segment.

### 2.2.2. Feature extraction

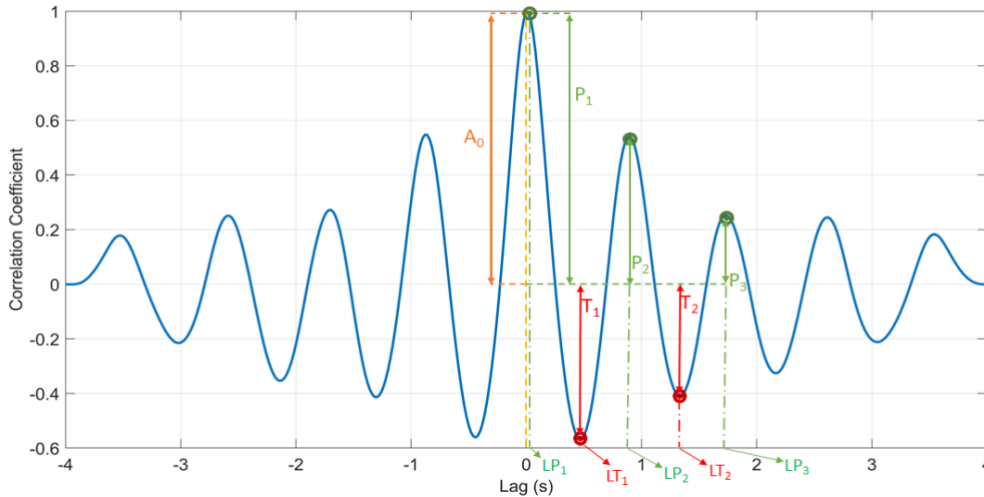
We extracted 14 morphological and temporal features based on the cardiac component present in fNIRS signals. The cardiac component manifests in fNIRS signals in the shape of semi-periodic pulses, of frequency within the heart rate range. The optical density signals OD1 (OD measured for the longer wavelength) and OD2 (OD measured for the shorter wavelength) capture this rhythmicity (see Figure 1). Hence, for high quality fNIRS signals, both OD1 and OD2 signals have similar shapes but with different amplitudes. Therefore, the normalized cross-correlation between the optical densities ( $OD\_corr$ ) from high quality fNIRS signals results in a damped sinusoidal signal, symmetrical around the y-axis (see Figure 1 (E)). The peaks of this normalized cross correlation signal correspond to the lags where the pulses of the shifted OD signal align with those of the fixed one. The troughs correspond to the lags where the peaks of one OD signal align with the troughs of the other. We extracted 12 features ( $A0 - LT2$ , see below) from the normalized cross correlation signal ( $OD\_corr$ ), which represent the pulsatility of the signal (see Figure 2). We used the  $OD\_corr$  signal from zero-lag onwards to extract these features. Two additional features were also extracted: one from the OD signals ( $sumOD\_ratio$ ) and the other from the O2Hb and HHb signals ( $Hb\_corr$ ). The extracted features are described in Table 1.

**Table 1:** The 14 features extracted in the feature extraction stage of the MLSQI algorithm from the normalized cross correlation ( $OD\_corr$ ), optical densities, and O2Hb and HHb signals.

Feature	Description
$A0$	Amplitude of $OD\_corr$ at lag zero
$P1$	Amplitude of the first peak in $OD\_corr$ signal
$LP1$	Lag of the first peak in $OD\_corr$ signal
$P2$	Amplitude of the second peak in $OD\_corr$ signal
$LP2$	Lag of the second peak in $OD\_corr$ signal
$P3$	Amplitude of the third peak in $OD\_corr$ signal
$LP3$	Lag of the third peak in $OD\_corr$ signal
$T1$	Amplitude of the trough between the first and second peaks in $OD\_corr$ signal
$LT1$	Lag of the trough between the first and second peaks in $OD\_corr$ signal
$T2$	Amplitude of the trough between the second and third peaks in $OD\_corr$ signal
$LT2$	Lag of the trough between the second and third peaks in $OD\_corr$ signal
$hr$	This feature considers the heart rate information in $OD\_corr$ signal. To compute this feature, we first obtained the lag intervals between the second and first peaks ( $LP2 - LP1$ ) and between the third and second peaks ( $LP3 - LP2$ ) and converted them to frequency in Hertz by inverting them. Then, we computed the estimated heart rate in beats per minute (BPM) by taking the mean of both frequencies multiplied by 60.
$sumOD\_ratio$	This feature considers the difference in scale between both optical density signals. To compute this feature, the ratio between the sum over samples of the absolute value of OD1 (optical density for the longer wavelength) and the sum over samples of the absolute value of OD2 (optical density for the shorter wavelength) was taken: $sumOD\_ratio = \sum OD1  / \sum OD2 $ , where $ x $ represents the absolute value of $x$ .
$Hb\_corr$	Pearson correlation coefficient between the O2Hb and HHb signals

**Figure 1:** Two sample signal segments, one having a very high quality (left column) and the other having a very low quality (right column), according to the ratings of all annotators. (A, B) Red and blue curves represent the detrended changes in O2Hb and HHb concentrations, respectively. (C, D) Orange and black curves represent the detrended optical density signals for the longer and shorter wavelengths, respectively. (E, F) These curves represent the cross-correlation signals calculated between the detrended optical density signals.





**Figure 2:** Cross correlation (OD\_corr) between the detrended optical density signals, and 11 of the 12 features extracted from it. Feature A0 corresponds to the amplitude of OD\_corr at lag zero. Features P1, P2, and P3 correspond to the amplitude of the first, second, and third peaks in the OD\_corr signal, respectively; while features LP1, LP2, and LP3 correspond to the lags of the first, second, and third peaks in the OD\_corr signal, respectively. Analogously, features T1, T2, and T3 correspond to the amplitude of the first, second, and third troughs in the OD\_corr signal, respectively; while features LT1, LT2, and LT3 correspond to the lags of the first, second, and third troughs in the OD\_corr signal, respectively.

### 2.2.3. Regression

After feature extraction, we trained a gaussian SVM regressor on the training dataset to estimate fNIRS signal quality, giving a quantitative signal quality rating between 1 (very low signal quality) and 5 (very high signal quality) as output. We used MATLAB's *fitrsvm* function to build and train the model, with parameter box constraint  $C = 0.7413$ , epsilon margin  $\epsilon = 0.0741$ , and kernel scale  $k = 3.7$ . All features were standardized to have zero mean and unit variance.

## 2.3. Performance assessment

We assessed the performance of the MLSQI algorithm in: 1) quantitatively rating the signal quality between 1 (very low signal quality) and 5 (very high signal quality); 2) binary classifying the signal quality (good or bad signal quality). We compared its performance with the performance of three existing fNIRS signal quality algorithms: SCI, PHOEBE, and SQI. All computation and analysis in this study were done using MATLAB R2019b (MathWorks, Natick, Massachusetts), on an ASUS workstation with Intel Core-i7-8565U @ 1.99 GHz CPU and 16 GB RAM.

SCI quantifies the similarity of the cardiac component in both optical densities to determine the strength of the coupling between the scalp and the optodes. The algorithm computes the zero-lag cross-correlation between both optical density signal segments as a quantitative measure of the signal-to-noise ratio of the signal segment. According to the original implementation of the SCI algorithm [10], signal segments with a zero-lag cross-correlation value above 0.75 are considered good quality signals, and all other signal segments are considered bad quality signals. The PHOEBE algorithm evaluates the similarity of both optical density signal segments by means of the SCI metric, as well as the spectral power of their cross-correlation to determine the strength of the cardiac component. According to the original implementation of the PHOEBE algorithm [11], signal segments with a zero-lag cross-

correlation value above 0.75 as well as a spectral value above 0.1 are considered good quality signals, otherwise they are considered bad quality signal segments. Both SCI and PHOEBE algorithms were designed for binary classifying fNIRS signal quality into good and bad quality signals. The SQI algorithm, in turn, is capable of quantitatively assessing the quality of fNIRS signals on a numeric scale from 1 (very low quality) to 5 (very high quality). It is composed of three different rating stages: identifying very low quality signals, identifying very high quality signals, and signal quality rating. Each of these stages includes one or more features that translate heuristics obtained by visual assessment of fNIRS signal quality into a numeric rating. All three algorithms were computed as implemented by the authors of the original papers [10–12].

### 2.3.1. Quantitative rating performance

To assess the performance of the MLSQI algorithm in quantitatively rating fNIRS signal quality, we compared the ratings obtained for the algorithm in the 1-5 scale and contrasted them with the mean annotators' ratings. We compared its performance with that of the SCI, PHOEBE, and SQI algorithms. These three algorithms were compared in a similar way as described in Sappia, Hakimi et al. [12]. The linear model for each of the algorithms was built on the training dataset used in this study, by computing a least squares regression between the mean annotators' ratings and the raw values of the rating feature. The ratings obtained for each of the algorithms in a 1-5 scale will be hereon referred to as estimated ratings.

Bland Altman limits of agreement (BLA [17]) were computed between the mean annotators ratings and the estimated ratings derived from the three algorithms: SCI, PHOEBE, SQI, and MLSQI. Bland Altman limits of agreement were computed by calculating 1.96 times the standard deviation of the error, representing the range in which 95% of the differences between the mean annotators ratings and estimated ratings fell. Three quantitative measures were calculated to further compare the performance of the evaluated algorithms: mean of error (ME), standard deviation of error (StdE), and coefficient of determination ( $r^2$ ) assessed by Pearson's correlation coefficient. To test for the significance of the correlation, we applied a significance level  $\alpha=5\%$  and computed the p-value using a Student's t-distribution as implemented in Matlab's *corr* function. In addition to the three quantitative measures, the computation time using our reference implementation in MATLAB R2019b (MathWorks, Natick, Massachusetts) for a single signal segment was measured five times and then averaged to compare the time required for each of the algorithms to obtain the estimated rating.

### 2.3.2. Binary classification performance

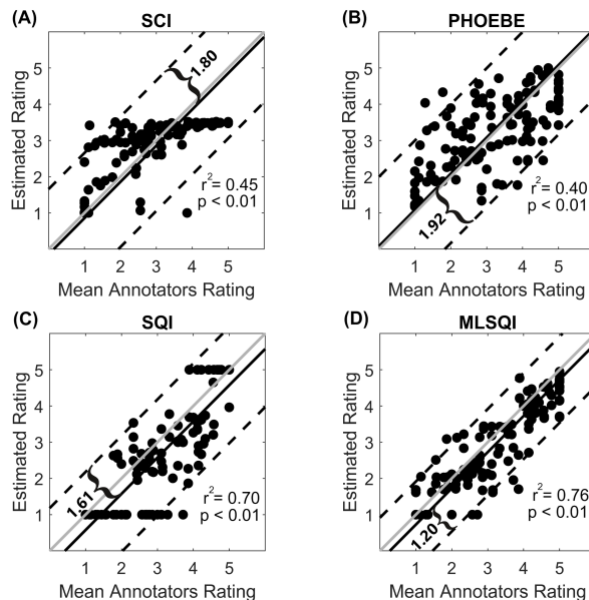
To conduct binary classification for the proposed MLSQI algorithm, we thresholded the quantitative ratings at a value of 3.5 (see Sappia, Hakimi et al. [12]). Signal segments with a value equal or higher than this threshold was classified as "good quality" signals. They were otherwise classified as "bad quality" signals. The same threshold was applied to the annotators ratings in order to obtain the reference binary ratings. For SCI, PHOEBE, and SQI, we used the thresholds proposed in the original papers [10–12]: a zero-lag cross-correlation value of 0.75 for the SCI algorithm, a peak spectral power of 0.1 for the PHOEBE algorithm, and a threshold of 1.478 for the rating feature in the signal quality rating stage for the SQI algorithm. For these algorithms, signal segments with values equal or higher than the considered thresholds were classified as "good quality" signals. They were otherwise classified as "bad quality" signals.

Accuracy, sensitivity, specificity, precision, and F1-Score were computed as performance measures in the binary classification of the signal quality. We applied McNemar's binomial test [18] with a significance level  $\alpha=5\%$  to assess whether the classification accuracies obtained for the different algorithms were significantly different from each other. We applied the Benjamini-Hochberg method [19] to correct for multiple comparisons.

## 3. Results

### 3.1. Quantitative rating performance

Figure 3 illustrates the scatter plots for SCI, PHOEBE, SQI and MLSQI algorithms, comparing their ratings with the mean annotators' ratings. MLSQI shows a better fit to the mean annotators' ratings than SCI, PHOEBE, and SQI, although a significant positive correlation can be observed for all algorithms (all  $p < 0.01$ ). MLSQI explains a higher percentage of variance, with a value of 76% in contrast to 45% for SCI, 40% for PHOEBE, and 70% for SQI. Data points in the scatter plots for the SCI, PHOEBE and SQI algorithms are more sparsely distributed around the  $y = x$  line than for MLSQI. This is reflected by the higher Bland Altman limits of agreement obtained for SCI (BLA=1.80), PHOEBE (BLA=1.92), and SQI (BLA=1.61) with respect to MLSQI (BLA=1.20). Quantitative measures showing the similarity between the mean annotators' ratings and estimated ratings for each of the algorithms are reported in Table 2. MLSQI showed a lower standard deviation of error (0.61) than SCI (0.92), PHOEBE (0.98), and SQI (0.82). The absolute mean of error of all considered algorithms was below 0.5, with PHOEBE obtaining the lowest value (0.08). The computation time took the longest (107 ms) for the MLSQI algorithm, approximately ten, five, and two times as much as for the SCI, PHOEBE, and SQI algorithms, respectively.



**Figure 3:** Scatter plots for the estimated ratings in the 1-5 continuous scale. The scatter plots show the estimated ratings (y axis) for each of the signal segments with respect to the mean annotators' ratings (x axis). Each dot represents a single signal segment in the validation dataset. The full gray and black lines represent the  $y=x$  line and mean of error, respectively. The dashed lines represent the Bland Altman limit of agreements. The figure shows the scatter plots for: SCI (A), PHOEBE (B), SQI (C), and MLSQI (D).

**Table 2:** Quantitative measures for comparing the performance of the considered algorithms in quantitatively rating fNIRS signal quality on the validation dataset.

Method	ME <sup>a</sup>	StdE <sup>b</sup>	r <sup>2</sup>	p-value of correlation <sup>c</sup>	Computation Time <sup>d</sup> (ms)
SCI	-0.14	0.92	0.45	<0.01	9.58
PHOEBE	0.08	0.98	0.40	<0.01	20.63
SQI	-0.43	0.82	0.70	<0.01	49.72
MLSQI	-0.28	0.61	0.76	<0.01	107.21

<sup>a</sup> Mean of error.

<sup>b</sup> Standard deviation of error.

<sup>c</sup> P-values were corrected for multiple comparisons applying the Benjamini-Hochberg method [19].

<sup>d</sup> The computation time was calculated for a single signal segment using our reference implementation in MATLAB R2019b (MathWorks, Natick, Massachusetts) without considering the time for the training of the model.

## 3.2. Binary classification performance

We compared the performance of the four considered algorithms on the validation dataset in binary classifying the signal segments into “bad quality” and “good quality” signals. We binarized the outputs of the algorithms as well as the mean annotators ratings, as detailed in Materials and Methods (section 2.3.2). The binary classification performance results are reported in Table 3. All considered algorithms performed significantly differently from each other in binary classifying fNIRS signal quality. All considered algorithms performed with an accuracy higher than chance level (50%). The accuracy for the MLSQI algorithm (accuracy = 90.2%) was approximately 30% higher (p-value < 0.01) than SCI (accuracy = 58.5%), 25% higher (p-value < 0.01) than PHOEBE (accuracy = 64.2%) and 5% higher (p-value < 0.05) than SQI (accuracy = 85.4%). Although both PHOEBE and SCI obtained values higher than or equal to 95% for both specificity and precision, they obtained low sensitivity values: 27.5% for SCI and 37.7% for PHOEBE. This shows that, despite correctly classifying the majority of good quality signals, they did not perform well at classifying bad quality signals. This is reflected in the lower accuracy and F1-score obtained for both algorithms compared to MLSQI and SQI. Conversely, both MLSQI and SQI algorithms showed high values for the three metrics, with MLSQI obtaining the highest. The results of McNemar’s binomial test, conducted to test the significance of the performance difference between the algorithms, are reported in Table 4. All considered algorithms performed significantly differently (p-value < 0.05) from each other at binary classifying fNIRS signal quality.

**Table 3:** Performance measures for comparing the performance of the considered algorithms in binary classification of fNIRS signal quality on the validation dataset.

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)
SCI	58.5	27.5	98.2	95	42.7
PHOEBE	64.2	37.7	98.2	96.3	54.2
SQI	85.4	95.7	72.2	81.5	88
MLSQI	90.2	100	77.8	85.2	92

**Table 4:** Z-scores and p-values calculated by applying McNemar’s binomial test to compare the classification accuracies between each pair of algorithms with respect to the binarized mean annotators ratings on the validation dataset. P-values were corrected for multiple comparisons applying the Benjamini-Hochberg method.

	SCI	PHOEBE	SQI	MLSQI
SCI	-	$z(123) = -2.35,$ $p < 0.01$	$z(123) = -3.98,$ $p < 0.01$	$z(123) = -4.65,$ $p < 0.01$
PHOEBE	$z(123) = -2.35,$ $p < 0.01$	-	$z(123) = -3.27,$ $p < 0.01$	$z(123) = -4.05,$ $p < 0.01$
SQI	$z(123) = -3.98,$ $p < 0.01$	$z(123) = -3.27,$ $p < 0.01$	-	$z(123) = -1.76,$ $p < 0.05$
MLSQI	$z(123) = -4.65,$ $p < 0.01$	$z(123) = -4.05,$ $p < 0.01$	$z(123) = -1.76,$ $p < 0.05$	-

## 4. Discussion

In this study, we showed the performance of our proposed machine learning based algorithm, MLSQI, for quantitatively rating fNIRS signal quality on a continuous scale from 1 (very low signal quality) to 5 (very high signal quality). The proposed algorithm relies on 14 temporal and morphological features extracted from the fNIRS signals, which are fed into a medium gaussian SVM regressor to provide a quantitative rating. We compared its performance with three state-of-the-art algorithms for fNIRS signal quality assessment: SCI, PHOEBE and SQI. We found that MLSQI significantly ( $p < 0.05$ ) outperformed all three algorithms in both quantitatively and binary rating fNIRS signal quality. In quantitative rating (see Table 2), MLSQI explained a higher percentage of variance ( $r^2 = 0.76$ ,  $p < 0.01$ ) than SQI ( $r^2 = 0.70$ ,  $p < 0.01$ ), SCI ( $r^2 = 0.45$ ,  $p < 0.01$ ), and PHOEBE ( $r^2 = 0.40$ ,  $p < 0.01$ ). In binary rating (see Tables 3 and 4), MLSQI performed significantly better ( $p < 0.05$ ) than SQI, SCI, and PHOEBE, in terms of accuracy (90.2%), sensitivity (100%), and F1-score (92%).

Different features extracted from fNIRS signals have been previously introduced for translating current heuristics in the visual assessment of fNIRS signal quality [10–12]. The morphological and temporal characteristics of the cardiac pulses in fNIRS signals provide useful information about their quality level. Whereas the ideal scenario would be to summarize this information in only one metric, as proposed by Pollonini et al. [10] for the SCI algorithm, extraction of different features for determining the quality of the signals has shown to improve the performance of the algorithms [11,12]. In this study, we extracted 14 features from fNIRS signals. These features reflect the strength of the cardiac pulses in the fNIRS signals and their stability in the short time window we considered (4 seconds). Our results showed that the combination of this greater number of features with a machine learning algorithm significantly outperforms algorithms which are based on fewer features.

In a previous study [12], we proposed the SQI algorithm for quantitatively rating the fNIRS signal quality of 10-second signal segments. In some cases, however, it is desirable to assess the quality of shorter signal segments in order to provide faster feedback. Therefore, in this study we proposed the MLSQI algorithm, which estimates the signal quality of 4-second signal segments. However, the computation time for the MLSQI algorithm (107 ms) is approximately twice as long as for the SQI algorithm (50 ms) because MLSQI extracts and uses more features than SQI. As a consequence, this higher computation time makes MLSQI less

suitable than SQI in applications where computation time is critical. It is worth mentioning that the computation time of MLSQI was estimated based on an unoptimized MATLAB implementation, which is considerably slower than an optimized implementation in a modern programming language. Furthermore, computation of many segments can be done independently from each other, for example by implementing the algorithm in a parallel computing system.

The MLSQI algorithm showed promising results for both offline and online implementations, but further validation is required using more extensive datasets. The datasets used in this study only included data from healthy young adults. However, fNIRS brain monitoring is also important in clinical applications such as the identification of cerebral ischemia [20], the study of atypical development on executive and language dysfunctions, and prognosis of neurodevelopmental disorders including autism and attention-deficit hyperactivity disorder [21–24]. Therefore, a dataset comprising signals from infants and older adults as well as from subjects with brain abnormalities is needed for further validation.

In addition, all algorithms evaluated in this study focus solely on long channel fNIRS signals (~ 3 - 3.5 cm source-detector separation). However, the use of short channels (~1.5 cm source-detector separation) is becoming increasingly popular in fNIRS experiments [25–28] and, to the best of our knowledge, no quality measures are currently available for them. Therefore, fNIRS signal quality measures should be further developed and validated in sets of data comprising both short and long channels. Furthermore, in the present study, we trained and evaluated the algorithm in signals free of motion. Further work should be conducted regarding motion recognition, since the main applicability of the proposed algorithm is for online use, and it might misperform in the presence of motion artifacts.

Moreover, further studies should be conducted on a dataset where signal segments are rated individually. Part of the 4-second signal segments in the training and validation datasets were extracted from 10-second signal segments included in a previously published set of data. We assigned the same ratings given to each 10-second signal segment to the two corresponding 4-second signal segments we extracted from it. We assumed that the quality does not greatly vary in the short 10-second time window. However, slight variations could occur in this period and, therefore, evaluating this algorithm on a set of data where signal segments are rated individually would provide more accurate results.

This study focused on the development of the MLSQI algorithm, which quantitatively rates fNIRS signal quality for 4-second signal segments, and outperformed current state-of-the-art algorithms: SCI, PHOEBE, and SQI. This algorithm not only helps in properly setting up fNIRS optodes prior to conducting experiments in online applications, but it can also be used for assessing fNIRS signal quality in offline applications. The widespread use of such an objective fNIRS quality assessment algorithm will contribute to eliminating researchers' subjective bias and to collecting good quality data. This will reduce misinterpretations and false findings in the analysis of fNIRS signals.

## References

1. Chance, B.; Zhuang, Z.; UnAh, C.; Alter, C.; Lipton, L. Cognition-Activated Low-Frequency Modulation of Light Absorption in Human Brain. *Proc Natl Acad Sci U S A* **1993**, *90*, 3770–3774, doi:10.1073/pnas.90.8.3770.
2. Villringer, A.; Planck, J.; Hock, C.; Schleinkofer, L.; Dirnagl, U. Near Infrared Spectroscopy (NIRS): A New Tool to Study Hemodynamic Changes during Activation of Brain Function in Human Adults. *Neurosci Lett* **1993**, *154*, 101–104, doi:10.1016/0304-3940(93)90181-J.
3. Colier, W.N.J.M.; Quaresima, V.; Oeseburg, B.; Ferrari, M. Human Motor Cortex Oxygenation Changes Induced by Cyclic Coupled Movements of Hand and Foot. *Exp Brain Res* **1999**, *129*, 457–461, doi:10.1007/s002210050913.
4. Ehlis, A.C.; Schneider, S.; Dresler, T.; Fallgatter, A.J. Application of Functional Near-Infrared Spectroscopy in Psychiatry. *Neuroimage* **2014**, *85*, 478–488.
5. Hakimi, N.; Setarehdan, S.K. Stress Assessment by Means of Heart Rate Derived from Functional Near-Infrared Spectroscopy. *J Biomed Opt* **2018**, *23*, 1, doi:10.1117/1.JBO.23.11.115001.
6. Hakimi, N.; Jodeiri, A.; Mirbagheri, M.; Setarehdan, S.K. Proposing a Convolutional Neural Network for Stress Assessment by Means of Derived Heart Rate from Functional near Infrared Spectroscopy. *Comput Biol Med* **2020**, *121*, 103810, doi:10.1016/j.compbiomed.2020.103810.
7. Pfeifer, M.D.; Scholkmann, F.; Labruyère, R. Signal Processing in Functional Near-Infrared Spectroscopy (FNIRS): Methodological Differences Lead to Different Statistical Results. *Front Hum Neurosci* **2018**, *11*, 641, doi:10.3389/fnhum.2017.00641.
8. Ferrari, M. Progress of Near-Infrared Spectroscopy and Topography for Brain and Muscle Clinical Applications. *Journal of Biomedical Optics* **2007**, *12*, 062104, doi:10.1117/1.2804899.
9. Tachtsidis, I.; Scholkmann, F. False Positives and False Negatives in Functional Near-Infrared Spectroscopy: Issues, Challenges, and the Way Forward. *Neurophotonics* **2016**, *3*, 031405, doi:10.1117/1.nph.3.3.031405.
10. Pollonini, L.; Olds, C.; Abaya, H.; Bortfeld, H.; Beauchamp, M.S.; Oghalai, J.S. Auditory Cortex Activation to Natural Speech and Simulated Cochlear Implant Speech Measured with Functional Near-Infrared Spectroscopy. *Hear Res* **2014**, *309*, 84–93, doi:10.1016/j.heares.2013.11.007.
11. Pollonini, L.; Bortfeld, H.; Oghalai, J.S. PHOEBE: A Method for Real Time Mapping of Optodes-Scalp Coupling in Functional near-Infrared Spectroscopy. *Biomed Opt Express* **2016**, *7*, 5104, doi:10.1364/boe.7.005104.
12. Sappia, M.S.; Hakimi, N.; Colier, W.N.J.M.; Horschig, J.M. Signal Quality Index: An Algorithm for Quantitative Assessment of Functional near Infrared Spectroscopy Signal Quality. *Biomedical Optics Express* **2020**, *11*, 6732, doi:10.1364/boe.409317.
13. Orihuela-Espina, F.; Leff, D.R.; James, D.R.C.; Darzi, A.W.; Yang, G.Z. Quality Control and Assurance in Functional near Infrared Spectroscopy (FNIRS) Experimentation.



- Physics in Medicine and Biology* **2010**, *55*, 3701–3724, doi:10.1088/0031-9155/55/13/009.
14. Delpy, D.T.; Cope, M.; Van Der Zee, P.; Arridge, S.; Wray, S.; Wyatt, J. Estimation of Optical Pathlength through Tissue from Direct Time of Flight Measurement. *Phys Med Biol* **1988**, *33*, 1433–1442, doi:10.1088/0031-9155/33/12/008.
  15. Ifeachor, E.C.; Jervis, B.W. *Digital Signal Processing: A Practical Approach*; Prentice Hall, Upper Saddle River, NJ, 2002;
  16. Oostenveld, R.; Fries, P.; Maris, E.; Schoffelen, J.M. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Comput Intell Neurosci* **2011**, *2011*, doi:10.1155/2011/156869.
  17. Altman, D.G.; Bland, J.M. *Measurement in Medicine: The Analysis of Method Comparison Studies †*; 1983; Vol. 32;.
  18. Dietterich, T.G. Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. *Neural Comput* **1998**, *10*, 1895–1923, doi:10.1162/089976698300017197.
  19. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **1995**, *57*, 289–300, doi:10.1111/j.2517-6161.1995.tb02031.x.
  20. Al-Rawi, P.G.; Kirkpatrick, P.J. Tissue Oxygen Index: Thresholds for Cerebral Ischemia Using near-Infrared Spectroscopy. *Stroke* **2006**, *37*, 2720–2725, doi:10.1161/01.STR.0000244807.99073.ae.
  21. Liu, T.; Liu, X.; Yi, L.; Zhu, C.; Markey, P.S.; Pelowski, M. Assessing Autism at Its Social and Developmental Roots: A Review of Autism Spectrum Disorder Studies Using Functional near-Infrared Spectroscopy. *NeuroImage* **2019**, *185*, 955–967.
  22. Keehn, B.; Wagner, J.; Tager-Flusberg, H.; Nelson, C.A. Functional Connectivity in the First Year of Life in Infants At-Risk for Autism: A Preliminary near-Infrared Spectroscopy Study. *Frontiers in Human Neuroscience* **2013**, *7*, doi:10.3389/fnhum.2013.00444.
  23. Ishii-Takahashi, A.; Takizawa, R.; Nishimura, Y.; Kawakubo, Y.; Kuwabara, H.; Matsubayashi, J.; Hamada, K.; Okuhata, S.; Yahata, N.; Igarashi, T.; et al. Prefrontal Activation during Inhibitory Control Measured by Near-Infrared Spectroscopy for Differentiating between Autism Spectrum Disorders and Attention Deficit Hyperactivity Disorder in Adults. *NeuroImage: Clinical* **2014**, *4*, 53–63, doi:10.1016/j.nicl.2013.10.002.
  24. Pinti, P.; Tachtsidis, I.; Hamilton, A.; Hirsch, J.; Aichelburg, C.; Gilbert, S.; Burgess, P.W. The Present and Future Use of Functional Near-Infrared Spectroscopy (fNIRS) for Cognitive Neuroscience. *Ann N Y Acad Sci* **2018**, *1464*, 5, doi:10.1111/nyas.13948.
  25. Gagnon, L.; Perdue, K.; Greve, D.N.; Goldenholz, D.; Kaskhedikar, G.; Boas, D.A. Improved Recovery of the Hemodynamic Response in Diffuse Optical Imaging Using Short Optode Separations and State-Space Modeling. *NeuroImage* **2011**, *56*, 1362–1371, doi:10.1016/j.neuroimage.2011.03.001.

26. Saager, R.B.; Berger, A.J. Direct Characterization and Removal of Interfering Absorption Trends in Two-Layer Turbid Media. *Journal of the Optical Society of America A* **2005**, *22*, 1874, doi:10.1364/josaa.22.001874.
27. Umeyama, S.; Yamada, T. Monte Carlo Study of Global Interference Cancellation by Multidistance Measurement of Near-Infrared Spectroscopy. *Journal of Biomedical Optics* **2009**, *14*, 064025, doi:10.1117/1.3275466.
28. Zhang, Q.; Brown, E.N.; Strangman, G.E. Adaptive Filtering for Global Interference Cancellation and Real-Time Recovery of Evoked Brain Activity: A Monte Carlo Simulation Study. *Journal of Biomedical Optics* **2007**, *12*, 044014, doi:10.1117/1.2754714.





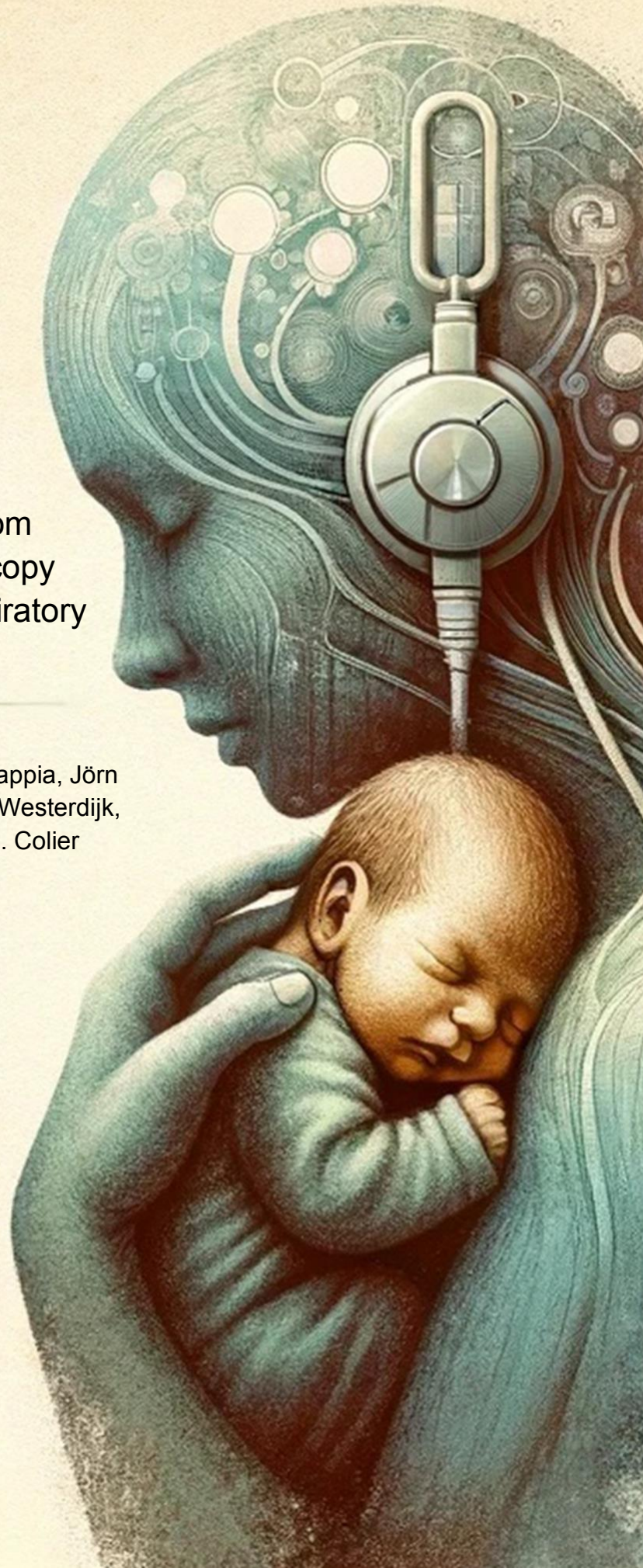
## CHAPTER 4

Estimation of respiratory rate from  
functional near-infrared spectroscopy  
(fNIRS): a new perspective on respiratory  
interference

Naser Hakimi\*, Mohammad Shahbakhti\*, Sofia Sappia, Jörn  
M. Horschig, Mathijs Bronkhorst, Marianne Floor-Westerdijk,  
Gaetano Valenza, Jeroen Dudink, Willy N.J.M. Colier

\* shared first author

doi: <https://doi.org/10.3390/bios12121170>



## Abstract

Respiration is recognized as a systematic physiological interference in functional near-infrared spectroscopy (fNIRS). However, it remains unanswered as to whether it is possible to estimate the respiratory rate (RR) from such interference. Undoubtedly, RR estimation from fNIRS can provide complementary information that can be used alongside the cerebral activity analysis, e.g., sport studies. Thus, the objective of this paper is to propose a method for RR estimation from fNIRS. Our primary presumption is that changes in the baseline wander of oxygenated hemoglobin concentration (O<sub>2</sub>Hb) signal are related to RR. fNIRS and respiratory signals were concurrently collected from subjects during controlled breathing tasks at a constant rate from 0.1 Hz to 0.4 Hz. Firstly, the signal quality index algorithm is employed to select the best O<sub>2</sub>Hb signal, and then a band-pass filter with cut-off frequencies from 0.05 to 2 Hz is used to remove very low- and high-frequency artifacts. Secondly, troughs of the filtered O<sub>2</sub>Hb signal are localized for synthesizing the baseline wander (S1) using cubic spline interpolation. Finally, the fast Fourier transform of the S1 signal is computed, and its dominant frequency is considered as RR. In this paper, two different datasets were employed, where the first one was used for the parameter adjustment of the proposed method, and the second one was solely used for testing. The low mean absolute error between the reference and estimated RRs for the first and second datasets (2.6 and 1.3 breaths per minute, respectively) indicates the feasibility of the proposed method for RR estimation from fNIRS. This paper provides a novel view on the respiration interference as a source of complementary information in fNIRS.



# 1. Introduction

Over the last 20 years, functional near-infrared spectroscopy (fNIRS) has arisen as an effective optical neuroimaging modality for measuring oxygenated (O<sub>2</sub>Hb) and deoxygenated (HHb) hemoglobin concentrations, associated with the neuronal activity [1,2]. Compared to other neuroimaging techniques such as electroencephalography (EEG) and functional magnetic resonance imaging (fMRI), fNIRS provides better spatial and temporal resolutions, respectively [3,4]. Thus, a wide range of studies in different cognitive tasks and clinical settings have employed fNIRS, e.g., [5–8].

Beside the EEG, due to its portable and non-invasive nature, fNIRS has been also used for outdoor applications [9]. In particular, the emergence of lightweight low-channel fNIRS equipment has provided a new possibility for non-laborious investigations, e.g., sport studies [10–13]. Nevertheless, the susceptibility of fNIRS to artifacts that are stemmed from various sources causes a great challenge for the accurate analysis of brain activity [14].

Generally, artifacts in fNIRS are classified into two categories: external and physiological interference [15,16]. The most prominent example of the former is motion artifacts, manifested by abrupt changes in the signal [17,18]. The latter is the interference originated from physiological systemic activities such as heart rate, blood pressure, Mayer waves, and respiration [19].

Nevertheless, despite a majority of studies that have considered physiological interference as source of artifacts in fNIRS, a few investigations have showed that such an interference can provide useful information for enhancing the accuracy of cerebral activity analysis. For example, Svinkunaite et al. [20] showed that using cardiac and respiratory features extracted from the fNIRS spectrum can enhance the accuracy of mental workload classification when employed alongside fNIRS temporal analysis. More interestingly, Hakimi et al. [21] showed the synergy of combining fNIRS temporal analysis alongside the extracted heart rate variability (HRV) for the stress assessment. According to the reported results by the authors, employing the extracted HRV from fNIRS improved the accuracy of classification by 10%.

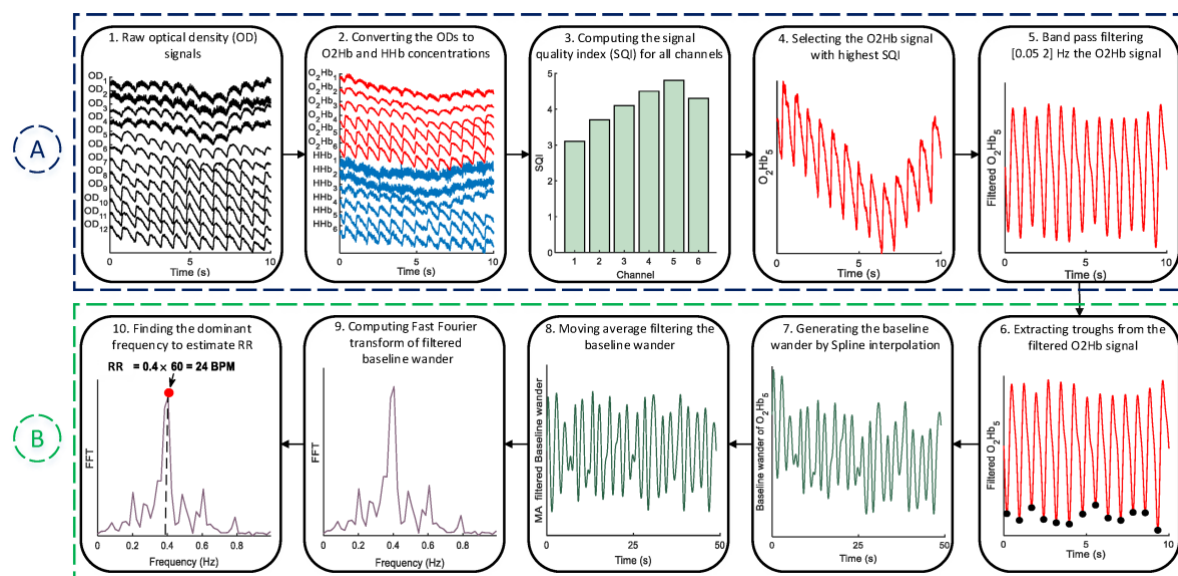
Besides the heart rate, respiration is another physiological interference that is vividly observable in the fNIRS spectrum (usually ranging from 0.2 to 0.4 Hz) [22]. According to the best of our knowledge, no research has yet considered the possibility of respiratory rate (RR) estimation from fNIRS. This is while fNIRS is being employed in several applications where RR can also play an important role, e.g., meditation [23], stress assessment [24], and exercise [10]. Unarguably, such an estimation can provide complementary information to be used in conjunction with the cerebral activity analysis.

Motivated accordingly, we propose a new method for estimating RR from fNIRS. Inspired by the studies that estimated RR from the photoplethysmography (PPG) [25], our hypothesis is that the baseline wander of fNIRS also might be related to alternations in the respiration. On the other hand, it has been shown that respiration has a stronger influence on the O<sub>2</sub>Hb than the HHb signal [26]. Hence, the basis of our method is to (i) extract the troughs of the O<sub>2</sub>Hb signal, (ii) synthesize the baseline wander using the cubic spline interpolation of the extracted troughs, and (iii) find the dominant frequency of the baseline wander to estimate RR. In order to assess the performance of the proposed method, two different datasets are used. The first

one is used to adjust the parameters of proposed method whereas the second one is only used for testing.

## 2. Methods

The block diagram of the proposed method for RR estimation from fNIRS signals is shown in Figure 1. It mainly consists of two stages: pre-processing (A) and RR estimation (B). In the subsections below, each step of the proposed method is explained in detail.



**Figure 1:** The block diagram of the proposed method. It should be noted that for the sake of clarity, fNIRS signals are shown only for 10 s.

### 2.1. Preprocessing

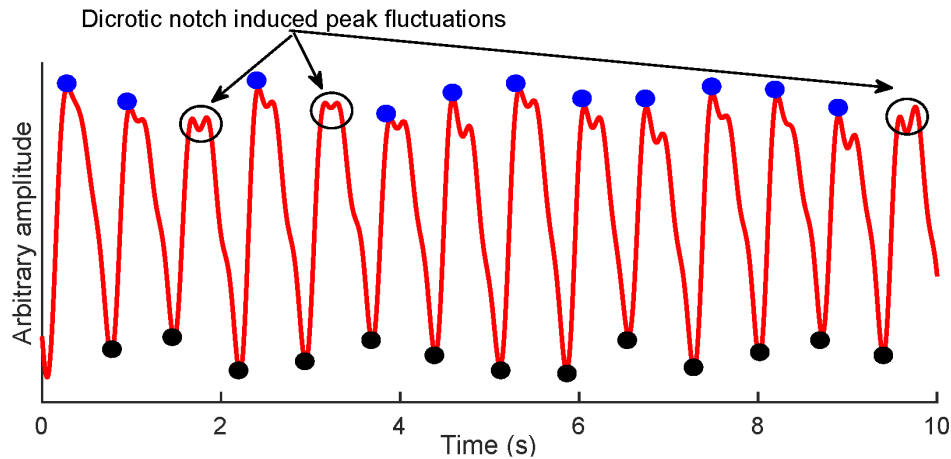
One the most important steps in every fNIRS-based study is to find high quality data for the analysis [27]. Thus, we have employed the signal quality index (SQI) [28] algorithm, which quantitatively scores fNIRS data in a numeric scale from 1 (very low quality) to 5 (very high quality). To compute the SQI, firstly, the modified Beer-Lambert law [29] is applied to convert the optical density (OD) signals into O<sub>2</sub>Hb and HHb changes in concentration. Secondly, OD, O<sub>2</sub>Hb, and HHb signals are detrended by subtracting the least-squares fit of a straight line to the data. Thirdly, a 208th-order zero-phase FIR band-pass filter with cutoff frequencies at 0.4 Hz and 3 Hz is applied on the signals from the previous step. Finally, several features are extracted from the filtered signals, and each channel is scored numerically between 1 and 5. It should be also noted that the SQI is computed based on 10 s windows. For a more detailed explanation, see [28]. After finding the highest quality channel, a zero-phase FIR band-pass filter with cut-off frequencies between 0.05 and 2 Hz is used for the removal of very low- and high-frequency artifacts from the selected O<sub>2</sub>Hb signal via the SQI algorithm.



## 2.2. RR Estimation

### 2.2.1. Trough Detection

Inspired from the PPG-based studies to estimate RR [25,30], our assumption is that alternations in the baseline wander of O2Hb signal can be related to RR. To this end, the fiducial points of O2Hb signal, i.e., the peaks and troughs, can be used. As shown in Figure 2, the synthesis of the baseline wander from troughs, compared to peaks, are more convenient as they are not subjected to dirotic notch-induced peak fluctuations.



**Figure 2:** An example of O2Hb signal with its corresponding peaks (blue) and troughs (black).

To localize the troughs, firstly, the filtered O2Hb signal,  $x(n)$ , is normalized to between  $-1$  and  $1$ . Then, the local minima that have a value lower than  $Th1 = A \times Mean(x(n))$  are considered as the potential troughs (Figure 3(a)). Yet, the emergence of motion artifacts can still jeopardize the accurate synthesis of baseline wander (Figure 3(b)). To overcome this problem, after localizing the troughs, their corresponding magnitudes are set on a vector,  $z(n)$ , and elements with values lower than  $Th2 = Mean(z(n)) + B \times std(z(n))$  are discarded (Figure 3(c)). The coefficients  $A$  and  $B$  are constants that will be regulated empirically (see Section 4.1.1). The main steps of trough detection are summarized in Algorithm 1.

---

**Algorithm 1** Localization of O2Hb signal troughs

---

**Input:** O2Hb signal  $x(n)$ , constants  $A, B$

**Output:** Troughs,  $K$

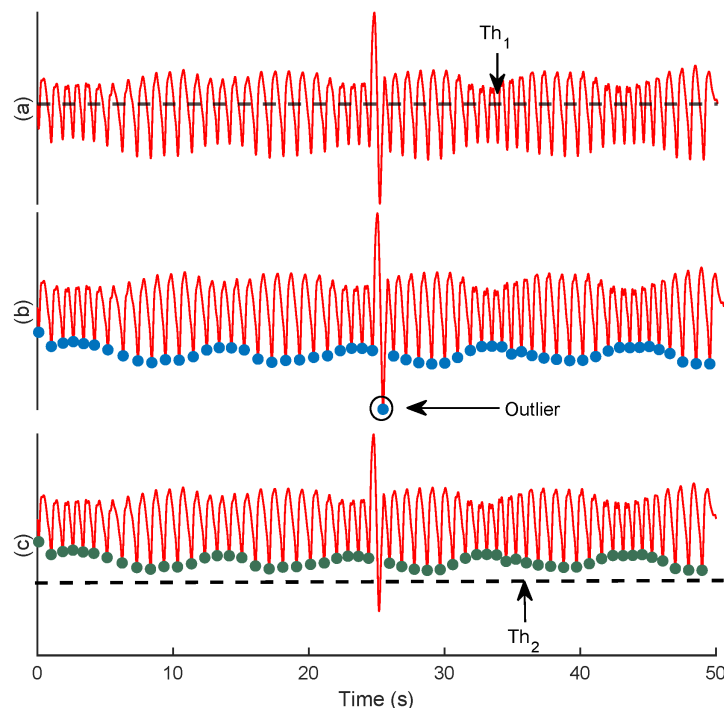
*Initialization*  $Th1 \leftarrow 0, Th2 \leftarrow 0, J \leftarrow [], K \leftarrow []$

- 1:  $x(n) \leftarrow \text{Normalize}(x(n))$
- 2:  $Th1 \leftarrow A \times \text{Mean}(x(n))$
- 3: **for**  $i=2$  to  $\text{length}(x(n))-1$  **do**
- 4:   **if**  $x(i) < x(i-1) \ \&\& \ x(i) < x(i+1) \ \&\& \ x(i) < Th1$  **then**
- 5:      $J \leftarrow [i]$
- 6:   **end if**
- 7: **end for**

```

8:   $z(n) \leftarrow x(J)$ 
9:   $Th2 \leftarrow \text{Mean}(z(n)) + B \times \text{std}(z(n))$ 
10: for  $i=1$  to  $\text{length}(z(n))$  do
11:   if  $z(i) > Th2$  then
12:     $K = [Ki]$ 
13:   end if
14: end for
15: return  $K$ 

```



**Figure 3:** An example of the trough detection. The filtered O2Hb signal (a), the selected troughs after employing  $Th1$  (b), and  $Th2$  (c).

### 2.2.2. Forming the Baseline Wander Signal

After extracting the troughs, the corresponding time series, i.e., baseline wander, needs to be re-sampled. This is a necessary step, as the baseline wander generated from the troughs is irregularly sampled, whereas the following analysis needs a regularly sampled signal [25]. For this aim, the cubic spline interpolation method is employed, which approximates a signal by connecting a series of points through a polynomial equation that passes through all of those points continuously.

### 2.2.3. FFT for RR Estimation

After synthesizing the baseline wander, non-respiratory oscillations, i.e., very-low-frequency components, should be filtered, as they can hinder the identification of the dominant frequency in the FFT domain (Figure 4(a)) [25,31]. For this aim, a moving average (MA) filter is used. After MA filtering, the dominant frequency of baseline wander (Figure 4(b)) is multiplied by 60 to estimate RR in breaths per minute (BPM). Although it can be argued that applying the MA

may also influence the baseline wander, the MA filtering removes very-low-frequency components (below 0.04 Hz) that are not in the frequency range of RR for healthy subjects. The regulation of the moving average filter's length is described in Section 4.1.2. The summary of RR estimation procedure after extracting the troughs of O2Hb signal is presented in Algorithm 2.

---

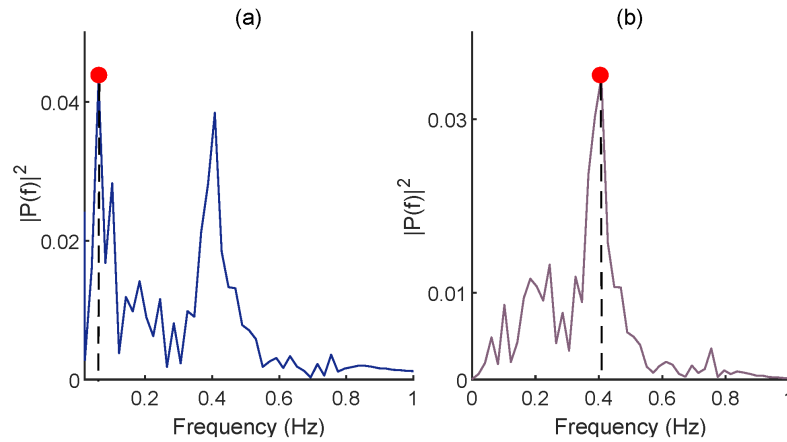
**Algorithm 2** Estimation of the RR from the O2Hb signal's baseline wander

---

**Input:** O2HB signal  $x(n)$ , Troughs of O2HB signal  $K$ , and the length of moving average filter  $L$

**Output:** RR

- 1:  $m(n) \leftarrow \text{Spline}(K, x(K), 1 \text{ to } \text{length}(x(n)))$
  - 2:  $MA \leftarrow 1L \times \text{ones}(L, 1)$
  - 3:  $S(n) \leftarrow \text{filtfilt}(MA, 1, m(n))$
  - 4:  $G(n) \leftarrow m(n) - S(n)$
  - 5:  $[P, F] \leftarrow \text{FFT}(G(n))$
  - 6:  $r \leftarrow \text{find}(P(F) \leftarrow \text{Max}(P(F)))$
  - 7:  $RR \leftarrow r \times 60$
  - 8: **return** RR
- 



**Figure 4:** The FFT of the baseline wander before (a) and after (b) employing the MA filtering. The red dot stands for dominant frequency in the FFT domain. Note that the reference RR is 0.4 Hz in this example.

## 2.3. Evaluation Criteria

The performance of trough and motion-induced artifact detection is assessed using the critical success index (CSI), defined as

$$CSI = \frac{T_P}{T_P + F_N + F_P} \quad (1)$$

where  $TP$ ,  $FN$ , and  $FP$  stand for correctly, missed, and wrongly detected trough and artifactual samples. Regarding RR estimation, the absolute error (AE) between the reference RR obtained from the respiratory signal and estimated RR from O2Hb signal is employed as follows:

$$AE = |Reference_{RR} - Estimated_{RR}| \quad (2)$$

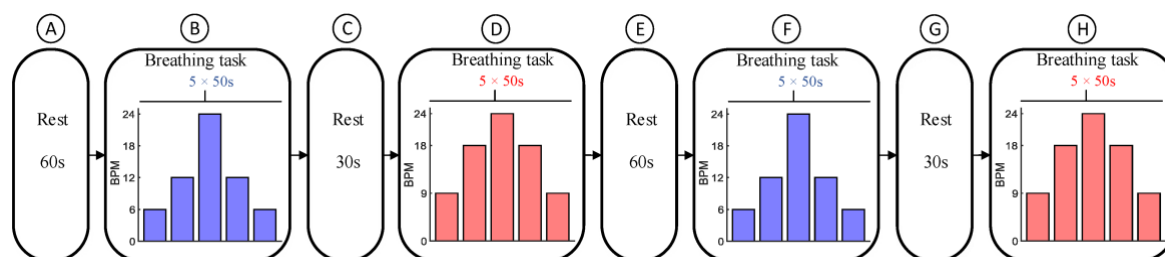
To investigate whether there is a significant difference between the estimated and reference RRs, the paired samples t-test with a significance level of 0.05 was performed for each subject.

### 3. Data

In this paper, two different datasets are employed. The first one is used to adjust the parameters of the proposed method, whereas the second one is only used for testing. Two different fNIRS devices were used for each dataset, and the reference respiratory signals were recorded simultaneously using a chest-band with a TMSi SAGA 32+/64+ amplifier (Twente Medical Systems International B.V., Oldenzaal, The Netherlands) at a sampling rate of 4000 Hz.

#### 3.1. Data Recording Protocol

Before starting the experiment, the subjects were briefed on the procedure and instructed on how to perform the tasks, in English. The data recording protocol for RR estimation, which was adapted from [32–34], is shown in Figure 5. As displayed, it consisted of one block of a resting period lasting for 60 s (A), followed by two blocks of breathing control tasks (B and D), separated by a 30 s rest period (C). Subsequently, the same blocks were repeated (E to H). The subjects were asked to inhale and exhale at a constant pace and at specific rates while watching a bar moving vertically together with a text showing inhale- or exhale phases on the screen. Each block of the breathing control task consisted of 5 steps with a constant RR over a period of 50 s. The RRs for the first and third blocks were 6, 12, 24, 12, and 6 BPM, and for the second and fourth blocks, they were 9, 18, 24, 18, and 9 BPM.



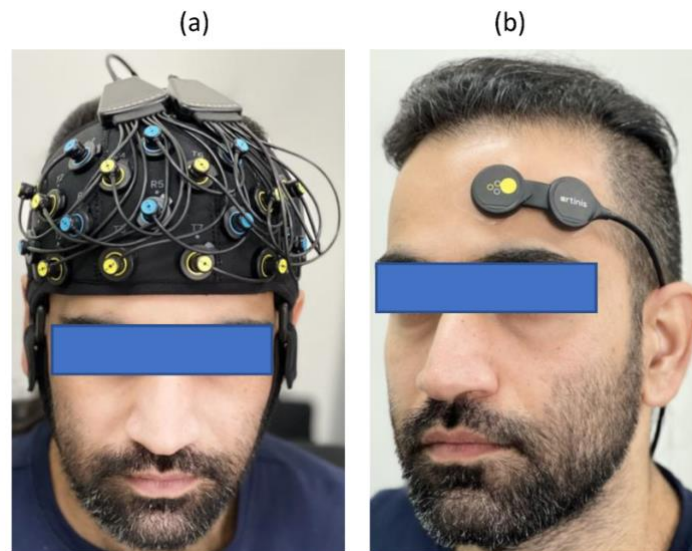
**Figure 5:** Data recording protocol. It consists of a resting period for 60 s (A), and two breathing control tasks lasting for 250 s (B, D), which are separated by a 30 s resting period (C). Subsequently, the same blocks were repeated (E–H).

The local ethics committee of Comitato Bioetico of the University of Pisa approved this study protocol with ref. num. 2/2020. Before starting the experiment, all subjects were informed about the experiment and signed the consent form. All methods were performed based on the guidelines and regulations required by the Declaration of Helsinki. Data were registered at Artinis Medical Systems B.V., Elst, The Netherlands.

## 3.2. fNIRS Systems for Data Collection

### 3.2.1. Dataset I

This dataset comprised fNIRS data from 8 healthy subjects (3 female) aged from 21 to 32 years recorded using a portable wireless 23 channel fNIRS system (Brite23, Artinis Medical Systems B.V., The Netherlands) covering the whole frontal cortex (Figure 6(a)). This device is supplied with a source–detector separation of 35 mm, nominal wavelengths of 760 and 850 nm, ambient light correction, and a sampling frequency of 50 Hz.



**Figure 6:** fNIRS optode placement for dataset I (a) and dataset II (b).

### 3.2.2. Dataset II

This dataset consisted of fNIRS data collected from 18 healthy subjects (9 female) aged from 24 to 37 years using a wireless multi-sensor fNIRS-system (PortaLite MKII, Artinis Medical Systems B.V., The Netherlands). This device is equipped with up to 2 sensors, each having 3 long channels (the source–detector distance being up to 41 mm), and 3 short-separation channels (with distances of 7.2 and 8.0 mm), nominal wavelengths of 760 and 850 nm, ambient light correction, and a sampling rate of 100 Hz. The sensors simultaneously recorded any movement using an IMU embedded within each sensor, and were designed to be placed on both hemispheres of the prefrontal cortex of the brain. In this paper, we have used only a single sensor placed on the left hemisphere of the prefrontal cortex (Figure 6(b)).

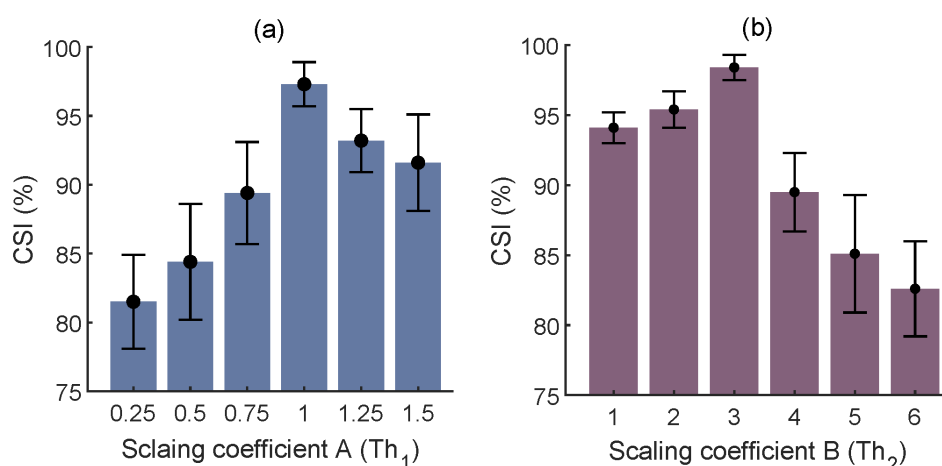
## 4. Experimental Results

In this section, the obtained results from both datasets are described. It is worth mentioning that the required parameters of the proposed method were first tuned based on the optimal results obtained from dataset I, then the adjusted parameters were used for the analysis of dataset II. In addition, the short separation channels of dataset II were discarded to have a similar data structure and analysis for both datasets.

## 4.1. Optimization of the Proposed Method's Parameters

### 4.1.1. Trough Detection

Two empirical thresholds require tuning for the trough detection:  $Th1$ , which is necessary for ignoring the local minima in the dicrotic notch, i.e., a small downward deflection between the peaks and troughs, and  $Th2$ , which is used to discard the troughs contaminated by the motion artifacts. Regarding the scaling coefficient of  $Th1$ , values from 0.25 to 1.5 with a step size of 0.25 were inspected. The best fit, i.e., the highest mean CSI, was  $A=1$  (Figure 7(a)). As for the scaling coefficient of  $Th2$ , values from 1 to 6 with a step size of 1 were investigated, and  $B=3$  was obtained as the best fit (Figure 7(b)).



**Figure 7:** Regulation of constants for trough detection in terms of mean±std of the CSI.  $Th1$  (a) and  $Th2$  (b).

### 4.1.2. The Length of MA Filtering

As displayed in Figure 4, removing very-low-frequency components of the generated baseline wander is of great importance for the accurate estimation of RR. For this aim, different lengths of the MA filter from 2 to 5 s with a stepping size of 0.5 s were investigated. According to the obtained average AE, although no noticeable difference was observed between different lengths, 3 s had the lowest error (Table 1).

**Table 1:** The Influence of MA Filter's Length on the RR estimation.

MA Filter Length (s)	Average AE ± Std (BPM)
2	3.2 ± 1.9
2.5	3.1 ± 1.8
3	2.6 ± 1.3
3.5	2.7 ± 1.4
4	2.9 ± 1.9
4.5	2.9 ± 2.1
5	3.1 ± 2.2

### 4.1.3. Results of RR Estimation from Dataset I

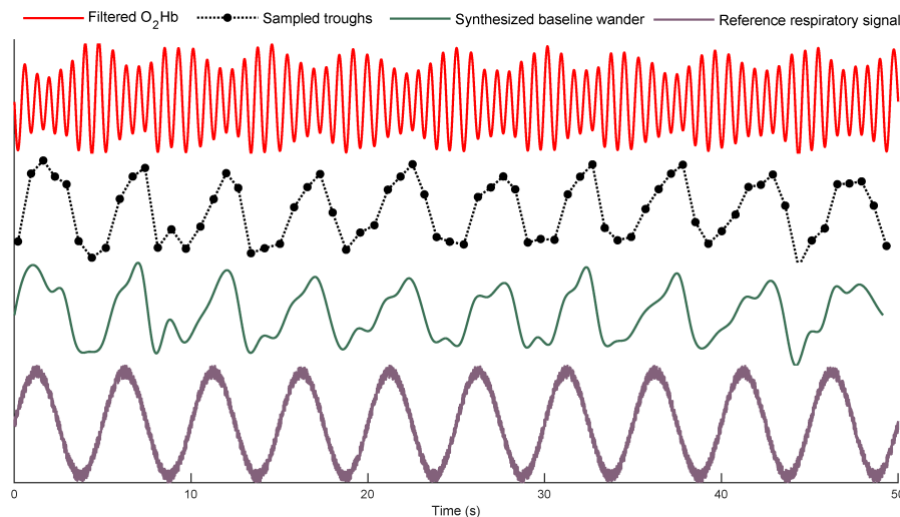
In total, 160 50 s trials of the concurrent O<sub>2</sub>Hb and respiratory signals were used, where each RR was repeated 4 times per subject. Table 2 displays the average AE between the reference and estimated RRs for each subject. The average AE of all trials was 2.6 BPM. Given the reported results from the PPG-based studies for RR estimation [32–34], the obtained results indicate the feasibility of the proposed method for RR estimation from fNIRS. According to the conducted statistical analysis between the reference and estimated RRs, except for the subject 8, there is no significant difference ( $p>0.05$ ).

**Table 2:** The average AE between the reference and estimated RRs for each subject of dataset I.

Subjects	Average AE (BPM)
1	0.9
2	2.7
3	2.7
4	1.1
5	2.2
6	1.9
7	2.1
8	5.2

### 4.2. Results of RR Estimation from Dataset II

An example of the filtered O<sub>2</sub>Hb signal, as well as its corresponding extracted baseline wanders and the reference respiratory signal are displayed in Figure 8. As it can be seen, the frequency of baseline wander is close to that of the reference respiratory signal.

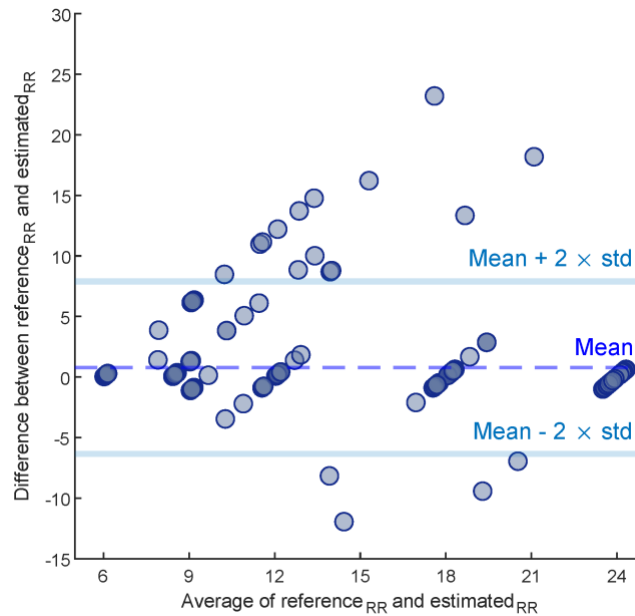


**Figure 8:** An example of the filtered O<sub>2</sub>Hb signal, the corresponding baseline wanders, and the reference respiratory signal.

The average AE between the reference and estimated RRs for each subject of dataset II is disclosed in Table 3. Except for subject 13, the statistical analysis shows no significant difference between the reference and estimated RRs ( $p>0.05$ ). The average AE for all trials is



1.3 BPM. To evaluate the overall performance of the proposed method on dataset II, a Bland–Altman plot (Figure 9) was used; this assesses the agreement between reference and estimated RRs by showing the difference between each estimate and the references against their mean. In this paper, the Limit of Agreement (LOA) is computed as  $[mean - 2 \times std, mean + 2 \times std]$ . In this range, 94% of the differences are inside.



**Figure 9:** The Bland–Altman plot of the estimated RRs on dataset II.

**Table 3:** The average of AE between the reference and estimated RRs for each subject of dataset II.

Subjects	Average AE (BPM)
1	1.7
2	0.3
3	0.3
4	0.5
5	1.8
6	0.8
7	1.5
8	2.7
9	2.1
10	0.3
11	0.4
12	0.7
13	3.6
14	0.5
15	1.8
16	1.8
17	0.4
18	2.1

## 5. Discussion

The aim of this paper was to investigate the possibility of RR estimation from fNIRS. Indisputably, this is a novel view on respiration interference in fNIRS, as almost all studies have considered such an interference as a detrimental phenomenon. The importance of this study is to derive an extra measure, i.e., RR, to the cerebral activity analysis without requiring a reference signal. From the research point of view, the proposed method facilitates the approximation of RR in applications where both cerebral and respiratory activities may synergize the analysis. For instance, there is solid evidence in the literature suggesting that RR alternation is an indicator of anxiety and mental workload levels [35,36].

### 5.1. Significance and Robustness of the Proposed Method

It is common knowledge that inhalation and exhalation can alter the blood flow within the body [37]. On the other hand, respiratory fluctuations affect the cerebral blood volume and flow [38]. Therefore, it can be expected that alternation in RR is revealed in fNIRS, and in particular, in O<sub>2</sub>Hb signals [26]. Our underlying assumption, which was inspired from PPG-based studies [25], was that fluctuations in the frequency of O<sub>2</sub>Hb signal's baseline wander can be related to RR.

The proposed method is based on the FFT obtained from the baseline wander of an O<sub>2</sub>Hb signal. To synthesize the baseline wander, either troughs or peaks can be used. In this paper, we have found the localization of troughs to be more convenient (Figure 2). Yet, synthesizing the baseline wander required an interpolation method due to the irregular sampling. Here, we used cubic spline interpolation, which was already proven as an effective method in PPG-based studies [30]. Afterwards, the dominant frequency of synthesized baseline wander in the FFT domain was considered as RR. The motivation behind using FFT rather than other spectral analyses such as Welch or MUSIC is its simplicity. i.e., the FFT is a non-parametric algorithm.

To assess the feasibility of the proposed method, two different datasets were used: one for adjustment of the required parameters of the proposed method and the other for testing. Indeed, the second dataset was used to investigate the robustness of the proposed method's parameters when data were recorded from another fNIRS equipment with different characteristics, e.g., the sampling rate, receiver gain, LED types, etc. The comparison between Table 2 and Table 3 confirms such robustness as there is no noticeable difference between the obtained results from both datasets. More surprisingly, even a lower mean of AE was achieved with the second dataset. One plausible explanation can be the weak performance of our method for subject 8 in dataset I.

### 5.2. Comparison with State-of-the-Art Methods

As this is the first research that proposes a method for RR estimation from fNIRS, there is no possibility for comparing the performance of the proposed method to other studies. Yet, a few investigations have aimed to regress out the respiratory components from fNIRS data. For example, Tong et al. [38] used a zero-phase band-pass filter with cut-off frequencies from 0.2 to 0.6 Hz to partition out the respiratory components from the O<sub>2</sub>Hb signals. In another study, Lühmann et al. [39] proposed a multimodal extension of the general linear model based on temporally embedded canonical correlation analysis to extract respiratory components from

fNIRS measurements. The former methodology may not be efficient, as such a band width also involves cerebral activities, and the latter requires the simultaneously recorded respiratory signal. Nonetheless, neither of the mentioned studies have considered RR estimation from the regressed respiratory components. Although it may be inequitable to compare the quality of estimated RR from fNIRS with PPG-based studies, our results are still comparable to [34], where both the finger and forehead PPG data were used.

On the other hand, the mentioned studies used a large amount of data, which is not appropriate for practical applications, as wearing a cap covering the whole head discomforts the user for long-term recording. Furthermore, such a configuration usually involves covering the hair-bearing areas of the head that are more subjected to noise. In contrast, we developed and tested our algorithm-based sensors placed only on the frontal region of head, which provide the user with more comfort, as it is mostly a hairless area. In addition, such a configuration reduces the complexity of wearable instrumentation as only one region of the brain is monitored.

### 5.3. Directions for Future Work

Regardless of the reported promising results, this research has several limitations that should be addressed in future works. Firstly, the employed SQI algorithm is not necessarily an optimal method for finding the best fNIRS data for the analysis. In particular, the SQI is not sensitive to the emergence of motion artifacts, which can significantly influence the synthesis of baseline wander. Yet, to the best of our knowledge, it is one of the best methods for monitoring the quality of data. Secondly, the data recording protocol was a simple breathing task where subjects needed to sit on a chair and perform the experiment. In future work, the subjects should be asked to perform more dynamic activities, e.g., cycling, to obtain more realistic RRs. Thirdly, due to the nature of the data recording protocol, fNIRS data were not significantly contaminated by motion artifacts. Thus, the performance of *Th1* and *Th2* for trough detection should be investigated further, with more artifactual data. Fourthly, the performance of the proposed method should be investigated by analyzing different time windows. Fifthly, it should be mentioned that the correction for multiple comparisons has not been conducted for statistical analysis. Lastly, the reliability of the proposed method was only evaluated on young healthy subjects. It is also of great importance to assess the robustness of the proposed method on more diverse cohorts (e.g., the elderly, neonates, and patients) as some studies showed that the performance of RR estimation algorithm can be affected by different factors such as age [40]. Nevertheless, it should be noted that this is the first research that has proposed a method for RR estimation from fNIRS; therefore, having the mentioned limitations was necessary for investigating the possibility of RR estimation.

## 6. Conclusions

In this paper, a method based on the spectrum analysis of the O<sub>2</sub>Hb signal's baseline wander was introduced to estimate RR, and its performance was assessed on two different datasets with distinctive data recording characteristics. The comparison between the obtained results from both datasets confirmed the robustness of the proposed method, which is of great importance for real-world applications. The milestone of this research was to consider the respiration interference in fNIRS as source of complementary information, rather than a source

of artifact. Indeed, the proposed method can provide extra information from fNIRS that can be used alongside the cerebral activity analysis.

## References

1. Paulmurugan, K.; Vijayaragavan, V.; Ghosh, S.; Padmanabhan, P.; Gulyás, B. Brain–Computer Interfacing Using Functional Near-Infrared Spectroscopy (FNIRS). *Biosensors* **2021**, Vol. 11, Page 389 **2021**, 11, 389, doi:10.3390/BIOS11100389.
2. Ferrari, M.; Quaresima, V. A Brief Review on the History of Human Functional Near-Infrared Spectroscopy (FNIRS) Development and Fields of Application. *Neuroimage* **2012**, 63, 921–935, doi:10.1016/J.NEUROIMAGE.2012.03.049.
3. Almajidy, R.K.; Mankodiya, K.; Abtahi, M.; Hofmann, U.G. A Newcomer’s Guide to Functional Near Infrared Spectroscopy Experiments. *IEEE Rev Biomed Eng* **2020**, 13, 292–308, doi:10.1109/RBME.2019.2944351.
4. Scholkmann, F.; Kleiser, S.; Metz, A.J.; Zimmermann, R.; Mata Pavia, J.; Wolf, U.; Wolf, M. A Review on Continuous Wave Functional Near-Infrared Spectroscopy and Imaging Instrumentation and Methodology. *Neuroimage* **2014**, 85 Pt 1, 6–27, doi:10.1016/J.NEUROIMAGE.2013.05.004.
5. Chao, J.; Zheng, S.; Wu, H.; Wang, D.; Zhang, X.; Peng, H.; Hu, B. FNIRS Evidence for Distinguishing Patients with Major Depression and Healthy Controls. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2021**, 29, 2211–2221, doi:10.1109/TNSRE.2021.3115266.
6. Borgheai, S.B.; Mclinden, J.; Zisk, A.H.; Hosni, S.I.; Deligani, R.J.; Abtahi, M.; Mankodiya, K.; Shahriari, Y. Enhancing Communication for People in Late-Stage ALS Using an FNIRS-Based BCI System. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2020**, 28, 1198–1207, doi:10.1109/TNSRE.2020.2980772.
7. Wang, Z.; Zhang, J.; Zhang, X.; Chen, P.; Wang, B. Transformer Model for Functional Near-Infrared Spectroscopy Classification. *IEEE J Biomed Health Inform* **2022**, 26, 2559–2569, doi:10.1109/JBHI.2022.3140531.
8. Sommer, N.M.; Kakillioglu, B.; Grant, T.; Velipasalar, S.; Hirshfield, L. Classification of FNIRS Finger Tapping Data With Multi-Labeling and Deep Learning. *IEEE Sens J* **2021**, 21, 24558–24568, doi:10.1109/JSEN.2021.3115405.
9. Joshi, S.; Herrera, R.R.; Springett, D.N.; Weedon, B.D.; Ramirez, D.Z.M.; Holloway, C.; Dawes, H.; Ayaz, H. Neuroergonomic Assessment of Wheelchair Control Using Mobile FNIRS. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2020**, 28, 1488–1496, doi:10.1109/TNSRE.2020.2992382.
10. Pellegrini-Laplagne, M.; Dupuy, O.; Sosner, P.; Bosquet, L. Effect of Simultaneous Exercise and Cognitive Training on Executive Functions, Baroreflex Sensitivity, and Pre-Frontal Cortex Oxygenation in Healthy Older Adults: A Pilot Study. *GeroScience* **2022**, 45:1 **2022**, 45, 119–140, doi:10.1007/S11357-022-00595-3.
11. Germain, C.; Perrot, A.; Tomasino, C.; Bonnal, J.; Ozsancak, C.; Auzou, P.; Prieur, F. Effect of the Level of Physical Activity on Prefrontal Cortex Hemodynamics in Older Adults During Single- and Dual-Task Walking. *J Aging Phys Act* **2022**, 31, 96–104, doi:10.1123/JAPA.2021-0410.

12. Goenarjo, R.; Dupuy, O.; Fraser, S.; Berryman, N.; Perrochon, A.; Bosquet, L. Cardiorespiratory Fitness and Prefrontal Cortex Oxygenation during Stroop Task in Older Males. *Physiol Behav* **2021**, *242*, 113621, doi:10.1016/J.PHYSBEH.2021.113621.
13. Koren, Y.; Mairon, R.; Sofer, I.; Parmet, Y.; Ben-Shahar, O.; Bar-Haim, S. Vision, Cognition, and Walking Stability in Young Adults. *Scientific Reports* **2022**, *12:1* **2022**, *12*, 1–10, doi:10.1038/s41598-021-04540-w.
14. Bizzego, A.; Neoh, M.; Gabrieli, G.; Esposito, G. A Machine Learning Perspective on FNIRS Signal Quality Control Approaches. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2022**, doi:10.1109/TNSRE.2022.3198110.
15. Patashov, D.; Menahem, Y.; Gurevitch, G.; Kameda, Y.; Goldstein, D.; Balberg, M. FNIRS: Non-Stationary Preprocessing Methods. *Biomed Signal Process Control* **2023**, *79*, 104110, doi:10.1016/J.BSPC.2022.104110.
16. Tachtsidis, I.; Scholkmann, F. False Positives and False Negatives in Functional Near-Infrared Spectroscopy: Issues, Challenges, and the Way Forward. <https://doi.org/10.1117/1.NPh.3.3.031405> **2016**, *3*, 031405, doi:10.1117/1.NPH.3.3.031405.
17. Zhou, L.; Chen, C.; Liu, Z.; Hu, Y.; Chen, M.; Li, Y.; Hu, Y.; Wang, G.; Zhao, J. A Coarse/Fine Dual-Stage Motion Artifacts Removal Algorithm for Wearable NIRS Systems. *IEEE Sens J* **2021**, *21*, 13574–13583, doi:10.1109/JSEN.2021.3069553.
18. Ivo, I. de A.; Horschig, J.M.; Gerakaki, S.; Wanrooij, M.M. van; Colier, W.N.J.M. Cerebral Oxygenation Responses to Head Movement Measured with Near-Infrared Spectroscopy. <https://doi.org/10.1117/12.2587168> **2021**, *11638*, 40–52, doi:10.1117/12.2587168.
19. Zhang, F.; Cheong, D.; Khan, A.F.; Chen, Y.; Ding, L.; Yuan, H. Correcting Physiological Noise in Whole-Head Functional near-Infrared Spectroscopy. *J Neurosci Methods* **2021**, *360*, doi:10.1016/J.JNEUMETH.2021.109262.
20. Svinkunaite, L.; Horschig, J.M.; Floor-Westerdijk, M.J. Employing Cardiac and Respiratory Features Extracted from FNIRS Signals for Mental Workload Classification. **2021**, *13*, doi:10.1117/12.2587155.
21. Hakimi, N.; Jodeiri, A.; Mirbagheri, M.; Setarehdan, S.K. Proposing a Convolutional Neural Network for Stress Assessment by Means of Derived Heart Rate from Functional near Infrared Spectroscopy. *Comput Biol Med* **2020**, *121*, 103810, doi:10.1016/J.COMPBIOMED.2020.103810.
22. Izzetoglu, M.; Holtzer, R. Effects of Processing Methods on FNIRS Signals Assessed during Active Walking Tasks in Older Adults. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2020**, *28*, 699–709, doi:10.1109/TNSRE.2020.2970407.
23. Bellissimo, G.; Leslie, E.; Maestas, V.; Zuhl, M. The Effects of Fast and Slow Yoga Breathing on Cerebral and Central Hemodynamics. *Int J Yoga* **2020**, *13*, 207, doi:10.4103/IJOY.IJOY\_98\_19.

24. Bak, S.; Shin, J.; Jeong, J. Subdividing Stress Groups into Eustress and Distress Groups Using Laterality Index Calculated from Brain Hemodynamic Response. *Biosensors* **2022**, *Vol. 12*, Page 33 **2022**, *12*, 33, doi:10.3390/BIOS12010033.
25. Charlton, P.H.; Birrenkott, D.A.; Bonnici, T.; Pimentel, M.A.F.; Johnson, A.E.W.; Alastruey, J.; Tarassenko, L.; Watkinson, P.J.; Beale, R.; Clifton, D.A. Breathing Rate Estimation from the Electrocardiogram and Photoplethysmogram: A Review. *IEEE Rev Biomed Eng* **2018**, *11*, 2–20, doi:10.1109/RBME.2017.2763681.
26. Reddy, P.; Izzetoglu, M.; Shewokis, P.A.; Sangobowale, M.; Diaz-Arrastia, R.; Izzetoglu, K. Evaluation of FNIRS Signal Components Elicited by Cognitive and Hypercapnic Stimuli. *Scientific Reports* **2021** *11:1* **2021**, *11*, 1–15, doi:10.1038/s41598-021-02076-7.
27. Pollonini, L.; Bortfeld, H.; Oghalai, J.S. PHOEBE: A Method for Real Time Mapping of Optodes-Scalp Coupling in Functional near-Infrared Spectroscopy. *Biomedical Optics Express*, *Vol. 7*, *Issue 12*, pp. 5104-5119 **2016**, *7*, 5104–5119, doi:10.1364/BOE.7.005104.
28. Sappia, M.S.; Hakimi, N.; Colier, W.N.J.M.; Horschig, J.M. Signal Quality Index: An Algorithm for Quantitative Assessment of Functional near Infrared Spectroscopy Signal Quality. *Biomedical Optics Express*, *Vol. 11*, *Issue 11*, pp. 6732-6754 **2020**, *11*, 6732–6754, doi:10.1364/BOE.409317.
29. Delpy, D.T.; Cope, M.; Van Der Zee, P.; Arridge, S.; Wray, S.; Wyatt, J. Estimation of Optical Pathlength through Tissue from Direct Time of Flight Measurement. *Phys Med Biol* **1988**, *33*, 1433, doi:10.1088/0031-9155/33/12/008.
30. Deriving Respiration from the Pulse Photoplethysmographic Signal | IEEE Conference Publication | IEEE Xplore Available online: <https://ieeexplore.ieee.org/abstract/document/6164665> (accessed on 25 March 2024).
31. Estañol, B.; Senties-Madrid, H.; Elías, Y.; Coyac, P.; Martínez-Memije, R.; Infante, Ó.; Téllez-Zenteno, J.F.; García-Ramos, G. Respiratory and Non Respiratory Oscillations of the Skin Blood Flow: A Window to the Function of the Sympathetic Fibers to the Skin Blood Vessels. *Arch Cardiol Mex* **2008**, *78*, 187–194.
32. Madhav, K.V.; Ram, M.R.; Krishna, E.H.; Komalla, N.R.; Reddy, K.A. Robust Extraction of Respiratory Activity from PPG Signals Using Modified MSPCA. *IEEE Trans Instrum Meas* **2013**, *62*, 1094–1106, doi:10.1109/TIM.2012.2232393.
33. Hernando, A.; Pelaez, M.D.; Lozano, M.T.; Aiger, M.; Gil, E.; Lazaro, J. Finger and Forehead PPG Signal Comparison for Respiratory Rate Estimation Based on Pulse Amplitude Variability. *25th European Signal Processing Conference, EUSIPCO 2017* **2017**, *2017-January*, 2076–2080, doi:10.23919/EUSIPCO.2017.8081575.
34. Hernando, A.; Peláez-Coca, M.D.; Lozano, M.T.; Lázaro, J.; Gil, E. Finger and Forehead PPG Signal Comparison for Respiratory Rate Estimation. *Physiol Meas* **2019**, *40*, 095007, doi:10.1088/1361-6579/AB3BE0.
35. Tipton, M.J.; Harper, A.; Paton, J.F.R.; Costello, J.T. The Human Ventilatory Response to Stress: Rate or Depth? *Journal of Physiology* **2017**, *595*, 5729–5752, doi:10.1113/JP274596.



36. Grassmann, M.; Vlemincx, E.; Von Leupoldt, A.; Mittelstädt, J.M.; Van Den Bergh, O. Respiratory Changes in Response to Cognitive Load: A Systematic Review. *Neural Plast* **2016**, *2016*, doi:10.1155/2016/8146809.
37. Iqbal, T.; Elahi, A.; Ganly, S.; Wijns, W.; Shahzad, A. Photoplethysmography-Based Respiratory Rate Estimation Algorithm for Health Monitoring Applications. *J Med Biol Eng* **2022**, *42*, 242–252, doi:10.1007/S40846-022-00700-Z/TABLES/6.
38. Tong, Y.; Lindsey, K.P.; Frederick, B.D. Partitioning of Physiological Noise Signals in the Brain with Concurrent Near-Infrared Spectroscopy and FMRI. *Journal of Cerebral Blood Flow and Metabolism* **2011**, *31*, 2352–2362, doi:10.1038/JCBFM.2011.100/ASSET/IMAGES/LARGE/10.1038\_JCBFM.2011.100-FIG6.JPEG.
39. von Lüthmann, A.; Li, X.; Müller, K.R.; Boas, D.A.; Yücel, M.A. Improved Physiological Noise Regression in FNIRS: A Multimodal Extension of the General Linear Model Using Temporally Embedded Canonical Correlation Analysis. *Neuroimage* **2020**, *208*, 116472, doi:10.1016/J.NEUROIMAGE.2019.116472.
40. Charlton, P.H.; Bonnici, T.; Tarassenko, L.; Alastruey, J.; Clifton, D.A.; Beale, R.; Watkinson, P.J. Extraction of Respiratory Signals from the Electrocardiogram and Photoplethysmogram: Technical and Physiological Determinants. *Physiol Meas* **2017**, *38*, 669, doi:10.1088/1361-6579/AA670E.



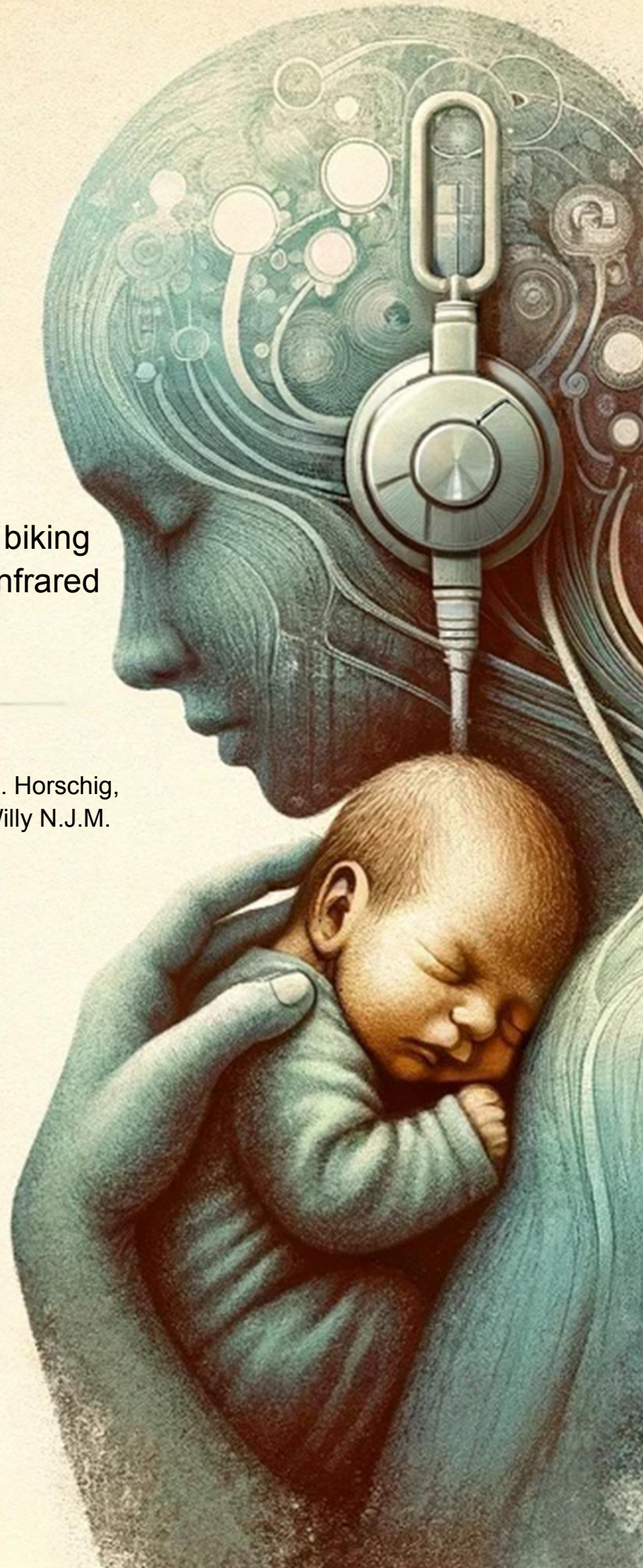
## CHAPTER 5

Estimation of respiratory rate during biking with a single sensor functional near-infrared spectroscopy (fNIRS) system

Mohammad Shahbakhti\*, Naser Hakimi\*, Jörn M. Horschig, Marianne Floor-Westerdijk, Jurgen Claassen, Willy N.J.M. Colier

\* shared first author

doi: <https://doi.org/10.3390/s23073632>



## Abstract

The employment of wearable systems for continuous monitoring of vital signs is increasing. However, due to substantial susceptibility of conventional bio-signals recorded by wearable systems to motion artifacts, estimation of the respiratory rate (RR) during physical activities is a challenging task. Alternatively, functional Near-Infrared Spectroscopy (fNIRS) can be used, which has been proven less vulnerable to the subject's movements. This paper proposes a fusion-based method for estimating RR during bicycling from fNIRS signals recorded by a wearable system. Firstly, five respiratory modulations are extracted, based on amplitude, frequency, and intensity of the oxygenated hemoglobin concentration (O<sub>2</sub>Hb) signal. Secondly, the dominant frequency of each modulation is computed using the fast Fourier transform. Finally, dominant frequencies of all modulations are fused, based on averaging, to estimate RR. The performance of the proposed method was validated on 22 young healthy subjects, whose respiratory and fNIRS signals were simultaneously recorded during a bicycling task, and compared against a zero delay Fourier domain band-pass filter. The comparison between results obtained by the proposed method and band-pass filtering indicated the superiority of the former, with a lower mean absolute error (3.66 vs. 11.06 breaths per minute,  $p < 0.05$ ). The proposed fusion strategy also outperformed RR estimations based on the analysis of individual modulation. This study orients towards the practical limitations of traditional bio-signals for RR estimation during physical activities.



# 1. Introduction

The advent of wearable systems for the measuring of vital signs has started a revolution in non-laborious healthcare and sports science practices [1]. Unarguably, such systems have provided the possibility of customized continuous monitoring of physiological parameters, such as heart rate (HR), blood pressure, and respiratory rate (RR), which are of great value for both clinical and research purposes [2,3].

A majority of wearable systems indirectly estimate the aforementioned parameters using bio-signals, such as the photoplethysmography (PPG) and electrocardiography (ECG) [4]. Despite the excellent performance during resting state, wearable systems usually do not show a satisfactory performance when the subject is engaging in physical activity [5]. The reason for such failure is the substantial susceptibility of bio-signals to motion artifacts originating from a subject's movements, which are inevitable during physical activity. Amongst the physiological parameters mentioned, estimation of RR from bio-signals contaminated with motion artifacts is the most challenging [6].

Alternatively, functional Near-Infrared Spectroscopy (fNIRS) can be employed and is less sensitive to a subject's movements. Specifically, fNIRS is an optical brain imaging technique that quantifies oxygenated (O<sub>2</sub>Hb) and deoxygenated (HHb) hemoglobin concentrations, related to the neural activities in the cerebral cortex. Not only does fNIRS provide a better spatial resolution than electroencephalography (EEG), it also outperforms functional magnetic resonance imaging (fMRI) in terms of temporal resolution [7]. More importantly, compared to other bio-signal monitoring techniques, fNIRS is more motion tolerant, i.e., the subject's movements do not necessarily affect the quality of the data [8,9]. Due to its noninvasive and mobile nature, fNIRS also has the advantage of being usable in outdoor applications [10]. Indeed, the advent of lightweight single sensor fNIRS equipment has opened new doors for outdoor studies, and, in particular, fitness applications [11–14].

Opportunely, fNIRS is contaminated by non-cerebral interference cited from physiological systemic activities, such as HR, blood pressure, Mayer waves, and respiration [15], providing the possibility of extraction of complementary physiological parameters. Although the HR estimation from fNIRS measurements has been addressed in several studies, no research, except for our previous study [16], has yet investigated the possibility of RR estimation from fNIRS. Respiration is one of the main physiological interferences in fNIRS, manifesting in its spectrum ( $\sim 0.2$  to  $0.4$  Hz) [17,18]. The reason for such contamination is twofold; changes of the blood flow within the body during inhaling and exhaling, and the influence of respiratory oscillations on cerebral blood volume and flow [19]. Therefore, it can be presumed that RR changes are disclosed in the fNIRS signals.

Due to the similar principles of PPG and fNIRS measurements, and, in particular, O<sub>2</sub>Hb signal, and considering the fact that cerebral blood volume and flow changes are proportional to fNIRS output [20], it can be expected that respiratory modulations exhibited in PPG behave similarly in O<sub>2</sub>Hb signals. Therefore, we hypothesize that respiration can influence the O<sub>2</sub>Hb signal in three aspects: frequency, intensity and amplitude.

In our previous research, [16], we showed the possibility of RR estimation from an O<sub>2</sub>Hb signal in a resting state. However, the main challenge is to estimate RR when subjects are performing physical activities. The objective of this paper is, thus, to propose a new method

for estimation of RR from O2Hb signals during a bicycling task. The basis of the proposed method involves the following: (i) extraction of five respiratory modulations, based on amplitude, frequency, and intensity of an O2Hb signal, (ii) computation of the dominant frequency of each modulation, and (iii) application of a mean-based fusion strategy on the computed dominant frequencies to estimate RR.

## 2. Methods

The block diagram of the proposed method is shown in Figure 1. As displayed, it mainly consists of two blocks: pre-processing and RR estimation.

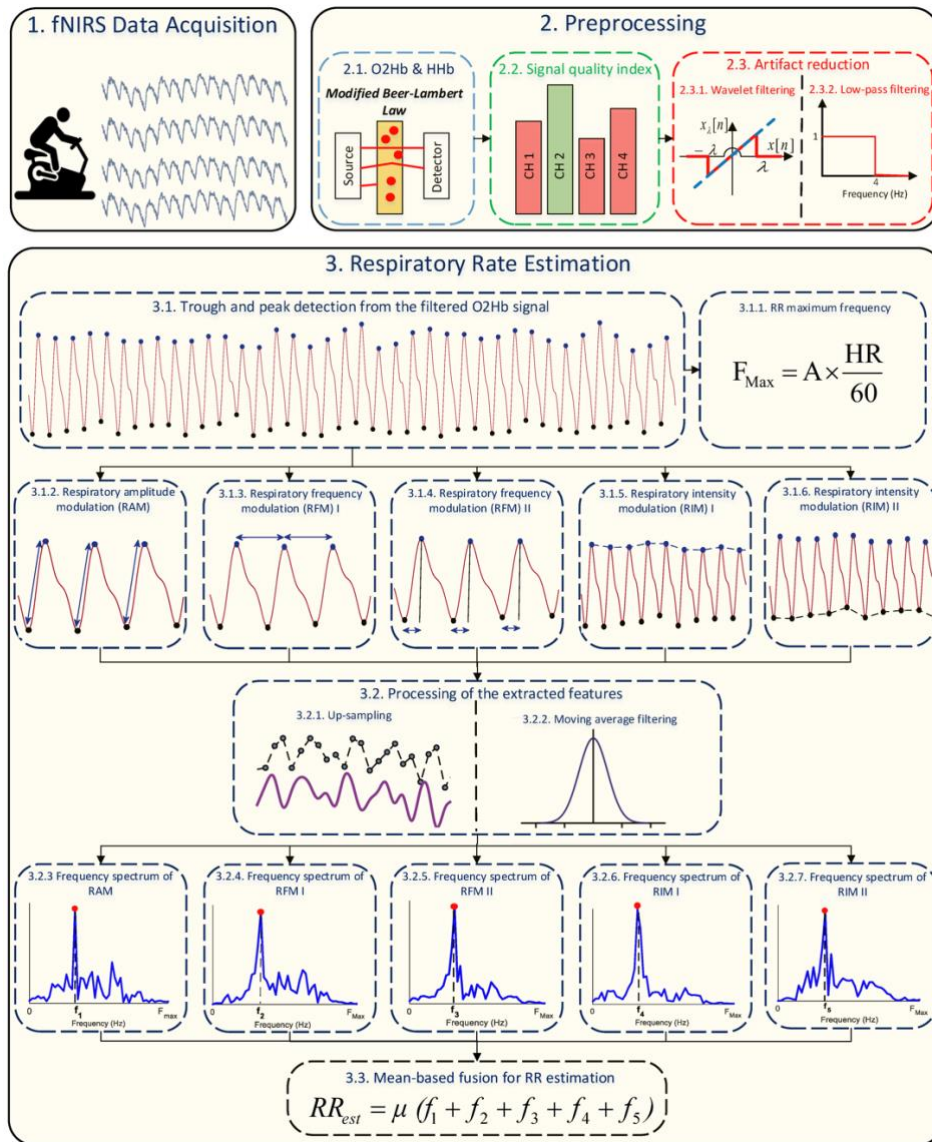


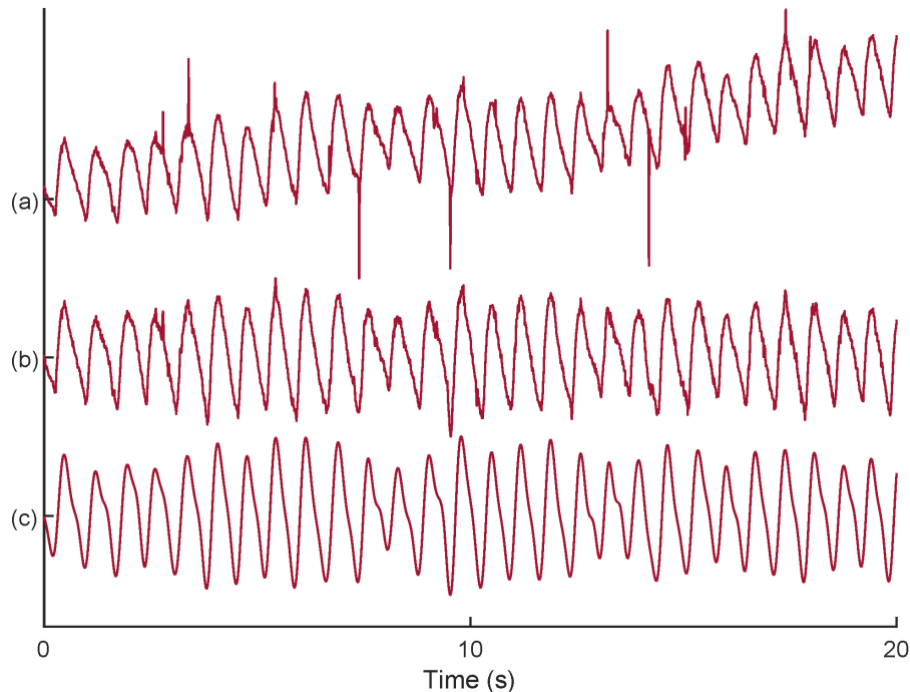
Figure 1: Block diagram of the proposed method for RR estimation using fNIRS.

## 2.1. Preprocessing

The first step of the proposed method is to find a high quality fNIRS channel for the analysis. Towards this aim, firstly, the modified Beer–Lambert law is employed to convert raw optical density signals to O2Hb and HHb changes in concentrations as follows [21]:

$$\Delta OD(\lambda) = \epsilon(\lambda) \cdot \Delta C \cdot L \cdot DPF \quad (1)$$

where  $OD$ ,  $\lambda$ ,  $\epsilon$ ,  $C$ ,  $L$ , and  $DPF$  stand for optical density, wavelength, absorption coefficient, concentration, physical path length, and differential path length factors, respectively. Then, the signal quality index (SQI) algorithm [22] is used, which numerically scores the signals from 1 (low) to 5 (high) quality. After finding the fNIRS channel with the highest signal quality, the corresponding O2Hb signal (Figure 2(a)) is filtered by a wavelet transform [23] for elimination of possible motion artifacts (Figure 2(b)). Then, an FIR low-pass filter with cut-off frequency at 4 Hz is used to remove high frequency noise originating from non-physiological sources (Figure 2(c)).



**Figure 2:** An example on noise reduction from an O2Hb signal: (a) selected O2Hb signal with the SQI; (b) motion artifact reduction using DWT algorithm; (c) low-pass filtering for high frequency noise removal.

## 2.2. RR Estimation from *O2Hb*

### 2.2.1. Background

As was already mentioned, the O2Hb signal originating from fNIRS functions on the same basis as PPG, and, thus, respiratory modulations in PPG can also appear in the O2Hb signal. It is common knowledge that HR increases during inhalation and decreases during exhalation, and, thus, given the fact that each cycle of the O2Hb signal is a heartbeat, respiration causes frequency modulation in the O2Hb signal [24]. On the one hand, increasing and decreasing of the intrathoracic pressure during inhalations and exhalations leads to a variation in blood volume, known as baseline wander, and, thus, respiration causes intensity modulation in the O2Hb signal [25]. On the other hand, reduction of the blood stroke volume during inhalation,



because of alternation in intrathoracic pressure, leads to changes in the O2Hb signal's strength, causing respiratory amplitude modulation [26].

### 2.2.2. Extraction of Respiratory Modulations

In this paper, two frequency and two intensity respiratory modulations were extracted from the O2Hb signal, along with one amplitude respiratory modulation, for RR estimation. Regarding the frequency modulations, the following were computed: (i) the distance between two consecutive peaks and (ii) the distance between troughs and peaks. For the intensity modulations, two baseline wanders were formed: (i) based on peaks and (ii) based on troughs. Regarding the amplitude modulation, the absolute difference between the magnitude of troughs and peaks were computed.

### 2.2.3. Trough and Peak Detection

In order to localize peaks and troughs, we used our previous method for trough detection from the O2Hb signal, wherein the O2Hb signal is first normalized between  $-1$  and  $1$ , and then, local minima, with values lower than average of the O2Hb signal, are considered as potential troughs. Yet, the existence of motion artifacts may lead to a fallacious analysis. To overcome this issue, the corresponding magnitude of potential troughs were set on a vector, e.g.,  $z[n]$ , and, then, troughs with a lower value than  $Mean(z[n]) + 3 \times std(z[n])$  were discarded [16]. Afterwards, the maxima between two consecutive troughs were considered the peaks.

### 2.2.4. Processing of the Extracted Modulations

After forming the extracted modulations, up-sampling is an essential step, as the modulations were formed based on irregular sampling, whereas regularly sampled signals are often required for the subsequent analysis [27]. To this end, the cubic spline interpolation method was used, which has already been shown to be promising in PPG studies [28].

Before computing the dominant frequency of the up-sampled extracted modulations, it is also necessary to remove non-respiratory fluctuations, e.g., Mayer waves [29], which can hamper the accurate detection of the corresponding RR frequency [27]. Non-respiratory fluctuations can have both very low and very high frequency components. Unfortunately, there is no assent on the ideal frequency range of respiration. Indeed, such a range might need to be regulated based on the subject's background, e.g., age [30]. In our previous study, it was shown that low-frequency components of non-respiratory fluctuations can be removed by a 3s moving average filter. For the maximum frequency of RR, we adopted a threshold based on the HR [31] as follows:

$$F_{Max} = A \times \frac{HR}{60} \quad (2)$$

where  $A$  is a constant. It is worth mentioning that HR was estimated based on the number of detected peaks on a 50s window length. The regulation of the constant  $A$  is described in Section 4.

After applying the moving average filter on the up-sampled modulations, the fast Fourier transform (FFT) algorithm was computed and the dominant frequency in the range of  $[0, F_{Max}]$  considered to be the RR.

### 2.2.5. Mean-Based Fusion Method for RR Estimation

According to several studies, there is no optimal modulation for RR estimation [32]. For example, it was shown that frequency modulations may not be efficient for elderly subjects [33]. Therefore, computing the RR based on a fusion strategy, where several modulations are combined, can be more effective. Here, we used a simple mean fusion method where the final RR is computed based on the average of computed dominant frequencies of all modulations.

## 2.3. Method under Comparison

To the best of our knowledge, there is no direct method to estimate RR from fNIRS. Yet, some studies tried to portion out respiratory components from fNIRS measurements [19,34]. In this paper, we compared performance of the proposed method to a band-pass filtering strategy to extract respiratory components from the O2Hb signal. As described in [19], a zero delay Fourier domain band-pass filter, with cut-off frequencies from 0.2 to 0.6 Hz, was used. In this paper, after finding the fNIRS channel with the highest signal quality, the corresponding O2Hb signal was filtered and the dominant frequency within the range [0.2 to 0.6] Hz considered to be the estimated RR.

## 2.4. Performance Index

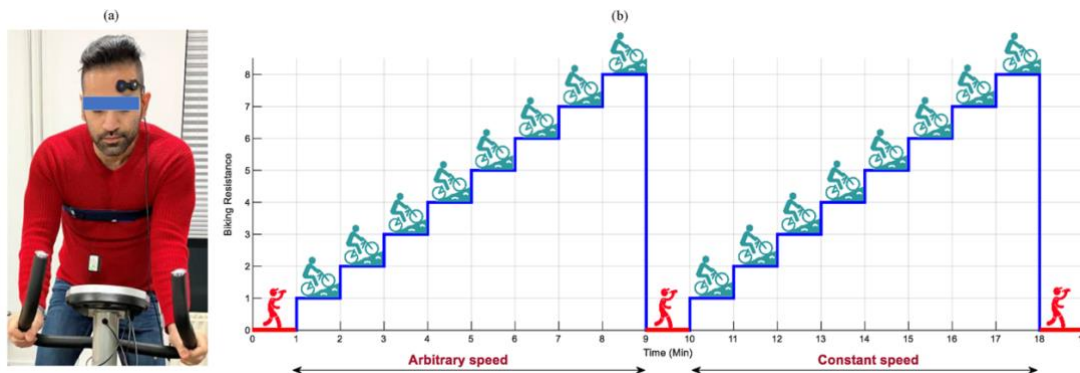
In order to assess the performance of RR estimation using both methods, the absolute error (AE) between the reference and estimated RRs was computed as follows:

$$AE = |RR_{ref} - RR_{est}| \quad (3)$$

## 3. Data Collection

### 3.1. The Experimental Setup

The experimental setup involved an ergometer bike, a wireless single-sensor fNIRS-system (PortaLite MKII, Artinis Medical Systems B.V., the Netherlands), placed on the left hemispheres of the prefrontal cortex, and a respiratory chest belt connected to a TMSi SAGA 32+/64+ amplifier (Twente Medical Systems International B.V., The Netherlands), with a sampling rate of 4000 Hz (Figure 3(a)). The fNIRS system was equipped with three long channels (transmitter–receiver distance up to 41 mm), recording data with nominal wavelengths of 760 and 850 nm, and a sampling rate of 100 Hz.



**Figure 3:** The data recording procedure. The experimental (a): setup and (b): paradigm.

## 3.2. The Experimental Paradigm for Data Recording

The data recording procedure is displayed in Figure 3(b). It comprised of a 60 s resting period block, followed by two bicycling blocks detached by a 60 s resting period, and the final resting period block lasting for 60 s. At each bicycling block, subjects bicycled for 8 min, wherein every minute the resistance of the ergometer increased by 1. In the first bicycling block, subjects were allowed to regulate their bicycling speed arbitrarily. In the second bicycling block, the subjects were asked to keep the speed constant around 18–22 km/h. In the whole procedure, subjects were allowed to freely move their heads and arms.

## 3.3. Participants

Twenty-two young healthy subjects with varied ethnicity and fitness endurance volunteered to take part in the experiment (Table 1). None of the subjects had any psychological and neurological disorders nor were they under any specific medications in the week of the experiment. Before starting the experiment, all participants were briefed about the study and signed a consent letter. A waiver for formal ethical approval was obtained from the medical ethics committee (CMO Arnhem-Nijmegen) with the ethical code of 2021-8284. All regulations and guidelines needed by the Declaration of Helsinki were followed. The data were collected at the premises of Artinis Medical Systems B.V., Elst, The Netherlands.

**Table 1:** Characteristics of subjects who participated in the experiment. The BMI stands for the body mass index.

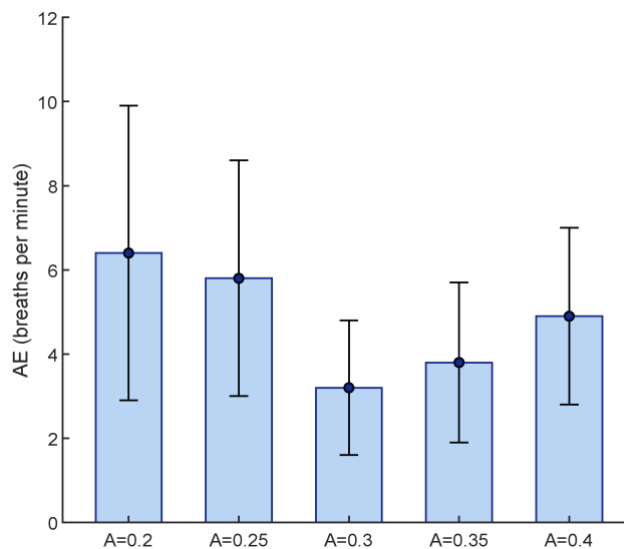
Subject	BMI	Age	Blood Pressure Status	Sex
1	23.4	37	Hypotension	Female
2	22.6	37	Hypotension	Male
3	20.4	24	Normotension	Male
4	18.6	33	Normotension	Female
5	25.6	22	Normotension	Male
6	19.3	22	Normotension	Male
7	19.8	26	Normotension	Male
8	21.4	31	Hypotension	Female
9	24.2	27	Hypotension	Female
10	22.3	24	Normotension	Female
11	26.8	24	Normotension	Male
12	25.4	32	Normotension	Female
13	23.8	22	Normotension	Male
14	26.2	28	Normotension	Male
15	23.4	23	Normotension	Male
16	21.8	22	Normotension	Male
17	26.7	24	Normotension	Female
18	22.9	22	Normotension	Female
19	20.8	25	Normotension	Female
20	21.5	24	Normotension	Female
21	26.4	24	Normotension	Male
22	24.6	33	Normotension	Female

## 4. Results

In this section, the experimental results obtained from the proposed method and the method under comparison are described. Before starting the analysis, the respiratory reference signals were manually inspected and trials with an unreliable RR were rejected. After the manual inspection, the data associated with subject 21 were completely discarded. It should be noted that after discarding unreliable respiratory signals, 418 trials with a length of 50 s were used for the analysis.

### 4.1. Regulation of RR Maximum Frequency

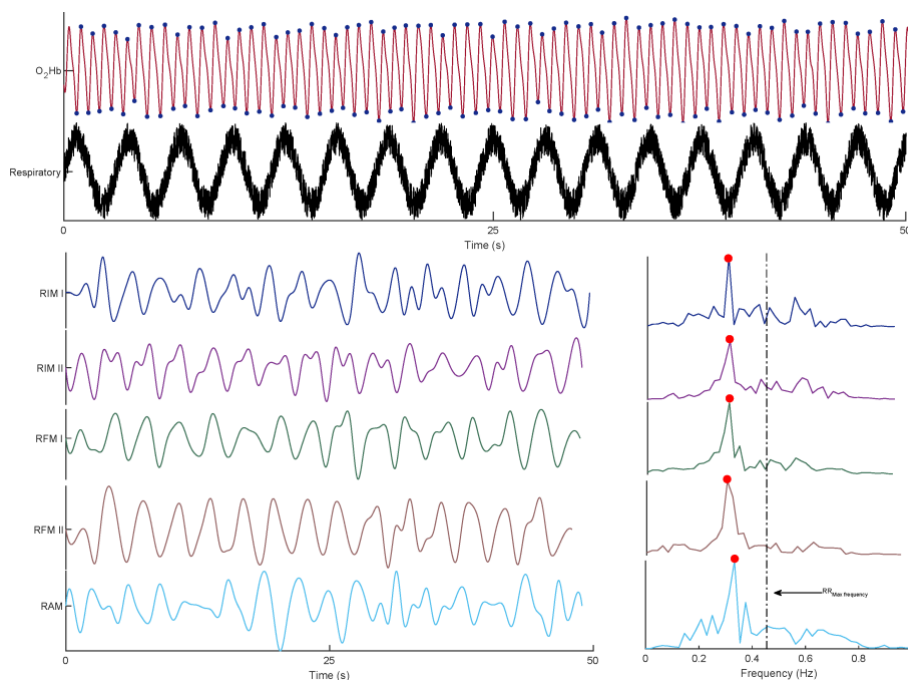
The only parameter of the proposed method that was required to be tuned was the constant  $A$ , which regulated the RR maximum frequency based on HR. For this purpose, 78 trials, having 50 s lengths, from five subjects were randomly selected and values from 0.2 to 0.4 with a step size of 0.05 were investigated. To find the best fit, the lowest mean AE between the reference and estimated RR was considered (Figure 4). Although there was no significant difference between the obtained results by  $A = 0.3$  and  $A = 0.35$ , we considered  $A = 0.3$  to be the best fit, due to it having the lowest mean AE.



**Figure 4:** Tuning the constant  $A$  for determination of RR maximum frequency, based on lowest mean AE obtained from 78 randomly selected trials. The AE is shown in terms of mean and standard deviations.

### 4.2. The Performance of Proposed Method

Figure 5 shows an example of the O<sub>2</sub>Hb signal and its corresponding modulations, followed by the frequency spectrum of each modulation. It should be noted that the reference RR was 0.32 Hz.



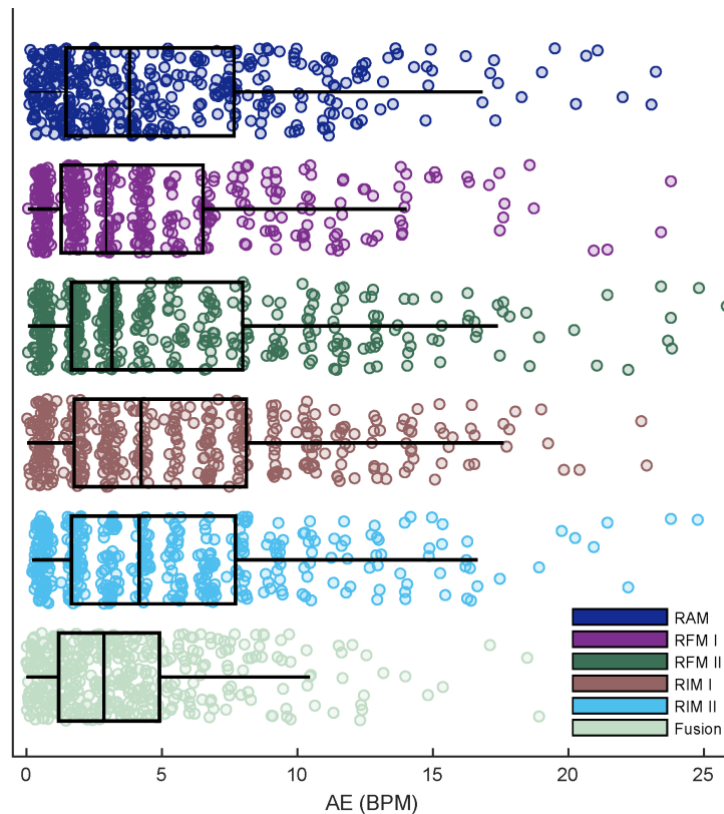
**Figure 5:** An example of the O<sub>2</sub>Hb signal and its corresponding respiratory signal, as well as its extracted respiratory modulations and spectra.

Table 2 displays the mean AE between the reference and estimated RRs for each subject, indicating the superiority of the fusion-based strategy over the individual modulation for the majority of the subjects.

**Table 2:** The mean AE between reference and estimated RRs (BPM) for each subject with individual and fusion features. The lowest mean AE is in bold.

Subject	RAM	RFM I	RFM II	RIM I	RMI II	Fusion
1	1.45	1.30	1.46	2.17	2.23	<b>1.05</b>
2	8.54	8.72	6.30	7.79	7.17	<b>6.17</b>
3	4.85	3.72	6.91	6.27	6.9	<b>3.83</b>
4	5.62	4.01	3.27	4.95	5.23	<b>3.08</b>
5	3.29	2.53	2.75	2.53	3.25	<b>1.35</b>
6	6.48	7.78	8.50	7.60	7.77	<b>5.35</b>
7	6.51	6.76	5.15	5.49	7.02	<b>5.11</b>
8	<b>3.09</b>	3.25	4.91	5.66	3.89	3.75
9	5.96	<b>3.99</b>	10.96	9.16	8.07	5.10
10	8.66	5.38	5.18	7.47	7.68	<b>5.00</b>
11	5.98	4.31	6.21	6.93	5.43	<b>4.31</b>
12	5.47	3.61	5.01	5.05	3.9	<b>2.81</b>
13	4.97	3.09	6.09	6.47	5.31	<b>3.35</b>
14	7.62	6.82	5.63	3.98	<b>3.92</b>	4.99
15	5.87	7.30	5.62	<b>3.38</b>	5.86	3.98
16	5.73	3.67	2.88	2.54	3.00	<b>2.00</b>
17	<b>2.87</b>	3.56	5.84	8.36	6.98	3.61
18	5.34	4.07	2.40	3.73	3.84	<b>3.13</b>
19	4.73	4.03	7.26	6.98	5.09	<b>3.47</b>
20	3.71	3.69	6.85	8.59	5.31	<b>3.67</b>
21	<b>2.72</b>	2.82	4.87	4.36	3.60	2.75

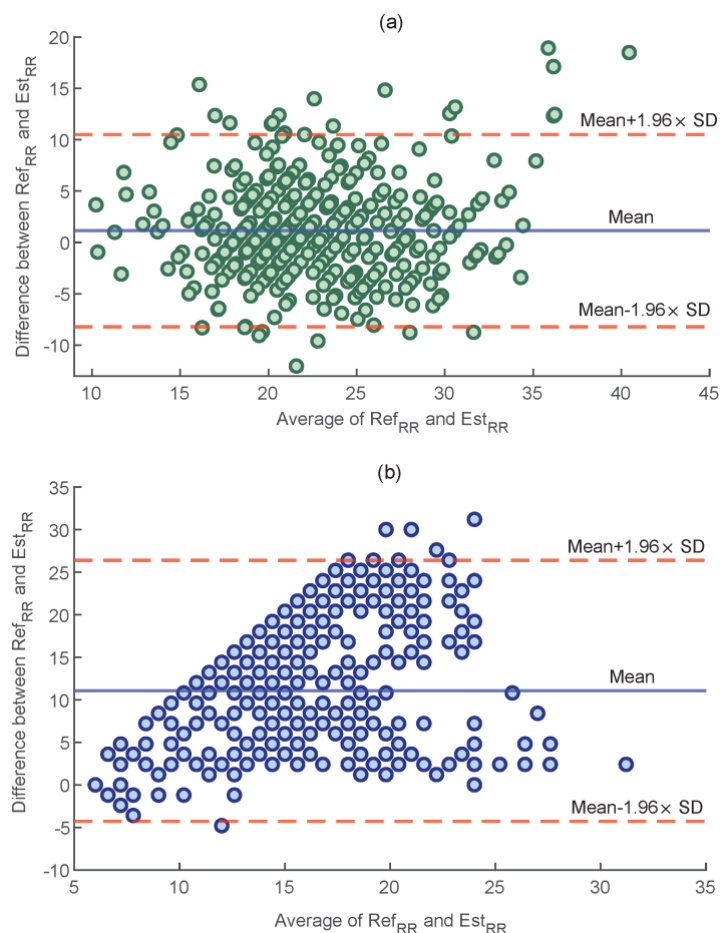
When comparing the performance of each individual modulation with the fusion-based strategy on all trials (Figure 6), the latter showed a lower mean AE (3.66) against those obtained from RAM (5.14), RFM I (4.45), RFM II (5.32), RIM I (5.61), and RIM II (5.22) BPM. According to the conducted t-test analysis, regarding the individual modulation, there was a significant difference between the AE of RFM I and others ( $p < 0.05$ ). Notably, there was also a significant difference between the fusion-based strategy and RFM I ( $p < 0.05$ ).



**Figure 6:** The overall performance of each individual feature and fusion-based strategy in terms of AE.

### 4.3. Comparison with the Band-Pass Filtering Method

To compare the performance of both methods, a Bland–Altman plot was used, which evaluates the agreement between two quantitative measurements by displaying their difference as a function of their average. Figure 7 shows the Bland–Altman plot of the AE values obtained by the proposed method and the method based on band-pass filtering, as used in [19] for all trials. As displayed, the proposed method outperformed the band-pass filtering with a lower average of the differences (3.66 vs. 11.06 BPM,  $p < 0.05$ ).



**Figure 7:** The Bland–Altman plots for: (a) the proposed method and (b) band–pass filtering [19].

## 5. Discussion

The objective of this paper was to propose a robust RR estimation method during bicycling using fNIRS signals recorded by portable equipment. This study, by taking advantage of the lower vulnerability of fNIRS to a subject’s movements, was oriented towards highlighting the practical limitations of the current state-of-the-art studies for RR estimation during physical activity. The proposed method provides an extra measure, i.e., RR, that can be used alongside brain activity analysis. Genuinely, RR estimation from fNIRS can synergize the cerebral activity analysis in applications, such as stress assessment and mental workload, where RR changes are important indicators [35,36].

### 5.1. The Significance of Proposed Fusion Strategy

According to the literature, the performance of each extracted respiratory modulation for RR estimation depends on several factors, such as age, gender, body position, and health status [27]. Thus, a fusion-based estimation, i.e., one employing multiple respiratory modulations, can lead to a more robust RR estimation. In this paper, a fusion-based strategy was provided with an average of five respiratory modulations, based on the amplitude, intensity and frequency of an O<sub>2</sub>Hb signal, over a 50 s window. The comparison between the individual and fusion-based RR estimations confirmed such robustness as the 15 subjects showed lower



mean AE (Table 2). Furthermore, the comparison also confirmed that the performance of an individual modulation may not be generalized for the whole population, as each modulation performed differently for different subjects. In terms of overall performance over all the trials, the fusion-based strategy outperformed each individual modulation with a lower mean of AE.

## 5.2. Comparison with State-of-the-Art

Except for our previous study, that estimated RR from an O<sub>2</sub>Hb signal in a resting state, there is no research estimating RR from fNIRS. Yet, a few studies have tried to portion out the respiratory components from fNIRS measurements. For instance, a multimodal extension of the general linear model using a temporal embedded canonical correlation method was employed to regress out respiratory components from fNIRS [34]. Firstly, the mentioned study required a reference respiratory signal to be recorded concurrently. Secondly, employing a general linear model usually requires a large number of fNIRS channels, which increases the complexity of wearable instrumentation. In contrast, we used a single sensor fNIRS system placed on the prefrontal cortex, which reduces the complexity of wearable instrumentation. Furthermore, the prefrontal cortex is a hairless area, and, therefore, provides the user with more comfort.

In other research, a zero delay Fourier domain band-pass filter, with cut-off frequencies from 0.2 to 0.6 Hz, was applied to the O<sub>2</sub>Hb signal to separate respiratory components from the cerebral brain activity [19]. The comparison between performance of the proposed method and band-pass filtering method (Figure 7) indicates the superiority of our proposed method, as a lower mean of difference between the reference and estimated RRs was obtained (3.66 vs. 11.06 BPM). The reason for such a high error could be the fact that the suggested bandwidth also includes other non-respiratory components, such as cerebral and cardiac activities.

## 5.3. Directions for Future Work

Although the results obtained by the proposed method are promising, there are some issues that should be addressed in future work. Firstly, performance of the proposed method was only investigated on young healthy subjects. It is important to consider a broader range of subjects (e.g., the elderly) to further assess the robustness of the proposed method. Furthermore, it may also be important to consider different characteristics of subjects, e.g., race, skin color, and BMI in developing an RR estimation algorithm. In particular, the employed maximum respiratory frequency ( $F_{max}$ ) might require further adjustment for a different cohort. Secondly, although the data were recorded during physical activity and the subjects were allowed to freely move their bodies throughout the experiment, performance of the proposed method should be further validated on data where the subjects are performing more intense exercise, e.g., squatting, with more movements. Lastly, employing a smart fusion strategy, such as [37], may also further improve the performance of the proposed method. However, we did not find enormous differences amongst the estimated RRs from individual modulations (Table 2).

## 6. Conclusion

This paper presented a fusion-based strategy for RR estimation during a bicycling task from O<sub>2</sub>Hb signals originating from fNIRS measurements. For this purpose, five respiratory modulations were extracted from an O<sub>2</sub>Hb signal and, then, the RR was estimated, based on the average of dominant frequencies of each modulation. The comparison between the obtained results from the proposed method and band-pass filtering showed the superiority of the proposed method with a lower mean of AE (3.66 vs. 11.03 BPM). The importance of this study was that it addressed RR estimation during physical activity, and, by taking advantage of the motion tolerance quality of fNIRS, was oriented towards overcoming the practical limitations of traditional bio-signals, i.e., vulnerability to the subject's movements. Furthermore, the proposed method derives an extra parameter from fNIRS signals, i.e., RR, which can be used alongside brain activity analysis.

## References

1. Wang, Y.; Miao, F.; An, Q.; Liu, Z.; Chen, C.; Li, Y. Wearable Multimodal Vital Sign Monitoring Sensor With Fully Integrated Analog Front End. *IEEE Sens J* **2022**, *22*, 13462–13471, doi:10.1109/JSEN.2022.3177205.
2. De Fazio, R.; Mastronardi, V.M.; De Vittorio, M.; Visconti, P. Wearable Sensors and Smart Devices to Monitor Rehabilitation Parameters and Sports Performance: An Overview. *Sensors* **2023**, *Vol. 23*, Page 1856 **2023**, *23*, 1856, doi:10.3390/S23041856.
3. Moshawrab, M.; Adda, M.; Bouzouane, A.; Ibrahim, H.; Raad, A. Smart Wearables for the Detection of Cardiovascular Diseases: A Systematic Literature Review. *Sensors* **2023**, *Vol. 23*, Page 828 **2023**, *23*, 828, doi:10.3390/S23020828.
4. Rathore, K.S.; Vijayarangan, S.; SP, P.; Sivaprakasam, M. A Multifunctional Network with Uncertainty Estimation and Attention-Based Knowledge Distillation to Address Practical Challenges in Respiration Rate Estimation. *Sensors* **2023**, *Vol. 23*, Page 1599 **2023**, *23*, 1599, doi:10.3390/S23031599.
5. Lanata, A.; Scilingo, E.P.; Nardini, E.; Loriga, G.; Paradiso, R.; De-Rossi, D. Comparative Evaluation of Susceptibility to Motion Artifact in Different Wearable Systems for Monitoring Respiratory Rate. *IEEE Transactions on Information Technology in Biomedicine* **2010**, *14*, 378–386, doi:10.1109/TITB.2009.2037614.
6. Alam, R.; Peden, D.B.; Lach, J.C. Wearable Respiration Monitoring: Interpretable Inference with Context and Sensor Biomarkers. *IEEE J Biomed Health Inform* **2021**, *25*, 1938–1948, doi:10.1109/JBHI.2020.3035776.
7. Mihajlovic, V.; Patki, S.; Xu, J. Noninvasive Wearable Brain Sensing. *IEEE Sens J* **2018**, *18*, 7860–7867, doi:10.1109/JSEN.2018.2844174.
8. Almajidy, R.K.; Mankodiya, K.; Abtahi, M.; Hofmann, U.G. A Newcomer's Guide to Functional Near Infrared Spectroscopy Experiments. *IEEE Rev Biomed Eng* **2020**, *13*, 292–308, doi:10.1109/RBME.2019.2944351.
9. Cui, X.; Baker, J.M.; Liu, N.; Reiss, A.L. Sensitivity of FNIRS Measurement to Head Motion: An Applied Use of Smartphones in the Lab. *J Neurosci Methods* **2015**, *245*, 37–43, doi:10.1016/J.JNEUMETH.2015.02.006.
10. Joshi, S.; Herrera, R.R.; Springett, D.N.; Weedon, B.D.; Ramirez, D.Z.M.; Holloway, C.; Dawes, H.; Ayaz, H. Neuroergonomic Assessment of Wheelchair Control Using Mobile FNIRS. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2020**, *28*, 1488–1496, doi:10.1109/TNSRE.2020.2992382.
11. Pellegrini-Laplagne, M.; Dupuy, O.; Sosner, P.; Bosquet, L. Effect of Simultaneous Exercise and Cognitive Training on Executive Functions, Baroreflex Sensitivity, and Pre-Frontal Cortex Oxygenation in Healthy Older Adults: A Pilot Study. *GeroScience* **2022**, *45:1* **2022**, *45*, 119–140, doi:10.1007/S11357-022-00595-3.
12. Germain, C.; Perrot, A.; Tomasino, C.; Bonnal, J.; Ozsancak, C.; Auzou, P.; Prieur, F. Effect of the Level of Physical Activity on Prefrontal Cortex Hemodynamics in Older Adults During Single- and Dual-Task Walking. *J Aging Phys Act* **2022**, *31*, 96–104, doi:10.1123/JAPA.2021-0410.

13. Goenarjo, R.; Dupuy, O.; Fraser, S.; Berryman, N.; Perrochon, A.; Bosquet, L. Cardiorespiratory Fitness and Prefrontal Cortex Oxygenation during Stroop Task in Older Males. *Physiol Behav* **2021**, *242*, 113621, doi:10.1016/J.PHYSBEH.2021.113621.
14. Koren, Y.; Mairon, R.; Sofer, I.; Parmet, Y.; Ben-Shahar, O.; Bar-Haim, S. Vision, Cognition, and Walking Stability in Young Adults. *Scientific Reports* **2022**, *12*, 1–10, doi:10.1038/s41598-021-04540-w.
15. Zhang, F.; Cheong, D.; Khan, A.F.; Chen, Y.; Ding, L.; Yuan, H. Correcting Physiological Noise in Whole-Head Functional near-Infrared Spectroscopy. *J Neurosci Methods* **2021**, *360*, doi:10.1016/J.JNEUMETH.2021.109262.
16. Hakimi, N.; Shahbakhti, M.; Sappia, S.; Horschig, J.M.; Bronkhorst, M.; Floor-Westerdijk, M.; Valenza, G.; Dudink, J.; Colier, W.N.J.M. Estimation of Respiratory Rate from Functional Near-Infrared Spectroscopy (FNIRS): A New Perspective on Respiratory Interference. *Biosensors (Basel)* **2022**, *12*, doi:10.3390/BIOS12121170.
17. Izzetoglu, M.; Holtzer, R. Effects of Processing Methods on FNIRS Signals Assessed during Active Walking Tasks in Older Adults. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2020**, *28*, 699–709, doi:10.1109/TNSRE.2020.2970407.
18. Reddy, P.; Izzetoglu, M.; Shewokis, P.A.; Sangobowale, M.; Diaz-Arrastia, R.; Izzetoglu, K. Evaluation of FNIRS Signal Components Elicited by Cognitive and Hypercapnic Stimuli. *Scientific Reports* **2021**, *11*, 1–15, doi:10.1038/s41598-021-02076-7.
19. Tong, Y.; Lindsey, K.P.; Frederick, B.D. Partitioning of Physiological Noise Signals in the Brain with Concurrent Near-Infrared Spectroscopy and fMRI. *Journal of Cerebral Blood Flow and Metabolism* **2011**, *31*, 2352–2362, doi:10.1038/JCBFM.2011.100/ASSET/IMAGES/LARGE/10.1038\_JCBFM.2011.100-FIG6.JPEG.
20. Afkhami, R.; Walker, F.R.; Ramadan, S.; Wong, R.; Johnson, S.J. Indexing Cerebrovascular Health Using Near-Infrared Spectroscopy. *Scientific Reports* **2021**, *11*, 1–9, doi:10.1038/s41598-021-94348-5.
21. Delpy, D.T.; Cope, M.; Van Der Zee, P.; Arridge, S.; Wray, S.; Wyatt, J. Estimation of Optical Pathlength through Tissue from Direct Time of Flight Measurement. *Phys Med Biol* **1988**, *33*, 1433, doi:10.1088/0031-9155/33/12/008.
22. Sappia, M.S.; Hakimi, N.; Colier, W.N.J.M.; Horschig, J.M. Signal Quality Index: An Algorithm for Quantitative Assessment of Functional near Infrared Spectroscopy Signal Quality. *Biomedical Optics Express*, Vol. 11, Issue 11, pp. 6732–6754 **2020**, *11*, 6732–6754, doi:10.1364/BOE.409317.
23. Chiarelli, A.M.; Maclin, E.L.; Fabiani, M.; Gratton, G. A Kurtosis-Based Wavelet Algorithm for Motion Artifact Correction of FNIRS Data. *Neuroimage* **2015**, *112*, 128–137, doi:10.1016/J.NEUROIMAGE.2015.02.057.
24. BERNTSON, G.G.; CACIOPPO, J.T.; QUIGLEY, K.S. Respiratory Sinus Arrhythmia: Autonomic Origins, Physiological Mechanisms, and Psychophysiological Implications. *Psychophysiology* **1993**, *30*, 183–196, doi:10.1111/J.1469-8986.1993.TB01731.X.

25. Nitzan, M.; Faib, I.; Friedman, H. Respiration-Induced Changes in Tissue Blood Volume Distal to Occluded Artery, Measured by Photoplethysmography. <https://doi.org/10.1117/1.2236285> **2006**, *11*, 040506, doi:10.1117/1.2236285.
26. Meredith, D.J.; Clifton, D.; Charlton, P.; Brooks, J.; Pugh, C.W.; Tarassenko, L. Photoplethysmographic Derivation of Respiratory Rate: A Review of Relevant Physiology. *J Med Eng Technol* **2012**, *36*, 1–7, doi:10.3109/03091902.2011.638965.
27. Charlton, P.H.; Birrenkott, D.A.; Bonnici, T.; Pimentel, M.A.F.; Johnson, A.E.W.; Alastruey, J.; Tarassenko, L.; Watkinson, P.J.; Beale, R.; Clifton, D.A. Breathing Rate Estimation from the Electrocardiogram and Photoplethysmogram: A Review. *IEEE Rev Biomed Eng* **2018**, *11*, 2–20, doi:10.1109/RBME.2017.2763681.
28. Deriving Respiration from the Pulse Photoplethysmographic Signal | IEEE Conference Publication | IEEE Xplore Available online: <https://ieeexplore.ieee.org/abstract/document/6164665> (accessed on 25 March 2024).
29. Lindberg, L.G.; Ugnell, H.; Öberg, P.Å. Monitoring of Respiratory and Heart Rates Using a Fibre-Optic Sensor. *Med Biol Eng Comput* **1992**, *30*, 533–537, doi:10.1007/BF02457833/METRICS.
30. Fleming, S.; Thompson, M.; Stevens, R.; Heneghan, C.; Plüddemann, A.; MacOnochie, I.; Tarassenko, L.; Mant, D. Normal Ranges of Heart Rate and Respiratory Rate in Children from Birth to 18 Years of Age: A Systematic Review of Observational Studies. *The Lancet* **2011**, *377*, 1011–1018, doi:10.1016/S0140-6736(10)62226-X.
31. Li, D.; Zhao, H.; Dou, S. A New Signal Decomposition to Estimate Breathing Rate and Heart Rate from Photoplethysmography Signal. *Biomed Signal Process Control* **2015**, *19*, 89–95, doi:10.1016/J.BSPC.2015.03.008.
32. Li, J.; Jin, J.; Chen, X.; Sun, W.; Guo, P. Comparison of Respiratory-Induced Variations in Photoplethysmographic Signals. *Physiol Meas* **2010**, *31*, 415, doi:10.1088/0967-3334/31/3/009.
33. Charlton, P.H.; Bonnici, T.; Tarassenko, L.; Alastruey, J.; Clifton, D.A.; Beale, R.; Watkinson, P.J. Extraction of Respiratory Signals from the Electrocardiogram and Photoplethysmogram: Technical and Physiological Determinants. *Physiol Meas* **2017**, *38*, 669, doi:10.1088/1361-6579/AA670E.
34. von Lümann, A.; Li, X.; Müller, K.R.; Boas, D.A.; Yücel, M.A. Improved Physiological Noise Regression in FNIRS: A Multimodal Extension of the General Linear Model Using Temporally Embedded Canonical Correlation Analysis. *Neuroimage* **2020**, *208*, 116472, doi:10.1016/J.NEUROIMAGE.2019.116472.
35. Sarlija, M.; Popovic, S.; Jagodic, M.; Jovanovic, T.; Ivkovic, V.; Zhang, Q.; Strangman, G.; Cosic, K. Prediction of Task Performance from Physiological Features of Stress Resilience. *IEEE J Biomed Health Inform* **2021**, *25*, 2150–2161, doi:10.1109/JBHI.2020.3041315.
36. Grassmann, M.; Vlemincx, E.; Von Leupoldt, A.; Mittelstädt, J.M.; Van Den Bergh, O. Respiratory Changes in Response to Cognitive Load: A Systematic Review. *Neural Plast* **2016**, *2016*, doi:10.1155/2016/8146809.

37. Karlen, W.; Raman, S.; Ansermino, J.M.; Dumont, G.A. Multiparameter Respiratory Rate Estimation from the Photoplethysmogram. *IEEE Trans Biomed Eng* **2013**, *60*, 1946–1953, doi:10.1109/TBME.2013.2246160.

# Part II

Physiological Information  
in Neonatal NIRS







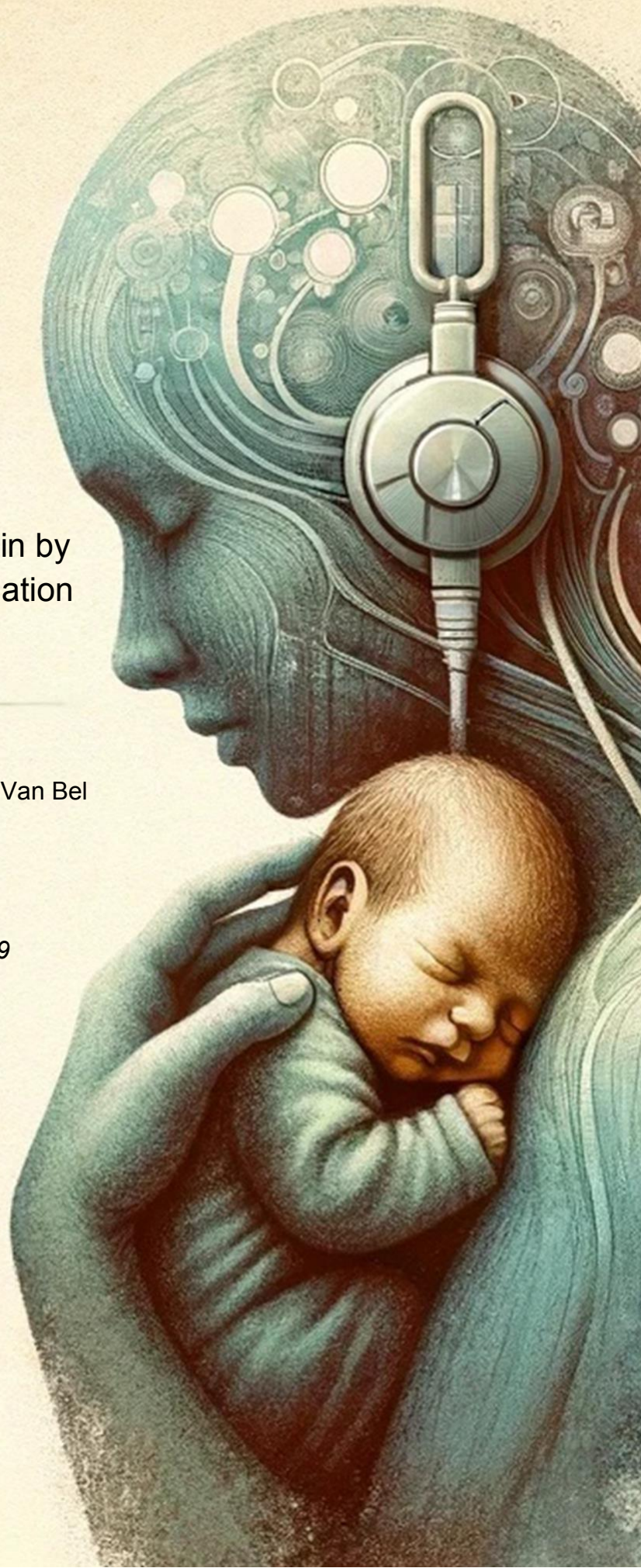
## CHAPTER 6

Neuroprotection of the perinatal brain by  
early information of cerebral oxygenation  
and perfusion patterns

Filipe Gonçalves Costa\*, Naser Hakimi\*, Frank Van Bel

\* shared first author

doi: <https://doi.org/10.3390/ijms22105389>



## Abstract

Abnormal patterns of cerebral perfusion/oxygenation are associated with neuronal damage. In preterm neonates, hypoxemia, hypo-/hypercapnia and lack of cerebral autoregulation are related to peri-intraventricular hemorrhages and white matter injury. Reperfusion damage after perinatal hypoxic ischemia in term neonates seems related with cerebral hyperoxygenation. Since biological tissue is transparent for near infrared (NIR) light, NIR-spectroscopy (NIRS) is a noninvasive bedside tool to monitor brain oxygenation and perfusion. This review focuses on early assessment and guiding abnormal cerebral oxygenation/perfusion patterns to possibly reduce brain injury. In term infants, early patterns of brain oxygenation help to decide whether or not therapy (hypothermia) and add-on therapies should be considered. Further NIRS-related technical advances such as the use of (functional) NIRS allowing simultaneous estimation and integrating of heart rate, respiration rate and monitoring cerebral autoregulation will be discussed.



# 1. Introduction

Reduction and prevention of perinatal brain damage focus, among other measures, mainly on stress reduction (preterm neonate), moderate hypothermia (term neonate) and pharmacological (add-on) therapy [1–3]. However, abnormal oxygenation and/or perfusion patterns often precede brain damage. The occurrence and extension of peri-intraventricular hemorrhages (PIVH) in very and extremely preterm infants (born between 28–32 weeks gestational age and below 28 weeks, respectively), whereas profound and persistent hypoxemia or hyperoxemia often precedes white matter injury (WMI) in this vulnerable group of neonates [4,5]. Likewise, in term neonates experiencing perinatal hypoxia-ischemia or birth asphyxia, (early) cerebral hyperoxygenation, loss of cerebral autoregulation and subsequent perfusion disturbances are related with additional brain damage and adverse 2-year outcome, or even death [6–8]. Early identification of erroneous cerebral oxygenation and/or perfusion patterns may therefore be important for timely corrective interventions, if possible, in order to prevent or at least reduce damage to the developing brain.

Biological tissue is transparent for the near infrared light part of the spectrum (700–1000 nm) and its subsequent absorption by oxygenated and deoxygenated hemoglobin makes (continuous-wave) near-infrared spectroscopy (NIRS) an ideal bedside tool to noninvasively monitor substantial changes in oxygenation and perfusion of the developing brain [9–12]. It actually measures the regional cerebral saturation (rScO<sub>2</sub>) in a mixture of arterial, capillary and venous blood (normally in a ratio of ~25~5~75%, respectively) within the reach of the NIRS sensor attached to the skin of the head [9]. Typically, a banana-shaped area between emitting and receiving optodes is measured. The maximal penetration depth of the near infrared light is between 2 and 3 cm, at least when the optodes distance is at least 4 cm [13]. It is expected that white and grey matter are both “insonated”, at least in the preterm infants. This may be less straight forward when dealing with full-term infants. Oxygenation of the different brain regions seem to be quite homogeneous and did not substantially differ as investigated by us [14], suggesting that a single sensor placement is reliable to determine global rScO<sub>2</sub>. It is important to state that the clinically used NIRS devices monitor “trends” or changes in rScO<sub>2</sub> rather than absolute values. Earlier studies showed that the precision of the commercially available devices is about 5.2% (ability to measure the same value after repeated measurement attempts) [9,10,15]. Moreover, there are sometimes quite significant differences in readings of rScO<sub>2</sub> by the various available devices based on different basic assumptions or algorithms [10,16]. It is widely accepted, though, that “normal” values of rScO<sub>2</sub> in neonates are between 55–85% [17,18]. In several experimental and clinical studies, it has been shown that sustained episodes of rScO<sub>2</sub> values below 40–45% compromise oxygen metabolism of the developing brain and may cause neuronal damage [19–22].

Despite the reservations with respect to the exactness of NIRS-determined cerebral oxygenation we are of the opinion that “trend monitoring” of the oxygenation of the immature brain at the bedside in this noninvasive manner can be relevant for clinical care. The early recognition of (persistent) abnormal patterns of rScO<sub>2</sub> values and NIRS-derived cerebral oxygen extraction may reflect derailment of cerebral oxygenation and perfusion [15,23,24] and energetic and quick correction, if possible, may protect the immature brain.

In the next sections we will concentrate on the possible benefits for clinical care of early evaluation of the NIRS-monitored oxygen saturation signal (pattern recognition) in extremely

preterm infants [15] and in term infants who experienced perinatal hypoxia and ischemia [8]. Further refinement such as integration with other monitoring parameters such as heart rate, respiratory rate and amplitude-integrated EEG (aEEG) will be briefly discussed, as well as the benefits of integration of the NIRS signal with blood pressure to assess (the lack of) cerebral autoregulation.

## 2. Mechanisms of Brain Injury in the Preterm Infant and Cerebral Oxygenation

In extremely preterm neonates, PIVH is one of the most important reasons for adverse (motor) developmental outcome and death, which has an incidence up to 45% in this population [25,26]. Mostly the bleedings are limited to the germinal matrix or with an additional moderate extension into the lateral cerebral ventricles (grade I and II, respectively, according to the Papile classification [27]). A minority evolves to larger PIVHs, grade III (large amount of blood in the cerebral ventricles which are dilated) or even extension of the blood into the brain parenchyma around the cerebral ventricles (grade IV or venous infarction). The etiology of PIVH is multifactorial and PIVHs developed perinatally and/or extended within 12 h after birth seem strongly associated with pro-inflammatory cytokines and formation of oxidative species [28,29]. However, most PIVHs develop and extend during the first 3 days of life or even later. The combination of vascular immaturity in the microcirculation of the germinal matrix [30] and often substantial increases (and fluctuations) in brain perfusion due to the respiratory distress syndrome (IRDS)-induced acutely increases in pCO<sub>2</sub> and/or hypoxia [31,32], setting the brain for the development of PIVHs and their extension to more severe stages (i.e., grade III and IV) [28,33]. Important in this respect is that intravascular CO<sub>2</sub> diffuses very quickly to the perivascular space of the cerebral resistance vasculature (contrary to bicarbonate: the blood–brain barrier is relatively impermeable to bicarbonate [34]), where it lowers the perivascular pH, inducing acute vasodilation [32]. The fragile autoregulatory ability of the cerebral vascular bed in these extremely immature neonates is further provoked by the vasodilating effect of hypercarbia. This could result in a blood pressure-passive brain circulation contributing to fluctuations in the cerebral perfusion, possible episodes of hypoxia during hypotension and to extensions of already existing bleedings [35,36].

It may be clear that prevention or early recognition and down tuning of cerebral hyper(hypo-)perfusion/-oxygenation and/or a fluctuating pattern of the cerebral blood flow potentially reduce or even prevent the occurrence and extension of PIVHs. In this respect it is especially important to aim for a stable arterial carbon dioxide level within normal limits (i.e., 35–50 mmHg). Monitoring the pattern of cerebral oxygenation using NIRS-determined rScO<sub>2</sub> as early after birth as possible, preferably already on the resuscitation table, at least up to postnatal day 4 in these vulnerable group of neonates can alert the clinician at an early point in time for hypercapnia-induced hyperoxygenation/hyperperfusion [25,35] and hypoxia- or hypocapnia-induced underoxygenation/hypoperfusion [32] of the immature brain. A very recent study from Tamura et al. [37] showed that hypercapnia in the delivery room was associated with the occurrence of all grades of PIVH in a large cohort of preterm infants. Alderliesten et al. [35] showed supranormal rScO<sub>2</sub> values [17] during 24 h preceding PIVH in a case–control study of 650 preterm infants below 32 gestational weeks. In the same study it was shown that the combination of hyperoxygenation (hyperperfusion) and loss of

autoregulation preceded the extension of PIVHs to a higher grade. It was also reported that inotropes were related with loss of cerebral autoregulation (defined as a moving correlation  $>0.5$  between rScO<sub>2</sub> and mean arterial pressure). Other potentially more precise mathematical models to describe presence or absence of cerebral autoregulation such as coherence and transfer function gain are intensively investigated on their clinical relevance [38], which will be discussed in more detail in another section of this paper. A follow-up study at 15 and 24 months corrected age of a cohort of 734 preterm infants showed an association between low rScO<sub>2</sub> values and adverse neurodevelopmental outcome at 2 years of age [39]. A recent prospective study in 185 extremely preterm neonates using NIRS-monitored rScO<sub>2</sub> showed that prolonged cerebral desaturation was associated with the occurrence of any grade of PIVH [40]. So, early intervention to quickly normalize the arterial carbon dioxide, cerebral oxygen saturation levels, and prevention of large fluctuations in cerebral perfusion due to lack of autoregulatory ability of the cerebral vascular bed (“hands off” policy and/or sedation) will probably lower the PIVH incidence and its severity [41]. Up to now, only one prospective randomized intervention study in which in 166 preterm infants born after less than 28 weeks cerebral oxygenation was monitored with NIRS (study group) or not (control group) aiming to reduce episodes of hypo- or hyperoxia during the first 3 days of life [42]. It was reported that monitored infants, contrary to those infants whose monitors were blinded, lowered the burden of hypoxia and/or hyperoxia and showed a more stable cerebral oxygen saturation which was mostly within normal limits (55–85%) [17]. This decreased the incidence of particularly severe PIVHs (grade III and IV) [43].

Although the incidence of cystic periventricular leukomalacia has greatly decreased in modern neonatal intensive care units, diffuse white matter injury (dWMI) occurs in about 50% of the extremely and very preterm infants according to MRI studies [44,45]. It has been suggested that an aborted maturation of the very vulnerable precursors of oligodendrocytes and by that arrest of myelination is responsible for the dWMI [46] and a very important reason for a suboptimal neurodevelopmental outcome of this patient group [47]. dWMI is often related to maternal and/or fetal infection or fetal hypoxia leading to pro-inflammatory cytokines and increased activity of microglia disturbing the development of oligodendrocyte precursors [48]. However, postnatal issues and in particular hypoxia with or without ischemia are also strongly related to dWMI. A relatively important role postnatally plays acute hypocapnia (and to a lesser extend hyperoxia) leading to instantaneous vasoconstriction of the cerebral vascular bed with consequent hypoperfusion, and thus lower oxygen delivery [5,32]. Especially abnormally low rScO<sub>2</sub> values in a stable but artificially ventilated infant is suspicious for a ventilation-induced hypocapnia leading to ischemic-hypoxic damage of the immature brain [49]. The combination of NIRS-monitored cerebral oxygenation, arterial oxygen saturation, and end-tidal CO<sub>2</sub> monitoring, increasingly used in standard clinical care in newborn intensive care units, might be a strong preventive policy against hypocapnia-induced WMI.

Some caution is also justified in case of a longstanding hemodynamically Patent Ductus Arteriosus (hsPDA), especially at present where an expectative policy with regard to (surgical) closure of the duct becomes mainstream [50,51]. A hsPDA gives rise to a ductal steal of blood originally destined for the brain with potential underperfusion/oxygenation of the developing immature brain and may not match metabolic demands [52]. In a study with a relatively large population of infants with hsPDA, we found abnormally low values of rScO<sub>2</sub>, often below 45% (so well below the lower limit of “normal” values of 55% [17]). Persistent values lower than 40–45% are reported to be related with brain damage [19,53]. We indeed found an association between low cerebral oxygenation and duration of a hsPDA on the one hand and a lower

volume of especially the cerebellum, a brain region with a remarkably high metabolism, on the other hand [53–55]. This finding confirmed earlier studies which reported negative effects of hsPDA on cerebellar growth in these immature neonates [22,56]. This suggests that longstanding hsPDAs may not always be as innocent as earlier believed. In our NICU all infants with signs of an hsPDA the cerebral oxygenation will be monitored for reasons as explained above.

The definition of “hypotension” of prematurity, mostly defined as a mean arterial pressure lower than the gestational age in weeks, gives way to overtreatment with positive inotropes. Several studies report the absence of signs of organ system dysfunction, metabolic acidosis or NIRS-monitored under-oxygenation of the brain during these “hypotensive” periods [35,57]. Moreover, long-term developmental outcome was not different compared to infants without hypotension [57,58]. At present our policy toward “hypotension” is expectative when NIRS-determined oxygenation of the brain is within normal limits with no signs of systemic under-perfusion or acidosis.

Other issues in which NIRS-monitored cerebral oxygenation can help to prevent injury to the immature brain are anemia and hypoglycemia: anemia with hemoglobin values below 6 mmol/L is reported to affect proper oxygenation of the preterm brain [59], whereas profound hypoglycemia can be recognized by spontaneously increasing levels of cerebral oxygenation due to depletion of glucose leading to a hampering of metabolism and hence oxygen utilization [60].

### 3. Mechanisms of Brain Injury of the Term Infant and Cerebral Oxygenation

Hypoxic-ischemic encephalopathy (HIE) due to perinatal asphyxia is a major cause of adverse neurodevelopmental outcome and neonatal mortality [61]. Although neuronal damage occurs during the actual (mostly fetal) hypoxic episode(s) due to excessive formation of excitatory neurotransmitters, profound calcium influx into the neuronal cells leading to acute oxidative damage to the cell membrane, a substantial part of birth asphyxia-induced brain injury occurs during the first 48 to 96 h of life and can even last for weeks and starts upon reperfusion and reoxygenation at birth [62]. Phosphorus nuclear magnetic resonance spectroscopy showed that during the first hours after reperfusion/reoxygenation upon birth cerebral oxidative metabolism is apparently normal, but secondary cerebral energy failure started after 10 to 12 h of age which seems to be maximal between 24 and 72 h of age leading to substantial additional brain injury [22,62]. Reoxygenation-induced production of free radicals, nitric monoxide and other toxic compounds give not only rise to acute damage to neuronal cells but also to activation of pro-inflammatory cytokine production leading to secondary cerebral energy failure and a decrease in trophic factors [63]. This delayed damage due to secondary energy failure, provides us with a “therapeutic” window in which reduction of brain damage can be achieved by preventing and/or scavenging free radical and nitrosative stress, important toxic substances contributing to secondary energy failure and concomitant brain damage. Until now, moderate hypothermia of the body down to 33.5 °C for 72 h is the only established therapy reducing reperfusion damage after birth asphyxia. However, add-on therapy with (combinations of) pharmacologic neuroprotective compounds are currently investigated: It



seems very likely that the combination of hypothermia with pharmacological compounds contribute to a further reduction of birth asphyxia related brain damage [63].

For an optimal reduction or even prevention of birth asphyxia-induced reoxygenation–reperfusion damage of the developing brain, it is crucial to select infants that will benefit from postnatal treatment. This selection needs to be done as early as possible in this “therapeutic” window (alleged span of time first 6 h of life).

Abnormally high values of rScO<sub>2</sub> (>80%) during the first 72 to 96 h of life in perinatally moderately and severely asphyxiated neonates, probably indicating less utilization of oxygen due to secondary energy failure and ongoing brain damage, are related with adverse long-term neurodevelopmental outcome with a high predictive value at as early as 3 h of life [8,64,65]. The pattern of rScO<sub>2</sub> can, therefore, be an attractive noninvasive means to determine whether or not (add-on) neuroprotective therapy is necessary to reduce brain damage: although not properly investigated yet, a normalization of rScO<sub>2</sub> values within normal limits upon (add-on) therapy may indicate a beneficial effect of treatment.

Since neuroimaging techniques such as cranial ultrasound, nowadays standard implemented in clinical care in the newborn intensive care unit, and (advanced) MRI of the neonatal brain improved substantially in the recent years, perinatal arterial ischemic stroke (PAIS) appeared to occur more often than previously reported and with an incidence of at least in 1:2300 newborns it is no longer a rare complication [66]. PAIS presents itself mostly in the first days after birth with seizures and can lead to cerebral palsy and to cognitive and motor impairments [67,68]. Neuroprotective/neurodegenerative therapeutic options were not existent up to recently and treatment was mostly limited to anti-epileptic therapy. However, nowadays potentially beneficial early neuroprotective intervention strategies are reported such as the use of neurotrophic substances such as erythropoietin and mesenchymal stem cells to boost neuro-regeneration [69,70].

Neuromonitoring, i.e., the combination of NIRS-monitored cerebral oxygenation and amplitude-integrated encephalography (aEEG), has an important prognostic value in PAIS [71]. Moreover, the early pattern of rScO<sub>2</sub> and electrical background pattern and sleep-wake cycling of the aEEG are probably useful in identifying neonates with PAIS who may benefit from neuroprotective interventions and for assessment of beneficial effects of these interventions [71]. rScO<sub>2</sub> is substantially higher in the affected brain half as compared to the contralateral side, probably indicating luxury perfusion in the affected hemisphere [72,73]. The magnitude and duration of this difference was related with neurodevelopmental outcome [71]. Likewise, time to normal aEEG-derived electrical background (continuous normal voltage) of the affected brain half was longer as compared to the healthy brain half. The time to normalization of the electrical background pattern of the affected side was also related to neurodevelopmental outcome [71]. It is conceivable that neuroprotective interventions can shorten the duration of these asymmetries indicating a therapeutic effect. This assumption, however, remains to be investigated.

Table 1 summarizes the association between the nature of preterm and term acquired brain injuries on the one hand and the pattern of cerebral oxygenation on the other.

**Table 1:** Summary of associations between pathological substrate and pattern of cerebral oxygenation.

<b>Development/Extension of PIVH</b>	
-	Hypercarbia-induced cerebral vasodilation: (abnormally) high rScO <sub>2</sub>
-	Lack of CAR: Blood pressure passive fluctuating pattern of rScO <sub>2</sub>
<b>(Cystic) Diffuse White Matter Injury</b>	
-	Hypocarbia-induced cerebral vasoconstriction: (abnormally) Low rScO <sub>2</sub>
-	Hyperoxia-induced formation of cytokines/free rad: Low-to-normal rScO <sub>2</sub>
-	hsPDA-related ductal steal of brain perfusion: prolonged episodes of low (<45%) rScO <sub>2</sub>
-	Anemia-induced hypoxemia of the preterm brain: Low rScO <sub>2</sub>
-	Hypoglycemia-related disturbance of glucose metabolism: increasing rScO <sub>2</sub> values
<b>Perinatal Ischemia-Hypoxia (Birth Asphyxia)</b>	
-	Secondary energy failure-induced: (abnormally) high rScO <sub>2</sub> (from birth up to 72–96 h of age)
<b>Perinatal Arterial Ischemic Stroke</b>	
-	Luxury perfusion-induced: higher rScO <sub>2</sub> in the ipsilateral hemisphere as compared to contralateral hemisphere

hsPDA: hemodynamically significant Ductus Arteriosus. PIVH: peri-intraventricular hemorrhages. rScO<sub>2</sub>: regional cerebral oxygen saturation.

## 4. Future Perspectives of Multimodal Cerebral Oxygenation Monitoring: What Can It Add to Neuroprotection of the Perinatal Brain in the Clinical Setting?

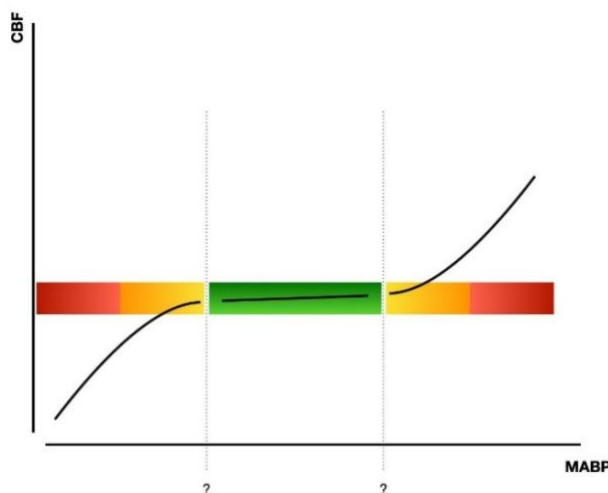
NIRS-derived monitoring of the cerebral oxygenation pattern in the preterm neonate and to a lesser extent also in the term neonate, has undeniable neuroprotective potential as showed above. However, applied research of NIRS-derived surveillance of oxygenation, cerebral perfusion and cerebral autoregulation is progressing quickly with the ultimate goal to improve clinical care.

In this perspective we want to discuss two important developments in perinatal NIRS research which may lead to clinical application in the near future: proper assessment of cerebral hemodynamics regulation which aims to couple systemic (hemodynamic and oxygenation) variables to NIRS signals; a more detailed analysis of NIRS-derived intensity signals for the surge to other usable physiological variables and artifacts detection is reviewed in Sections 5 and 6 below.

## 5. Assessment of Cerebral Hemodynamics Regulation Using NIRS Signals

In regard to cerebral hemodynamics regulation, one must take into account the mechanism of Cerebral Vascular Autoregulation (CVAR)—or pressure-flow reactivity. This property of the brain has the end goal of maintaining the cerebral blood flow at a constant level, regardless of the systemic blood pressure (BP)—but still within a limited range of BP values [74,75]. In the case of the BP going too low (or too high), the CVAR mechanism, through the arterioles,

can no longer regulate the transmural pressure and the blood flow in the brain will, respectively, decrease or increase [76] (see Figure 1).



**Figure 1:** Autoregulatory curve of CBF variations according to mean arterial BP (MABP). In green, the ideal/healthy autoregulatory plateau, and in red/orange the extreme, and therefore dangerous areas of the plot.

Term infants have functional autoregulation and vasoreactivity, but preterm infants are born with an underdeveloped cerebral vascular system [74,77]. This is why the study of CVAR becomes relevant, when one takes into consideration the fact that (very) preterm infants are exposed to a higher risk of hemodynamic instability—and, therefore, brain damage—leading to PIVH or white matter injury [78] as also stated above. For this reason, online assessment and monitoring of the brain tissue oxygenation, combined with the state of the autoregulatory mechanism, is of utmost importance in the frontline of preventing brain injury and further neurodevelopmental disorders [39,79].

By evaluating the relation between BP and Cerebral Blood Flow (CBF) in a continuous way, it is possible to assess the mechanism of CVAR [75]. Since there is not a direct and non-invasive way to measure CBF, and assuming a constant arterial oxygen saturation ( $\text{SaO}_2$ ), the NIRS signal can be used as a surrogate for CBF changes, providing information on the oxygen saturation of the brain ( $\text{rScO}_2$ ) [39,80,81]. By observing the  $\text{rScO}_2$  signal it becomes possible to detect sudden changes and eventually relate them to variations in BP, therefore identifying the presence or absence of the dynamics of autoregulation and providing us the possibility to interfere as already suggesting above [74].

The first steps into this analysis focus on the pre-processing stage, including filtering of the data, artifact removal and  $\text{SaO}_2$  correction [39]. Once this step is performed, there are two main approaches possible: analysis in the time domain or in the frequency domain. Time domain analysis calculates the correlation between the BP and  $\text{rScO}_2$  over short continuous segments, which can then be averages and distributed according to the corresponding BP values, thus having one correlation value per “bin” of blood pressure value [80,82–84]. Frequency domain analysis also explores the relation between these two signals, but with a focus on specific frequency bands [39]. This might come as an advantage when one considers that CVAR may be composed of responses with different time lags, therefore taking into account the delay between changes in BP and  $\text{rScO}_2$  [39]. One of these methods focuses on the Transfer Function (TF), analyzing the power spectral density (PSD) of the signals, which is a representation of the presence of each different frequency in the segments considered. From this, one can assess the TF gain, which indicates the impact that a change in the input

signal (the BP) had in the output signal (the rScO<sub>2</sub>), and therefore assessing if CVAR was indeed present or not [39,79]. Another method of analysis is Coherence: this function measures the linearity of the two signals considered (BP and rScO<sub>2</sub>) in the frequency domain, indicating if there is a linear relationship between the input and the output of the system in focus [39,76,77,80,84].

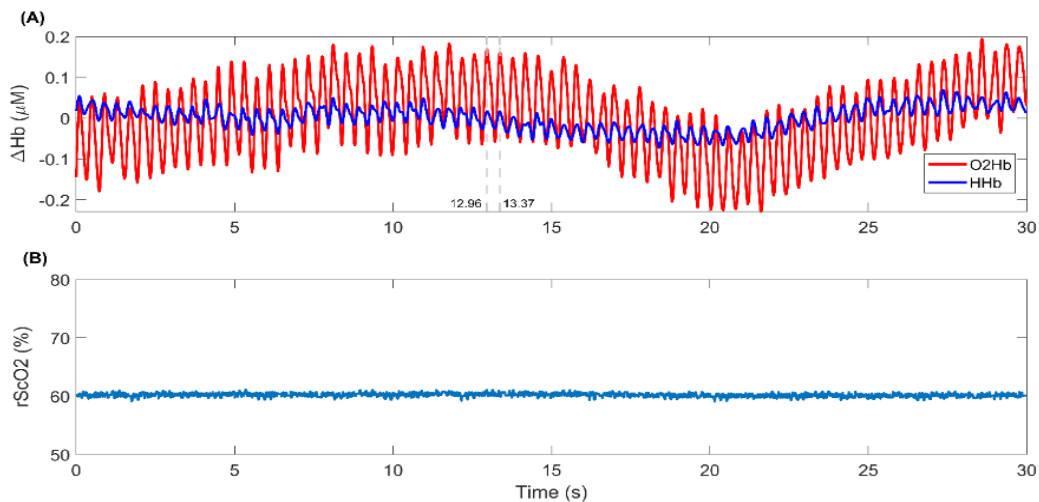
In all the methods described above, one must take into account the length of the epochs considered for analysis, as shorter epochs might produce noisier values—due to a smaller number of sub windows in the analysis—but epochs that are too long may indicate the presence of more nonstationary segments than the ones regarding CVAR, and, therefore, producing lower values and hindering the results [84]. Another factor from the clinical care point of view is that it is important to have in mind that these methods require signals that are continuous and stable enough to provide trustworthy results [85].

We, therefore, must always aim for a stable CVAR mechanism, even though there is still a lack of agreement as to what the best method for analysis is, as well as a proper online monitoring method at the bedside.

Despite the methods presented above, a global consensus is still lacking, as the ideal (online) system should be easily reproducible regardless of the population in focus and the preterm age. Ongoing research of CVAR assessment should also focus on the demonstration of a globally accepted tool or method that monitors CBF and rScO<sub>2</sub>, confirming the relation of the changes in those signals with neurodevelopmental outcome [76].

## 6. NIRS-Derived Intensity Signals and the Importance of Signal Quality Assessment

Using the NIRS technique, different information can be acquired from the intensity of the light measured by the receivers. As stated above cerebral oxygenation (rScO<sub>2</sub>) is the most common information acquired when using NIRS in clinical applications. However, the NIRS-derived intensity signals provide, apart from estimating the cerebral oxygenation, some useful additional functional and physiological information. The intensity signals can be converted into concentration changes in oxygenated hemoglobin (O<sub>2</sub>Hb) and deoxygenated hemoglobin (HHb) [86], which can be used to determine brain hemodynamic response in different brain regions [86–88]. The intensity and concentration signals contain some physiological information including cardiac pulsation, respiration, and vasomotor action [89–92]. Figure 2 shows an example of the O<sub>2</sub>Hb and HHb signals, in which the heartbeats are visibly seen with approximately 0.4-s time intervals, together with the rScO<sub>2</sub> measured by Artinis cerebral oximetry device (TOM, Artinis Medical Systems B.V., Elst, The Netherlands).



**Figure 2:** An example of the NIRS signals recorded by the TOM device (Artinis Medical Systems B.V., Elst, the Netherlands). **(A)** Concentration changes in oxygenated hemoglobin (O2Hb) and deoxygenated hemoglobin (HHb) in red and blue, respectively. These signals have been detrended to have a similar baseline. The heartbeats are visible in both signals as pulsatile changes with approximately 0.4-s time interval. **(B)** The cerebral tissue oxygenation, rScO2, computed by the TOM device in percentage. The signals are sampled at 100 Hz.

## 6.1. NIRS-Derived Heartbeat Pattern

The most prominent physiological information in NIRS signals are the pulsatile fluctuations caused by heartbeats and will be discussed here. This information in the NIRS signals allows for the extraction of heart rate (HR). Several studies have shown the compatibility of the HR derived from NIRS with the HR extracted from electrocardiography and photoplethysmography [93–95]. Moreover, the HR derived from NIRS has shown to be a richer source of information than HR derived from PPG or ECG under physical and mental stressful conditions [92,94–96]. Deriving HR from NIRS signals is highly useful in neonates as it can eliminate the need to use extra electrode for recording this information [97]. Eliminating extra electrodes can relieve neonates from physical stresses induced by electrode placement and potentially reduces injury risks [98].

The HR and brain hemodynamic information in the NIRS signals can be fused in different applications where both brain and heart functions are altered. For instance, they can be combined to accurately assess (physical) stress [92,94,97]. Furthermore, extracting the HR information from NIRS can reduce the ambiguity of the hemodynamic response model [96,99–101]. The extracted HR information can be used as a known physiological term in a part of the model representing different physiological components in the NIRS signal. This can ameliorate the hemodynamic response model and enhance the accuracy of brain hemodynamic response estimation in the clinical setting.

## 6.2. Signal Quality Assessment

In NIRS techniques, the NIR light travels through different tissue layers including intra- and extracranial layers before it strikes the brain. Some factors in the extracranial layer directly affect the optical pathway of the transmitted light. For example, scalp thickness, skin properties, and the presence and color of hair can disrupt the light passing through the head [102,103]. As a consequence, the coupling between NIRS sensors and scalp is decreased, and the NIRS signal quality is then deteriorated [104]. This may result in wrong interpretations and consequent false findings in the analyses. Therefore, assessing the quality of the NIRS

signals before analyzing the signals is important that can increase the reliability of the statistical analysis and decrease misinterpretations.

To assess NIRS signal quality, the presence of a strong cardiac component has been often used as the main indicator of a reliable sensor-scalp coupling [105–108]. A reliable sensor-scalp coupling guarantees that NIR light travels through both intra- and extracranial layers. As a result, the received light (NIRS signals) contains brain hemodynamic and physiological information from both layers. Therefore, the presence of heartbeat, the most prominent systemic information in NIRS signals, indicates that sensors have been well coupled to the scalp and that enough light strikes the brain and is backscattered to the receiver. This knowledge will greatly improve the reliability of the NIRS signal and hence its clinical use.

## 7. Conclusion

Wide-scale utilization of NIRS-derived monitoring of cerebral oxygenation based on timely recognition of an abnormal pattern of cerebral oxygenation seems a relatively straightforward means for prevention and/or reduction of brain injury in very and extremely preterm neonates (i.e., PIVH and dWMI) and to a lesser extent also in the term asphyxiated neonate. Despite this it is very important that the precision of NIRS monitored cerebral oxygen saturation further increases. Equally important is the extended research of NIRS signals to reliably monitor CVAR and the use of the NIRS signal (quality and diversity) to monitor also hemodynamic functions, which contributes to a more informed decision on whether or not to handle the infant at a certain time. These developments may further contribute to the prevention of perinatal brain damage. International collaboration and unified definitions are important to increase utilization of NIRS as a potentially neuroprotective strategy.

## References

1. Van Bel, F.; Vaes, J.; Groenendaal, F. Prevention, Reduction and Repair of Brain Injury of the Preterm Infant. *Front Physiol* **2019**, *10*, 426097, doi:10.3389/FPHYS.2019.00181/BIBTEX.
2. Jacobs, S.; Hunt, R.; Tarnow-Mordi, W.; Inder, T.; Davis, P. Cooling for Newborns with Hypoxic Ischaemic Encephalopathy. *The Cochrane Database of Systematic Reviews (Protocol)* **2003**, doi:10.1002/14651858.CD003311/INFORMATION/EN.
3. van Bel, F. V.; Groenendaal, F. Drugs for Neuroprotection after Birth Asphyxia: Pharmacologic Adjuncts to Hypothermia. *Semin Perinatol* **2016**, *40*, 152–159, doi:10.1053/J.SEMPERI.2015.12.003.
4. Perlman, J.M.; McMenamin, J.B.; Volpe, J.J. Fluctuating Cerebral Blood-Flow Velocity in Respiratory-Distress Syndrome. <http://dx.doi.org/10.1056/NEJM198307283090402> **2010**, *309*, 204–209, doi:10.1056/NEJM198307283090402.
5. Shankaran, S.; Langer, J.C.; Kazzi, S.N.; Laptook, A.R.; Walsh, M. Cumulative Index of Exposure to Hypocarbica and Hyperoxia as Risk Factors for Periventricular Leukomalacia in Low Birth Weight Infants. *Pediatrics* **2006**, *118*, 1654–1659, doi:10.1542/PEDS.2005-2463.
6. Greisen, G. Cerebral Blood Flow and Oxygenation in Infants after Birth Asphyxia. Clinically Useful Information? *Early Hum Dev* **2014**, *90*, 703–705, doi:10.1016/J.EARLHUMDEV.2014.06.007.
7. Howlett, J.A.; Northington, F.J.; Gilmore, M.M.; Tekes, A.; Huisman, T.A.G.M.; Parkinson, C.; Chung, S.E.; Jennings, J.M.; Jamrogowicz, J.J.; Larson, A.C.; et al. Cerebrovascular Autoregulation and Neurologic Injury in Neonatal Hypoxic–Ischemic Encephalopathy. *Pediatric Research* *2013* **74:5** **2013**, *74*, 525–535, doi:10.1038/pr.2013.132.
8. Lemmers, P.M.A.; Zwanenburg, R.J.; Benders, M.J.N.L.; De Vries, L.S.; Groenendaal, F.; Van Bel, F.; Toet, M.C. Cerebral Oxygenation and Brain Activity after Perinatal Asphyxia: Does Hypothermia Change Their Prognostic Value? *Pediatric Research* *2013* **74:2** **2013**, *74*, 180–185, doi:10.1038/pr.2013.84.
9. Van Bel, F.; Lemmers, P.; Naulaers, G. Monitoring Neonatal Regional Cerebral Oxygen Saturation in Clinical Practice: Value and Pitfalls. *Neonatology* **2008**, *94*, 237–244, doi:10.1159/000151642.
10. Greisen, G.; Andresen, B.; Plomgaard, A.M.; Hyttel-Sørensen, S. Cerebral Oximetry in Preterm Infants: An Agenda for Research with a Clear Clinical Goal. <https://doi.org/10.1117/1.NPh.3.3.031407> **2016**, *3*, 031407, doi:10.1117/1.NPH.3.3.031407.
11. Skov, L.; Pryds, O.; Greisen, G. Estimating Cerebral Blood Flow in Newborn Infants: Comparison of Near Infrared Spectroscopy and <sup>133</sup>Xe Clearance. *Pediatric Research* *1991* **30:6** **1991**, *30*, 570–573, doi:10.1203/00006450-199112000-00016.
12. Alderliesten, T.; De Vis, J.B.; Lemmers, P.M.; Hendrikse, J.; Groenendaal, F.; Van Bel, F.; Benders, M.J.; Petersen, E.T. Brain Oxygen Saturation Assessment in Neonates



- Using T 2-Prepared Blood Imaging of Oxygen Saturation and near-Infrared Spectroscopy. *Journal of Cerebral Blood Flow and Metabolism* **2017**, *37*, 902–913, doi:10.1177/0271678X16647737/ASSET/IMAGES/LARGE/10.1177\_0271678X16647737-FIG5.JPEG.
13. Greisen, G. Is Near-Infrared Spectroscopy Living up to Its Promises? *Semin Fetal Neonatal Med* **2006**, *11*, 498–502, doi:10.1016/j.siny.2006.07.010.
  14. Wijbenga, R.G.; Lemmers, P.M.A.; Van Bel, F. Cerebral Oxygenation During the First Days of Life in Preterm and Term Neonates: Differences Between Different Brain Regions. *Pediatric Research* **2011**, *70*, 389–394, doi:10.1203/pdr.0b013e31822a36db.
  15. van Bel, F.; Mintzer, J.P. Monitoring Cerebral Oxygenation of the Immature Brain: A Neuroprotective Strategy? *Pediatric Research* **2018**, *84*, 159–164, doi:10.1038/s41390-018-0026-8.
  16. Dix, L.M.L.; Van Bel, F.; Baerts, W.; Lemmers, P.M.A. Comparing Near-Infrared Spectroscopy Devices and Their Sensors for Monitoring Regional Cerebral Oxygen Saturation in the Neonate. *Pediatric Research* **2013**, *74*, 557–563, doi:10.1038/pr.2013.133.
  17. Alderliesten, T.; Dix, L.; Baerts, W.; Caicedo, A.; Van Huffel, S.; Naulaers, G.; Groenendaal, F.; Van Bel, F.; Lemmers, P. Reference Values of Regional Cerebral Oxygen Saturation during the First 3 Days of Life in Preterm Neonates. *Pediatric Research* **2015**, *79*, 55–64, doi:10.1038/pr.2015.186.
  18. Pellicer, A.; Greisen, G.; Benders, M.; Claris, O.; Dempsey, E.; Fumagalli, M.; Gluud, C.; Hagmann, C.; Hellström-Westas, L.; Hyttel-Sorensen, S.; et al. The SafeBoosC Phase II Randomised Clinical Trial: A Treatment Guideline for Targeted Near-Infrared-Derived Cerebral Tissue Oxygenation versus Standard Treatment in Extremely Preterm Infants. *Neonatology* **2013**, *104*, 171–178, doi:10.1159/000351346.
  19. Hou, X.; Ding, H.; Teng, Y.; Zhou, C.; Tang, X.; Li, S.; Ding, H. Research on the Relationship between Brain Anoxia at Different Regional Oxygen Saturations and Brain Damage Using Near-Infrared Spectroscopy. *Physiol Meas* **2007**, *28*, 1251, doi:10.1088/0967-3334/28/10/010.
  20. Kurth, C.D.; McCann, J.C.; Wu, J.; Miles, L.; Loepke, A.W. Cerebral Oxygen Saturation-Time Threshold for Hypoxic-Ischemic Injury in Piglets. *Anesth Analg* **2009**, *108*, 1268–1277, doi:10.1213/ANE.0B013E318196AC8E.
  21. Dent, C.L.; Spaeth, J.P.; Jones, B. V.; Schwartz, S.M.; Glauser, T.A.; Hallinan, B.; Pearl, J.M.; Khoury, P.R.; Kurth, C.D. Brain Magnetic Resonance Imaging Abnormalities after the Norwood Procedure Using Regional Cerebral Perfusion. *Journal of Thoracic and Cardiovascular Surgery* **2006**, *131*, 190–197, doi:10.1016/j.jtcvs.2005.10.003.
  22. Kusaka, T.; Ueno, M.; Miki, T.; Kuboi, T.; Nakamura, S.; Koyano, K.; Ijichi, S.; Yasuda, S.; Okubo, K.; Kawada, K.; et al. Relationship Between Cerebral Oxygenation and Phosphorylation Potential During Secondary Energy Failure in Hypoxic-Ischemic Newborn Piglets. *Pediatric Research* **2009**, *65*, 317–322, doi:10.1203/pdr.0b013e318194fa73.

23. Wintermark, P.; Hansen, A.; Warfield, S.K.; Dukhovny, D.; Soul, J.S. Near-Infrared Spectroscopy versus Magnetic Resonance Imaging to Study Brain Perfusion in Newborns with Hypoxic–Ischemic Encephalopathy Treated with Hypothermia. *Neuroimage* **2014**, *85*, 287–293, doi:10.1016/J.NEUROIMAGE.2013.04.072.
24. Arman, D.; Sancak, S.; Gürsoy, T.; Topcuoğlu, S.; Karatekin, G.; Ovalı, F. The Association between NIRS and Doppler Ultrasonography in Preterm Infants with Patent Ductus Arteriosus. *The Journal of Maternal-Fetal & Neonatal Medicine* **2020**, *33*, 1245–1252, doi:10.1080/14767058.2019.1639661.
25. Volpe, J. Hypoxic Ischemic Encephalopathy. In *Clinical Aspect in Volpe J J Neurology of Newborn*. **2008**, 40–80.
26. Hollebrandse, N.L.; Spittle, A.J.; Burnett, A.C.; Anderson, P.J.; Roberts, G.; Doyle, L.W.; Cheong, J.L.Y. School-Age Outcomes Following Intraventricular Haemorrhage in Infants Born Extremely Preterm. *Arch Dis Child Fetal Neonatal Ed* **2021**, *106*, 4–8, doi:10.1136/ARCHDISCHILD-2020-318989.
27. Papile, L.A.; Burstein, J.; Burstein, R.; Koffler, H. Incidence and Evolution of Subependymal and Intraventricular Hemorrhage: A Study of Infants with Birth Weights Less than 1,500 Gm. *J Pediatr* **1978**, *92*, 529–534, doi:10.1016/S0022-3476(78)80282-0.
28. Krediet, T.G.; Kavelaars, A.; Vreman, H.J.; Heijnen, C.J.; van Bel, F. Respiratory Distress Syndrome-Associated Inflammation Is Related to Early but Not Late Peri/Intraventricular Hemorrhage in Preterm Infants. *Journal of Pediatrics* **2006**, *148*, 740–746, doi:10.1016/j.jpeds.2006.01.037.
29. Villamor-Martinez, E.; Fumagalli, M.; Rahim, O.M.; Passera, S.; Cavallaro, G.; Degraeuwe, P.; Mosca, F.; Villamor, E. Chorioamnionitis Is a Risk Factor for Intraventricular Hemorrhage in Preterm Infants: A Systematic Review and Meta-Analysis. *Front Physiol* **2018**, *9*, 401570, doi:10.3389/FPHYS.2018.01253/BIBTEX.
30. Ballabh, P. Pathogenesis and Prevention of Intraventricular Hemorrhage. *Clin Perinatol* **2014**, *41*, 47–67, doi:10.1016/J.CLP.2013.09.007.
31. Kontos, H.A.; Wei, E.P.; Jarrell Raper, A.; Patterson, J.L. Local Mechanism of CO<sub>2</sub> Action of Cat Pial Arterioles. *Stroke* **1977**, *8*, 226–229, doi:10.1161/01.STR.8.2.226.
32. Dix, L.M.L.; Weeke, L.C.; de Vries, L.S.; Groenendaal, F.; Baerts, W.; van Bel, F.; Lemmers, P.M.A. Carbon Dioxide Fluctuations Are Associated with Changes in Cerebral Oxygenation and Electrical Activity in Infants Born Preterm. *Journal of Pediatrics* **2017**, *187*, 66-72.e1, doi:10.1016/j.jpeds.2017.04.043.
33. Van Bel, F.; Van de Bor, M.; Stijnen, T.; Baan, J.; Ruys, J.H. Aetiological Rôle Of Cerebral Blood-Flow Alterations In Development And Extension Of Peri-Intraventricular Haemorrhage. *Dev Med Child Neurol* **1987**, *29*, 601–614, doi:10.1111/J.1469-8749.1987.TB08502.X.
34. Maren, T.H. Effect of Varying CO<sub>2</sub> Equilibria on Rates of HCO<sub>3</sub><sup>-</sup> Formation in Cerebrospinal Fluid. <https://doi.org/10.1152/jappl.1979.47.3.471> **1979**, *47*, 471–477, doi:10.1152/JAPPL.1979.47.3.471.
35. Alderliesten, T.; Lemmers, P.M.A.; Smarius, J.J.M.; Van De Vosse, R.E.; Baerts, W.; Van Bel, F. Cerebral Oxygenation, Extraction, and Autoregulation in Very Preterm

- Infants Who Develop Peri-Intraventricular Hemorrhage. *Journal of Pediatrics* **2013**, *162*, 698-704.e2, doi:10.1016/j.jpeds.2012.09.038.
36. Hoffman, S.B.; Cheng, Y.J.; Magder, L.S.; Shet, N.; Viscardi, R.M. Cerebral Autoregulation in Premature Infants during the First 96 Hours of Life and Relationship to Adverse Outcomes. *Arch Dis Child Fetal Neonatal Ed* **2019**, *104*, F473–F479, doi:10.1136/ARCHDISCHILD-2018-315725.
  37. Tamura, K.; Williams, E.E.; Dassios, T.; Pahuja, A.; Hunt, K.A.; Murthy, V.; Bhat, P.; Bhat, R.; Milner, A.; Greenough, A. End-Tidal Carbon Dioxide Levels during Resuscitation and Carbon Dioxide Levels in the Immediate Neonatal Period and Intraventricular Haemorrhage. *Eur J Pediatr* **2020**, *179*, 555–559, doi:10.1007/S00431-019-03543-0/METRICS.
  38. Thewissen, L.; Caicedo, A.; Lemmers, P.; Van Bel, F. V.; Van Huffel, S. V.; Naulaers, G. Measuring Near-Infrared Spectroscopy Derived Cerebral Autoregulation in Neonates: From Research Tool toward Bedside Multimodal Monitoring. *Front Pediatr* **2018**, *6*, 341349, doi:10.3389/FPED.2018.00117/BIBTEX.
  39. Alderliesten, T.; van Bel, F.; van der Aa, N.E.; Steendijk, P.; van Haastert, I.C.; de Vries, L.S.; Groenendaal, F.; Lemmers, P. Low Cerebral Oxygenation in Preterm Infants Is Associated with Adverse Neurodevelopmental Outcome. *Journal of Pediatrics* **2019**, *207*, 109-116.e2, doi:10.1016/j.jpeds.2018.11.038.
  40. Vesoulis, Z.A.; Whitehead, H. V.; Liao, S.M.; Mathur, A.M. The Hidden Consequence of Intraventricular Hemorrhage: Persistent Cerebral Desaturation after IVH in Preterm Infants. *Pediatric Research* **2020**, *89*, 869–877, doi:10.1038/s41390-020-01189-5.
  41. Hoffman, S.B.; Cheng, Y.J.; Magder, L.S.; Shet, N.; Viscardi, R.M. Cerebral Autoregulation in Premature Infants during the First 96 Hours of Life and Relationship to Adverse Outcomes. *Arch Dis Child Fetal Neonatal Ed* **2019**, *104*, F473–F479, doi:10.1136/ARCHDISCHILD-2018-315725.
  42. Hyttel-Sorensen, S.; Pellicer, A.; Alderliesten, T.; Austin, T.; Van Bel, F.; Benders, M.; Claris, O.; Dempsey, E.; Franz, A.R.; Fumagalli, M.; et al. Cerebral near Infrared Spectroscopy Oximetry in Extremely Preterm Infants: Phase II Randomised Clinical Trial. *BMJ* **2015**, *350*, doi:10.1136/BMJ.G7635.
  43. Plomgaard, A.M.; Alderliesten, T.; Austin, T.; Van Bel, F.; Benders, M.; Claris, O.; Dempsey, E.; Fumagalli, M.; Gluud, C.; Hagmann, C.; et al. Early Biomarkers of Brain Injury and Cerebral Hypo- and Hyperoxia in the SafeBoosC II Trial. *PLoS One* **2017**, *12*, e0173440, doi:10.1371/JOURNAL.PONE.0173440.
  44. Volpe, J.J. Brain Injury in Premature Infants: A Complex Amalgam of Destructive and Developmental Disturbances. *Lancet Neurol* **2009**, *8*, 110–124, doi:10.1016/S1474-4422(08)70294-1.
  45. Chau, V.; Synnes, A.; Grunau, R.E.; Poskitt, K.J.; Brant, R.; Miller, S.P. Abnormal Brain Maturation in Preterm Neonates Associated with Adverse Developmental Outcomes. *Neurology* **2013**, *81*, 2082–2089, doi:10.1212/01.WNL.0000437298.43688.B9/SUPPL\_FILE/TABLES\_AND\_E-METHODS.DOCX.

46. van Tilborg, E.; de Theije, C.G.M.; van Hal, M.; Wagenaar, N.; de Vries, L.S.; Benders, M.J.; Rowitch, D.H.; Nijboer, C.H. Origin and Dynamics of Oligodendrocytes in the Developing Brain: Implications for Perinatal White Matter Injury. *Glia* **2018**, *66*, 221–238, doi:10.1002/GLIA.23256.
47. Brouwer, M.J.; Kersbergen, K.J.; Van Kooij, B.J.M.; Benders, M.J.N.L.; Van Haastert, I.C.; Koopman-Esseboom, C.; Neil, J.J.; De Vries, L.S.; Kidokoro, H.; Inder, T.E.; et al. Preterm Brain Injury on Term-Equivalent Age MRI in Relation to Perinatal Factors and Neurodevelopmental Outcome at Two Years. *PLoS One* **2017**, *12*, e0177128, doi:10.1371/JOURNAL.PONE.0177128.
48. Favrais, G.; Van De Looij, Y.; Fleiss, B.; Ramanantsoa, N.; Bonnin, P.; Stoltenberg-Didinger, G.; Lacaud, A.; Saliba, E.; Dammann, O.; Gallego, J.; et al. Systemic Inflammation Disrupts the Developmental Program of White Matter. *Ann Neurol* **2011**, *70*, 550–565, doi:10.1002/ANA.22489.
49. Greisen, G.; Vannucci, R.C. Is Periventricular Leucomalacia a Result of Hypoxic-Ischaemic Injury? Hypocapnia and the Preterm Brain. *Biol Neonate* **2001**, *79*, 194–200, doi:10.1159/000047090.
50. Benitz, W.E. Treatment of Persistent Patent Ductus Arteriosus in Preterm Infants: Time to Accept the Null Hypothesis? *Journal of Perinatology* **2010** *30:4* **2010**, *30*, 241–252, doi:10.1038/jp.2010.3.
51. Noori, S. Patent Ductus Arteriosus in the Preterm Infant: To Treat or Not to Treat? *Journal of Perinatology* **2010** *30:1* **2010**, *30*, S31–S37, doi:10.1038/jp.2010.97.
52. Giesinger, R.E.; McNamara, P.J. Hemodynamic Instability in the Critically Ill Neonate: An Approach to Cardiovascular Support Based on Disease Pathophysiology. *Semin Perinatol* **2016**, *40*, 174–188, doi:10.1053/J.SEMPERI.2015.12.005.
53. Lemmers, P.M.A.; Benders, M.J.N.L.; D’Ascenzo, R.; Zethof, J.; Alderliesten, T.; Kersbergen, K.J.; Isgum, I.; De Vries, L.S.; Groenendaal, F.; Van Bel, F. Patent Ductus Arteriosus and Brain Volume. *Pediatrics* **2016**, *137*, doi:10.1542/PEDS.2015-3090/81366.
54. Volpe, J.J. Cerebellum of the Premature Infant: Rapidly Developing, Vulnerable, Clinically Important. <http://dx.doi.org/10.1177/0883073809338067> **2009**, *24*, 1085–1104, doi:10.1177/0883073809338067.
55. Witter, L.; Rudolph, S.; Pressler, R.T.; Lahlaf, S.I.; Regehr, W.G. Purkinje Cell Collaterals Enable Output Signals from the Cerebellar Cortex to Feed Back to Purkinje Cells and Interneurons. *Neuron* **2016**, *91*, 312–319, doi:10.1016/j.neuron.2016.05.037.
56. Limperopoulos, C.; Benson, C.B.; Bassan, H.; Disalvo, D.N.; Kinnamon, D.D.; Moore, M.; Ringer, S.A.; Volpe, J.J.; Du Plessis, A.J. Cerebellar Hemorrhage in the Preterm Infant: Ultrasonographic Findings and Risk Factors. *Pediatrics* **2005**, *116*, 717–724, doi:10.1542/PEDS.2005-0556.
57. Alderliesten, T.; Lemmers, P.M.A.; Van Haastert, I.C.; De Vries, L.S.; Bonestroo, H.J.C.; Baerts, W.; Van Bel, F. Hypotension in Preterm Neonates: Low Blood Pressure Alone Does Not Affect Neurodevelopmental Outcome. *Journal of Pediatrics* **2014**, *164*, 986–991, doi:10.1016/j.jpeds.2013.12.042.

58. Dempsey, E.M.; Al Hazzani, F.; Barrington, K.J. Permissive Hypotension in the Extremely Low Birthweight Infant with Signs of Good Perfusion. *Arch Dis Child Fetal Neonatal Ed* **2009**, *94*, F241–F244, doi:10.1136/ADC.2007.124263.
59. Van Hoften, J.C.R.; Verhagen, E.A.; Keating, P.; Ter Horst, H.J.; Bos, A.F. Cerebral Tissue Oxygen Saturation and Extraction in Preterm Infants before and after Blood Transfusion. *Arch Dis Child Fetal Neonatal Ed* **2010**, *95*, F352–F358, doi:10.1136/ADC.2009.163592.
60. Vanderhaegen, J.; Vanhaesebrouck, S.; Vanhole, C.; Casaer, P.; Naulaers, G. The Effect of Glycaemia on the Cerebral Oxygenation in Very Low Birthweight Infants as Measured by Near-Infrared Spectroscopy. *Adv Exp Med Biol* **2010**, *662*, 461–466, doi:10.1007/978-1-4419-1241-1\_66/FIGURES/66\_1\_117031\_1\_EN.
61. Black, R.E.; Cousens, S.; Johnson, H.L.; Lawn, J.E.; Rudan, I.; Bassani, D.G.; Jha, P.; Campbell, H.; Walker, C.F.; Cibulskis, R.; et al. Global, Regional, and National Causes of Child Mortality in 2008: A Systematic Analysis. *Lancet* **2010**, *375*, 1969–1987, doi:10.1016/S0140-6736(10)60549-1.
62. Lorek, A.; Takei, Y.; Cady, E.B.; Wyatt, J.S.; Penrice, J.; Edwards, A.D.; Peebles, D.; Wylezinska, M.; Owen-Reece, H.; Kirkbride, V.; et al. Delayed (“Secondary”) Cerebral Energy Failure after Acute Hypoxia-Ischemia in the Newborn Piglet: Continuous 48-Hour Studies by Phosphorus Magnetic Resonance Spectroscopy. *Pediatric Research* **1994** *36:6* **1994**, *36*, 699–706, doi:10.1203/00006450-199412000-00003.
63. Kazma, J.M.; van den Anker, J.N.; Allegaert, K.; Dallmann, A.; Ahmadzia, H.K. Role of Placenta in Drug Metabolism and Drug Transfer. *Neonatal and Pediatric Pharmacology: Therapeutic Principles in Practice* **2020**.
64. Toet, M.C.; Lemmers, P.M.A.; Van Schelven, L.J.; Van Bel, F. Cerebral Oxygenation and Electrical Activity After Birth Asphyxia: Their Relation to Outcome. *Pediatrics* **2006**, *117*, 333–339, doi:10.1542/PEDS.2005-0987.
65. Ancora, G.; Maranella, E.; Grandi, S.; Sbravati, F.; Coccolini, E.; Savini, S.; Faldella, G. Early Predictors of Short Term Neurodevelopmental Outcome in Asphyxiated Cooled Infants. A Combined Brain Amplitude Integrated Electroencephalography and near Infrared Spectroscopy Study. *Brain Dev* **2013**, *35*, 26–31, doi:10.1016/j.braindev.2011.09.008.
66. Lynch, J.K. Epidemiology and Classification of Perinatal Stroke. *Semin Fetal Neonatal Med* **2009**, *14*, 245–249, doi:10.1016/j.siny.2009.07.001.
67. Fernández-López, D.; Natarajan, N.; Ashwal, S.; Vexler, Z.S. Mechanisms of Perinatal Arterial Ischemic Stroke. *Journal of Cerebral Blood Flow and Metabolism* **2014**, *34*, 921–932, doi:10.1038/JCBFM.2014.41/ASSET/IMAGES/LARGE/10.1038\_JCBFM.2014.41-FIG4.JPEG.
68. Wagenaar, N.; Martinez-Biarge, M.; Van Der Aa, N.E.; Van Haastert, I.C.; Groenendaal, F.; Benders, M.J.N.L.; Cowan, F.M.; De Vries, L.S. Neurodevelopment after Perinatal Arterial Ischemic Stroke. *Pediatrics* **2018**, *142*, doi:10.1542/PEDS.2017-4164/38586.
69. Benders, M.J.; Van Der Aa, N.E.; Roks, M.; Van Straaten, H.L.; Isgum, I.; Viergever, M.A.; Groenendaal, F.; De Vries, L.S.; Van Bel, F. Feasibility and Safety of

- Erythropoietin for Neuroprotection after Perinatal Arterial Ischemic Stroke. *Journal of Pediatrics* **2014**, *164*, 481–486.e2, doi:10.1016/j.jpeds.2013.10.084.
70. Wagenaar, N.; De Theije, C.G.M.; De Vries, L.S.; Groenendaal, F.; Benders, M.J.N.L.; Nijboer, C.H.A. Promoting Neuroregeneration after Perinatal Arterial Ischemic Stroke: Neurotrophic Factors and Mesenchymal Stem Cells. *Pediatric Research* **2018** *83:1* **2017**, *83*, 372–384, doi:10.1038/pr.2017.243.
71. Wagenaar, N.; van den Berk, D.J.M.; Lemmers, P.M.A.; van der Aa, N.E.; Dudink, J.; van Bel, F.; Groenendaal, F.; de Vries, L.S.; Benders, M.J.N.L.; Alderliesten, T. Brain Activity and Cerebral Oxygenation after Perinatal Arterial Ischemic Stroke Are Associated with Neurodevelopment. *Stroke* **2019**, *50*, 2668–2676, doi:10.1161/STROKEAHA.119.025346.
72. Wintermark, P.; Warfield, S.K. New Insights in Perinatal Arterial Ischemic Stroke by Assessing Brain Perfusion. *Transl Stroke Res* **2012**, *3*, 255–262, doi:10.1007/S12975-011-0122-0/FIGURES/4.
73. Van Der Aa, N.E.; Porsius, E.D.; Hendrikse, J.; Van Kooij, B.J.M.; Benders, M.J.N.L.; De Vries, L.S.; Groenendaal, F. Changes in Carotid Blood Flow after Unilateral Perinatal Arterial Ischemic Stroke. *Pediatric Research* **2012** *72:1* **2012**, *72*, 50–56, doi:10.1038/pr.2012.39.
74. Greisen, G. Autoregulation of Cerebral Blood Flow in Newborn Babies. *Early Hum Dev* **2005**, *81*, 423–428, doi:10.1016/J.EARLHUMDEV.2005.03.005.
75. Caicedo, A.; De Smet, D.; Vanderhaegen, J.; Naulaers, G.; Wolf, M.; Lemmers, P.; Van Bel, F.; Ameye, L.; Van Huffel, S. Impaired Cerebral Autoregulation Using Near-Infrared Spectroscopy and Its Relation to Clinical Outcomes in Premature Infants. *Adv Exp Med Biol* **2011**, *701*, 233–239, doi:10.1007/978-1-4419-7756-4\_31/COVER.
76. El-Dib, M.; Soul, J.S. Monitoring and Management of Brain Hemodynamics and Oxygenation. *Handb Clin Neurol* **2019**, *162*, 295–314, doi:10.1016/B978-0-444-64029-1.00014-X.
77. Rhee, C.J.; da Costa, C.S.; Austin, T.; Brady, K.M.; Czosnyka, M.; Lee, J.K. Neonatal Cerebrovascular Autoregulation. *Pediatric Research* **2018** *84:5* **2018**, *84*, 602–610, doi:10.1038/s41390-018-0141-6.
78. Chock, V.Y.; Kwon, S.H.; Ambalavanan, N.; Batton, B.; Nelin, L.D.; Chalak, L.F.; Tian, L.; Van Meurs, K.P. Cerebral Oxygenation and Autoregulation in Preterm Infants (Early NIRS Study). *Journal of Pediatrics* **2020**, *227*, 94–100.e1, doi:10.1016/j.jpeds.2020.08.036.
79. Massaro, A.N.; Govindan, R.B.; Vezina, G.; Chang, T.; Andescavage, N.N.; Wang, Y.; Al-Shargabi, T.; Metzler, M.; Harris, K.; du Plessis, A.J. Impaired Cerebral Autoregulation and Brain Injury in Newborns with Hypoxic-Ischemic Encephalopathy Treated with Hypothermia. *J Neurophysiol* **2015**, *114*, 818–824, doi:10.1152/JN.00353.2015/ASSET/IMAGES/LARGE/Z9K0131532000002.JPEG.
80. Vesoulis, Z.A.; Mathur, A.M. Cerebral Autoregulation, Brain Injury, and the Transitioning Premature Infant. *Front Pediatr* **2017**, *5*, 242901, doi:10.3389/FPED.2017.00064/BIBTEX.

81. Tsuji, M.; Saul, J.P.; Du Plessis, A.; Eichenwald, E.; Sobh, J.; Crocker, R.; Volpe, J.J. Cerebral Intravascular Oxygenation Correlates With Mean Arterial Pressure in Critically Ill Premature Infants. *Pediatrics* **2000**, *106*, 625–632, doi:10.1542/PEDS.106.4.625.
82. Tasker, R.C. Brain Vascular and Hydrodynamic Physiology. *Semin Pediatr Surg* **2013**, *22*, 168–173, doi:10.1053/J.SEMPEDSURG.2013.10.003.
83. Vesoulis, Z.A.; Mintzer, J.P.; Chock, V.Y. Neonatal NIRS Monitoring: Recommendations for Data Capture and Review of Analytics. *Journal of Perinatology* **2021**, *41:4* **2021**, *41*, 675–688, doi:10.1038/s41372-021-00946-6.
84. Detection of Cerebral Autoregulation by Near-Infrared Spectroscopy in Neonates: Performance Analysis of Measurement Methods Available online: [https://www.spiedigitallibrary.org/journals/journal-of-biomedical-optics/volume-17/issue-11/117003/Detection-of-cerebral-autoregulation-by-near-infrared-spectroscopy-in-neonates/10.1117/1.JBO.17.11.117003.full#\\_=\\_](https://www.spiedigitallibrary.org/journals/journal-of-biomedical-optics/volume-17/issue-11/117003/Detection-of-cerebral-autoregulation-by-near-infrared-spectroscopy-in-neonates/10.1117/1.JBO.17.11.117003.full#_=_) (accessed on 25 March 2024).
85. Liem, K.D.; Greisen, G. Monitoring of Cerebral Haemodynamics in Newborn Infants. *Early Hum Dev* **2010**, *86*, 155–158, doi:10.1016/J.EARLHUMDEV.2010.01.029.
86. Delpy, D.T.; Cope, M.; Van Der Zee, P.; Arridge, S.; Wray, S.; Wyatt, J. Estimation of Optical Pathlength through Tissue from Direct Time of Flight Measurement. *Phys Med Biol* **1988**, *33*, 1433, doi:10.1088/0031-9155/33/12/008.
87. Villringer, A.; Planck, J.; Hock, C.; Schleinkofer, L.; Dirnagl, U. Near Infrared Spectroscopy (NIRS): A New Tool to Study Hemodynamic Changes during Activation of Brain Function in Human Adults. *Neurosci Lett* **1993**, *154*, 101–104, doi:10.1016/0304-3940(93)90181-J.
88. Ferrari, M.; Quaresima, V. A Brief Review on the History of Human Functional Near-Infrared Spectroscopy (FNIRS) Development and Fields of Application. *Neuroimage* **2012**, *63*, 921–935, doi:10.1016/J.NEUROIMAGE.2012.03.049.
89. Measurement of Layer-like Hemodynamic Trends in Scalp and Cortex: Implications for Physiological Baseline Suppression in Functional near-Infrared Spectroscopy Available online: [https://www.spiedigitallibrary.org/journals/journal-of-biomedical-optics/volume-13/issue-03/034017/Measurement-of-layer-like-hemodynamic-trends-in-scalp-and-cortex/10.1117/1.2940587.full#\\_=\\_](https://www.spiedigitallibrary.org/journals/journal-of-biomedical-optics/volume-13/issue-03/034017/Measurement-of-layer-like-hemodynamic-trends-in-scalp-and-cortex/10.1117/1.2940587.full#_=_) (accessed on 25 March 2024).
90. Kirilina, E.; Yu, N.; Jelzow, A.; Wabnitz, H.; Jacobs, A.M.; Tachtsidis, L. Identifying and Quantifying Main Components of Physiological Noise in Functional near Infrared Spectroscopy on the Prefrontal Cortex. *Front Hum Neurosci* **2013**, *7*, 72520, doi:10.3389/FNHUM.2013.00864/BIBTEX.
91. Gregg, N.M.; White, B.R.; Zeff, B.W.; Berger, A.J.; Culver, J.P. Brain Specificity of Diffuse Optical Imaging: Improvements from Superficial Signal Regression and Tomography. *Front Neuroenergetics* **2010**, *2*, 1537, doi:10.3389/FNENE.2010.00014.
92. Hakimi, N.; Jodeiri, A.; Mirbagheri, M.; Setarehdan, S.K. Proposing a Convolutional Neural Network for Stress Assessment by Means of Derived Heart Rate from Functional near Infrared Spectroscopy. *Comput Biol Med* **2020**, *121*, 103810, doi:10.1016/J.COMPBIOMED.2020.103810.



93. Extraction of Heart Rate from Functional Near-Infrared Spectroscopy in Infants Available online: [https://www.spiedigitallibrary.org/journals/journal-of-biomedical-optics/volume-19/issue-06/067010/Extraction-of-heart-rate-from-functional-near-infrared-spectroscopy-in/10.1117/1.JBO.19.6.067010.full#\\_=\\_](https://www.spiedigitallibrary.org/journals/journal-of-biomedical-optics/volume-19/issue-06/067010/Extraction-of-heart-rate-from-functional-near-infrared-spectroscopy-in/10.1117/1.JBO.19.6.067010.full#_=_) (accessed on 25 March 2024).
94. Short-Term Pulse Rate Variability Is Better Characterized by Functional near-Infrared Spectroscopy than by Photoplethysmography Available online: [https://www.spiedigitallibrary.org/journals/journal-of-biomedical-optics/volume-21/issue-09/091308/Short-term-pulse-rate-variability-is-better-characterized-by-functional/10.1117/1.JBO.21.9.091308.full#\\_=\\_](https://www.spiedigitallibrary.org/journals/journal-of-biomedical-optics/volume-21/issue-09/091308/Short-term-pulse-rate-variability-is-better-characterized-by-functional/10.1117/1.JBO.21.9.091308.full#_=_) (accessed on 25 March 2024).
95. Hakimi, N.; Kamaledin Setarehdan, S. Stress Assessment by Means of Heart Rate Derived from Functional Near-Infrared Spectroscopy. <https://doi.org/10.1117/1.JBO.23.11.115001> **2018**, *23*, 115001, doi:10.1117/1.JBO.23.11.115001.
96. Mirbagheri, M.; Hakimi, N.; Ebrahimzadeh, E.; Setarehdan, S.K. Quality Analysis of Heart Rate Derived from Functional Near-Infrared Spectroscopy in Stress Assessment. *Inform Med Unlocked* **2020**, *18*, 100286, doi:10.1016/J.IMU.2019.100286.
97. Perdue, K.L.; Edwards, L.A.; Tager-Flusberg, H.; Nelson, C.A. Differing Developmental Trajectories in Heart Rate Responses to Speech Stimuli in Infants at High and Low Risk for Autism Spectrum Disorder. *J Autism Dev Disord* **2017**, *47*, 2434–2442, doi:10.1007/S10803-017-3167-4/FIGURES/4.
98. Lund, C. Medical Adhesives in the NICU. *Newborn and Infant Nursing Reviews* **2014**, *14*, 160–165, doi:10.1053/J.NAINR.2014.10.001.
99. Scarpa, F.; Brigadoi, S.; Cutini, S.; Scatturin, P.; Zorzi, M.; Dellracqua, R.; Sparacino, G. A Methodology to Improve Estimation of Stimulus-Evoked Hemodynamic Response from FNIRS Measurements. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* **2011**, 785–788, doi:10.1109/IEMBS.2011.6090180.
100. Hemmati Berivanlou, N.; Setarehdan, S.K.; Ahmadi Noubari, H. Evoked Hemodynamic Response Estimation Using Ensemble Empirical Mode Decomposition Based Adaptive Algorithm Applied to Dual Channel Functional near Infrared Spectroscopy (FNIRS). *J Neurosci Methods* **2014**, *224*, 13–25, doi:10.1016/J.JNEUMETH.2013.12.007.
101. Kamran, M.A.; Jeong, M.Y.; Mannan, M.M.N. Optimal Hemodynamic Response Model for Functional Near-Infrared Spectroscopy. *Front Behav Neurosci* **2015**, *9*, 146711, doi:10.3389/FNBEH.2015.00151/BIBTEX.
102. Scholkmann, F.; Kleiser, S.; Metz, A.J.; Zimmermann, R.; Mata Pavia, J.; Wolf, U.; Wolf, M. A Review on Continuous Wave Functional Near-Infrared Spectroscopy and Imaging Instrumentation and Methodology. *Neuroimage* **2014**, *85 Pt 1*, 6–27, doi:10.1016/J.NEUROIMAGE.2013.05.004.
103. Caldwell, M.; Scholkmann, F.; Wolf, U.; Wolf, M.; Elwell, C.; Tachtsidis, I. Modelling Confounding Effects from Extracerebral Contamination and Systemic Factors on Functional Near-Infrared Spectroscopy. *Neuroimage* **2016**, *143*, 91–105, doi:10.1016/J.NEUROIMAGE.2016.08.058.

104. Orihuela-Espina, F.; Leff, D.R.; James, D.R.C.; Darzi, A.W.; Yang, G.Z. Quality Control and Assurance in Functional near Infrared Spectroscopy (FNIRS) Experimentation. *Phys Med Biol* **2010**, *55*, 3701, doi:10.1088/0031-9155/55/13/009.
105. Pollonini, L.; Olds, C.; Abaya, H.; Bortfeld, H.; Beauchamp, M.S.; Oghalai, J.S. Auditory Cortex Activation to Natural Speech and Simulated Cochlear Implant Speech Measured with Functional Near-Infrared Spectroscopy. *Hear Res* **2014**, *309*, 84–93, doi:10.1016/J.HEARES.2013.11.007.
106. Pollonini, L.; Bortfeld, H.; Oghalai, J.S. PHOEBE: A Method for Real Time Mapping of Optodes-Scalp Coupling in Functional near-Infrared Spectroscopy. *Biomedical Optics Express*, Vol. 7, Issue 12, pp. 5104-5119 **2016**, 7, 5104–5119, doi:10.1364/BOE.7.005104.
107. Sappia, M.S.; N.; Hakimi, N.; Colier, W.N.J.M.; Horschig, J.M. Signal Quality Index: An Algorithm for Quantitative Assessment of Functional near Infrared Spectroscopy Signal Quality. *Biomedical Optics Express*, Vol. 11, Issue 11, pp. 6732-6754 **2020**, 11, 6732–6754, doi:10.1364/BOE.409317.
108. Sappia, M.S.; Hakimi, N.; Svinkunaite, L.; Alderliesten, T.; Horschig, J.M.; Colier, W.N.J.M. FNIRS Signal Quality Estimation by Means of a Machine Learning Algorithm Trained on Morphological and Temporal Features. <https://doi.org/10.1117/12.2587188> **2021**, 11638, 29–39, doi:10.1117/12.2587188.

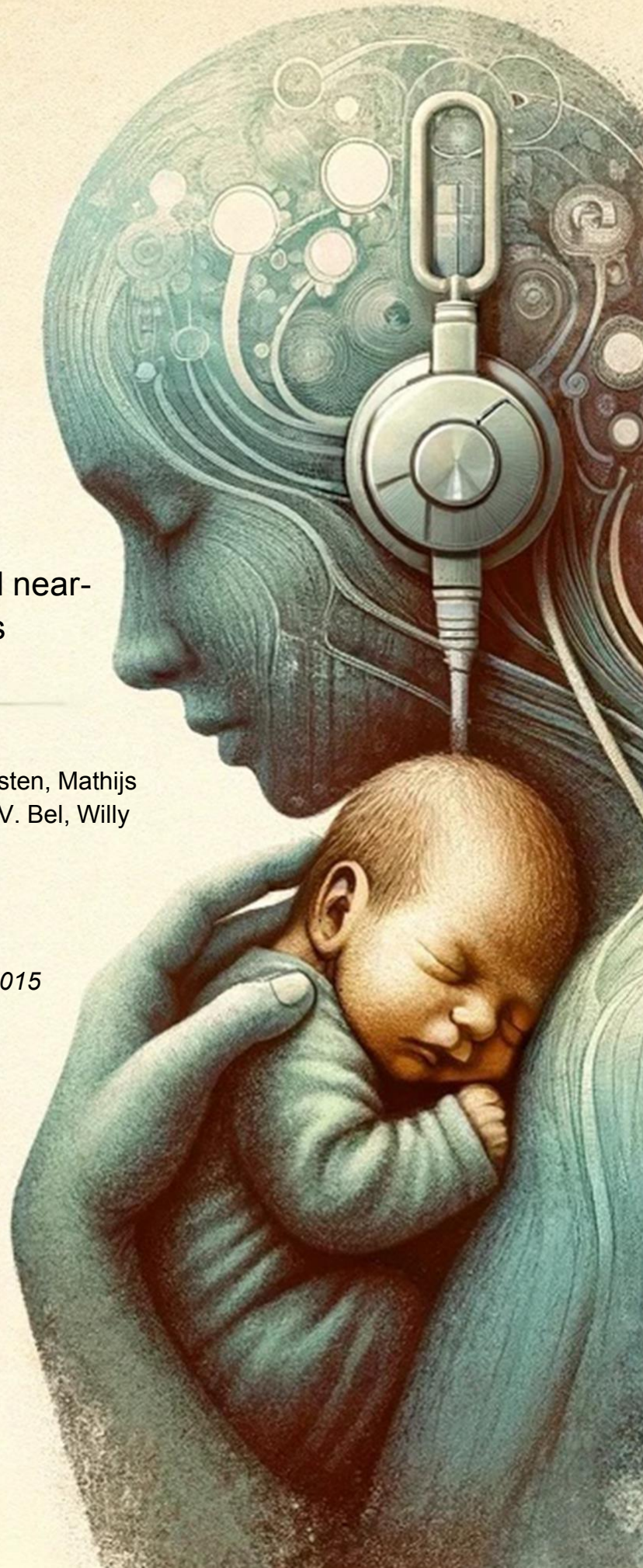


# CHAPTER 7

Heart rate extraction from neonatal near-infrared spectroscopy signals

Naser Hakimi, Jörn M. Horschig, Thomas Alderliesten, Mathijs Bronkhorst, Marianne J. Floor-Westerdijk, Frank V. Bel, Willy N.J.M. Colier, Jeroen Dudink

doi: <https://doi.org/10.1109/TIM.2023.3295015>



## Abstract

Near-infrared spectroscopy (NIRS) intensity signals provide useful additional physiological information of which the most prominent one is the pulsatile fluctuation by heartbeats. This allows for the extraction of heart rate (HR), one of the primary clinical indicators of health in neonates. In this study, we propose a novel algorithm, NHR (NIRS HR), for extracting HR from NIRS signals acquired from neonates admitted to the neonatal intensive care unit. After parental consent, we recorded synchronously NIRS, at 100 Hz, and reference HR, at 1 Hz, from ten newborn infants (gestational age =  $38 \pm 5$  weeks;  $3092 \pm 990$  grams). The NHR algorithm consists of two main parts. The first part includes four steps implemented once on the whole NIRS measurement: preprocessing, HR frequency bandwidth determination, interquartile range computation, and segmentation. The second part includes three steps implemented on each signal segment: motion artifact detection, signal quality assessment, and HR computation. We compared the NHR algorithm with two existing algorithms. The results showed that the proposed NHR algorithm provides a significantly ( $p < 0.05$ ) higher correlation ( $r=99.5\%$ ) and lower Bland-Altman ratio ( $BAR=3.6\%$ ) between the extracted and reference HRs, compared to the existing algorithms. Extracting HR from NIRS in a clinical setting of critically ill neonates admitted to neonatal intensive care is feasible. With NIRS and HR combined in a single monitoring system, it is possible to have a perfectly time-synced integrated analysis of cerebral hemodynamics as well as systemic hemodynamics and autonomic nervous system tone.



# 1. Introduction

Near-infrared spectroscopy (NIRS) is a non-invasive system for the continuous monitoring of brain hemodynamics [1,2]. It is an ideal system to measure changes in oxygenation and perfusion of the neonatal brain [3,4]. Although NIRS is not yet part of the standard clinical routine in most neonatal intensive care units (NICUs), it is increasingly used across various clinical settings [5]. NIRS-monitored cerebral oxygenation has been shown to be of clinical benefit in infants during (non-) cardiac surgery as well as for the assessment of some clinical conditions such as asphyxia, hypoxic-ischemic encephalopathy, anemia, hypoglycemia, hypotension, and hypocapnia [6–11]. Using NIRS, different types of information can be acquired, where cerebral oxygenation is the most common one used in clinical settings [3]. Another source of information is the raw NIRS-derived intensity signals, which are widely disregarded for clinical purposes. The NIRS intensity signals are contaminated by physiological artifacts such as cardiac pulsation, respiration, and blood pressure [12], which can potentially be harnessed to yield additional physiological information. The most prominent physiological information in the NIRS signals is the cardiac pulsation which allows for the extraction of heart rate (HR) from NIRS [3,13].

Extracting HR from NIRS signals offers several potential advantages in neonates over traditional methods for HR monitoring, i.e., electrocardiography (ECG) and photoplethysmography (PPG). NIRS is a non-invasive technique that does not require sensors to be adhesively placed on the skin [3]. This could be especially advantageous for extremely preterm infants with vulnerable skin [17], preventing epidermal stripping [14]. In extremely preterm infants during the first days of life, when the skin is at its most fragile state, often no ECG sensors are placed due to the risk of skin injury, and heart rate (or pulse rate) is extracted from oximetry measurements [15]. The NIRS signals recorded on the forehead are less prone to motion artifacts than the PPG signals recorded in the periphery. This is because head movements are less frequent in neonates than hand movements [16]. Moreover, as reported in the previous studies [17,18], the HR derived from NIRS might be a richer source of information on the autonomic nervous system than the HR derived from PPG and ECG under stressful conditions. Additionally, NIRS also enables continuous monitoring of cerebral blood flow and oxygenation. Measuring HR with the same system would allow for a better understanding of physiological mechanisms and interventions in newborns with concurrent analysis of cerebral oxygenation and HR [3]. Lastly, NIRS devices are portable and easy to use [10], making them convenient for monitoring HR and other physiological variables in resource-limited settings or when rapid, non-invasive monitoring is necessary.

Several studies have reported the compatibility of the HR extracted from NIRS with the standard reference HR derived from ECG and the HR derived from PPG in healthy adults [17–20]. In 2012, Scholkmann et al. proposed an algorithm named AMPD (Automatic Multiscale-based Peak Detection) for peak detection in noisy periodic and quasi-periodic signals [21]. Thereafter in 2016, Holper et al. employed the AMPD algorithm for extracting heart rate from NIRS data recorded from healthy adults [17]. Moreover, in the same study, they showed the usability of NIRS-derived HR in the assessment of physical stress in healthy adults. In 2018, Hakimi and Setarehdan showed the outperformance of the AMPD algorithm in HR extraction from NIRS signals recorded in healthy adults with respect to the existing algorithms proposed for extracting HR from PPG [18]. Furthermore, in the same study, they showed the usability of the NIRS-derived HR in the assessment of mental stress.

Extraction of HR from infant NIRS is more challenging than adult NIRS due to a lower signal-to-noise ratio, a higher level of motion artifacts, and a greater physiological HR range [22]. The latter makes that there is a need for a system with a high sampling rate. In 2014, Perdue et al. reported an algorithm for the extraction of HR from NIRS signals recorded from healthy infants (~7 months old) in a controlled setting, sampled at 10 Hz [22]. In 2017, Perdue et al. conducted a longitudinal study investigating the applicability of the extracted HR in infants aged 3-12 months old with familial risk for autism spectrum disorder to observe the HR developmental trajectories to speech stimuli [23].

Although the existing studies have shown the feasibility of extracting HR from NIRS signals recorded in both infants and adults, the existing algorithms have only been validated on data recorded from healthy subjects in controlled environments. There are however challenges existing in analyzing data acquired in a clinical environment in comparison with a controlled environment, including lower data quality, higher contamination by artifacts due to patient movement, intermittent sensor replacement, and ventilator use, more variability of the signal source, and more restriction imposed by patient safety regulations in the hospital. Due to these challenges, therefore, the existing algorithms might not be properly accurate in the extraction of HR from clinical NIRS data. Therefore, there is a lack of a study to assess the performance of the existing algorithms in extracting heart rate from NIRS data recorded in clinical settings, especially in neonatal patients in the NICU. The issue of non-invasive monitoring of critical functions of the body is pre-eminent in very tiny and vulnerable (preterm) neonates with very vulnerable skin [14]. Especially, in this patient group one wants to have as few sensors and lines as possible, and in this issue, to obtain as much critical information of body functions with minimal wiring and burden on the patients [24]. A neonatal NIRS system with a robust HR extraction algorithm validated in clinical settings would allow for having a complementary view of the autonomic nervous system and systemic circulation, and cerebral hemodynamics within a single device.

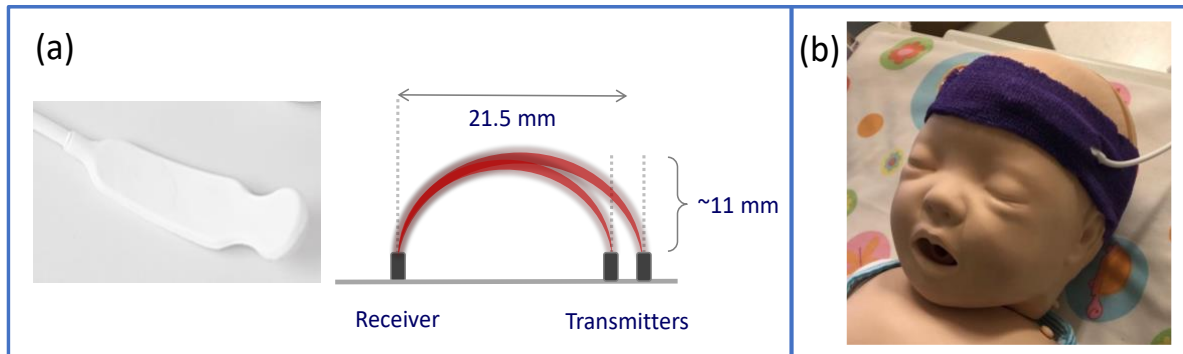
In this study, a novel algorithm, NHR (NIRS HR), is proposed for extracting HR from NIRS signals acquired from neonates admitted to the NICU. The algorithm is validated with respect to the reference HR measured with a clinically approved patient monitor system. In addition, we compare the performance of the NHR algorithm with two state-of-art algorithms: the Perdue algorithm [22], introduced by Perdue et al., validated on NIRS data recorded from healthy infants; the AMPD algorithm [21], introduced by Scholkmann et al., validated on NIRS data recorded from healthy adults. Since the proposed algorithm has some processing steps involved for signal quality assessment and motion artifact detection, we also investigate whether these processing steps would improve the HR extraction accuracy with respect to a simple version of the algorithm.

## 2. Data

### 2.1. Data acquisition

In this study, NIRS signals were recorded by using a cerebral oximetry system (TOM, Artinis Medical Systems B.V., Elst, the Netherlands). This device has a sensor with two transmitters firing at two nominal wavelengths of 760 and 850 nm and a receiver at a 21.5 mm distance from the transmitters. It provides optical densities as well as cerebral oxygenation, commonly known as rSO<sub>2</sub> or StO<sub>2</sub>, with a 100 Hz sampling rate. We placed the sensor on the neonate's

forehead on the opposite side the neonate was lying on. We used a clinical self-adhesive elastic bandage to keep placed the sensor on the neonate's forehead. Figure 1 shows a schematic of the configuration of the two transmitters and the receiver providing two NIRS channels in the sensor and a photograph of the sensor placed on a baby manikin's forehead, covered by the bandage. We recorded the reference HR (RHR) with a Philips IntelliVue MP70 (Philips Medical Systems, Best, the Netherlands) patient monitor, at a 1 Hz sampling rate.



**Figure 1:** (a) Schematic of the layout of the two transmitters and one receiver in the TOM sensor (Artinis Medical Systems B.V., the Netherlands), with the banana-shaped light pathways. The transmitters-receiver distance is 21.5 mm which provides an approximate penetration depth of 11 mm (half of 21.5 mm) [25,26]. (b) Photograph of the TOM sensor placed on a baby manikin's forehead, covered with a clinical self-adhesive elastic bandage.

## 2.2. Participants

All participants were admitted to the NICU of the Wilhelmina Children's Hospital, Utrecht, the Netherlands. This study was approved by the local ethics committee of the University Medical Center Utrecht, with the ethics code 21-098/C. We included ten newborn infants (gestational age =  $38 \pm 5$  weeks; birth weight =  $3092 \pm 990$  grams, 3 female) who already had an indication for NIRS monitoring, as determined by the caring physician or local neuromonitoring protocols. Parents were informed about the study protocol and gave written consent prior to the data acquisition. We recorded a total of 19 measurements with an average length of 1.5 hours from the included neonates, without interrupting clinical routines and with no specific experimental paradigm.

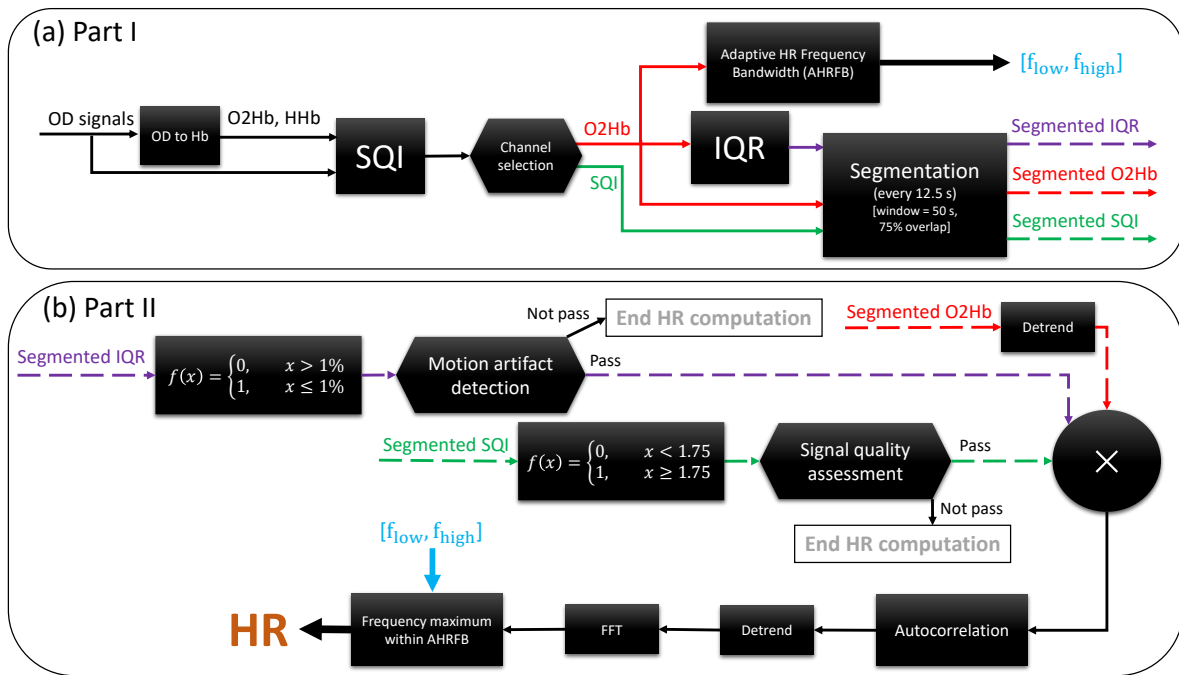
## 3. Heart rate extraction algorithm

Figure 2 illustrates the block diagram of the proposed HR extraction algorithm, NHR (NIRS HR). The algorithm is divided into two parts, Part I (Figure 2(a)), and Part II (Figure 2(b)). Each part of the algorithm is explained in detail in the next sections. The analyses in this study were implemented in Python using *numpy*, *scipy*, and *matplotlib* modules [27,28].

### 3.1. Part I

Part I of the NHR algorithm consists of four steps that are implemented once at the beginning of the analyses on the whole NIRS measurement (Figure 2(a)): Preprocessing, Adaptive HR Frequency Bandwidth (AHRFB), Interquartile range (IQR), and Segmentation.





**Figure 2:** The block diagram of the heart rate extraction algorithm proposed in this study, NHR (NIRS HR). The NHR algorithm consists of two parts: (a) Part I, consisting of four steps: Preprocessing (including OD to Hb, SQI, and Channel selection); Adaptive HR Frequency Bandwidth (AHRFB); Interquartile range (IQR); Segmentation. (b) Part II, consisting of three steps: Motion artifact detection; Signal quality assessment; HR computation (including Detrend, multiplication operator, Autocorrelation, FFT, and Frequency maximum within AHRFB).

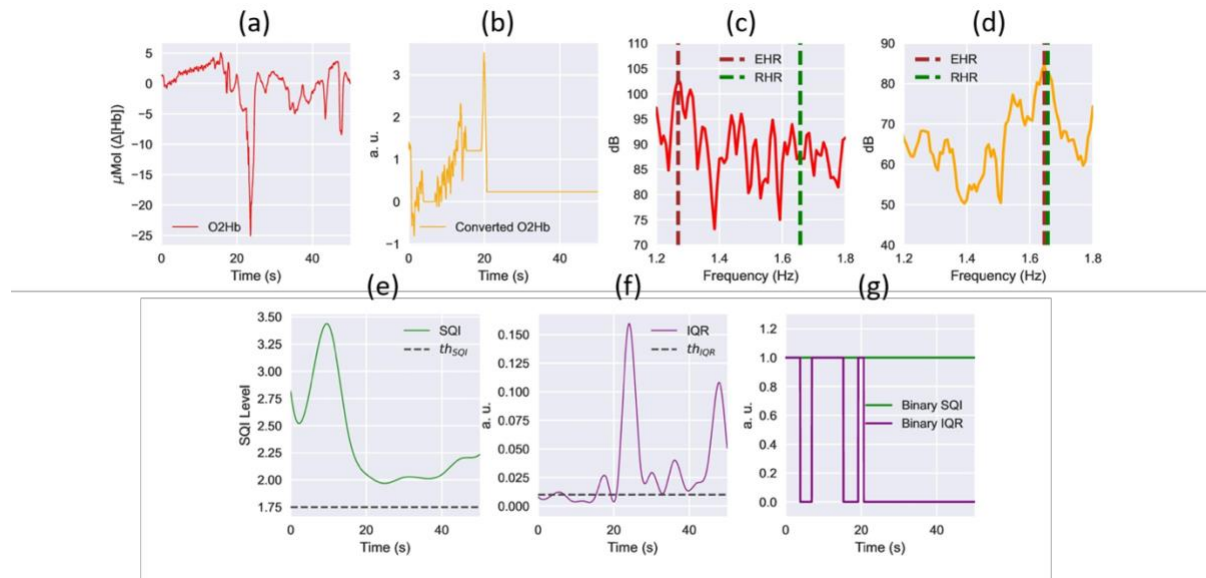
1) Preprocessing: First, the optical densities (ODs) of the two NIRS channels are converted to concentration changes in oxygenated hemoglobin (O2Hb) and deoxygenated hemoglobin (HHb) by using the modified Beer-Lambert law [29] (OD to Hb in Figure 2 (a)). Second, using the SQI (Signal Quality Index) algorithm [30], the SQI signals of the two channels are computed (SQI in Figure 2(a)), based on sliding windows of 10 seconds, overlapping by 50%. Third, the average of the SQI signals is computed separately, and finally, the channel with the greater average SQI is selected (Channel selection in Figure 2(a)). In the subsequent steps of the algorithm, the O2Hb signal and the SQI signal of the selected channel are used. To have an equal sampling rate, the SQI ratings of all 10-second windows are finally interpolated to 100 Hz using cubic interpolation.

2) Adaptive HR Frequency Bandwidth (AHRFB): The HR frequency bandwidth is adaptively adjusted per measurement according to a predefined bandwidth between 1.25 Hz and 3.5 Hz. The predefined bandwidth was selected according to the range of the RHR ( $75 \text{ BPM} < \text{RHR} < 210$ ). First, the O2Hb signal is low-pass filtered by using a moving average filter of 1 second, and then subtracted from the original O2Hb signal to filter out low-frequency components ( $\sim < 1 \text{ Hz}$ ). Second, the FFT (Fast Fourier Transformation) magnitude of the filtered signal multiplied by a Hamming window with the same length as the signal is computed. Third, the FFT magnitude within the predefined bandwidth ( $[1.25, 3.5] \text{ Hz}$ ) is sorted in descending order. Then the average frequency of the first fifty components ( $\sim 50\%$  of all components) is computed. Finally, the average frequency minus 0.5 Hz and plus 0.5 Hz are considered as the minimum and maximum thresholds of the adaptive HR frequency bandwidth that will be referred respectively to as  $f_{low}$  and  $f_{high}$  in the subsequent analyses.

3) Interquartile range (IQR): First, the difference between the third and first percentiles, known as the interquartile range, of the selected O2Hb signal is computed in sliding windows of 3

seconds. Second, the IQR values of all 3-second windows are interpolated to 100 Hz using cubic interpolation to have an equal sampling rate. Finally, the IQR signal is normalized by dividing it by the median of the O2Hb signal.

4) Segmentation: On a sliding window approach, the resulting O2Hb, SQI, and IQR signals are segmented into windows of 50 seconds, overlapping by 75% (shifted every 12.5 seconds). It is noteworthy that the first HR value will be computed at earliest at 50 seconds of measurement time. Figure 3(a) illustrates a segment of the O2Hb signal acquired from the data recorded from one of the participants (Measurement 6).



**Figure 3:** (a) A segment of the O2Hb signal acquired from one of the participants (Measurement 6). (b) The converted O2Hb signal calculated by multiplying the O2Hb signal with the binary SQI signal multiplied by the binary IQR signal. (c) The FFT magnitude of the autocorrelation computed from the O2Hb signal. Dashed green and brown lines depict the reference (RHR) and extracted (EHR) heart rates, respectively. (d) the FFT magnitude of the autocorrelation computed from the converted O2Hb signal. (e) The SQI signal calculated based on a sliding window approach of 10 seconds overlapping 50% and resampled to 100 Hz. The black dashed line depicts the SQI predefined threshold ( $th_{SQI} = 1.75$ ). (f) The IQR signal calculated from the O2Hb signal based on a sliding window approach of 3 seconds and resampled to 100 Hz. The dashed black line depicts the IQR predefined threshold ( $th_{IQR} = 1\%$ ). (g) The binary or thresholded SQI and IQR signals computed, shown in green and purple, respectively.

### 3.2. Part II

Part II of the NHR algorithm consists of three steps that are implemented every 12.5 seconds on each 50-second segment of the O2Hb, SQI, and IQR signals, computing or updating HR (Figure 2(b)): (1) Motion artifact detection, (2) Signal quality assessment, and (3) HR computation.

1) Motion artifact detection: First, the segmented (normalized) IQR signal is thresholded with respect to a predefined threshold ( $th_{IQR}$ ) of 1% and is then converted to a binary signal. The threshold was obtained empirically. The converted IQR signal is zero when the IQR is above the threshold, meaning presumably contaminated with motion artifact. Second, the percentage of the zeros in the binary signal is calculated. If the percentage is above 80%, meaning that more than 80% of the signal in the 50-second window is contaminated with motion artifact, no HR will be computed for the considered window. Otherwise, the algorithm passes to the next step to assess the signal quality of the signal segment. Figure 3(f) illustrates a segment of the

IQR signal computed from the 50-second signal segment shown in Figure 3(a). After comparing the IQR signal with the corresponding predefined threshold ( $th_{IQR}$ ), the binary IQR signal is computed, which is shown in purple in Figure 3(g). It is observed that the motion artifacts that exist in about the second half of the signal segment have been correctly captured.

2) Signal quality assessment: In this step, first, the segmented SQI signal is thresholded with respect to a predefined threshold ( $th_{SQI}$ ) of 1.75 and is then converted to a binary signal. The threshold was obtained empirically. The converted SQI signal is zero when the SQI is below the threshold. Second, the percentage of the zeros in the binary signal is calculated. If the percentage is above 25%, meaning that more than 25% of the signal in the considered window has a low signal quality (lower than an SQI of 1.75), then no HR for the considered window is computed. Otherwise, the algorithm passes to compute the HR. Figure 3(e) illustrates a segment of the SQI signal corresponding to the signal shown in Figure 3(a). It is observed that the signal quality in the second half of the segment is lower than that in the first half. After comparing the SQI signal with the corresponding predefined threshold ( $th_{SQI}$ ), the binary SQI signal is computed, which is shown in green in Figure 3(g).

3) HR computation: The segmented O2Hb signal is detrended (Detrend in Figure 2(b)) to force its mean to zero as well as to reduce its overall linear variation (trend). The detrending is conducted by subtracting the least-squares fit of a straight line to the data. Following, the detrended O2Hb signal is multiplied by the product of the binary (or thresholded) IQR and SQI signals (multiplication operator in Figure 2(b)). Figure 3(b) shows a segment of the converted O2Hb signal which is corresponding to the O2Hb signal shown in Figure 3(a). Next, the autocorrelation of the converted signal is computed and is then detrended (Autocorrelation and Detrend in Figure 2(b)). Then, the FFT of the autocorrelation signal multiplied by a Hamming window with the same length is calculated (FFT in Figure 2(b)). Finally, the HR in the considered window is computed by finding the frequency with the maximum magnitude within the AHRFB ( $[f_{low}, f_{high}]$ ). We multiply the maximum frequency by 60 to obtain the result in beats per minute, bpm (Frequency maximum within AHRFB in Figure 2(b)). Figure 3(d) illustrates the FFT magnitude of the autocorrelation computed from the converted O2Hb signal that was shown in Figure 3(b). The dashed green and brown lines depict the RHR and extracted HR (EHR), respectively. To show the impact of the conversion on the HR extraction, Figure 3(c) shows the FFT magnitude of the autocorrelation computed from the original O2Hb signal shown in Figure 3(a), so without any motion artefact detection and signal quality assessment. It is observed that there is a huge error of approximately 23 bpm (0.38 Hz) when the original O2Hb signal is used.

## 4. Algorithm performance assessment

We assessed the performance of the NHR algorithm by calculating four quantitative measures quantifying the agreement between the EHR and RHR. The quantitative measures employed are: mean of error (ME), root mean square error (RMSE), Bland-Altman limits of agreement (LoA) [31], and Bland-Altman ratio (BAR). LoA was computed by calculating 1.96 times the standard deviation of the error. It represents the range wherein 95% of the differences between the EHR and RHR are falling. BAR was computed by calculating the ratio of the LoA to the mean of the pairwise HRs (RHR and EHR). It is described as an efficient way to quantify the agreement between the reference and extracted HR signals [32]: the agreement is ranked

as good ( $\text{BAR} < 10\%$ ), moderate ( $10\% < \text{BAR} < 20\%$ ), or insufficient agreement ( $\text{BAR} > 20\%$ ). Furthermore, we computed Pearson's correlation to determine the linear association between the RHR and EHR, wherein the aforementioned quantitative measures fail in representing. We assessed the significance of the correlation with a Student's t-test at a significance level  $\alpha = 5\%$ , by using the Python *scipy.stats* module [28]. In addition, we counted the number of segments that have successfully passed the steps of Motion artifact detection and Signal quality assessment to compute the percentage of included segments. The mentioned quantitative measures were computed for each measurement separately, and the average and standard deviation for each measure were also calculated. To have a fair comparison between the EHR and RHR, we have used the average of the RHR in each 50 s segment.

We compared the performance of the NHR algorithm with the performance of two existing algorithms: the "Perdue" algorithm, proposed by Perdue et al. [22], used for extracting HR from NIRS in infants; the "AMPD" algorithm, proposed by Scholkmann et al. [17,21], used for extracting HR from NIRS in adults. Furthermore, to investigate the efficacy of the processing steps involved in the NHR algorithm for adaptively setting HR frequency bandwidth, detecting motion artifacts, and assessing signal quality, we also conducted a comparison with a simple version of the NHR algorithm. The simple version was obtained by excluding steps 2 and 3 in part I and steps 1 and 2 in part II of the NHR algorithm (see section 3), which will be referred to as the "Spectrum" algorithm. In the Perdue algorithm, the NIRS signals recorded with 10 Hz were upsampled to 100 Hz as a preprocessing step [22]. We exempted this step as the NIRS signals here were already recorded at a 100 Hz sampling rate. In both Perdue and AMPD algorithms, we took the average of the extracted HR in each 50 s segment, as it is implemented for the RHR signal (see previous paragraph). This was performed to have an accurate comparison between the performance of the existing algorithms and our proposed NHR algorithm.

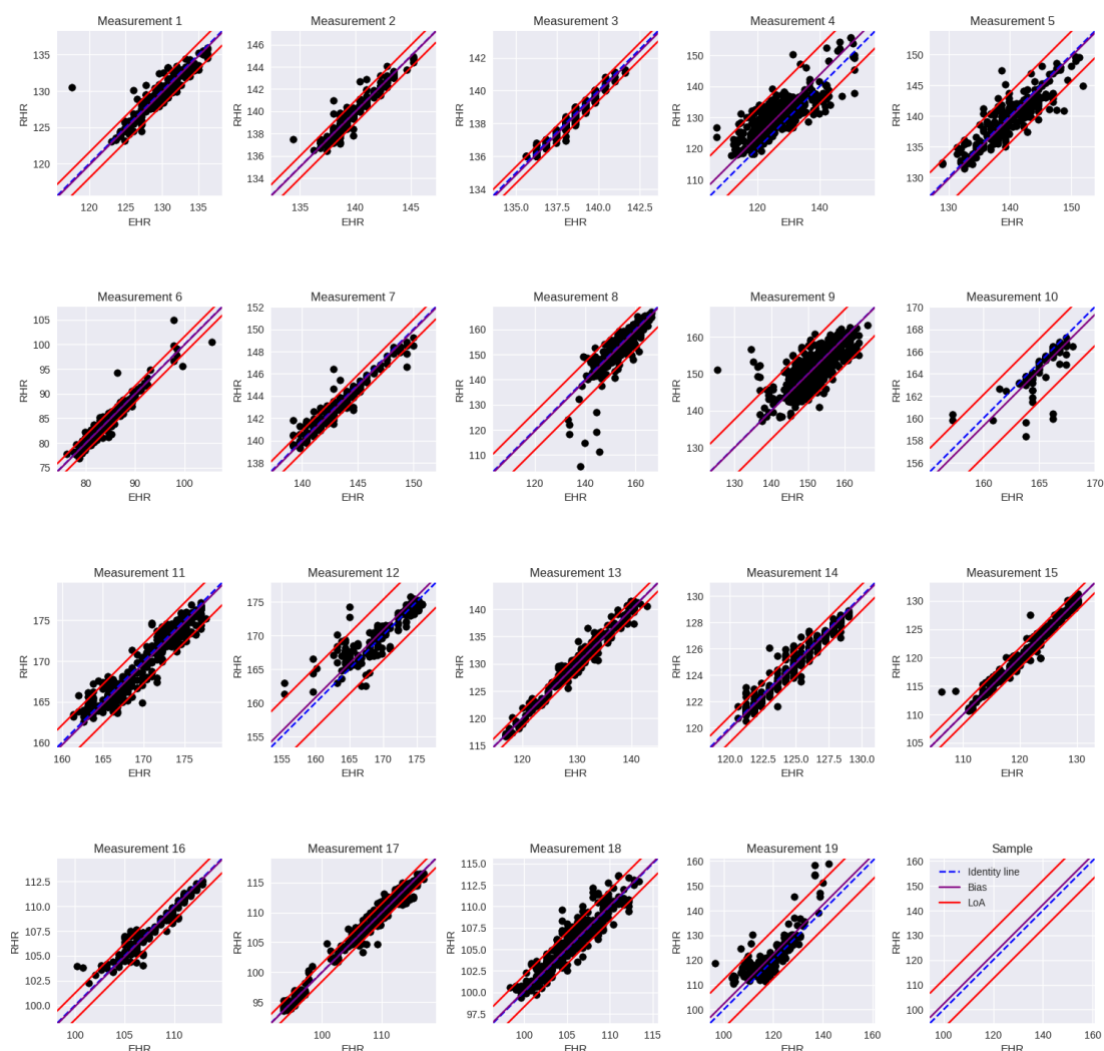
To quantitatively compare the performance of the algorithms, we concatenated all measurements together and then computed the mentioned quantitative measures, except the included segments' percentage. In addition, for a detailed comparison, we computed the quantitative metrics per measurement for the Perdue, AMPD, and Spectrum algorithms that are reported in the Appendix. As another comparison measure between the algorithms, we defined an agreement limit representing the range in which the difference between RHR and the respective EHR is less than 20% of the mean of the pairwise HRs (RHR and EHR) [33]. Furthermore, we applied Student's t-test with a significance level  $\alpha = 5\%$  to statistically compare the performance of the algorithms against each other.

We conducted a sensitivity analysis of the window length parameter for the NHR algorithm. We changed the window length from 10 s to 200 s (10 to 100 in steps of 10 s, and 100 to 200 in steps of 25 s) and computed EHR for all measurements by each window length selected. Starting from the shortest window length, we compared the EHR signals computed by each window length with the ones computed by the next greater window length (e.g., 10 s with 20 s). We used Student's t-test ( $\alpha = 5\%$ ) to find the significant cases when increasing the window length leads to a significantly different performance compared with the next shorter window length. With this analysis, we determined the optimal window length that provides less delay as well as statistically better performance than the other ones.

## 5. Results

### 5.1. Results of the proposed algorithm

Figure 4 illustrates the scatter plots comparing the RHR with the EHR computed by the proposed NHR algorithm for each measurement, separately. It is observed that the measurements had different HR ranges that have been properly captured by the algorithm. As an example, Measurement 6 has HRs ranging from  $\sim 75$  bpm to  $\sim 100$  bpm, while Measurement 12 has HRs ranging from  $\sim 160$  bpm to  $\sim 175$  bpm. From the scatter plots, it is observed that there is a significant correlation (i.e., linear association) between the RHR and EHR in all measurements ( $p < 0.05$ ). Looking at the bias (ME) in the scatter plots, we observe the bias in HR extraction to be negligible in all measurements. Looking at the LoA in the scatter plots, we observe it to be small in all measurements (exceptions: Measurements 4, 5, 9, 12, 19).



**Figure 4:** Scatter plots showing the association between the reference heart rate (RHR) and the extracted heart rate (EHR) by the NHR algorithm per measurement. The y and x axes represent the RHR and EHR in bpm (beats per minute), respectively. Each dot represents the HR in a 50s segment. The blue dashed line represents the identity line ( $y = x$  line) while the full purple and red lines represent the bias (mean of error, ME) and Bland-Altman limits of agreement (LoA), respectively. The correlation between RHR and EHR in all measurements is significant ( $p < 0.05\%$ ). The plot in the right-bottom corner is a sample plot to illustrate the identity, bias, and LoA lines.

The quantitative measures showing the agreement and linear association between the RHR and EHR signals for each measurement are reported in Table 1. The ME (or bias) is lower than 1 bpm in all measurements, except for Measurement 4 (ME = 3.7 bpm). The average and standard deviation of ME computed over all measurements are 0.2 bpm and 1.1 bpm, respectively. The RMSE is lower than 4 bpm in all measurements, except for Measurements 4 and 19 which is approximately 6 bpm, with an average and standard deviation of both approximately 2 bpm. The LoA has an average and standard deviation of approximately  $3 \pm 3$  bpm in all measurements, with the largest ones in Measurement 4 (9.2 bpm) and Measurement 19 (9.8 bpm). Likewise, the BAR has an average and standard deviation of  $2.5 \pm 2.2$  bpm in all measurements, with the highest ones in Measurement 4 (7.2%) and Measurement 19 (8.2%). The BAR percentage of each measurement falls within the good agreement category (BAR<10%) [32] between RHR and EHR. The EHR signal of each measurement showed a positive and significant ( $p < 0.05$ ) correlation of more than 74% with the RHR signal. Across all measurements, the average and standard deviation of Pearson's correlation are  $91.1 \pm 8.2\%$ . The included segments' percentage of each measurement is more than 95%, except for Measurements 8, 15, and 19 (88%, 82%, and 67%, respectively). Across all measurements, the average and standard deviation of the percentage of included segments are  $95.2 \pm 8.2\%$ , meaning that the NHR algorithm, on average, excluded approximately 5% of the signal segments, during which no HR computation could be performed.

**Table 1:** Quantitative measures for assessing the performance of the proposed NHR algorithm per measurement.

Measurement	ME <sup>a</sup> (bpm <sup>b</sup> )	RMSE <sup>c</sup> (bpm)	LoA <sup>d</sup> (bpm)	BAR <sup>e</sup> (%)	Pearson's r (%)	IS <sup>f</sup> (%)
1	-0.2	0.9	1.8	1.4	93.3	98.8
2	0.0	0.5	1.0	0.7	95.8	100.0
3	-0.1	0.3	0.5	0.4	97.5	99.6
4	3.7	6.0	9.2	7.2	81.2	95.0
5	-0.3	2.1	4.1	2.9	88.3	95.2
6	-0.1	0.9	1.7	2.0	96.4	95.2
7	-0.1	0.5	1.0	0.7	96.7	100.0
8	-0.5	3.9	7.5	4.9	85.7	88.4
9	-0.1	3.9	7.7	5.0	75.2	98.8
10	-0.7	1.6	2.8	1.7	74.2	100.0
11	-0.3	1.3	2.4	1.4	95.4	100.0
12	0.8	2.4	4.5	2.7	84.5	95.5
13	-0.1	0.8	1.5	1.2	98.3	98.0
14	-0.1	0.5	1.0	0.8	97.4	100.0
15	0.0	0.9	1.7	1.4	98.6	82.0
16	-0.1	0.7	1.3	1.2	96.5	100.0
17	-0.1	0.8	1.5	1.4	99.3	96.3
18	0.1	0.9	1.8	1.8	95.2	98.9
19	2.3	5.5	9.8	8.2	81.7	67.1
Average	0.2	1.8	3.3	2.5	91.1	95.2
Std <sup>g</sup>	1.1	1.7	3.0	2.2	8.2	8.2

<sup>a</sup> Mean of error.

<sup>b</sup> beats per minute.

<sup>c</sup> Root mean square error.

<sup>d</sup> Band-Altman limits of agreement.

<sup>e</sup> Bland-Altman ratio.

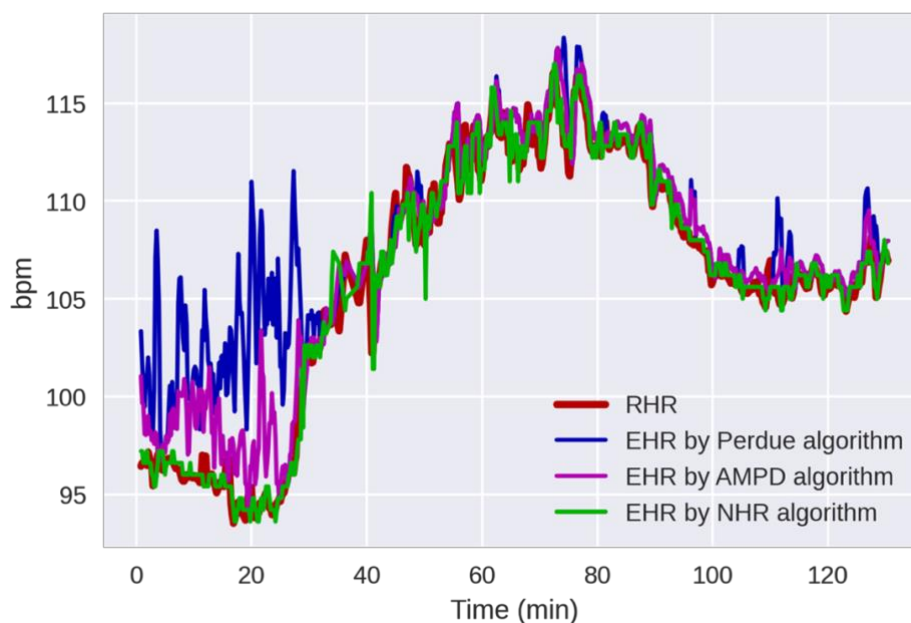
<sup>f</sup> Standard deviation.

<sup>g</sup> Included segments

## 5.2. Comparison with the existing algorithms

Figure 5 shows an example from one of the NIRS measurements (Measurement 17) of the EHR signals extracted by using our proposed NHR algorithm (in green), the Perdue algorithm

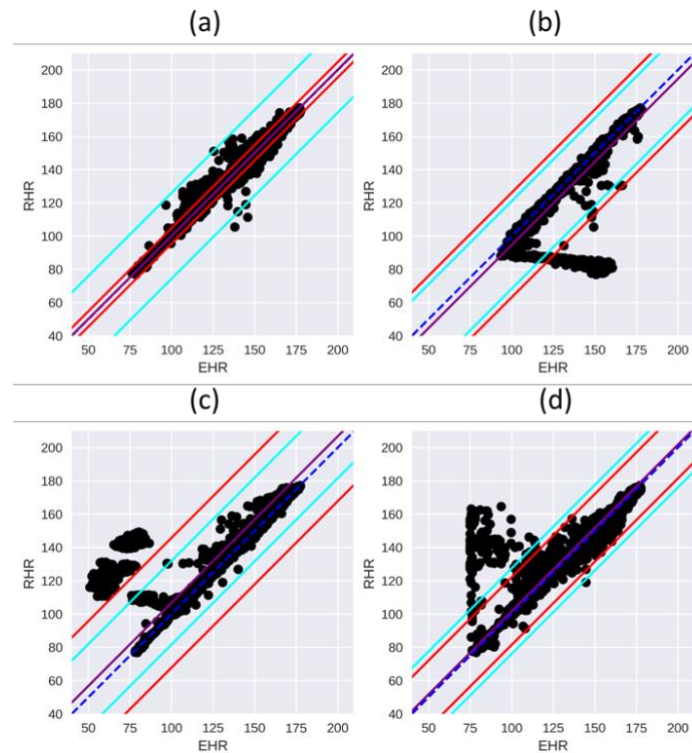
[22] (in blue), and the AMPD algorithm [21] (in purple). Our derived EHR signal is a better representation of the RHR (in red) that has properly followed the trend of the RHR signal from the beginning to the end of the measurement. In this measurement, Pearson's correlation between the RHR and the EHR computed by our proposed NHR algorithm is 99.3% ( $p < 0.05$ ) while that is 87.5% ( $p < 0.05$ ) by using the Perdue algorithm, and 93.2% ( $p < 0.05$ ) by using the AMPD algorithm.



**Figure 5:** An example of the extracted HR (EHR) signals by using the NHR algorithm (in green), the Perdue algorithm [22] (in blue), and the AMPD algorithm [21] (in purple), compared with the reference heart rate (RHR) signal (in red). The x and y axes represent the time in minutes (min) and the HR in beats per minute (bpm), respectively. In this measurement (Measurement 17), Pearson's correlation between the RHR and EHR by the NHR algorithm is 99.3% ( $p < 0.05$ ), whereas it is 87.5% ( $p < 0.05$ ) by the Perdue algorithm, and 93.2% ( $p < 0.05$ ) by the AMPD algorithm.

Figure 6 illustrates the scatterplots between the RHR and EHR computed by the NHR algorithm (Figure 6(a)), the Perdue algorithm (Figure 6(b)), the AMPD algorithm (Figure 6(c)), and the Spectrum algorithm (Figure 6(d)) when all the measurements were concatenated together. Data points in the scatter plots for the Perdue, AMPD, and Spectrum algorithms are more sparsely distributed around the identity line ( $y=x$  line) than the one for the proposed NHR algorithm in this study. Likewise, compared with the NHR algorithm (4 points), the Perdue (574 points), AMPD (847 points), and Spectrum (255 points) algorithms have much more data points outside the defined agreement limit (cyan lines), indicating errors of more than 20% of the pairwise HRs mean. In addition, the Perdue, AMPD, and Spectrum algorithms are observed to have a bias in computing HR, e.g., the Perdue algorithm toward a higher value (i.e., more data points below the identity line) and the AMPD and Spectrum algorithms toward a lower value (i.e., more data points above the identity line) than the RHR. On the contrary, this is not the case for the NHR algorithm as the bias seems to be negligible. Furthermore, the Perdue, AMPD, and Spectrum algorithms are observed to be unable to adaptively capture different HR ranges, e.g., the Perdue algorithm in the RHR lower than approximately 100 bpm and AMPD in the RHR between approximately 100 and 150 bpm. Conversely, the NHR algorithm seems to be adaptable to the different ranges of HR, i.e., leaving the dots close to the identity line in the scatterplot for the whole HR range (Figure 6(a)).





**Figure 6:** Scatter plots between the reference heart rate (RHR) and the extracted heart rate (EHR) by (a) the NHR algorithm, (b) the Perdue algorithm [22], (c) the AMPD algorithm [21], and (d) the spectrum algorithm, for all the measurements concatenated together. The y and x axes in each plot represent the RHR and EHR in bpm (beats per minute), respectively. Each dot represents the HR in a 50s segment. The blue dashed line represents the identity line ( $y = x$  line) while the full purple and red lines represent the bias (ME) and Bland-Altman limits of agreement (LoA), respectively. The cyan lines represent the defined agreement limit. They are in 20% of the mean of the pairwise HRs (RHR and EHR) distance from the identity line.

The quantitative measures showing the agreement and linear association between RHR and EHR computed by the four algorithms in all measurements concatenated together are reported in Table 2. The ME (or bias) is higher for the Perdue (ME=-5.3 bpm) and AMPD (6.8 bpm) algorithms than that for the Spectrum algorithm (1.8 bpm) and NHR algorithm (0.2 bpm). The RMSE is approximately 17 bpm, 21 bpm, and 10.5 bpm by the Perdue, AMPD, and Spectrum algorithms, respectively, whereas it is low as 2.4 bpm by the NHR algorithm. The LoA is higher for the Perdue (LoA=31.5 bpm), AMPD (39.0 bpm), and Spectrum (20.3 bpm) algorithms than that for the NHR algorithm (4.6 bpm), this was also observed in the scatter plots shown in Figure 6. Looking at the BAR percentages, we obtained a good agreement (BAR<10%) between RHR and EHR for the NHR algorithm, whereas the agreement between RHR and EHR for the Perdue and AMPD algorithms was insufficient (BAR>20%) and for the Spectrum algorithm was moderate (10%<BAR<20%). Pearson's correlation between the RHR and EHR for all algorithms was significant ( $p < 0.05$ ); however, the correlation obtained by the NHR algorithm was approximately 28%, 26%, and 15.8% greater than the one obtained by the Perdue (71.8%), AMPD (73.2%), and Spectrum (83.7%) algorithms, respectively. Figure 7 illustrates the raincloud plot of the absolute error between RHR and EHR signals obtained for the four algorithms. In this figure, we observe that the majority of the errors by the NHR algorithm are about zero, whereas the errors made by the Perdue, AMPD, and Spectrum algorithms are more sparsely distributed toward higher values, leaving an additional peak close to the greatest error. This additional peak in the raincloud plot for the Perdue algorithm is due to the huge error in computing HRs lower than approximately 100 bpm as was observed in Figure 6. Conversely, the additional peak in the raincloud plots for the AMPD and Spectrum

algorithms is due to the error observed in computing HRs above approximately 100 bpm. The result of the statistical comparison (Student's t-test with  $\alpha = 5\%$ ) between the performance of four algorithms exhibits that there is a significant difference ( $p < 0.05$ ) between the performance of the algorithms against each other.

**Table 2:** Quantitative measures for assessing the performance of the proposed NHR algorithm, the Perdue [22], the AMPD [21], and the Spectrum algorithms in all measurements concatenated together.

Algorithm	ME <sup>a</sup> (bpm <sup>b</sup> )	RMSE <sup>c</sup> (bpm)	LoA <sup>d</sup> (bpm)	BAR <sup>e</sup> (%)	Pearson's r (%)
NHR	0.2	2.4	4.6	3.6	99.5
Perdue	-5.4	16.9	31.5	23.9	71.8
AMPD	6.8	21.0	39.0	31.0	73.2
Spectrum	1.8	10.5	20.3	15.8	83.7

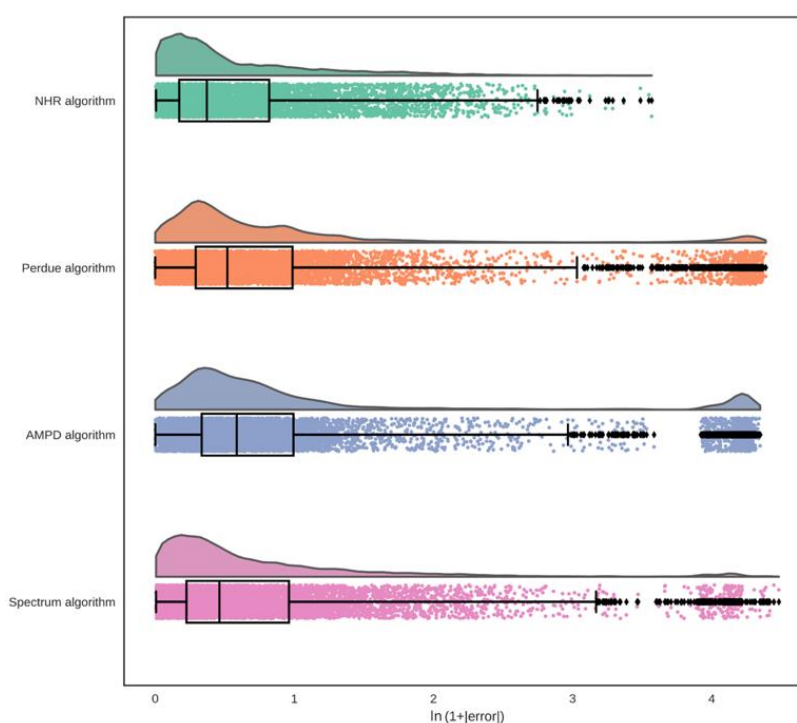
<sup>a</sup> Mean of error.

<sup>b</sup> beats per minute.

<sup>c</sup> Root mean square error.

<sup>d</sup> Band-Altman limits of agreement.

<sup>e</sup> Bland-Altman ratio.

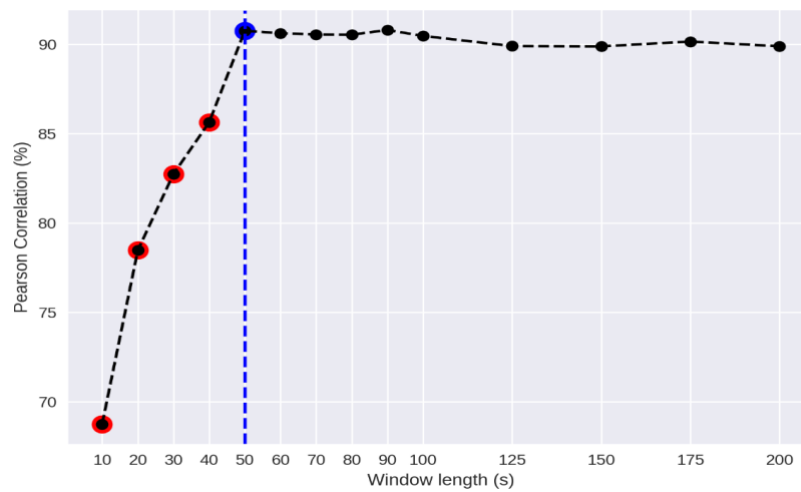


**Figure 7:** Raincloud plot of the absolute error between RHR and EHR computed by the NHR algorithm (first panel), the Perdue algorithm [22] (second panel), the AMPD algorithm [21] (third panel), and the Spectrum algorithm (fourth panel). For the sake of having better visualization in the distributions, the natural logarithm of the absolute error plus one (x-axis) was computed.

### 5.3. Determining the optimal window length

The optimal window length in the NHR algorithm was determined by changing the window length of the sliding window approach implemented in the Segmentation step. Figure 8 illustrates the Pearson's correlation computed between RHR and EHR with respect to the window length. It is generally observed that by selecting a greater window length, a greater correlation was achieved. Any window length greater than or equal to 30 seconds resulted in a correlation higher than 80%. Nevertheless, there was no statistically significant difference in the performance of the NHR algorithm when comparing window lengths greater than 50

seconds. Therefore, we defined the shortest optimal window length to be 50 seconds, because it statistically performs better than shorter window lengths ( $p < 0.05$ ), but not worse than longer window lengths ( $p > 0.05$ ).



**Figure 8:** Pearson’s correlation between the RHR and EHR when changing the window length in the sliding window approach implemented in the NHR algorithm. X and Y axes represent the window length in seconds and Pearson’s correlation in percentage, respectively. From the shortest window length upwards, the outcome of each window was compared statistically (Student’s t-test with  $\alpha = 5\%$ ) with the outcome of the next greater window length. The red and blue dots represent the significant cases ( $p < 0.05$ ) and the first insignificant case, respectively. The blue dashed line depicts the window length corresponding to the first insignificant case, indicating the shortest optimum window length.

## 6. Discussion

In this study, we developed a novel algorithm for extracting HR from the NIRS signals recorded from the critically ill neonates admitted to the NICU. To the best of our knowledge, the proposed NHR algorithm is the first HR extraction algorithm developed on neonatal NIRS signals acquired in a clinical environment. The existing AMPD algorithm, introduced by Scholkmann et al. [21], was validated on signals recorded from healthy adults in a controlled environment [17,18]. The existing Perdue algorithm, introduced by Perdue et al. [22], was developed on signals recorded from infants older than 3 months and in a controlled environment, whereas our proposed algorithm was developed on signals recorded from newborn infants admitted to NICU, i.e., in real-life clinical settings. Our results showed that the NHR algorithm provides a significantly ( $p < 0.05$ ) better performance in HR extraction from neonatal NIRS signals than the Perdue and AMPD algorithms (see Table 2 and Figures 6, 7). The NHR algorithm numerically showed a greater correlation ( $r=99.5\%$ ,  $p < 0.05$ ) between the EHR and RHR than the Perdue algorithm ( $r=71.8\%$ ,  $p < 0.05$ ), and the AMPD algorithm ( $r=73.2\%$ ,  $p < 0.05$ ). Furthermore, the NHR algorithm showed a higher agreement between the EHR and RHR than the Perdue and AMPD algorithms, in terms of ME, RMSE, LoA, and BAR (see Table 2).

Data acquisition in clinical settings has some challenges that have led us to use some novelties in the NHR algorithm by adding some processing steps in order to achieve a robust algorithm in HR extraction from neonatal NIRS. Compared with a controlled environment, the data recorded in a clinical environment contains more motion artifacts induced by patient movement, sensor replacement, and ventilator use. Hence, in this study, we have performed

steps in the NHR algorithm to detect motion artifacts per segment and reform the contaminated NIRS signal (see section 3.2). Figure 3 shows an example of a 50-second segment of the NIRS signal which is contaminated with motion artifacts (see Figure 3(a)). This figure illustrates the IQR (Figure 3(f)), thresholded IQR (Figure 3(g)), the converted NIRS signal (Figure 3(b)), and the spectra of the converted signal (Figure 3(d)) and the original signal (Figure 3(c)), i.e., without any motion artifact detection and signal quality assessment. With the conversion, the spectrum was enhanced in the vicinity of the reference HR frequency (1.66 Hz), and thereby a more accurate HR was computed (1.65 Hz vs 1.27 Hz). By using the proposed algorithm, we excluded only those signal segments where motion artifacts were detected in over 80% of the segment. Looking at the results reported in Table 1, we are including the majority of the signal segments for the HR computation, i.e., on average only excluding 5% of the 50-second segments which were highly (>80%) contaminated with motion artifacts. Another challenge in clinical data is the low data quality, which is partly due to the restrictions of data acquisition in clinical settings. To overcome this challenge, we have performed three steps in the NHR algorithm to select the highest quality NIRS channel (see section 3.1), assess the NIRS signal quality per segment, and include the good quality signals in the analyses (section 3.2). Another challenge in clinical data is the variability of the signal source, here the variability of HR (see Figure 4). To overcome this challenge, we have adaptively set the HR frequency bandwidth in each measurement in the NHR algorithm (see section 3.1). Our results of the comparison between the Spectrum and NHR algorithms showed the outperformance ( $p < 0.05$ ) of the NHR algorithm and therefore confirmed the efficacy of the processing steps involved in the NHR algorithm in the HR extraction with respect to the simple version of the NHR algorithm (i.e., Spectrum algorithm). For example, the additional processing steps not only improved the linear correlation (99.5% vs 83.7%) but also enhanced the agreement between the EHR and RHR in terms of ME (0.2 vs 1.8 bpm), RMSE (2.4 vs 10.5 bpm), LoA (4.6 vs 20.3 bpm), and BAR (3.6 vs 15.8%).

When an HR extraction algorithm is based on peak detection, such as the Perdue [22] and AMPD [17,21] algorithms, there will be an unavoidable discretization error in HR computation. The NIRS signals recorded in this study were sampled at 100 Hz, whereas the NIRS signals recorded in the Perdue et al. study [22] and Scholkmann et al. study [21] were sampled at 10 Hz. The maximum error in peak detection with a 10 Hz sampling rate due to discretization error is 0.05 s, whereas it is 0.005 s with a 100 Hz sampling rate. The peak detection error yields an HR computation error depending on the HR range. For instance, if the HR is 65 bpm ( $= \frac{60}{\text{interbeat interval} = 0.92 \text{ s}}$ ), which is approximately the average resting state HR in male adults reported in [34], then the maximum error in HR computation due to discretization error with a 10 Hz sampling rate will be approximately 3 bpm ( $\frac{60}{0.92} - \frac{60}{0.92 + 0.05}$ ), whereas it will be approximately 0.4 bpm with a 100 Hz sampling rate. Although this might not be a considerable error, it will be more pronounced in infant data. As an example, if the HR is 130 bpm, which is the average RHR in the neonates admitted in this study, the maximum error in HR computation with a 10 Hz sampling rate will be approximately 13 bpm, whereas it will be approximately 1 bpm with a 100 Hz sampling rate. This explains the importance of having a high sampling rate in the NIRS system when extracting HR signals in infants.

Another advantage of having the NIRS signal recorded with a high sampling rate (100 Hz) is that it provides the opportunity to have an accurate assessment of the signal quality as well as motion artifacts. In this study, we have employed the SQI algorithm [30] to assess the NIRS signal quality. This algorithm quantifies the strength of the heartbeat components in the NIRS

signals, representing the strength of the optode-scalp coupling. This assessment requires signals recorded with a high sampling rate, e.g., 50 Hz as has been used for developing SQI. As an additional step, in this study, we have employed IQR to detect the occurrence of motion artifacts. A high sampling rate provides more samples in the window where the IQR is computed rather than a low sampling rate, thereby a more accurate representation of motion artifacts (with different frequency components) is obtained. The impact of employing these two assessments is shown in Figure 3 which has led to a more accurate HR computation.

Although the results we obtained for the proposed NHR algorithm are promising, further improvements are possible. The NHR algorithm updates HR every 12.5 seconds, whereas the Perdue and AMPD algorithms could compute HR beat-by-beat. This delay in HR computation makes this algorithm ineffective in applications where short-term characteristics of HR are analyzed [35,36]. The proposed NHR algorithm has been designed to be implemented offline on the NIRS measurements; however, an online version of the algorithm could be established with relatively few adjustments. These adjustments would be in part I of the algorithm in steps of Preprocessing and Adaptive HR Frequency Bandwidth. For instance, in the Preprocessing step when selecting the best channel, instead of computing the SQI for the whole measurement, we could compute it from the beginning of the measurement up to the current moment. This could be repeated at several-minute intervals, which depends on the variability of the signals. Moreover, instead of using SQI values from the beginning of the measurement, we could use the values in the last several minutes. Also, it is possible to store the SQI values computed in each segment within Part II of the algorithm and then we could reuse them when selecting the best channel. Similarly, in the Adaptive HR Frequency Bandwidth step, the adaptive bandwidth could be calculated using the O2Hb signal either from the start of the measurement up until the present moment or during the last several minutes.

The NHR algorithm was validated in this study on a dataset of in total 7834 HR samples (19 measurements), shown as dots in Figs 4, 6. However, eliminating excessive HR monitoring sensors used in clinical settings requires further validation of the algorithm on a more extensive dataset acquired from neonates in the clinical environment, especially extremely preterm infants. Looking at the public datasets, to the best of our knowledge, there is no public dataset including raw NIRS data with a high sampling rate as well as reference HR that all were recorded from newborn infants in a clinical environment. Therefore, we were unable to validate the proposed algorithm on a larger and more extensive dataset than the one employed in this study. Furthermore, we were unable to use the public PPG datasets as the proposed algorithm with the defined processing steps cannot be applied to the PPG data, i.e., we have used methods in the proposed algorithm for the assessment of signal quality and motion artifacts that were proposed and validated on NIRS data.

Extracting HR from NIRS signals could have several advantages over traditional ECG and PPG methods. First, NIRS is a non-invasive technique that does not require electrodes or sensors to be adhesively attached to the skin. This could be particularly beneficial for fragile newborns, especially extremely preterm infants, in the NICU who have very delicate skin, especially in the first few days after birth [17]. Routinely placing and removing the adhesive sensors on the skin in these fragile newborns are known to induce discomfort, stress, and pain, which are reported to be linked with poor neurodevelopmental outcomes [14]–[16]. Second, in neonates, the PPG signals recorded in the periphery are more prone to motion artifacts than NIRS signals recorded on the forehead. This is because hand movements in newborn infants are more frequent than head movements [16]. So that some studies have

shown that newborn infants in the supine position tend to lie with their heads turned to the right, accounting for 70% to 85% of their lying time [37–39]. Moreover, it has been shown that the pulse oximeter sensor placed over the forehead has more sensitivity to arterial pulsatile changes under low perfusion conditions than the sensor placed over peripheral body locations [40]. This is due to the thin-skin layer of the forehead, which is coupled with a prominent bone structure, that helps to direct light back to the photodetector [41]. Third, two existing NIRS studies of stress assessment on adults have shown that the HR (i.e., pulse rate) derived from NIRS is a richer source of information than the HR derived from PPG and ECG under physical and mental stress conditions [17,18]. Hence, the HR signal derived from NIRS might be more beneficial in both research and clinical environments for the assessment of the autonomic nervous system; however, more studies need to be conducted to confirm the greater richness of the NIRS-derived HR. Fourth, NIRS provides continuous monitoring of cerebral blood flow and oxygenation [3]. Therefore, extracting HR from NIRS could help clinicians to simultaneously measure changes in perfusion, oxygenation, and HR and therefore better understand the physiological mechanisms underlying different health conditions and interventions in newborns. Lastly, NIRS devices are generally portable and easy to use [10], which could make them a convenient option for monitoring HR and other physiological variables in the NICU. This could be particularly beneficial in resource-limited settings or in situations where rapid, non-invasive monitoring is necessary.

This study aimed to harness the richness of high sampling rate NIRS signals that are recorded in clinical settings, i.e., taking advantage of the heartbeat information in NIRS, and proposing a robust algorithm for the HR extraction from clinical neonatal data. Although the proposed NHR algorithm has more delay in computing HR than the existing Perdue and AMPD algorithms, i.e., 12.5 seconds with respect to HR computation beat-by-beat, the NHR algorithm showed significantly ( $p < 0.05$ ) a greater accuracy in extracting HR from neonatal NIRS data. A NIRS system with the NHR algorithm provides the opportunity to have synchronized HR and cerebral oxygenation. NIRS-derived cerebral oxygenation is known as a surrogate for cerebral perfusion [42], and HR has been proposed as a considerable alternative for blood pressure in preterm infants [43]. Therefore, a high sampling rate NIRS system with the NHR algorithm could be used potentially as a future standalone system for assessing cerebrovascular autoregulation in preterm infants. Furthermore, the feasibility of the NIRS cerebral oxygenation and HR have been supported by a growing body of research in optimizing monitoring in newborns with hypoxic-ischemic encephalopathy, adding to neurodevelopmental outcome prediction and also in assessing sleep stages in infants [6,7,44–54]. Therefore, the fusion of the cerebral oxygenation and HR provided by the standalone NIRS system could be investigated in infants as a future study.

## 7. Conclusion

In this study, we developed a novel algorithm, NHR (NIRS HR), for extracting HR from clinical NIRS signals recorded in neonates. The results numerically showed the compatibility of the HR extracted from NIRS with the reference HR, measured with ECG, in terms of Pearson's correlation ( $r > 0.9$ ) and Bland-Altman ratio ( $< 5\%$ ). The NHR algorithm significantly outperformed two existing algorithms, i.e., Perdue and AMPD algorithms. The promising results achieved suggest that the feasibility of the NHR algorithm could be exploited in clinical

applications where a combination of NIRS and HR monitoring can contribute to monitoring the vulnerable newborn brain.

## Appendix

**Table 3:** Quantitative measures for assessing the performance of the Perdue, AMPD, and Spectrum algorithms per measurement.

Perdue Algorithm					
Measurement	ME <sup>a</sup> (bpm <sup>b</sup> )	RMSE <sup>c</sup> (bpm)	LoA <sup>d</sup> (bpm)	BAR <sup>e</sup> (%)	Pearson's r (%)
1	-0.7	3.1	5.9	4.5	62.8
2	-0.6	1.0	1.5	1.0	91.9
3	-0.6	1.0	1.5	1.1	82.2
4	-0.0	1.2	2.3	1.7	98.4
5	-0.4	1.1	2.0	1.4	96.4
6	-57.8	60.6	35.9	31.9	-83.3
7	-0.8	1.4	2.3	1.6	89.3
8	-1.4	4.0	7.3	4.8	86.9
9	-1.3	2.5	4.2	2.8	91.8
10	-1.2	3.6	6.7	4.1	6.1
11	-0.7	1.2	1.9	1.1	97.0
12	0.3	1.1	2.1	1.2	95.3
13	-0.9	1.8	3.2	2.5	93.4
14	-0.6	1.0	1.6	1.2	93.4
15	-1.1	1.6	2.3	1.9	97.4
16	-2.4	3.2	4.1	3.8	60.8
17	-2.2	3.9	6.4	6.0	87.5
18	-1.8	2.2	2.7	2.6	91.4
19	-0.1	1.1	2.2	1.8	99.2
Average	-4.0	5.1	5.1	4.1	75.7
Std <sup>f</sup>	12.7	13.1	7.5	6.7	43.0
AMPD Algorithm					
1	-0.1	2.8	5.4	4.2	62.0
2	-0.5	0.7	1.0	0.7	96.1
3	-0.2	0.3	0.5	0.3	98.0
4	0.6	1.7	3.1	2.4	96.8
5	-0.0	0.8	1.6	1.2	97.7
6	-0.5	1.0	1.6	1.9	96.8
7	68.4	68.5	6.1	5.6	45.2
8	0.3	3.1	6.1	4.0	90.7
9	-0.4	1.5	2.7	1.8	96.6
10	-0.2	1.4	2.8	1.7	83.2
11	0.7	2.1	3.8	2.2	90.4
12	4.4	5.9	7.8	4.6	78.9
13	-0.0	2.5	5.0	3.8	80.3
14	-0.9	1.1	1.2	1.0	96.6
15	60.7	61.0	11.4	12.4	44.6
16	7.7	13.1	20.8	20.1	-18.4
17	-0.5	1.2	2.2	2.0	93.2
18	-0.9	1.6	2.6	2.5	90.2
19	1.2	4.5	8.5	7.0	88.7
Average	7.3	9.2	5.0	4.2	79.3
Std	19.8	19.3	4.7	4.7	28.9
Spectrum Algorithm					
1	0.6	6.0	11.8	9.1	26.7
2	-0.0	0.6	1.3	0.9	93.5
3	-0.1	0.4	0.7	0.5	95.2
4	3.2	7.4	13.0	10.1	75.6
5	1.2	10.6	20.6	14.8	27.8
6	0.2	3.0	5.9	7.0	82.3
7	15.2	30.7	52.5	38.8	-31.2



## Chapter 7

---

8	2.3	15.5	30.0	19.8	51.2
9	-0.9	4.1	7.8	5.2	74.8
10	-0.7	1.7	3.1	1.9	66.5
11	-0.5	1.4	2.6	1.5	94.6
12	-0.2	1.7	3.4	2.0	87.8
13	1.6	9.1	17.5	13.6	13.4
14	-0.1	0.6	1.2	0.9	96.4
15	5.8	17.2	31.9	26.8	-15.0
16	-0.1	0.7	1.3	1.2	96.7
17	0.0	1.1	2.2	2.0	98.5
18	0.2	2.2	4.3	4.1	78.0
19	5.9	17.2	31.7	25.2	56.5
Average	1.8	6.9	12.8	9.8	61.6
Std	3.7	7.9	14.1	10.6	38.5

<sup>a</sup> Mean of error.

<sup>b</sup> beats per minute.

<sup>c</sup> Root mean square error.

<sup>d</sup> Band-Altman limits of agreement.

<sup>e</sup> Bland-Altman ratio.

<sup>f</sup> Standard deviation.

## References

1. Vliet, F.F.J. Discovery of the Near-Infrared Window into the Body and the Early Development of near-Infrared Spectroscopy. *https://doi.org/10.1117/1.429952* **1999**, *4*, 392–396, doi:10.1117/1.429952.
2. Frigo, G.; Brigadoi, S.; Giorgi, G.; Sparacino, G.; Narduzzi, C. Measuring Cerebral Activation From FNIRS Signals: An Approach Based on Compressive Sensing and Taylor-Fourier Model. *IEEE Trans Instrum Meas* **2016**, *65*, 1310–1318, doi:10.1109/TIM.2016.2518363.
3. Costa, F.G.; Hakimi, N.; van Bel, F. Neuroprotection of the Perinatal Brain by Early Information of Cerebral Oxygenation and Perfusion Patterns. *International Journal of Molecular Sciences* **2021**, *Vol. 22*, *Page 5389* **2021**, *22*, 5389, doi:10.3390/IJMS22105389.
4. Wang, L.; Izzetoglu, M.; Du, J.; Ayaz, H. Phantom and Model-Based Near Infrared Spectroscopy Measurements of Intracranial Hematoma From Infants to Adults. *IEEE Trans Instrum Meas* **2022**, *71*, doi:10.1109/TIM.2022.3157343.
5. Hansen, M.L.; Hyttel-Sørensen, S.; Jakobsen, J.C.; Gluud, C.; Kooi, E.M.W.; Mintzer, J.; de Boode, W.P.; Fumagalli, M.; Alarcon, A.; Alderliesten, T.; et al. Cerebral Near-Infrared Spectroscopy Monitoring (NIRS) in Children and Adults: A Systematic Review with Meta-Analysis. *Pediatric Research* **2022**, *1–12*, doi:10.1038/s41390-022-01995-z.
6. Peng, S.; Boudes, E.; Tan, X.; Saint-Martin, C.; Shevell, M.; Wintermark, P. Does Near-Infrared Spectroscopy Identify Asphyxiated Newborns at Risk of Developing Brain Injury during Hypothermia Treatment? *Am J Perinatol* **2015**, *32*, 555–563, doi:10.1055/S-0034-1396692/ID/JR140192-35.
7. Mitra, S.; Bale, G.; Meek, J.; Tachtsidis, I.; Robertson, N.J. Cerebral near Infrared Spectroscopy Monitoring in Term Infants with Hypoxic Ischemic Encephalopathy—a Systematic Review. *Front Neurol* **2020**, *11*, 1–17, doi:10.3389/FNEUR.2020.00393/BIBTEX.
8. Dani, C.; Pratesi, S.; Fontanelli, G.; Barp, J.; Bertini, G. Blood Transfusions Increase Cerebral, Splanchnic, and Renal Oxygenation in Anemic Preterm Infants. *Transfusion (Paris)* **2010**, *50*, 1220–1226, doi:10.1111/J.1537-2995.2009.02575.X.
9. Pavlek, L.R.; Mueller, C.; Jebbia, M.R.; Kielt, M.J.; Fathi, O. Near-Infrared Spectroscopy in Extremely Preterm Infants. *Front Pediatr* **2021**, *8*, 973, doi:10.3389/FPED.2020.624113/BIBTEX.
10. Bozkurt, A.; Rosen, A.; Rosen, H.; Onaral, B. A Portable near Infrared Spectroscopy System for Bedside Monitoring of Newborn Brain. *Biomed Eng Online* **2005**, *4*, 1–11, doi:10.1186/1475-925X-4-29/FIGURES/11.
11. Kussman, B.D.; Wypij, D.; Dinardo, J.A.; Newburger, J.W.; Mayer, J.E.; Del Nido, P.J.; Bacha, E.A.; Pigula, F.; McGrath, E.; Laussen, P.C.; et al. Cerebral Oximetry During Infant Cardiac Surgery: Evaluation of and Relationship to Early Postoperative Outcome. *Anesth Analg* **2009**, *108*, 1122, doi:10.1213/ANE.0B013E318199DCD2.

12. Izzetoglu, M.; Bunce, S.; Izzetoglu, K.; Onaral, B.; Pourrezaei, A. Functional Brain Imaging Using Near-Infrared Technology. *IEEE Engineering in Medicine and Biology Magazine* **2007**, *26*, 38–46, doi:10.1109/MEMB.2007.384094.
13. Tong, Y.; Lindsey, K.P.; Frederick, B.D. Partitioning of Physiological Noise Signals in the Brain with Concurrent Near-Infrared Spectroscopy and fMRI. *Journal of Cerebral Blood Flow & Metabolism* **2011**, *31*, 2352, doi:10.1038/JCBFM.2011.100.
14. Afsar, F.S. Skin Care for Preterm and Term Neonates. *Clin Exp Dermatol* **2009**, *34*, 855–858, doi:10.1111/J.1365-2230.2009.03424.X.
15. Chiera, M.; Cerritelli, F.; Casini, A.; Barsotti, N.; Boschiero, D.; Caviglioli, F.; Corti, C.G.; Manzotti, A. Heart Rate Variability in the Perinatal Period: A Critical and Conceptual Review. *Front Neurosci* **2020**, *14*, 999, doi:10.3389/FNINS.2020.561186/BIBTEX.
16. Kurjak, A.; Stanojevic, M.; Andonotopo, W.; Salihagic-Kadic, A.; Carrera, J.M.; Azumendi, G. Behavioral Pattern Continuity from Prenatal to Postnatal Life - A Study by Four-Dimensional (4D) Ultrasonography. *J Perinat Med* **2004**, *32*, 346–353, doi:10.1515/JPM.2004.065/MACHINEREADEABLECITATION/RIS.
17. Holper, L.; Seifritz, E.; Scholkmann, F. Short-Term Pulse Rate Variability Is Better Characterized by Functional near-Infrared Spectroscopy than by Photoplethysmography. *J Biomed Opt* **2016**, *21*, 091308, doi:10.1117/1.JBO.21.9.091308.
18. Hakimi, N.; Kamaledin Setarehdan, S. Stress Assessment by Means of Heart Rate Derived from Functional Near-Infrared Spectroscopy. <https://doi.org/10.1117/1.JBO.23.11.115001> **2018**, *23*, 115001, doi:10.1117/1.JBO.23.11.115001.
19. Trajkovic, I.; Scholkmann, F.; Wolf, M. Estimating and Validating the Interbeat Intervals of the Heart Using Near-Infrared Spectroscopy on the Human Forehead. <https://doi.org/10.1117/1.3606560> **2011**, *16*, 087002, doi:10.1117/1.3606560.
20. Hakimi, N.; Jodeiri, A.; Mirbagheri, M.; Setarehdan, S.K. Proposing a Convolutional Neural Network for Stress Assessment by Means of Derived Heart Rate from Functional near Infrared Spectroscopy. *Comput Biol Med* **2020**, *121*, 103810, doi:10.1016/J.COMPBIOMED.2020.103810.
21. Scholkmann, F.; Boss, J.; Wolf, M. An Efficient Algorithm for Automatic Peak Detection in Noisy Periodic and Quasi-Periodic Signals. *Algorithms 2012, Vol. 5, Pages 588-603* **2012**, *5*, 588–603, doi:10.3390/A5040588.
22. Perdue, K.L.; Westerlund, A.; McCormick, S.A.; Nelson, C.A.; Iii, C.A.N. Extraction of Heart Rate from Functional Near-Infrared Spectroscopy in Infants. <https://doi.org/10.1117/1.JBO.19.6.067010> **2014**, *19*, 067010, doi:10.1117/1.JBO.19.6.067010.
23. Perdue, K.L.; Edwards, L.A.; Tager-Flusberg, H.; Nelson, C.A. Differing Developmental Trajectories in Heart Rate Responses to Speech Stimuli in Infants at High and Low Risk for Autism Spectrum Disorder. *J Autism Dev Disord* **2017**, *47*, 2434–2442, doi:10.1007/S10803-017-3167-4/FIGURES/4.

24. Tarar, A.A.; Mohammad, U.; Srivastava, S.K. Wearable Skin Sensors and Their Challenges: A Review of Transdermal, Optical, and Mechanical Sensors. *Biosensors* **2020**, *Vol. 10*, Page 56 **2020**, *10*, 56, doi:10.3390/BIOS10060056.
25. Quaresima, V.; Bisconti, S.; Ferrari, M. A Brief Review on the Use of Functional Near-Infrared Spectroscopy (fNIRS) for Language Imaging Studies in Human Newborns and Adults. *Brain Lang* **2012**, *121*, 79–89, doi:10.1016/J.BANDL.2011.03.009.
26. Izzetoglu, M.; Pourrezaei, K.; Du, J.; Shewokis, P.A. Evaluation of Cerebral Tissue Oximeters Using Multilayered Dynamic Head Models. *IEEE Trans Instrum Meas* **2021**, *70*, doi:10.1109/TIM.2021.3066469.
27. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Comput Sci Eng* **2007**, *9*, 90–95, doi:10.1109/MCSE.2007.55.
28. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **2020** *17:3* **2020**, *17*, 261–272, doi:10.1038/s41592-019-0686-2.
29. Delpy, D.T.; Cope, M.; Van Der Zee, P.; Arridge, S.; Wray, S.; Wyatt, J. Estimation of Optical Pathlength through Tissue from Direct Time of Flight Measurement. *Phys Med Biol* **1988**, *33*, 1433, doi:10.1088/0031-9155/33/12/008.
30. Sappia, M.S.; N.; Hakimi, N.; Colier, W.N.J.M.; Horschig, J.M. Signal Quality Index: An Algorithm for Quantitative Assessment of Functional near Infrared Spectroscopy Signal Quality. *Biomedical Optics Express*, *Vol. 11*, Issue 11, pp. 6732-6754 **2020**, *11*, 6732–6754, doi:10.1364/BOE.409317.
31. Altman, D.G.; Bland, J.M. Measurement in Medicine: The Analysis of Method Comparison Studies. *Journal of the Royal Statistical Society: Series D (The Statistician)* **1983**, *32*, 307–317, doi:10.2307/2987937.
32. Schäfer, A.; Vagedes, J. How Accurate Is Pulse Rate Variability as an Estimate of Heart Rate Variability?: A Review on Studies Comparing Photoplethysmographic Technology with an Electrocardiogram. *Int J Cardiol* **2013**, *166*, 15–29, doi:10.1016/J.IJCARD.2012.03.119.
33. Peng, R.C.; Zhou, X.L.; Lin, W.H.; Zhang, Y.T. Extraction of Heart Rate Variability from Smartphone Photoplethysmograms. *Comput Math Methods Med* **2015**, *2015*, doi:10.1155/2015/516826.
34. Dantas, E.M.; Kemp, A.H.; Andreão, R.V.; da Silva, V.J.D.; Brunoni, A.R.; Hoshi, R.A.; Bensenor, I.M.; Lotufo, P.A.; Ribeiro, A.L.P.; Mill, J.G. Reference Values for Short-Term Resting-State Heart Rate Variability in Healthy Adults: Results from the Brazilian Longitudinal Study of Adult Health-ELSA-Brasil Study. *Psychophysiology* **2018**, *55*, doi:10.1111/PSYP.13052.
35. De Souza Filho, L.F.M.; De Oliveira, J.C.M.; Ribeiro, M.K.A.; Moura, M.C.; Fernandes, N.D.; De Sousa, R.D.; Pedrino, G.R.; Rebelo, A.C.S. Evaluation of the Autonomic Nervous System by Analysis of Heart Rate Variability in the Preterm Infants. *BMC Cardiovasc Disord* **2019**, *19*, 1–6, doi:10.1186/S12872-019-1166-4/FIGURES/1.

36. Latremouille, S.; Lam, J.; Shalish, W.; Sant'Anna, G. Neonatal Heart Rate Variability: A Contemporary Scoping Review of Analysis Methods and Clinical Applications. *BMJ Open* **2021**, *11*, e055209, doi:10.1136/BMJOPEN-2021-055209.
37. Ronnqvist, L.; Hopkins, B. Head Position Preference in the Human Newborn: A New Look. *Child Dev* **1998**, *69*, 13, doi:10.2307/1132066.
38. Coryell, J.F.; Michel, G.F. How Supine Postural Preferences of Infants Can Contribute toward the Development of Handedness. *Infant Behav Dev* **1978**, *1*, 245–257, doi:10.1016/S0163-6383(78)80036-8.
39. Turkewitz, G.; Gordon, E.W.; Birch, H.G. Head Turning in the Human Neonate: Spontaneous Patterns. <http://dx.doi.org/10.1080/00221325.1965.10532772> **2012**, *107*, 143–158, doi:10.1080/00221325.1965.10532772.
40. Mendelson, Y.; Pujary, C. Measurement Site and Photodetector Size Considerations in Optimizing Power Consumption of a Wearable Reflectance Pulse Oximeter. *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings* **2003**, *4*, 3016–3019, doi:10.1109/IEMBS.2003.1280775.
41. Tamura, T.; Maeda, Y.; Sekine, M.; Yoshida, M. Wearable Photoplethysmographic Sensors—Past and Present. *Electronics* **2014**, Vol. 3, Pages 282-302 **2014**, *3*, 282–302, doi:10.3390/ELECTRONICS3020282.
42. Kooi, E.M.W.; Richter, A.E. Cerebral Autoregulation in Sick Infants: Current Insights. *Clin Perinatol* **2020**, *47*, 449–467, doi:10.1016/J.CLP.2020.05.003.
43. Mitra, S.; Czosnyka, M.; Smielewski, P.; O'Reilly, H.; Brady, K.; Austin, T. Heart Rate Passivity of Cerebral Tissue Oxygenation Is Associated with Predictors of Poor Outcome in Preterm Infants. *Acta Paediatr* **2014**, *103*, e374–e382, doi:10.1111/APA.12696.
44. Sullivan, B.A.; Grice, S.M.; Lake, D.E.; Moorman, J.R.; Fairchild, K.D. Infection and Other Clinical Correlates of Abnormal Heart Rate Characteristics in Preterm Infants. *J Pediatr* **2014**, *164*, 775–780, doi:10.1016/J.JPEDI.2013.11.038.
45. Cardoso, S.; Silva, M.J.; Guimarães, H. Autonomic Nervous System in Newborns: A Review Based on Heart Rate Variability. *Child's Nervous System* **2017**, *33*, 1053–1063, doi:10.1007/S00381-017-3436-8/TABLES/5.
46. Aliefendioğlu, D.; Doğru, T.; Albayrak, M.; Misirlioğlu, E.D.; Şanlı, C. Heart Rate Variability in Neonates with Hypoxic Ischemic Encephalopathy. *Indian J Pediatr* **2012**, *79*, 1468–1472, doi:10.1007/S12098-012-0703-2.
47. Bersani, I.; Piersigilli, F.; Gazzolo, D.; Campi, F.; Savarese, I.; Dotta, A.; Tamborrino, P.P.; Auriti, C.; Di Mambro, C. Heart Rate Variability as Possible Marker of Brain Damage in Neonates with Hypoxic Ischemic Encephalopathy: A Systematic Review. *Eur J Pediatr* **2021**, *180*, 1335–1345, doi:10.1007/S00431-020-03882-3/FIGURES/1.
48. Anton, O.; Fernandez, R.; Rendon-Morales, E.; Aviles-Espinosa, R.; Jordan, H.; Rabe, H. Heart Rate Monitoring in Newborn Babies: A Systematic Review. *Neonatology* **2019**, *116*, 199–210, doi:10.1159/000499675.

49. Shellhaas, R.A.; Burns, J.W.; Hassan, F.; Carlson, M.D.; Barks, J.D.E.; Chervin, R.D. Neonatal Sleep–Wake Analyses Predict 18-Month Neurodevelopmental Outcomes. *Sleep* **2017**, *40*, doi:10.1093/SLEEP/ZSX144.
50. Dawson, J.A.; Morley, C.J. Monitoring Oxygen Saturation and Heart Rate in the Early Neonatal Period. *Semin Fetal Neonatal Med* **2010**, *15*, 203–207, doi:10.1016/J.SINY.2010.03.004.
51. Green, M.S.; Sehgal, S.; Tariq, R. Near-Infrared Spectroscopy: The New Must Have Tool in the Intensive Care Unit? *Semin Cardiothorac Vasc Anesth* **2016**, *20*, 213–224, doi:10.1177/1089253216644346/ASSET/IMAGES/LARGE/10.1177\_1089253216644346-FIG1.JPEG.
52. Chen, J.; Ke, L.; Du, Q.; Zheng, Y.; Liu, Y. Cerebral Blood Flow Autoregulation Measurement via Bioimpedance Technology. *IEEE Trans Instrum Meas* **2022**, *71*, doi:10.1109/TIM.2022.3169544.
53. Chen, X.; Cheng, J.; Song, R.; Liu, Y.; Ward, R.; Wang, Z.J. Video-Based Heart Rate Measurement: Recent Advances and Future Prospects. *IEEE Trans Instrum Meas* **2019**, *68*, 3600–3615, doi:10.1109/TIM.2018.2879706.
54. Wu, Q.; Xu, G.; Wei, F.; Kuang, J.; Zhang, X.; Chen, L.; Zhang, S. Automatically Measure the Quality of Infants' Spontaneous Movement via Videos to Predict the Risk of Cerebral Palsy. *IEEE Trans Instrum Meas* **2021**, *70*, doi:10.1109/TIM.2021.3125980.



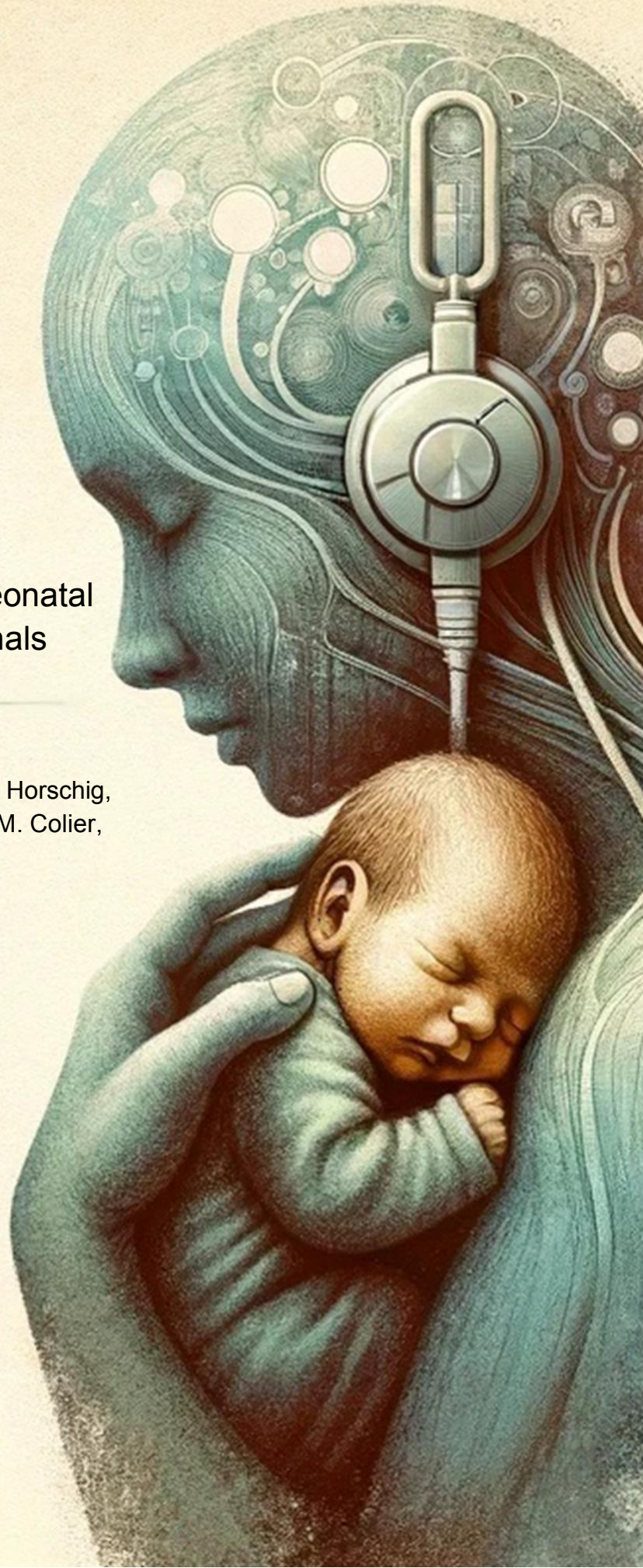


## CHAPTER 8

Respiratory rate extraction from neonatal  
near-infrared spectroscopy signals

Naser Hakimi, Mohammad Shahbakhti, Jörn M. Horschig,  
Thomas Alderliesten, Frank V. Bel, Willy N. J. M. Colier,  
Jeroen Dudink

doi: <https://doi.org/10.3390/s23094487>



## Abstract

Near-infrared spectroscopy (NIRS) relative concentration signals contain ‘noise’ from physiological processes such as respiration and heart rate. Simultaneous assessment of NIRS and respiratory rate (RR) using a single sensor would facilitate a perfectly time-synced assessment of (cerebral) physiology. Our aim was to extract respiratory rate from cerebral NIRS intensity signals in neonates admitted to a neonatal intensive care unit (NICU). A novel algorithm, NRR (NIRS RR), is developed for extracting RR from NIRS signals recorded from critically ill neonates. In total, 19 measurements were recorded from ten neonates admitted to the NICU with a gestational age and birth weight of  $38 \pm 5$  weeks and  $3092 \pm 990$  g, respectively. We synchronously recorded NIRS and reference RR signals sampled at 100 Hz and 0.5 Hz, respectively. The performance of the NRR algorithm is assessed in terms of the agreement and linear correlation between the reference and extracted RRs, and it is compared statistically with that of two existing methods. The NRR algorithm showed a mean error of 1.1 breaths per minute (BPM), a root mean square error of 3.8 BPM, and Bland–Altman limits of agreement of 6.7 BPM averaged over all measurements. In addition, a linear correlation of 84.5% ( $p < 0.01$ ) was achieved between the reference and extracted RRs. The statistical analyses confirmed the significant ( $p < 0.05$ ) outperformance of the NRR algorithm with respect to the existing methods. We showed the possibility of extracting RR from neonatal NIRS in an intensive care environment, which showed high correspondence with the reference RR recorded. Adding the NRR algorithm to a NIRS system provides the opportunity to record synchronously different physiological sources of information about cerebral perfusion and respiration by a single monitoring system. This allows for a concurrent integrated analysis of the impact of breathing (including apnea) on cerebral hemodynamics.



# 1. Introduction

Monitoring changes in the respiratory rate (RR) of term and preterm infants admitted to the neonatal intensive care unit (NICU) plays a vital role in the timely detection of abnormal respiratory events such as tachypnea and apnea [1–5]. According to the literature, almost half of the neonates in the NICU were admitted due to respiratory morbidity [6]. Respiratory events and associated desaturation events affect cerebral physiology and contain relevant information about cerebral hemodynamics [7–10]. Hence, continuous monitoring of RR changes in hospitalized neonates and the impact on cerebral hemodynamics is of importance in clinical practice [11–13].

For monitoring cerebral hemodynamics, near-infrared spectroscopy (NIRS) can be employed, which is a non-invasive optical neuroimaging modality for measuring oxygenated (O<sub>2</sub>Hb) and deoxygenated (HHb) hemoglobin concentrations associated with neural activities in the cerebral cortex [14]. NIRS is becoming part of routine clinical practice in an increasing number of NICUs worldwide [15,16]. Aside from cerebral tissue oxygen saturation measured with NIRS, the raw NIRS signals (i.e., O<sub>2</sub>Hb and HHb) are contaminated by physiological noise arising from, e.g., heart rate, blood pressure, Mayer waves, and respiration [17]. While—in most studies—either the raw NIRS signals are not used or the physiological noise is filtered out, some studies have investigated the possibility of extracting extra physiological parameters [18–21].

Respiration is a major source of physiological noise in NIRS that is vividly observable in the spectra of both O<sub>2</sub>Hb and HHb signals [22]. The emergence of respiration in NIRS measurements is due to (i) the alternation of blood flow within the whole body during the course of inspiration and expiration and (ii) the effect of respiratory fluctuations on cerebral blood volume and flow [23,24]. Some studies have proposed algorithms for separating respiratory components from NIRS measurements, e.g., in [24] by proposing a band-pass filtering (BPF) method. The possibility of RR extraction from NIRS measurements has only been investigated in our previous study [21]. The presented baseline wandering (BW) method, however, was only validated on adult subjects in resting-state measurements which were free from motion artifacts. The presented method was not ideally designed for clinical purposes, i.e., to be implemented in NICU. There are several challenges that exist in data acquisition from hospitalized neonates such as low data quality and patient movements.

Extracting RR from neonatal cerebral NIRS, on the one side, facilitates the concurrent and integrated analysis of breathing and cerebral oxygenation in the NICU. For instance, it could provide a perfectly time-synced analysis of the effect of respiration, including apnea and desaturation events, on cerebral hemodynamics. Deriving RR from neonatal NIRS, on the other side, could potentially reduce the need for adhesive electrodes in the future, which could reduce discomfort, stress, and epidermal stripping, and would promote parent–neonate interaction [25–30]. Therefore, the aim of the current study is to develop a novel algorithm, NRR (NIRS RR), for deriving RR from neonatal NIRS measurements recorded in NICUs. With respect to the reference RR recorded with a clinical patient monitor system, we assess the performance of the NRR algorithm in terms of agreement and linear correlation between the reference and extracted RRs. We also compare the NRR algorithm's performance against the performance of two state-of-the-art methods, i.e., BPF and BW.

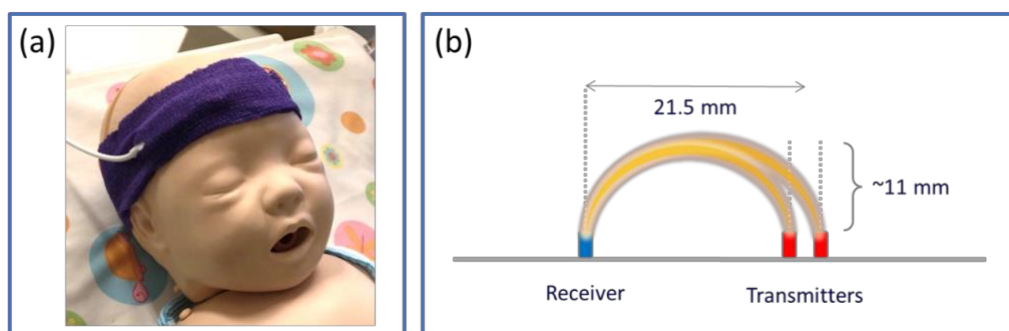
## 2. Materials and Methods

### 2.1. Participants

This study was approved by the local ethics committee of the University Medical Center Utrecht (21-098/C). We included ten newborn infants (three females) with gestational age  $38 \pm 5$  weeks and birth weight  $3092 \pm 990$  g that were admitted to the NICU of the Wilhelmina Children's Hospital, Utrecht, the Netherlands. Among them, three babies were preterm infants (GA < 37 weeks) and one baby was an extremely preterm infant (GA < 28 weeks). They had an indication for clinical NIRS monitoring which was determined by the caring physician or local neuromonitoring protocols. Parents gave informed consent to participate in the study before data acquisition. A total of 19 measurements were recorded from the included neonates without interrupting the clinical routines. After assessing the signal quality as explained in Section 2.3.1, we excluded three measurements due to very low signal quality, leaving 16 measurements from nine neonates (gestational age =  $39 \pm 3$  weeks; weight =  $3368 \pm 473$  g; three females) with a total length of approximately 20.5 h.

### 2.2. Data Acquisition

NIRS signals were recorded at 100 Hz with a cerebral oximetry system (TOM, Artinis Medical Systems B.V., Elst, the Netherlands). The infant-neonatal sensor of the system consists of two transmitters with nominal wavelengths of 760 and 850 nm and a receiver with a 21.5 mm distance from the transmitters, providing an approximate penetration depth of 11 mm, i.e., half the transmitter–receiver distance [31]. The sensor was placed on the neonate's forehead under a clinical elastic bandage. Figure 1 illustrates the infant-neonatal sensor placed on a baby manikin's forehead covered by the bandage (Figure 1(a)) with a schematic of the layout of the transmitters and the receiver, which provides two NIRS channels. The reference RR was recorded with a patient monitor system, Philips IntelliVue MP70 (Philips Medical Systems, Best, the Netherlands), sampled at 0.5 Hz.

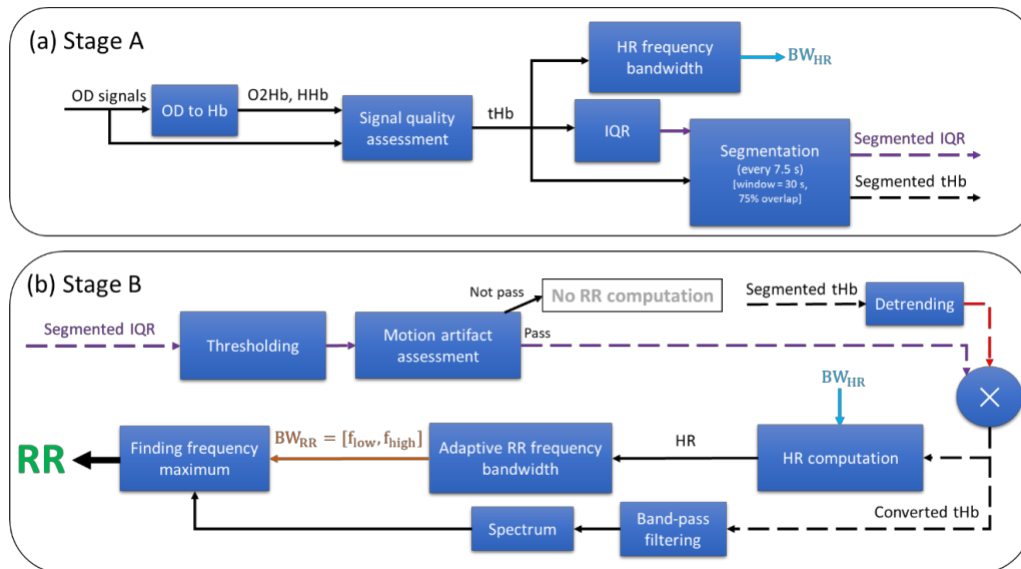


**Figure 1:** (a) A photograph of the infant-neonatal TOM sensor (Artinis Medical Systems B.V., the Netherlands) placed on a baby manikin's forehead, covered with a clinical self-adhesive elastic bandage. (b) Schematic of the configuration of the two transmitters and one receiver embedded in the TOM sensor, illustrated in red and blue, respectively.

### 2.3. Respiratory Rate Extraction Algorithm

The block diagram of the proposed RR extraction algorithm, NRR (NIRS RR), is illustrated in Figure 2. The algorithm comprises two main stages: Stage A, illustrated in Figure 2(a); and Stage B, in Figure 2(b). Stage A of the NRR algorithm consists of four steps: Preprocessing,

HR frequency bandwidth, interquartile range (IQR), and segmentation. These steps are executed once at the beginning of the complete NIRS measurement. Stage B of the NRR algorithm consists of three steps: Motion artifact assessment, HR computation, and RR computation. These are implemented on each 30-second segment of the total hemoglobin concentration (tHb) and IQR signals, computing RR per segment. As respiration influences both O2Hb and HHb signals [22], we have employed the tHb, i.e., the sum of the O2Hb and HHb signals, in order to have a better representation of RR information. We implemented the analyses in this study in Python using *numpy*, *scipy*, and *matplotlib* modules [32,33].



**Figure 2:** The block diagram of the proposed NRR (NIRS RR) algorithm, comprising Stage A and Stage B. (a) Stage A consists of four steps: preprocessing (including OD to Hb and signal quality assessment), HR frequency bandwidth, interquartile range (IQR), and segmentation. (b) Stage B consists of three steps: motion artifact assessment (including thresholding and motion artifact assessment), HR computation (including detrending, multiplication operator, and HR computation), and RR computation (including band-pass filtering, spectrum, adaptive RR frequency bandwidth, and finding frequency maximum).

### 2.3.1. Preprocessing

Using the modified Beer–Lambert law [34], the optical densities (ODs) of the two NIRS channels are first converted to concentration changes in oxygenated hemoglobin (O2Hb) and deoxygenated hemoglobin (HHb) (OD to Hb in Figure 2(a)). Then, the signal quality of the NIRS channels is assessed using the Signal Quality Index (SQI) algorithm [35] (signal quality assessment in Figure 2(a)). The SQI signal is computed in sliding windows of 10 s, overlapping by 50%. Next, the average of the SQI signal across time is computed for each channel, and the channel with the greater average SQI is selected. The whole measurement is excluded in the subsequent analyses if more than 75% of the samples of the SQI of the selected channel are less than an SQI level of 2, which represents low signal quality. Otherwise, the algorithm continues to the subsequent steps.

### 2.3.2. HR Frequency Bandwidth

First, the tHb signal is low-pass filtered by applying a moving average filter of 1 s and then subtracting it from the original tHb signal. Second, the spectrum of the filtered tHb signal is computed using the multitaper power spectral density estimation [36] based on Slepian sequences [37]. In this study, we used the Python method *scipy.signal.windows.dpss* in order to compute the Slepian sequences by setting the parameters of standardized half bandwidth

and the number of windows at 2.5 and 5, respectively. Next, a predefined HR frequency bandwidth between 1.25 Hz and 3.5 Hz is defined according to the HR range of the included patients ( $75 < HR < 210$  beats per minute). This bandwidth is computed once per measurement according to the dominant frequency of the spectrum as follows: the spectrum within the predefined bandwidth is sorted in descending order; the average frequency of the first fifty components (~50% of all components) is then computed; finally, the HR frequency bandwidth ( $BW_{HR}$ ) is set as the average frequency minus 0.5 Hz and plus 0.5 Hz.

### 2.3.3. Interquartile Range (IQR)

In sliding windows of 1 s, we compute the interquartile range (IQR)—the difference between the third and first quartiles—of the selected tHb signal. Afterwards, the computed IQR values are interpolated to 100 Hz using cubic interpolation. Subsequently, the IQR signal is normalized by dividing it by the median of the tHb signal.

### 2.3.4. Segmentation

The selected tHb and normalized IQR signals are segmented by sliding windows of 30 s overlapping by 75% (i.e., shifting 7.5 s). Examples of the segmented tHb and IQR are shown in Figure 3(a) and Figure 3(e), respectively. The subsequent steps are implemented on each segmented 30-second tHb and IQR signal, computing RR in each window (i.e., every 7.5 s).

### 2.3.5. Motion Artifact Assessment

The segmented IQR is thresholded with a threshold of 1%. The thresholded IQR is logically one if the IQR is below the threshold, meaning presumably free from motion artifacts. Figure 3(f) shows an example of the thresholded IQR signal which has been obtained from the segmented IQR illustrated in Figure 3(e). Next, the percentage of samples which are one in the thresholded IQR signal is computed. The segment will be dismissed for RR computation if the percentage is less than 50%. Otherwise, at least 50% of the segment is presumably free from motion artifacts and the algorithm passes to the next steps to compute RR.

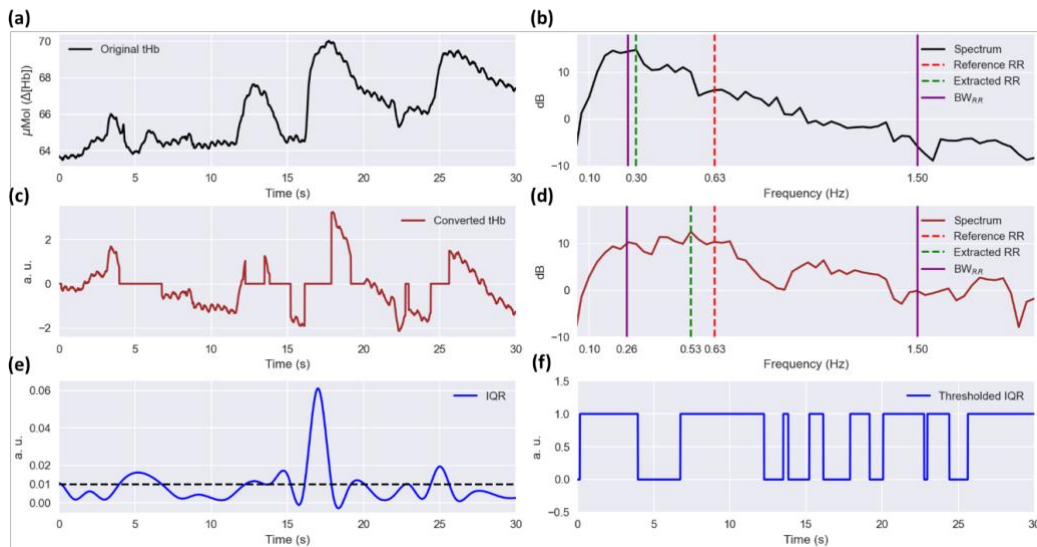
### 2.3.6. HR Computation

The segmented tHb signal is detrended by calculating a least-squares fit of a straight line to the data and then subtracting it (detrending in Figure 2(b)). Next, the detrended tHb signal is multiplied by the thresholded IQR (multiplication operator in Figure 2(b)). Figure 3(c) shows an example of this converted tHb signal, which corresponds to the tHb signal shown in Figure 3(a). Then, the spectrum of the converted tHb is computed by using the multitaper power spectral density estimation as explained in Section 2.3.2. Finally, the HR in the considered window is computed by finding the dominant frequency within the  $BW_{HR}$ . The HR is further used in the next step to limit the allowed RR range.

### 2.3.7. RR Computation

First, the converted tHb is filtered by using an FIR (finite impulse response) zero-phase band-pass filter implemented with a Kaiser window (band-pass filtering in Figure 2(b)). The lower and higher cutoff frequencies of the filter are set at 0.1 times the computed HR and 2 Hz, respectively. The higher cutoff frequency was chosen with a 0.5 Hz gap from the maximum reference RR frequency of the patients included (~1.5 Hz). The gap was considered in order to prevent filtering the RR components within the transition bandwidth of the filter. Second, the spectrum of the filtered signal is computed as explained in Sections 2.3.2 and 2.3.6. Afterward,

an adaptive RR frequency bandwidth ( $BW_{RR}$ ) is defined as 0.15 and 0.85 times the computed HR (i.e.,  $BW_{RR} = [0.15 * HR, 0.85 * HR]$ ). Finally, the RR is computed by finding the dominant frequency in the filtered signal within the  $BW_{RR}$ . Figure 3(d) illustrates the spectrum of the filtered converted tHb signal corresponding to the converted tHb shown in Figure 3(c). The dashed red and green lines and the solid magenta lines depict the reference and extracted RRs and the  $BW_{RR}$ , respectively. To illustrate the impact of the multiplication by thresholded IQR on the RR computation, Figure 3(b) shows the spectrum computed from the original tHb signal displayed in Figure 3(a), i.e., without any motion artifact assessment and multiplication by the thresholded IQR. It is observed that there is a high error of approximately 0.33 Hz (i.e., 20 breaths per minute, BPM) when the original tHb signal is used, while the error is 0.1 Hz (i.e., 6 BPM) when using the converted signal.



**Figure 3:** (a) A 30-second segment of the tHb signal recorded from one of the subjects (Measurement 15). (b) The spectrum computed for the original tHb signal, i.e., without motion artifact assessment and signal conversion. The dashed red and green lines and the solid magenta line depict the reference RR, extracted RR, and adaptive RR frequency bandwidth ( $BW_{RR}$ ), respectively. (c) The converted tHb signal after motion artifact assessment and multiplication by the thresholded IQR signal. (d) The spectrum of the converted tHb signal. (e) The IQR signal computed based on sliding windows of 1 s, interpolated to 100 Hz, and normalized by the median of the tHb signal. The dashed black line depicts the IQR threshold considered for motion artifact assessment (i.e., 1%). (f) The thresholded IQR signal. It is zero (i.e., contaminated with motion artifacts) if the IQR is above the defined threshold; otherwise, it is one.

## 2.4. Algorithm Performance Assessment

We employed three quantitative measures for quantifying the agreement between the reference and extracted RRs, including the mean of error (ME) or bias, root mean square error (RMSE), and Bland–Altman limits of agreement (LoA), which is defined as 1.96 times the standard deviation of the error [38]. In addition, as a quantitative measure to determine the linear association between the reference and extracted RRs, we computed Pearson’s correlation coefficient. We assessed the significance of the correlation by performing a Student’s t-test with  $\alpha = 1\%$ . Furthermore, we calculated the percentage of the segments wherein RR was computed in each measurement. The mentioned quantitative measures were computed for each measurement separately, and the average and standard deviation for each measure over all measurements were also calculated. Since the reference RR was recorded with 0.5 Hz, each segment of 30 s contains a total of 15 reference RR samples. To have an



equal number of samples between the reference and extracted RRs, we have taken the average of the reference RR samples in each 30 s segment.

We compared the performance of the NRR algorithm with two existing methods proposed for extracting RR from NIRS in adults, i.e., band-pass filtering (BPF) [24] and baseline wander (BW) [21]. In these two methods, we set the window length at 30 s, as used in the NRR algorithm. In addition, the cutoff frequencies of the band-pass filter used in the BPF method were set to [0.15, 2] Hz according to the range of the reference RR in this study. To quantitatively compare the performance of the three methods, we computed the average and standard deviation of the mentioned quantitative measures, determining the agreement and linear association between the reference and extracted RRs, over all measurements. As another comparison measure, we defined a boundary wherein the absolute error in RR computation is less than 30% of the mean of the pairwise reference and extracted RRs. We refer to this boundary as the 30% boundary. Furthermore, we performed a Student's t-test with  $\alpha = 5\%$  to determine whether there is a significant difference in performance between the methods.

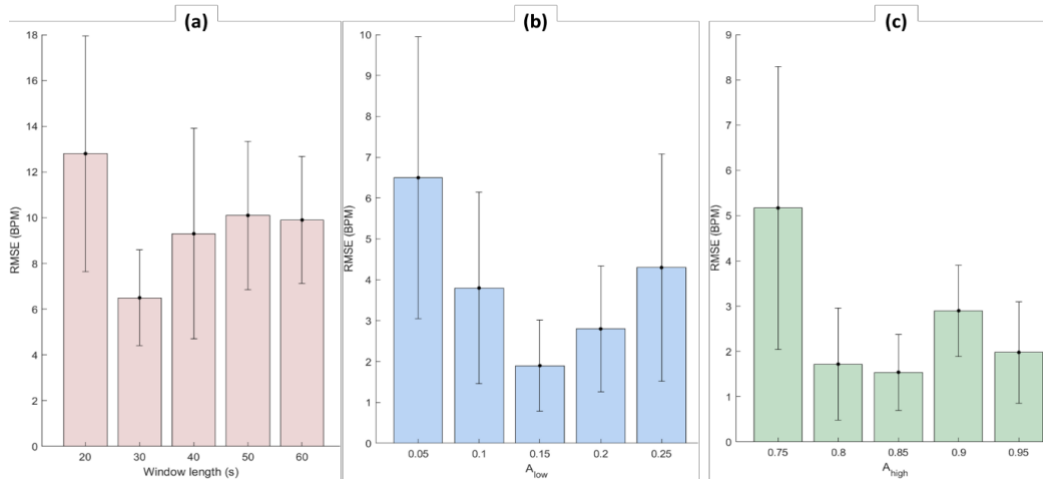
## 2.5. Optimization of the Parameters

We conducted a sensitivity analysis of the main parameters used in the NRR algorithm, i.e., the window length and the two constants 0.15 ( $A_{low}$ ) and 0.85 ( $A_{high}$ ) used in defining the adaptive RR frequency bandwidth ( $BW_{RR} = [A_{low} * HR, A_{high} * HR]$ ). Among the 19 recorded measurements, 3 measurements were excluded due to having very low signal quality which was assessed as explained in Section 2.3.1. A total of 20% (i.e., a total of 4.1 h) of the included 16 measurements were randomly selected to optimize the parameters of the NRR algorithm. We computed the RMSE as a measure of error and performed Student's t-test with  $\alpha = 5\%$  in order to find the statistically optimal parameters.

# 3. Results

## 3.1. Optimization of the Proposed Algorithm's Parameters

To find a suitable window length, firstly, constants  $A_{low}$  and  $A_{high}$  were set at 0.05 and 0.95, respectively. Then, different window lengths from 20 to 60 s with steps of 10 s were used. As displayed in Figure 4(a), the significantly ( $p < 0.05$ ) lowest RMSE between the reference and estimated RRs was obtained with a window length of 30 s. Different values of  $A_{low}$  ranging from 0.05 to 0.2 with steps of 0.05 were investigated to adjust this constant, given the fact that the window length and the  $A_{high}$  were set at the optimal window length obtained (30 s) and 0.95, respectively. As shown in Figure 4(b), 0.15 and 0.2 are the optimal values for the constant  $A_{low}$  which provide significantly lower RMSEs than the other values. However, we selected the lower one (0.15) which makes the bandwidth larger. In the last step, by setting the window length and the constant  $A_{low}$  at the optimal ones computed, we varied  $A_{high}$  from 0.75 to 0.95 with steps of 0.05 to find the optimal value providing the lowest RMSE. As illustrated in Figure 4(c), we found 0.8 and 0.85 as the optimal values for  $A_{high}$  providing significantly ( $p < 0.05$ ) lower RMSEs than the other values. However, similarly to the previous step, we selected the higher one (0.85) to set a larger bandwidth for RR.



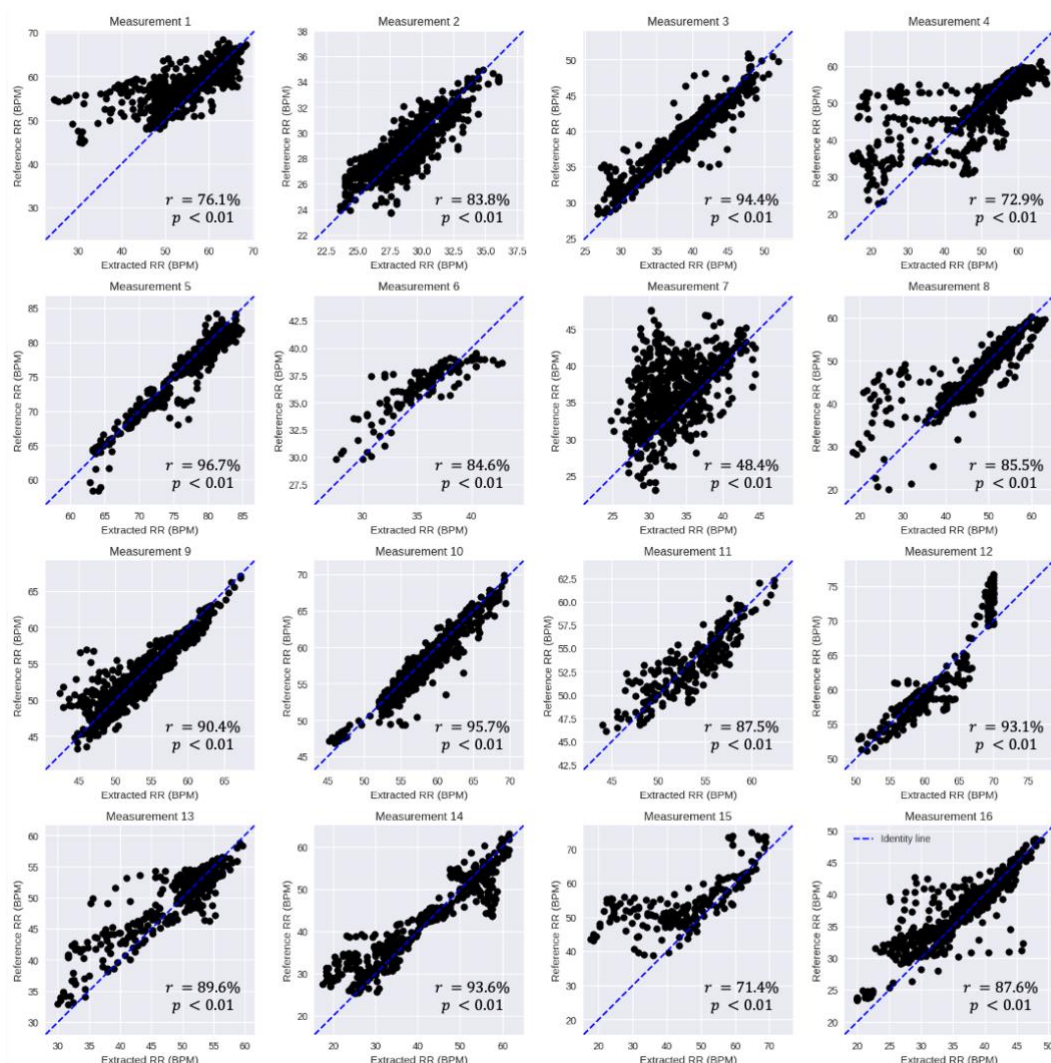
**Figure 4:** Sensitivity analysis of the main parameters of the NRR algorithm, i.e., window length and constants (i.e.,  $A_{low}$ ,  $A_{high}$ ) regulating the lower and higher sides of the adaptive RR frequency bandwidth ( $BW_{RR} = [A_{low} * HR, A_{high} * HR]$ ). The bar chart of RMSE computed between the reference and extracted RRs with respect to (a) the window length, (b) the constant  $A_{low}$ , and (c) the constant  $A_{high}$  selected, averaged over all measurements. The error bars and bar heights depict the standard deviation and average of RMSE, respectively.

## 3.2. Results of the Proposed Algorithm

After optimizing the parameters of the algorithm, the remaining 80% (i.e., not included in the sensitivity analysis, see Section 3.1) of the 16 included measurements, a total of 16.4 h or a total of 7872 30-s signal segments, were used for the validation of the proposed algorithm. Figure 5 illustrates the scatter plots of the measurements comparing the reference and extracted RRs. Looking at the reference RR, we observe that the range of reference RR differs between the measurements. As an example, the reference RR in Measurement 2 ranges approximately from 24 to 36 BPM while in Measurement 5 it ranges from approximately 60 to 85 BPM. In addition, it is observed that the patients in some measurements such as Measurements 8, 14, and 15 experienced more variation ( $\sim 40$  BPM) in RR than others. Looking at the scatterplots, we observe that there is a significant linear association ( $p < 0.01$ ) between the reference RR and the extracted RR by the NRR algorithm in all measurements.

Table 1 summarizes the quantitative measures computed in this study to determine the agreement and linear association between the reference and extracted RRs. The ME or bias was lower than approximately 1 BPM ( $\sim 2\%$  of the average reference RR) in all measurements, except for Measurements 1 (2.4 BPM), 7 (2.7 BPM), and 15, which had the highest bias (8.1 BPM). The average and standard deviation of bias calculated on all measurements are  $1.1$  and  $2.1$  BPM (i.e.,  $1.1 \pm 2.1$ ), respectively. The RMSE is lower than 4 BPM in all measurements (exceptions: 1, 4, 7, 15), with the highest one computed for Measurement 15 (12.2 BPM) and average and standard deviation of  $3.8 \pm 3.0$  BPM. The LoA has an average and standard deviation of  $6.7 \pm 4.7$  over all measurements. The highest LoA was obtained in Measurement 15 (18.1 BPM) which had also the highest bias and RMSE. However, looking at Pearson's correlation, we observe the lowest correlation obtained in Measurement 7 (48%) while it was 71% in Measurement 15. The greatest correlation was obtained in Measurement 5 (96.7%) with an average and standard deviation of  $84.5 \pm 12.3$  computed on all measurements. Looking at the included segments' percentages, we observe that the RR has been extracted on average for about 94% of the signal segments in all measurements, i.e., leaving on average about 6% of the signal segments with no RR computed due to being highly contaminated by motion artifacts. The lowest percentage of included segments was obtained in Measurement

15 (69.6%). This shows that Measurement 15 was the one in this study with the highest level of contamination by motion artifacts.



**Figure 5:** Scatter plots showing the linear association between the reference RR and the extracted RR by the NRR algorithm per measurement. The x and y axes represent the RR in breaths per minute (BPM). Each black dot corresponds to the reference and extracted RRs of each 30-second signal segment. The blue dashed line depicts the identity line ( $y = x$ ). The correlation between the reference and extracted RRs is significant in each measurement ( $p < 0.01$ ).

**Table 1:** Quantitative measures for assessing the performance of the proposed NRR algorithm per measurement.

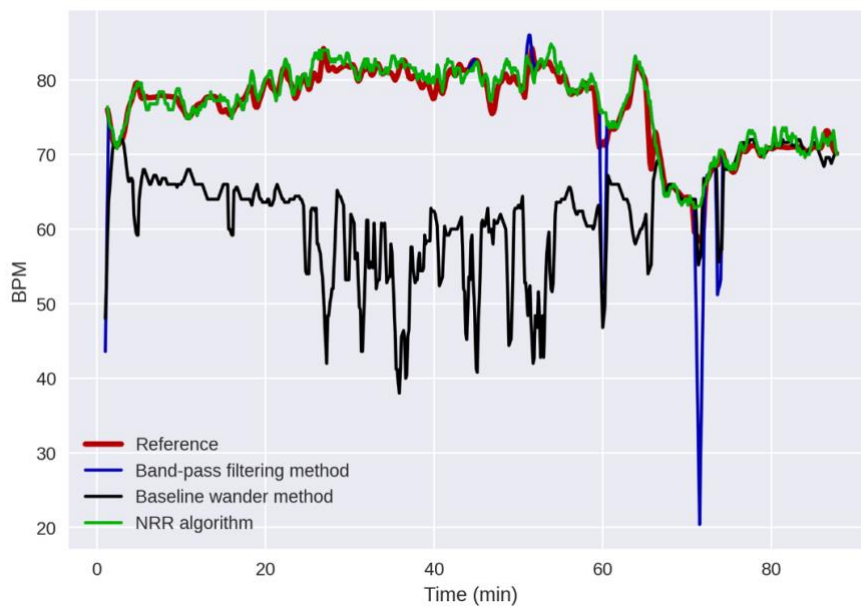
Measurement	ME <sup>1</sup> (BPM <sup>2</sup> )	RMSE <sup>3</sup> (BPM)	LoA <sup>4</sup> (BPM)	Pearson's r (%)	Included Segments (%)
1	2.4	5.9	10.6	76.1	92.1
2	0.2	1.3	2.5	83.8	99.8
3	0.3	1.7	3.2	94.4	99.9
4	0.4	8.1	15.9	72.9	79.2
5	-0.6	1.5	2.6	96.7	99.0
6	0.9	2.0	3.4	84.6	100
7	2.7	5.1	8.4	48.4	99.1
8	0.3	3.8	7.5	85.5	95.5
9	-0.1	2.0	4.0	90.4	96.2
10	-0.5	1.5	2.7	95.7	94.5
11	-0.1	1.8	3.6	87.5	98.8

12	0.0	2.5	4.9	93.1	87.0
13	1.1	3.5	6.5	89.6	96.0
14	0.8	4.4	8.6	93.6	96.5
15	8.1	12.2	18.1	71.4	69.6
16	1.0	2.9	5.7	87.6	97.4
Average	1.1	3.8	6.7	84.5	93.8
Std <sup>5</sup>	2.1	3.0	4.7	12.3	8.5

<sup>1</sup> Mean of error. <sup>2</sup> Breaths per minute. <sup>3</sup> Root mean square error. <sup>4</sup> Bland–Altman limits of agreement. <sup>5</sup> Standard deviation.

### 3.3. Comparison with the Existing Algorithms

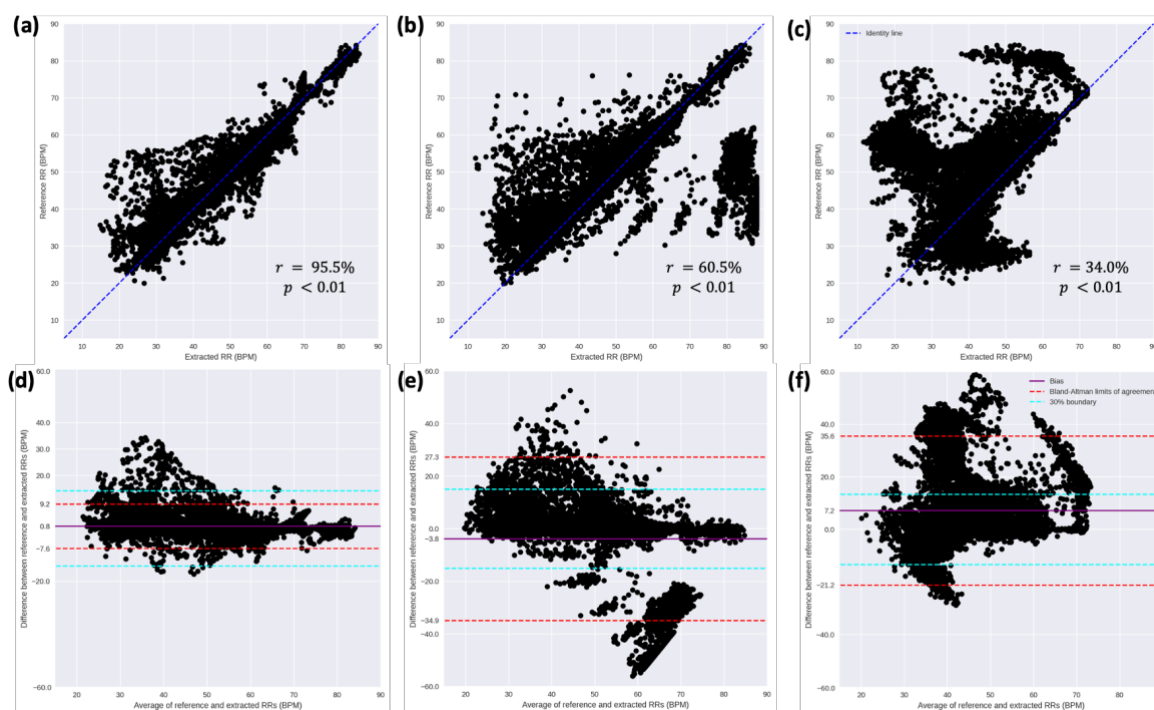
Figure 6 shows an example (Measurement 5) of the reference RR and the RR extracted by the proposed NRR algorithm compared with the ones obtained by the BPF and BW methods. It is observed that the BW method works poorly in this measurement in capturing the trend of the reference RR (except for the last ~20 min). The BPF method has provided a better RR extraction than the BW method, providing a Pearson's correlation of approximately 87%. However, it has some outliers that have errors of more than approximately 20 BPM. On the contrary, the RR extracted by the NRR algorithm has no outliers and follows the trend of the reference RR from the beginning of the measurement to the end, providing a correlation of approximately 97%.



**Figure 6:** An example of the reference RR signal (in red) and the extracted RR signals obtained by using the NRR algorithm (in green), band-pass filtering (BPF) method (in blue), and baseline wander (BW) method (in black). The x and y axes represent the time and RR in minutes and breaths per minute (BPM), respectively. In this measurement (Measurement 5), Pearson's correlation magnitude between the reference and extracted RRs is 97% ( $p < 0.01$ ) using the NRR algorithm, 87% ( $p < 0.01$ ) using the BPF method, and 56% ( $p < 0.01$ ) using the BW method, indicating the superior performance of the NRR algorithm than BPF and BW methods.

Figure 7 shows the scatterplot and the Bland–Altman plot for the three methods between the reference and extracted RRs when all the measurements are concatenated together. From the scatterplots, it is observed that the RR extracted by the NRR algorithm linearly follows the reference RR in different RR ranges with a Pearson's correlation coefficient of 95.5% (Figure 7(a)). This confirms the adaptability of the NRR algorithm to different ranges of RR. Conversely, the scatterplots obtained for the BPF (Figure 7(b)) and BW (Figure 7(c)) methods are more sparsely distributed around the identity line than the one obtained for the NRR algorithm. In the scatterplot of the BW method (Figure 7(c)), the majority of the points are

above the identity line, showing the bias of the method to result in a lower RR than the reference RR. Although a significant ( $p < 0.01$ ) correlation was obtained by the BPF and BW methods between the reference RR and the extracted RR, the correlation (60.5% and 34.0%, respectively) is substantially lower than the one obtained by the NRR algorithm. Looking at the Bland–Altman plots, we observe that the BPF and BW methods result in a greater bias in computing RR than the NRR algorithm. Looking at the Bland–Altman limits of agreement (LoA) and comparing it with the 30% boundary (explained in Section 2.4), we observe that the LoA of the NRR algorithm fell within the 30% boundary. This shows that 95% of the error between the reference and extracted RRs, represented by LoA, is lower than 30% of the average RRs. Looking at the Bland–Altman plots of the BPF and BW methods, we observe that the LoA of both methods fell outside the 30% boundary, indicating a higher error, i.e., a wider distribution of error, in RR computation by these methods than the NRR algorithm. The percentage of data points outside of the 30% boundary in the Bland–Altman plots of the NRR, BPF, and BW methods is 2.7%, 22.2%, and 27.4%, respectively. Looking at the bias in the plots, we observe that the BW method has a considerable bias of 7.2 BPM in computing RR, while it is  $-3.8$  and  $0.8$  BPM for the BPF and NRR methods, respectively.



**Figure 7:** Scatter plots between the reference RR and the extracted RRs using (a) the NRR algorithm, (b) the band-pass filtering (BPF) method, and (c) the baseline wander (BW) method when all measurements are pooled together. Each black dot corresponds to each 30-second signal segment. The Bland–Altman plots of the reference RR and the extracted RRs using (d) the NRR algorithm, (e) the BPF method, and (f) the BW method. The solid magenta line and the dashed red lines depict the bias and the Bland–Altman limits of agreement, respectively. The cyan lines represent the 30% boundary, wherein the error between the reference and extracted RRs is lower than 30% of the mean of the pairwise RRs.

Table 2 summarizes the quantitative measures determining the agreement and linear association between the reference and extracted RRs by each method, averaged over all measurements. The average of ME or bias computed over all measurements by the NRR (1.1 BPM) and BPF ( $-1.4$  BPM) methods is lower than the one in the BW method (8.4 BPM). However, the standard deviation of ME obtained in the NRR algorithm (2.1 BPM) is much lower than the ones obtained by the BPF (11.6 BPM) and BW (14.1 BPM) methods. The

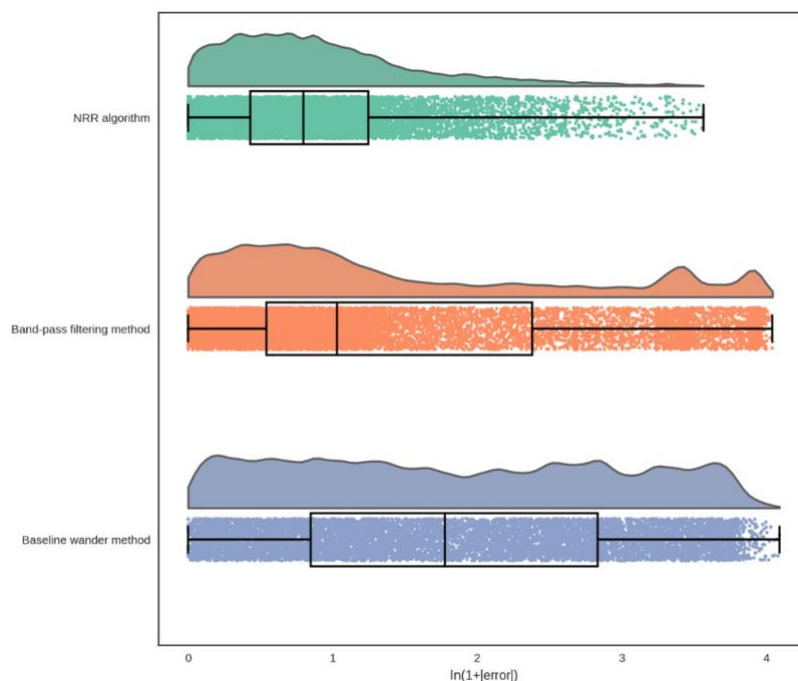
average and standard deviation of RMSE and LoA obtained in the NRR algorithm are  $3.8 \pm 3.0$  BPM and  $6.7 \pm 4.7$  BPM, respectively, which are substantially lower than the ones obtained in the BPF method (RMSE =  $10.0 \pm 10.7$  BPM; LoA =  $14.0 \pm 11.1$  BPM). Although the BW method had a higher average and standard deviation of RMSE ( $12.6 \pm 12.2$  BPM) compared with the BPF method, it has a lower average and standard deviation of LoA ( $11.4 \pm 5.8$  BPM), which are still higher than the ones obtained in the NRR algorithm. The average and standard deviation of Pearson's correlation obtained in the NRR algorithm ( $84.5 \pm 12.3\%$ ) are greater than the ones obtained by the BPF ( $70.8 \pm 15.0\%$ ) and BW ( $51.1 \pm 26.0\%$ ) methods. Figure 8 shows the raincloud plot of the absolute error between the reference and extracted RRs for each method. Here, the natural logarithm, i.e.,  $\ln()$ , of the absolute error plus one has been shown for the sake of having a proper visualization of the error. It is observed that the majority (75%, represented by the third quartile in the box plot) of the error in the NRR algorithm fell approximately between 0 and 1, in  $\ln(1 + \text{BPM})$ , whereas in the BPF and BW methods the error is more distributed and the majority fell up to approximately 2.5. Looking at the distribution plots, we observe the error was distributed mainly about zero in the NRR algorithm. Conversely, in the distribution plot for the BPF method, there are two additional peaks at about the greatest error, which is due to the outliers as shown in Figure 6. However, the distribution plot of error obtained for the BPF method is less sparse than the one obtained for the BW method, which is roughly uniformly distributed. The result of the statistical comparison between the performance of the three methods in RR extraction against each other showed all three methods performed significantly differently from each other ( $p < 0.05$ ). Therefore, the NRR algorithm performed significantly better than the BPF and BW methods in computing RR.

**Table 2:** Quantitative measures for assessing the performance of the proposed NRR algorithm and the two existing methods, i.e., band-pass filtering (BPF) and baseline wander (BW), averaged over all measurements.

Algorithm	ME <sup>1</sup> (BPM <sup>2</sup> )	RMSE <sup>3</sup> (BPM)	LoA <sup>4</sup> (BPM)	Pearson's r (%)
NRR	$1.1 \pm 2.1$	$3.8 \pm 3.0$	$6.7 \pm 4.7$	$84.5 \pm 12.3$
BPF	$-1.4 \pm 11.6$	$10.0 \pm 10.7$	$14.0 \pm 11.1$	$70.8 \pm 15.0$
BW	$8.4 \pm 14.1$	$12.6 \pm 12.2$	$11.4 \pm 5.8$	$51.1 \pm 26.0$

<sup>1</sup> Mean of error. <sup>2</sup> Breaths per minute. <sup>3</sup> Root mean square error. <sup>4</sup> Bland–Altman limits of agreement.





**Figure 8:** Raincloud plot of the absolute error between the reference RR and the extracted RR by using the NRR algorithm (first panel), the band-pass filtering (BPF) method (second panel), and the baseline wander (BW) method (third panel). The x axis represents the natural logarithm of the absolute error (in BPM) plus one. This was done to have a better visualization of the error distribution.

## 4. Discussion

In this study, we developed a novel algorithm, the NRR algorithm, for extracting RR from neonatal NIRS signals recorded in NICUs. To the best of our knowledge, the proposed NRR algorithm is the first algorithm in the NIRS literature introduced for extracting RR from clinical neonatal NIRS data. The existing methods of RR extraction from NIRS were validated on adult data recorded during the resting stage in a controlled environment (i.e., laboratory environment), whereas the NRR algorithm was validated on neonatal data recorded in a neonatal intensive care unit. We assessed the performance of the NRR algorithm in terms of agreement and linear correlation between the reference and extracted RRs. The results showed a high degree of agreement between the reference and extracted RRs in terms of ME ( $1.1 \pm 2.1$  BPM), RMSE ( $3.8 \pm 3.0$  BPM), and LoA ( $6.7 \pm 4.7$  BPM). In addition, a high linear correlation ( $r = 84.5 \pm 12.3\%$ ;  $p < 0.01$ ) was achieved between the reference and extracted RRs.

We compared the performance of the NRR algorithm with two existing methods, i.e., BPF [24] and BW [21]. Compared with the BPF and BW methods, the NRR algorithm showed a stronger agreement between the reference and extracted RRs in terms of ME, RMSE, and LoA. Likewise, the NRR algorithm showed a greater Pearson's correlation ( $84.5 \pm 12.3\%$ ) between the reference and extracted RRs compared with the BPF ( $70.8 \pm 15.0\%$ ) and BW ( $51.1 \pm 26.0\%$ ) methods. The results of the statistical test implemented confirmed the significant ( $p < 0.05$ ) outperformance of the NRR algorithm over the BPF and BW methods in RR extraction from neonatal NIRS data.



There are several challenges when working with clinical neonatal data compared with healthy adult data. First, neonatal HR and RR have a greater range than the adult ones, and therefore a NIRS system with a higher sampling rate is required. Hence, we used a NIRS system with a 100 Hz sampling rate, whereas the sampling rate was 12.5 Hz in [24] wherein the BPF method was introduced. Second, clinical data, especially the data recorded in the intensive care unit, typically has a lower data quality than the data recorded in a controlled environment due to electromagnetic interference and some limitations in data acquisition which could increase the error in the analyses. Thus, in this study, we used the SQI (Signal Quality Index) algorithm [35] to assess the NIRS signal quality and exclude the signal segments which have very low signal quality, i.e., lower than the SQI level of 2 (low signal quality). Third, clinical data generally has a higher level of motion artifacts due to voluntary and unintentional patient motion than the data recorded in a controlled environment where the subjects' movements are restricted. In the existing BPF [24] and BW [21] methods, motion artifacts were not considered in the analyses and the signals were recorded during the resting state. Conversely, in this study, we used the IQR (interquartile range) to detect motion artifacts and exclude the signal segments that were highly contaminated with motion artifacts. Fourth, there is more variation in the signal source in clinical recordings than in recordings from healthy subjects in the resting state. Therefore, in this study, we introduced an adaptive RR frequency bandwidth for computing RR that is updated in each segment according to the computed HR as explained in Section 2.3.7. All in all, taking into account the clinical data challenges, we have developed a robust algorithm, NRR, that results in a superior performance than the existing methods in extracting RR from neonatal and clinical NIRS data.

The NIRS system used in this study provides raw NIRS signals (optical densities) with a high sampling rate, i.e., 100 Hz, whereas the existing clinical NIRS systems are incapable of such. High sampling rate NIRS has advantages when extracting physiological information such as RR. One advantage is that it gives the opportunity to assess the signal quality. The NIRS signal quality can be assessed by determining the strength of the heartbeat component in the signals as proposed in [35,39,40], which requires a high sampling rate in order to capture the heartbeats. Another advantage is that, compared with a low sampling rate, it provides a larger number of samples in a specified window; so, with IQR computed in sliding windows of, e.g., 1 s, a better representation of the motion artifacts in the window is achieved.

This study has some limitations that could be addressed in future studies. Due to limited resources and access to participants, especially in an intensive care environment, we were only able to enroll 10 neonates in the study. We chose to enroll 10 neonates based on previous experience with similar studies [19,41–43] and the practical constraints of conducting research in a clinical setting. The algorithm proposed in this study was validated on a dataset with a total number of 7872 signal segments. However, a validation of the algorithm is needed on a more extensive dataset acquired from a larger population and a more diverse group of patients in order to recognize this RR monitoring approach as suitable for current clinical routines of RR monitoring in hospitals. The NRR algorithm computes RR every 7.5 s; however, this delay in RR computation could be problematic if analyzing short characteristics of RR is desired. In addition, the NRR algorithm was used as an offline algorithm in this study, but it would be more beneficial in clinical practice to develop an online version of the algorithm. This could be achieved by implementing a few adjustments in Stage A of the algorithm. For instance, the signal quality could be assessed in each 30-s window, and the HR frequency bandwidth could be updated during the measurement every 10 min.

Extracting RR from neonatal NIRS provides the opportunity to have two perfectly synchronized clinical biomarkers—i.e., RR as well as cerebral oxygen saturation—using a single clinical system. RR, on the one side, is known as one of the main vital signs recorded in a standard clinical routine that is an early predictor of clinical deterioration in children [44–47]. Furthermore, RR monitoring is of great importance in neonates admitted to NICUs as respiratory distress is one of the leading causes of morbidity in the first days of life [48–51]. NIRS-monitored cerebral oxygen saturation, on the other side, is of great value for determining brain oxygen perfusion, so it has been widely incorporated into the standard clinical routine in NICUs [18]. NIRS monitoring has shown feasibility in diagnosing different clinical conditions such as hypotension, hypoxia, hypocapnia, hypercapnia, anemia, apnea, and asphyxia [10,52–57]. Therefore, a NIRS system with the extracted RR opens up the possibility of having a concurrent and complementary assessment of respiration and cerebral perfusion in neonates admitted to the NICU. As a potential use, the system could facilitate the concurrent analysis of RR and cerebral oxygenation for detecting abnormal respiratory events such as tachypnea and apnea and analyzing their effect on cerebral hemodynamics [7,12]. In addition to RR and cerebral oxygen saturation, while applying the NRR algorithm to the raw NIRS signals we could also compute the HR signal (see Section 2.3.6). As a result, we could potentially measure three facets of physiological information, i.e., cerebral oxygen saturation, HR, and RR, using only a single cerebral NIRS system. The portability and easy-to-use features of such a system would make it ideal to be used in the first hours or days of life in clinics, i.e., imposing a minimal burden on the patient and the nurse. In addition, it could prospectively eliminate the need for extra electrodes for RR and HR monitoring with adhesive electrodes. The elimination of excessive and adhesive electrodes would reduce discomfort, stress, and the risk of epidermal stripping, and also could facilitate parent–neonate physical interaction [25–29,58,59].

## 5. Conclusion

In this study, we developed a novel algorithm, NRR (NIRS RR), for extracting RR from clinical NIRS signals recorded in neonates. The results showed a high degree of agreement and a high linear correlation between the reference RR and NIRS-extracted RR. The NRR algorithm outperformed two existing algorithms, i.e., BPF and BW. Therefore, simultaneous RR and cerebral oximetry using a single sensor in a neonatal intensive care setting is feasible. Combining neonatal cerebral NIRS with the extracted RR in a single monitoring system allows for a perfectly time-synced integrated analysis of the impact of abnormal respiratory events (e.g., apnea) on cerebral hemodynamics.

## References

1. Fuerch, J.H.; Thio, M.; Halamek, L.P.; Liley, H.G.; Wyckoff, M.H.; Rabi, Y. Respiratory Function Monitoring during Neonatal Resuscitation: A Systematic Review. *Resusc Plus* **2022**, *12*, 100327, doi:10.1016/J.RESPLU.2022.100327.
2. Fitzgerald, D.; Van Asperen, P.; O'leary, P.; Feddema, P.; Leslie, G.; Arnold, J.; Sullivan, C. Sleep, Respiratory Rate, and Growth Hormone in Chronic Neonatal Lung Disease. **1998**, doi:10.1002/(SICI)1099-0496(199810)26:4.
3. Sullivan, B.A.; Fairchild, K.D. Vital Signs as Physiometers of Neonatal Sepsis. *Pediatric Research* **2021** *91:2* **2021**, *91*, 273–282, doi:10.1038/s41390-021-01709-x.
4. Vavrinsky, E.; Esfahani, N.E.; Hausner, M.; Kuzma, A.; Rezo, V.; Donoval, M.; Kosnacova, H. The Current State of Optical Sensors in Medical Wearables. *Biosensors* **2022**, *Vol. 12*, *Page 217* **2022**, *12*, 217, doi:10.3390/BIOS12040217.
5. Lorato, I.; Stuijk, S.; Meftah, M.; Kommers, D.; Andriessen, P.; van Pul, C.; de Haan, G. Towards Continuous Camera-Based Respiration Monitoring in Infants. *Sensors* **2021**, *Vol. 21*, *Page 2268* **2021**, *21*, 2268, doi:10.3390/S21072268.
6. Reuter, S.; Moser, C.; Baack, M. Respiratory Distress in the Newborn. *Pediatr Rev* **2014**, *35*, 417–428, doi:10.1542/PIR.35-10-417.
7. Jenni, O.G.; Wolf, M.; Hengartner, M.; von Siebenthal, K.; Keel, M.; Bucher, H.U. Impact of Central, Obstructive and Mixed Apnea on Cerebral Hemodynamics in Preterm Infants. *Neonatology* **1996**, *70*, 91–100, doi:10.1159/000244353.
8. Horne, R.S.C.; Fung, A.C.H.; McNeil, S.; Fyfe, K.L.; Odoi, A.; Wong, F.Y. The Longitudinal Effects of Persistent Apnea on Cerebral Oxygenation in Infants Born Preterm. *J Pediatr* **2017**, *182*, 79–84, doi:10.1016/J.JPEDI.2016.11.081.
9. Payer, C.; Urlesberger, B.; Pauger, M.; Müller, W. Apnea Associated with Hypoxia in Preterm Infants: Impact on Cerebral Blood Volume. *Brain Dev* **2003**, *25*, 25–31, doi:10.1016/S0387-7604(02)00121-3.
10. Yamamoto, A.; Yokoyama, N.; Yonetani, M.; Uetani, Y.; Nakamura, H.; Nakao, H. Evaluation of Change of Cerebral Circulation by SpO<sub>2</sub> in Preterm Infants with Apneic Episodes Using near Infrared Spectroscopy. *Pediatrics International* **2003**, *45*, 661–664, doi:10.1111/J.1442-200X.2003.01803.X.
11. Coleman, J.; Ginsburg, A.S.; Macharia, W.M.; Ochieng, R.; Chomba, D.; Zhou, G.; Dunsmuir, D.; Karlen, W.; Ansermino, J.M. Assessment of Neonatal Respiratory Rate Variability. *J Clin Monit Comput* **2022**, *36*, 1869–1879, doi:10.1007/S10877-022-00840-2/FIGURES/4.
12. Choi, S.H.; Lee, J.; Nam, S.K.; Jun, Y.H. Cerebral Oxygenation during Apnea in Preterm Infants: Effects of Accompanying Peripheral Oxygen Desaturation. *Neonatal Medicine* **2021**, *28*, 14–21, doi:10.5385/NM.2021.28.1.14.
13. Zhu, Z.; Liu, T.; Li, G.; Li, T.; Inoue, Y. Wearable Sensor Systems for Infants. *Sensors* **2015**, *Vol. 15*, *Pages 3721-3749* **2015**, *15*, 3721–3749, doi:10.3390/S150203721.

14. Villringer, A.; Planck, J.; Hock, C.; Schleinkofer, L.; Dirnagl, U. Near Infrared Spectroscopy (NIRS): A New Tool to Study Hemodynamic Changes during Activation of Brain Function in Human Adults. *Neurosci Lett* **1993**, *154*, 101–104, doi:10.1016/0304-3940(93)90181-J.
15. Sood, B.G.; McLaughlin, K.; Cortez, J. Near-Infrared Spectroscopy: Applications in Neonates. *Semin Fetal Neonatal Med* **2015**, *20*, 164–172, doi:10.1016/J.SINY.2015.03.008.
16. Hansen, M.L.; Hyttel-Sørensen, S.; Jakobsen, J.C.; Gluud, C.; Kooi, E.M.W.; Mintzer, J.; de Boode, W.P.; Fumagalli, M.; Alarcon, A.; Alderliesten, T.; et al. Cerebral Near-Infrared Spectroscopy Monitoring (NIRS) in Children and Adults: A Systematic Review with Meta-Analysis. *Pediatric Research* **2022**, *2022*, 1–12, doi:10.1038/s41390-022-01995-z.
17. Zhang, F.; Cheong, D.; Khan, A.F.; Chen, Y.; Ding, L.; Yuan, H. Correcting Physiological Noise in Whole-Head Functional near-Infrared Spectroscopy. *J Neurosci Methods* **2021**, *360*, doi:10.1016/J.JNEUMETH.2021.109262.
18. Costa, F.G.; Hakimi, N.; Van Bel, F. Neuroprotection of the Perinatal Brain by Early Information of Cerebral Oxygenation and Perfusion Patterns. *Int J Mol Sci* **2021**, *22*, doi:10.3390/IJMS22105389.
19. Hakimi, N.; Kamaledin Setarehdan, S. Stress Assessment by Means of Heart Rate Derived from Functional Near-Infrared Spectroscopy. <https://doi.org/10.1117/1.JBO.23.11.115001> **2018**, *23*, 115001, doi:10.1117/1.JBO.23.11.115001.
20. Hakimi, N.; Jodeiri, A.; Mirbagheri, M.; Setarehdan, S.K. Proposing a Convolutional Neural Network for Stress Assessment by Means of Derived Heart Rate from Functional near Infrared Spectroscopy. *Comput Biol Med* **2020**, *121*, 103810, doi:10.1016/J.COMPBIOMED.2020.103810.
21. Hakimi, N.; Shahbakhti, M.; Sappia, S.; Horschig, J.M.; Bronkhorst, M.; Floor-Westerdijk, M.; Valenza, G.; Dudink, J.; Colier, W.N.J.M. Estimation of Respiratory Rate from Functional Near-Infrared Spectroscopy (FNIRS): A New Perspective on Respiratory Interference. *Biosensors (Basel)* **2022**, *12*, doi:10.3390/BIOS12121170.
22. Reddy, P.; Izzetoglu, M.; Shewokis, P.A.; Sangobowale, M.; Diaz-Arrastia, R.; Izzetoglu, K. Evaluation of FNIRS Signal Components Elicited by Cognitive and Hypercapnic Stimuli. *Scientific Reports* **2021**, *11*, 1–15, doi:10.1038/s41598-021-02076-7.
23. Iqbal, T.; Elahi, A.; Ganly, S.; Wijns, W.; Shahzad, A. Photoplethysmography-Based Respiratory Rate Estimation Algorithm for Health Monitoring Applications. *J Med Biol Eng* **2022**, *42*, 242–252, doi:10.1007/S40846-022-00700-Z/TABLES/6.
24. Tong, Y.; Lindsey, K.P.; Frederick, B.D. Partitioning of Physiological Noise Signals in the Brain with Concurrent Near-Infrared Spectroscopy and fMRI. *Journal of Cerebral Blood Flow and Metabolism* **2011**, *31*, 2352–2362, doi:10.1038/JCBFM.2011.100/ASSET/IMAGES/LARGE/10.1038\_JCBFM.2011.100-FIG6.JPEG.

25. Kuller, J.M.M. Skin Breakdown: Risk Factors, Prevention, and Treatment. *Newborn and Infant Nursing Reviews* **2001**, *1*, 35–42, doi:10.1053/NBIN.2001.22874.
26. Murray, J.S.; Noonan, C.; Quigley, S.; Curley, M.A.Q. Medical Device-Related Hospital-Acquired Pressure Ulcers in Children: An Integrative Review. *J Pediatr Nurs* **2013**, *28*, 585–595, doi:10.1016/J.PEDN.2013.05.004.
27. Anand, K.J.S.; Scalzo, F.M. Can Adverse Neonatal Experiences Alter Brain Development and Subsequent Behavior? *Biol Neonate* **2000**, *77*, 69–82, doi:10.1159/000014197.
28. Afsar, F.S. Skin Care for Preterm and Term Neonates. *Clin Exp Dermatol* **2009**, *34*, 855–858, doi:10.1111/J.1365-2230.2009.03424.X.
29. Ludington-Hoe, S.M.; Anderson, G.C.; Swinth, J.Y.; Thompson, C.; Hadeed, A.J. Randomized Controlled Trial of Kangaroo Care: Cardiorespiratory and Thermal Effects on Healthy Preterm Infants. *Neonatal Network* **2004**, *23*, 39–48, doi:10.1891/0730-0832.23.3.39.
30. Pereira, C.B.; Czaplik, M.; Blazek, V.; Leonhardt, S.; Teichmann, D. Monitoring of Cardiorespiratory Signals Using Thermal Imaging: A Pilot Study on Healthy Human Subjects. *Sensors* **2018**, *Vol. 18, Page 1541* **2018**, *18*, 1541, doi:10.3390/S18051541.
31. Quaresima, V.; Bisconti, S.; Ferrari, M. A Brief Review on the Use of Functional Near-Infrared Spectroscopy (fNIRS) for Language Imaging Studies in Human Newborns and Adults. *Brain Lang* **2012**, *121*, 79–89, doi:10.1016/J.BANDL.2011.03.009.
32. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Comput Sci Eng* **2007**, *9*, 90–95, doi:10.1109/MCSE.2007.55.
33. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **2020** *17:3* **2020**, *17*, 261–272, doi:10.1038/s41592-019-0686-2.
34. Delpy, D.T.; Cope, M.; Van Der Zee, P.; Arridge, S.; Wray, S.; Wyatt, J. Estimation of Optical Pathlength through Tissue from Direct Time of Flight Measurement. *Phys Med Biol* **1988**, *33*, 1433, doi:10.1088/0031-9155/33/12/008.
35. Sappia, M.S.; Hakimi, N.; Colier, W.N.J.M.; Horschig, J.M. Signal Quality Index: An Algorithm for Quantitative Assessment of Functional near Infrared Spectroscopy Signal Quality. *Biomedical Optics Express*, *Vol. 11, Issue 11*, pp. 6732-6754 **2020**, *11*, 6732–6754, doi:10.1364/BOE.409317.
36. Treitel, S. Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques. *Am Sci* **1995**, *83*, 195–197.
37. Slepian, D. Prolate Spheroidal Wave Functions, Fourier Analysis, and Uncertainty—V: The Discrete Case. *Bell System Technical Journal* **1978**, *57*, 1371–1430, doi:10.1002/J.1538-7305.1978.TB02104.X.
38. Altman, D.G.; Bland, J.M. Measurement in Medicine: The Analysis of Method Comparison Studies. *Journal of the Royal Statistical Society: Series D (The Statistician)* **1983**, *32*, 307–317, doi:10.2307/2987937.

39. Pollonini, L.; Olds, C.; Abaya, H.; Bortfeld, H.; Beauchamp, M.S.; Oghalai, J.S. Auditory Cortex Activation to Natural Speech and Simulated Cochlear Implant Speech Measured with Functional Near-Infrared Spectroscopy. *Hear Res* **2014**, *309*, 84–93, doi:10.1016/J.HEARES.2013.11.007.
40. Sappia, M.S.; Hakimi, N.; Svinkunaite, L.; Alderliesten, T.; Horschig, J.M.; Colier, W.N.J.M. FNIRS Signal Quality Estimation by Means of a Machine Learning Algorithm Trained on Morphological and Temporal Features. <https://doi.org/10.1117/12.2587188> **2021**, *11638*, 29–39, doi:10.1117/12.2587188.
41. Holper, L.; Seifritz, E.; Scholkmann, F. Short-Term Pulse Rate Variability Is Better Characterized by Functional near-Infrared Spectroscopy than by Photoplethysmography. *J Biomed Opt* **2016**, *21*, 091308, doi:10.1117/1.JBO.21.9.091308.
42. Mirbagheri, M.; Hakimi, N.; Ebrahimzadeh, E.; Setarehdan, S.K. Quality Analysis of Heart Rate Derived from Functional Near-Infrared Spectroscopy in Stress Assessment. *Inform Med Unlocked* **2020**, *18*, 100286, doi:10.1016/J.IMU.2019.100286.
43. Perdue, K.L.; Westerlund, A.; McCormick, S.A.; Nelson, C.A.; Iii, C.A.N. Extraction of Heart Rate from Functional Near-Infrared Spectroscopy in Infants. <https://doi.org/10.1117/1.JBO.19.6.067010> **2014**, *19*, 067010, doi:10.1117/1.JBO.19.6.067010.
44. Daw, W.; Kaur, R.; Delaney, M.; Elphick, H. Respiratory Rate Is an Early Predictor of Clinical Deterioration in Children. *Pediatr Pulmonol* **2020**, *55*, 2041–2049, doi:10.1002/PPUL.24853.
45. Braun, S.R. Respiratory Rate and Pattern. *Clinical Methods: The History, Physical, and Laboratory Examinations* **1990**.
46. Cretikos, M.A.; Bellomo, R.; Hillman, K.; Chen, J.; Finfer, S.; Flabouris, A. Respiratory Rate: The Neglected Vital Sign. *Med J Aust* **2008**, *188*, 657–659, doi:10.5694/J.1326-5377.2008.TB01825.X.
47. Elliott, M.; Coventry, A. Critical Care: The Eight Vital Signs of Patient Monitoring. <http://dx.doi.org/10.12968/bjon.2012.21.10.621> **2013**, *21*, 621–625, doi:10.12968/BJON.2012.21.10.621.
48. Clark, R.H. The Epidemiology of Respiratory Failure in Neonates Born at an Estimated Gestational Age of 34 Weeks or More. *Journal of Perinatology* **2005** *25:4* **2004**, *25*, 251–257, doi:10.1038/sj.jp.7211242.
49. Szabo, S.M.; Gooch, K.L.; Korol, E.E.; Bradt, P.; Vo, P.; Levy, A.R. Respiratory Distress Syndrome at Birth Is a Risk Factor for Hospitalization for Lower Respiratory Tract Infections in Infancy. *Pediatric Infectious Disease Journal* **2012**, *31*, 1245–1251, doi:10.1097/INF.0B013E3182737349.
50. Liu, J.; Sorantin, E. Neonatal Respiratory Distress Syndrome. *Neonatal Lung Ultrasonography* **2020**, 17–39, doi:10.1007/978-94-024-1549-0\_3.
51. Ahmed, W.; Veluthandath, A.V.; Rowe, D.J.; Madsen, J.; Clark, H.W.; Postle, A.D.; Wilkinson, J.S.; Murugan, G.S. Prediction of Neonatal Respiratory Distress Biomarker Concentration by Application of Machine Learning to Mid-Infrared Spectra. *Sensors* **2022**, *22*, 1744, doi:10.3390/S22051744/S1.

52. Dasgupta, S.J.; Gill, A.B. Hypotension in the Very Low Birthweight Infant: The Old, the New, and the Uncertain. *Arch Dis Child Fetal Neonatal Ed* **2003**, *88*, F450–F454, doi:10.1136/FN.88.6.F450.
53. Petrova, A.; Mehta, R. Near-Infrared Spectroscopy in the Detection of Regional Tissue Oxygenation during Hypoxic Events in Preterm Infants Undergoing Critical Care. *Pediatric Critical Care Medicine* **2006**, *7*, 449–454, doi:10.1097/01.PCC.0000235248.70482.14.
54. Owen-Reece, H.; Elwell, C.E.; Goldstone, J.; Smith, M.; Delpy, D.T.; Wyatt, J.S. Investigation of the Effects of Hypocapnia upon Cerebral Haemodynamics in Normal Volunteers and Anaesthetised Subjects by near Infrared Spectroscopy (NIRS). *Adv Exp Med Biol* **1994**, *361*, 475–482, doi:10.1007/978-1-4615-1875-4\_85/COVER.
55. Schopfer, L.; Habre, W.; Pichon, I.; Fodor, G.H. Effect of Permissive Mild Hypercapnia on Cerebral Vasoreactivity in Infants: A Randomized Controlled Crossover Trial. *Anesth Analg* **2021**, 976–983, doi:10.1213/ANE.0000000000005325.
56. Sandal, G.; Oguz, S.S.; Erdeve, O.; Akar, M.; Uras, N.; Dilmen, U. Assessment of Red Blood Cell Transfusion and Transfusion Duration on Cerebral and Mesenteric Oxygenation Using Near-Infrared Spectroscopy in Preterm Infants with Symptomatic Anemia. *Transfusion (Paris)* **2014**, *54*, 1100–1105, doi:10.1111/TRF.12359.
57. Korček, P.; Straňák, Z.; Širc, J.; Naulaers, G. The Role of Near-Infrared Spectroscopy Monitoring in Preterm Infants. *Journal of Perinatology 2017 37:10* **2017**, *37*, 1070–1077, doi:10.1038/jp.2017.60.
58. Russo, C.; Vincenzo, |; Senese, P.; Mazaheri, A. Functional Near-Infrared Spectroscopy Is a Useful Tool for Multi-Perspective Psychobiological Study of Neurophysiological Correlates of Parenting Behaviour. *European Journal of Neuroscience* **2023**, *57*, 258–284, doi:10.1111/EJN.15890.
59. Nguyen, T.; Abney, D.H.; Salamander, D.; Bertenthal, B.I.; Hoehl, S. Proximity and Touch Are Associated with Neural but Not Physiological Synchrony in Naturalistic Mother-Infant Interactions. *Neuroimage* **2021**, *244*, 118599, doi:10.1016/J.NEUROIMAGE.2021.118599.





## CHAPTER 9

### Near-infrared spectroscopy for neonatal sleep classification

Naser Hakimi\*, Emad Arasteh\*, Maren Zahn, Jörn M. Horschig, Willy N. J. M. Colier, Jeroen Dudink\*\*, Thomas Alderliesten\*\*

\* shared first author

\*\* shared last author

*Under submission*



## Abstract

Sleep, notably active sleep (AS) and quiet sleep (QS), plays a pivotal role in the brain development and gradual maturation of (pre-) term infants. Monitoring their sleep patterns is imperative, as it serves as a vital tool in tracking neurological growth and well-being, particularly crucial for preterm infants who are at heightened risk of immature brain development. Accurate classification of neonatal sleep stages can significantly contribute to optimizing treatment for high-risk infants, with respiratory rate (RR) and heart rate (HR) serving as key components in sleep assessment systems for neonates. Recent studies have demonstrated the feasibility of extracting both RR and HR from near-infrared spectroscopy (NIRS) in neonates. This study introduces a comprehensive sleep classification approach leveraging high-frequency NIRS signals, recorded at a sampling rate of 100 Hz, from a cohort of nine preterm infants admitted to a neonatal intensive care unit. Eight distinct features were extracted from raw NIRS signals, including HR, RR, motion-related parameters, and proxies for neural activity. These features served as inputs for a deep convolutional neural network (CNN) model designed for the classification of AS and QS sleep stages. Performance evaluation of the proposed CNN model was done using two cross-validation approaches: data pooling and leave-two-subjects-out (LTSO). The accuracy, balanced accuracy, F1-score, and Kappa, were employed to assess classifier performance. In addition, comparative analyses against six benchmark classifiers, comprising K-Nearest Neighbors, Naive Bayes, Support Vector Machines, Random Forest (RF), AdaBoost, and XGBoost (XGB), were conducted. Our results reveal the CNN model's superior performance, achieving an average accuracy of 88%, balanced accuracy of 94%, F1-score of 91%, and Kappa of 95% in data pooling cross-validation. Furthermore, in both cross-validation methods, RF and XGB demonstrated accuracy levels closely comparable to the CNN classifier. These findings underscore the feasibility of leveraging high-frequency NIRS data, coupled with NIRS-based HR and RR extraction, for assessing sleep stages in neonates, even in an intensive care setting. The user-friendliness, portability, and reduced sensor complexity of the approach suggest its potential applications in various and less demanding settings. This research thus presents a promising avenue for advancing neonatal sleep assessment and its implications for infant health and development.



## Introduction

Infants born prematurely spend their early weeks in an incubator instead of the protective environment of the mother's womb [1]. This period is crucial for their brain development, because the latter part of the second, and the third trimester are critical for creating new connections in the brain [2]. Fetal sleep, which is believed to be the primary driver of neural activity, is vital for neuronal survival, axonal guidance, and synapse maturation [3,4]. However, infants in the neonatal intensive care unit (NICU) receive hands-on care and might be exposed to various external stimuli that can significantly disrupt their sleep [5].

The significance of sleep for neonatal brain development is being increasingly acknowledged [6,7]. Disrupted sleep during the neonatal period has already proved to have negative developmental consequences [8]. Monitoring neonatal sleep continuously in the NICU could aid care in two ways. First, as neonatal sleep is considered to reflect typical brain maturation, it can also serve as a biomarker for future outcomes, such as the severity of an illness or the infant's maturational state [9]. Second, if sleep is monitored in real-time, elective care could be individually customized to sleep stages to ensure optimal brain development [10,11].

The primary method of assessing sleep in very preterm infants is behavioral observation [12]. Sleep comprises three behavioral stages: active sleep (AS), which is important for creating new brain connections; quiet sleep (QS), which is critical for consolidating connections and recovery; and intermediate sleep, which is a transitional stage [13]. Ideally, elective care should correspond with neonatal sleep stages, but this requires continuous sleep observation by trained observers, which is time-consuming and costly and therefore not feasible [14]. However, sleep stages in (pre-) term infants can also be determined using typical patterns in their heart rate and breathing [15]. This is different from adults which are categories are usually determined by polysomnography (PSG) [14].

Numerous studies have delved into the utilization of machine learning and neural network classification techniques to automatically detect sleep-wake states in infants. Sentner et al. [16] and Werth et al. [17] integrated physiological characteristics such as HR, RR, and peripheral oxygen saturation (SpO<sub>2</sub>) from regular monitor recordings in the NICU to classify sleep stages in preterm infants. Other studies utilized electroencephalography (EEG) features for sleep classification. Specifically, Hermans et al. [18] developed a convolutional neural network (CNN) classifier that fused multiple EEG-derived features to differentiate between sleep stages in term infants. Besides, Wang et al. [19] exclusively utilized EEG features in preterm infants. Koolen et al. [20] proposed a Support Vector Machines (SVM) classifier trained on EEG features in infants, and Ansari et al. [21] used exclusively EEG features to classify QS vs non-QS in preterm infants. These studies demonstrate the potential of machine learning and neural network classification techniques in accurately identifying sleep stages in neonates.

Despite the promising results of the automated sleep staging methods, they have some limitations that make them difficult to be used for continuous daily monitoring in the NICUs. For instance, some methods need multiple electrodes that can potentially damage the delicate skin of preterm infants [20,22,23]. Additionally, most of the existing studies, including those mentioned, have only classified sleep stages using pooled datasets, which lacks subject-specific analysis of preterm infants' sleep.

NIRS is a non-invasive technique that utilizes varying wavelengths of infrared light to monitor changes in oxygenated- and deoxygenated-hemoglobin in tissues, with venous circulation contributing significantly to the signal [24]. Therefore, the effects of many physiological signals of human body are embedded in recorded NIRS signals [25]. In this regard, recently, NIRS has demonstrated its ability to capture various physiological parameters such as heartbeats, respiration, and blood pressure [24–26].

Both HR and RR have been successfully derived from high sampling rate NIRS signals in adults and neonates with good accuracy [27–29]. Extracting HR and RR from NIRS not only allows for the measurement of diverse physiological information using a single sensor but also opens avenues for exploring multi-modal feature extraction to assess sleep stages, particularly in infants, a domain yet unexplored. This innovative approach holds the potential to serve as a simplified and non-intrusive alternative to traditional PSG, minimizing discomfort for preterm infants by employing only a single head-mounted sensor.

This study presents a novel approach for sleep assessment among the hospitalized preterm infants utilizing high sampling rate NIRS. In this approach, we took advantage of the high sampling rate NIRS by extracting multiple features including HR, RR, motion-related, and neuronal activity features. These features were then processed through a CNN-based classifier to classify the sleep stages AS versus QS. The performance of the classifier was evaluated in two cross-validation approaches and compared with six benchmark classifiers. Here we demonstrate the potential of high sampling rate NIRS for in-depth sleep analysis but also highlight the practicality of our single-sensor approach in clinical settings, offering a promising avenue for enhancing neonatal sleep stage assessment.

## 2. Materials and Methods

### 2.1. Participants

This study was carried out within the NICU at the Wilhelmina Children’s Hospital, Utrecht, the Netherlands. The inclusion criterion was a gestational age (GA) of less than 37 weeks. Exclusion criteria were genetic anomalies, congenital malformations, historical occurrences of seizures, evident brain injuries (e.g., intraventricular hemorrhage exceeding grade 2 according to the classification by Papile [30]), and maternal utilization of recreational drugs during pregnancy. The study involved the participation of 13 infants, resulting in a total of 17 measurements, with four infants contributing two measurements each. However, we excluded data from four infants due to factors such as recent administration of sedatives within 24 hours before recording, sensor malfunction, or the absence of reference data. Consequently, our analysis was performed on a dataset comprising 10 measurements obtained from nine infants. Prior to participation in the study, written parental informed consent was obtained. This research was approved by the local Medical Ethical Review Committee (METC number 21-098/C). An overview of patient demographics is presented in Table 1.

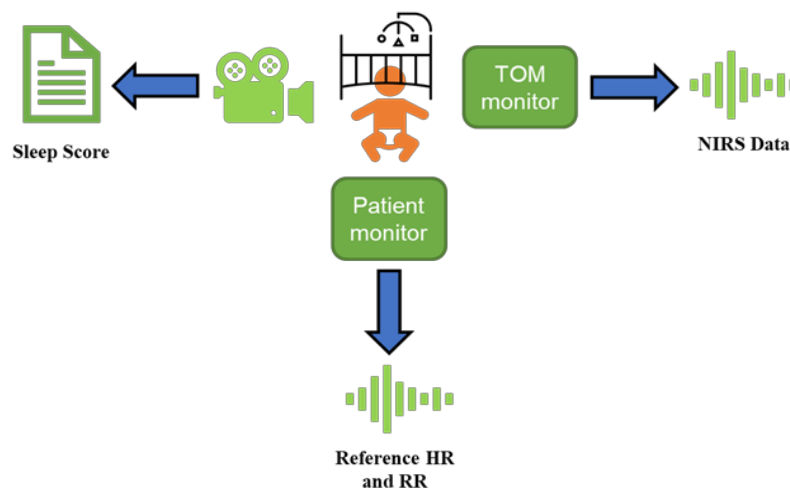
**Table 1:** Patient demographics.

Variable	Subjects (N=9)
Gender, male, % (n)	44% (4)
Gestational age (GA), weeks, mean $\pm$ SD <sup>1</sup>	31.4 $\pm$ 2.7
Birth weight, g, mean $\pm$ SD	1520 $\pm$ 462
Delivery method % (n)	
• Spontaneous	11% (1)
• Cesarean section	89% (8)
Apgar score, mean $\pm$ SD	
• At 5 min	7.4 $\pm$ 1.1
• At 10 min	8.8 $\pm$ 0.8

<sup>1</sup> Standard deviation

## 2.2. Data acquisition

A schematic overview of the data acquisition procedure is presented in Figure 1. NIRS data, encompassing measurements of oxygenated hemoglobin (O<sub>2</sub>Hb), deoxygenated hemoglobin (HHb), and tissue saturation index (TSI), were acquired using the TOM device (Artinis Medical Systems B.V., Elst, the Netherlands), at a sampling rate of 100 Hz. The TOM sensor, equipped with two NIRS channels, was positioned on the neonate's forehead, precisely above and parallel to the eyebrows, and secured beneath a hat or headband worn by the infant. Reference HR and RR values were obtained using a bedside Philips IntelliVue MP70 (Philips Medical Systems, Best, the Netherlands), stored at 0.4 Hz. This reference data was collected to facilitate the evaluation of the accuracy of the algorithms employed for HR and RR extraction from the NIRS data. Throughout the entire recording period, the facial and bodily expressions of the infants were concurrently captured by one or two cameras. These cameras, specifically the Intel® RealSense Depth Camera D435i (Intel Corporation, Santa Clara, USA) and the 1SEE VDO360 camera (VDO360, Maryland, USA), provided a frame rate of 30 frames per second and a resolution of 1920x1080 pixels. Data synchronization was achieved in a two-step process. First, the time settings on each device were synchronized with internet time. Second, both the internet time and the device's individual measurement start time were adjusted to achieve accurate synchronization of the collected data.



**Figure 1:** Schematic representation of the data acquisition protocol in this study.

## 2.3. Sleep observation

We conducted sleep observations in adherence to a validated sleep stage classification approach in preterm infants, known as the Behavioral Sleep Stage classification for Preterm Infants (BeSSPI) [12]. A trained annotator (Maren Zahn) used video data to categorize sleep stages, with a specific focus on two primary stages: AS and QS. These classifications were performed at one-minute intervals.

## 2.4. NIRS data analysis

The NIRS optical densities were initially converted into O<sub>2</sub>Hb and HHb signals by applying the modified Beer-Lambert law [31]. Subsequently, the Signal Quality Index (SQI) algorithm [32] was employed to assess the signal quality of each NIRS channel. The channel with the highest signal quality was selected for further analysis.

To extract HR and RR from the NIRS signals, we utilized the algorithms developed by Hakimi et al. [28,29]. This involved segmenting the total hemoglobin concentration (tHb) signal, which is the addition of O<sub>2</sub>Hb and HHb signals, and the normalized interquartile range (IQR) signal using sliding windows of 30 seconds with a 75% overlap (i.e., shifted 7.5 seconds). Subsequently, we computed the spectrum of the converted tHb signal and identified the dominant frequencies within the adaptive HR and RR frequency bandwidths. These dominant frequencies represented the extracted HR and RR values derived from the NIRS signals, respectively. The performance of the algorithms for extracting HR and RR was assessed using various metrics, including Mean Error (ME), Root Mean Square Error (RMSE), Pearson's correlation coefficient, and Bland-Altman Limits of Agreement (LoA), providing a comprehensive evaluation of their accuracy and effectiveness.

## 2.5. Sleep classification

### 2.5.1. Multi-modal feature extraction

We extracted eight features, treated as modalities, as time series data sampled at 1 Hz from the NIRS signals. By the methodology detailed in section 2.4, the extraction of HR and RR from the NIRS signals occurred at 7.5-second intervals. Subsequently, we applied spline



interpolation to convert the extracted HR and RR values into time series with a 1 Hz sampling rate.

For features related to motion artifacts present in the NIRS signals [29,33–35], we calculated the standard deviation and IQR based on the O2Hb signal within sliding windows of 1-second duration, with a 50% overlap. These computed values were then subjected to spline interpolation to achieve a 1 Hz sampling rate. To normalize these features, we divided the standard deviation by the signal's mean and the IQR by the signal's median. Additionally, we computed the Pearson's correlation between the O2Hb and HHb signals within the same sliding window and interpolation procedure. This feature serves a dual purpose [36,37], potentially indicating brain activation when exhibiting a strong negative correlation, or conversely, motion artifacts when demonstrating a strong positive correlation. Furthermore, for the remaining features, we applied a 1-second moving average filter to the O2Hb, HHb, and TSI signals. Subsequently, these signals were downsampled to 1 Hz.

Table 2 provides a comprehensive summary of the eight features derived from the NIRS signals. Notably, since the sleep scores were assigned in 1-minute intervals, and considering the 1 Hz sampling rate, each modality consisted of 60 data samples for analysis.

**Table 2:** List of the features (i.e., modalities) extracted from the NIRS signals accompanied by description for each feature.

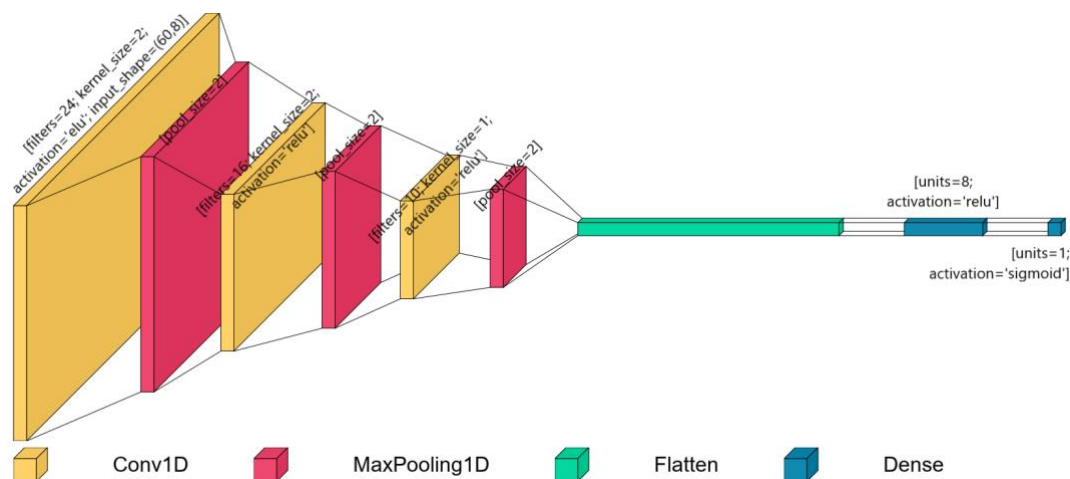
Feature	Description
O2Hb	Filtered O2Hb by using a moving average filter of 1 second and downsampled to 1 Hz.
HHb	Filtered HHb by using a moving average filter of 1 second and downsampled to 1 Hz.
TSI	Filtered TSI by using a moving average filter of 1 second and downsampled to 1 Hz.
HR	Extracted heart rate from NIRS and interpolated to 1 Hz.
RR	Extracted respiratory rate from NIRS and interpolated to 1 Hz.
Moving standard deviation	Normalized standard deviation on O2Hb in sliding windows of 1 second overlapping 50% and interpolated to 1 Hz.
Moving IQR	Normalized IQR on O2Hb in sliding windows of 1 second overlapping 50% and interpolated to 1 Hz.
Moving correlation	Pearson's correlation between O2Hb and HHb in sliding windows of 1 second overlapping 50% and interpolated to 1 Hz.

## 2.5.2. Classification

We employed six benchmark classifiers: K-Nearest Neighbors (KNN), Naive Bayes (NB), SVM, Random Forest (RF), AdaBoost (AdaB), and XGBoost (XGB) [38], alongside a novel classifier based on CNN. A visual depiction of the proposed CNN model can be seen in Figure 2.

The classifiers underwent training and testing utilizing the features extracted as described in Section 2.5.1, employing two distinct approaches. Initially, a data pooling approach was executed, where all measurements were consolidated, and the classifier was trained and tested through 10-fold cross-validation on the merged dataset. Subsequently, a 'leave-two-subjects-out' (LTSO) cross-validation approach was adopted to ensure the model's generalizability across individual subject-specific variations. In this approach, the 10

measurements (from 9 subjects) were partitioned into 5 groups (referred to as Group 1 through Group 5), ensuring that each subject was exclusively represented in one group, and minimizing the difference in the ratio of AS to QS events between the 5 groups. Consequently, the classifier underwent training on data from four groups while its performance was assessed for the remaining group in a repeatable loop. Table 3 provides a summary of the optimized hyperparameters for the benchmark classifiers, determined through hyperparameter optimization utilizing the GridSearchCV package in Python [39]. The classifiers' performance was evaluated in terms of accuracy, balanced accuracy, F1-score, and Kappa.



**Figure 2:** Graphical representation of the CNN model proposed in this study.

**Table 3:** Optimized hyperparameters for benchmark classifiers.

Classifier	Hyperparameters
KNN <sup>1</sup>	{'Distance weight': Squared inverse}; {'Number of Neighbors'=12}; {'NS method': Exhaustive}; {'Break ties': Smallest}
NB <sup>2</sup>	{'Kernel': Normal – 'Width'=0.028}
SVM <sup>3</sup>	{'Kernel': Linear}; {'Box constraints'=0.072}; {'Solver': SMO}; {'Kernel scale'=1}; {'Decision function shape': Ovo}
RF <sup>4</sup>	{'n_estimators': 100}; {'max_depth': 20}; {'min_samples_split': 2}; {'criterion': 'gini'}
AdaB <sup>5</sup>	{'Number of learning cycles'=479}; {'Learning rate'= 0.8}; {'Minimum Leaf Size'=4}; {'Max Splits'=58}
XGB <sup>6</sup>	{'Learning rate'= 0.3}; {'Min split loss'=0}; {'Max depth'=6}

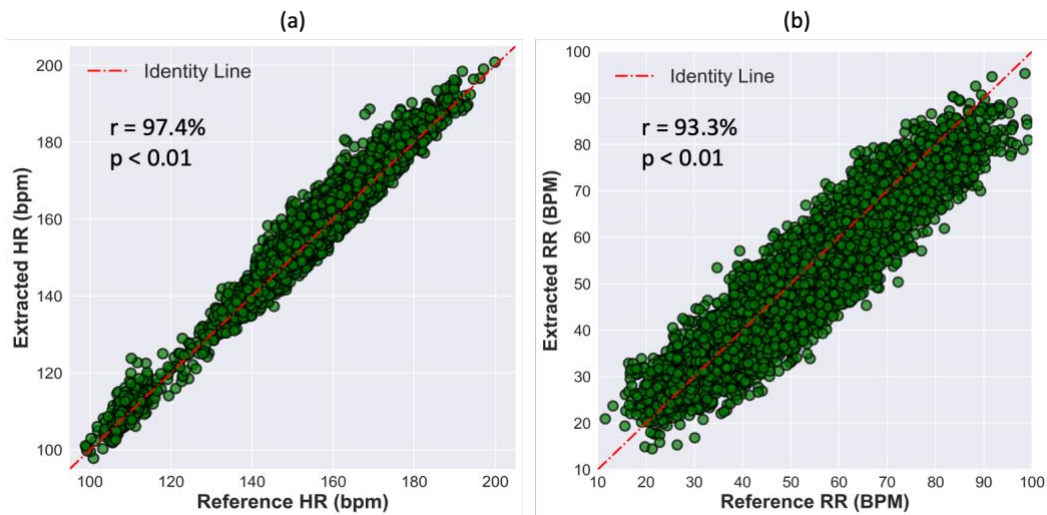
<sup>1</sup> K-Nearest Neighbors. <sup>2</sup> Naive Bayes. <sup>3</sup> Support Vector Machine. <sup>4</sup> Random Forest. <sup>5</sup> AdaBoost. <sup>6</sup> XGBoost.

## 3. Results

### 3.1. Results of HR and RR extraction

Figure 3 depicts scatter plots showcasing the relationship between the extracted HR from NIRS and the reference HR in Figure 3(a), and similarly, the extracted RR from NIRS in Figure 3(b) against the reference RR. Notably, both scatter plots reveal a statistically significant ( $p <$

0.01) linear correlation between the extracted physiological metrics (HR and RR) and their corresponding reference values. Specifically, Pearson's correlation analysis reveals a strong correlation of 97.4% for HR and 93.3% for RR between the reference and extracted data. Looking at the scatter plots, we observe that the algorithms utilized for HR and RR extraction from NIRS effectively estimate the reference physiological parameters across a range of values. Table 4 summarizes the quantitative metrics averaged on all measurements that were used for validating the performance of the algorithm extracting HR and RR from NIRS. For HR extraction, the ME and RMSE between the reference and extracted HRs were computed at  $-1.4 \pm 0.2$  bpm and  $3.4 \pm 0.2$  bpm, respectively. In the case of RR extraction, the ME and RMSE were slightly higher in magnitude, calculated at  $-2.1 \pm 0.5$  BPM and  $6.1 \pm 0.2$  BPM, respectively. Additionally, the LoA for HR extraction was narrower, at  $6.0 \pm 0.3$  bpm, compared to RR extraction, where they were  $11.1 \pm 0.5$  BPM. Furthermore, Pearson's correlation analysis yielded a correlation of  $93.9 \pm 4.2\%$  ( $p < 0.01$ ) for HR extraction and  $89.3 \pm 6.6\%$  ( $p < 0.01$ ) for RR extraction, reaffirming the robustness of the algorithm in accurately estimating the reference physiological information.



**Figure 3:** (a) Scatter plot between the reference HR and the extracted HR (in beats per minute (bpm)) from NIRS when all measurements are concatenated. (b) scatter plot between the reference RR and extracted RR from NIRS when all measurements are concatenated together, in BPM, which stands for breaths per minute. Each green dot corresponds to a 30-second signal segment.

**Table 4:** Quantitative metrics for assessing the performance of the algorithms used for extracting the physiological information, i.e., HR and RR, from NIRS, averaged across all measurements.

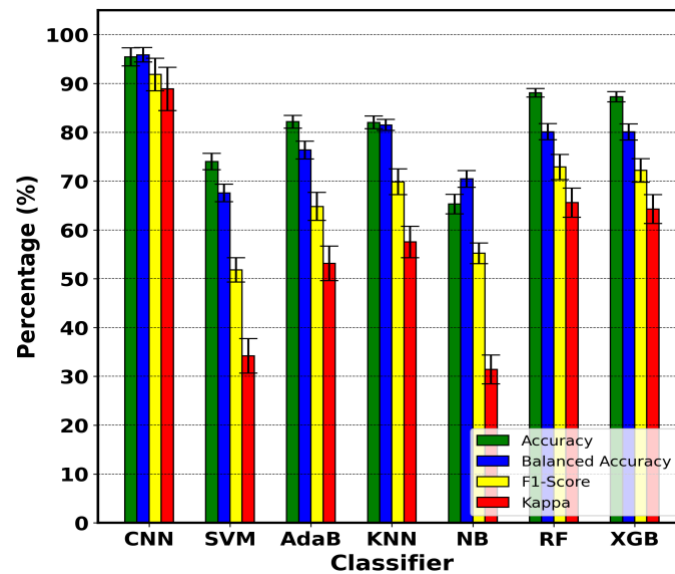
Physiological information	ME <sup>1</sup> (bpm <sup>2</sup> or BPM <sup>3</sup> )	RMSE <sup>4</sup> (bpm or BPM)	LoA <sup>5</sup> (bpm or BPM)	Pearson's r (%) ( $p < 0.01$ )
HR	$-1.4 \pm 0.2$	$3.4 \pm 0.2$	$6.0 \pm 0.3$	$93.9 \pm 4.2$
RR	$-2.1 \pm 0.5$	$6.1 \pm 0.2$	$11.1 \pm 0.5$	$89.3 \pm 6.6$

<sup>1</sup> Mean of error. <sup>2</sup> Beats per minute. <sup>3</sup> Breaths per minute. <sup>4</sup> Root mean square error. <sup>5</sup> Bland-Altman limits of agreement.

### 3.2. Results of sleep classification with data pooling cross-validation approach

Figure 4 illustrates a bar chart depicting various quantitative performance metrics, including accuracy, balanced accuracy, F1-score, and Kappa, employed for evaluating classifier performance within the context of the data pooling cross-validation approach. Error bars are

incorporated into each bar chart to convey metric variability. Notably, the CNN classifier consistently outperforms the six benchmark classifiers (SVM, AdaB, KNN, NB, RF, and XGB) across all four metrics.



**Figure 4:** Bar chart of the quantitative metrics (i.e., accuracy, balanced accuracy, F1-score, and Kappa) used for assessing the performance of the classifiers with error bars averaged across all measurements, obtained using the pooling cross-validation approach.

Among the benchmark classifiers, RF and XGB exhibit higher accuracy, F1-score, and Kappa values compared to the other benchmarks, while KNN attains the highest balanced accuracy. A noteworthy observation emerges when comparing accuracy to balanced accuracy: in most instances, balanced accuracy falls below accuracy. However, exceptions arise for CNN and NB, where balanced accuracy surpasses accuracy, suggesting superior performance on the minority class (i.e., QS) by these classifiers. Furthermore, it is consistently observed that Kappa values are lower than their corresponding F1-scores, which, in turn, are lower than both accuracy and balanced accuracy.

Table 5 summarizes numerically the average and standard deviation values for each quantitative metric across all measurements. Employing the data pooling cross-validation approach, our proposed CNN classifier consistently outperforms benchmark classifiers, yielding an average accuracy of 88%, balanced accuracy of 94%, F1-score of 91%, and Kappa of 95%. In contrast, benchmark classifiers, such as SVM and NB, consistently demonstrate performance levels lower than the best-performing classifiers. AdaB and KNN classifiers, while not reaching the highest levels of performance, also exhibit notable performance, while RF and XGB classifiers excel in terms of accuracy, averaging 88% and 86%, respectively. However, it's important to note that their superior accuracy comes at the expense of performance across other metrics, which remains notably lower than that of the proposed CNN classifier.

**Table 5:** Average and standard deviation of quantitative metrics employed to assess classifier performance using the data pooling cross-validation approach, calculated across all measurements.

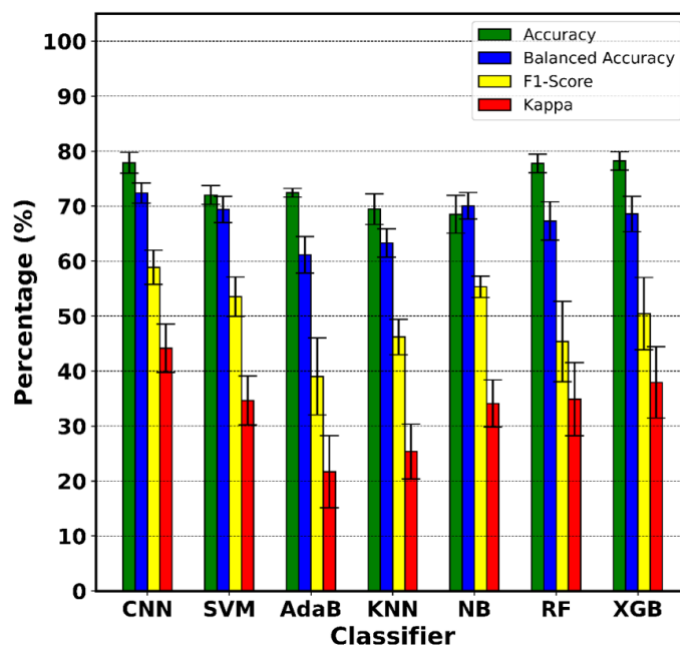
Classifier	Accuracy (%)	Balanced Accuracy (%)	F1-score (%)	Kappa (%)
CNN <sup>1</sup>	88 ± 12	94 ± 7	91 ± 9	95 ± 5
SVM <sup>2</sup>	74 ± 4	68 ± 4	53 ± 7	35 ± 9
AdaB <sup>3</sup>	83 ± 4	77 ± 5	66 ± 9	54 ± 11
KNN <sup>4</sup>	82 ± 5	83 ± 6	71 ± 9	59 ± 12
NB <sup>5</sup>	63 ± 5	67 ± 5	51 ± 6	26 ± 8
RF <sup>6</sup>	88 ± 5	79 ± 6	72 ± 10	65 ± 12
XGB <sup>7</sup>	86 ± 3	79 ± 5	71 ± 8	62 ± 9

<sup>1</sup> Convolutional Neural Network. <sup>2</sup> Support Vector Machine. <sup>3</sup> AdaBoost. <sup>4</sup> K-Nearest Neighbors. <sup>5</sup> Naive Bayes. <sup>6</sup> Random Forest. <sup>7</sup> XGBoost.

### 3.3. Results of sleep classification with LTSO cross-validation approach

Figure 5 presents a bar chart illustrating various quantitative metrics, including accuracy, balanced accuracy, F1-score, and Kappa, utilized to evaluate classifier performance within the context of the Leave-Two-Subjects-Out (LTSO) cross-validation approach. The chart demonstrates that the CNN classifier consistently outperformed benchmark classifiers across all performance metrics, with the exception of accuracy where RF and XGB classifiers exhibited comparable performance. Among the benchmark classifiers, RF and XGB classifiers demonstrated superior accuracy and Kappa values, while NB and SVM classifiers excelled in terms of balanced accuracy and F1-score. Notably, Kappa values for each classifier consistently registered lower values compared to their respective F1-scores, which, in turn, were lower than the attained accuracy and balanced accuracy. Furthermore, it is worth highlighting that, in most cases, balanced accuracy fell short of accuracy across all classifiers, except for the NB classifier.

Table 6 summarizes a numerical summary of average and standard deviation values for each quantitative metric across all measurements, employing the LTSO cross-validation approach. Regarding accuracy, the proposed CNN, RF, and XGB classifiers demonstrated comparable performance, achieving the highest average accuracy among all classifiers at 78%. The proposed CNN classifier showcased superior performance in terms of balanced accuracy (72 ± 6%), F1-score (59 ± 10%), and Kappa (44 ± 14%). Among the benchmark classifiers, the NB classifier recorded the highest values for balanced accuracy (70 ± 8%) and F1-score (55 ± 6%), albeit with a lower accuracy (68 ± 11%). The XGB classifier exhibited the highest Kappa (38 ± 21%), although it displayed a relatively higher standard deviation.



**Figure 5:** Bar chart of the quantitative metrics (i.e., accuracy, balanced accuracy, F1-score, and Kappa) used for assessing the performance of the classifiers with error bars averaged across all measurements, obtained using the LTSO cross-validation approach.

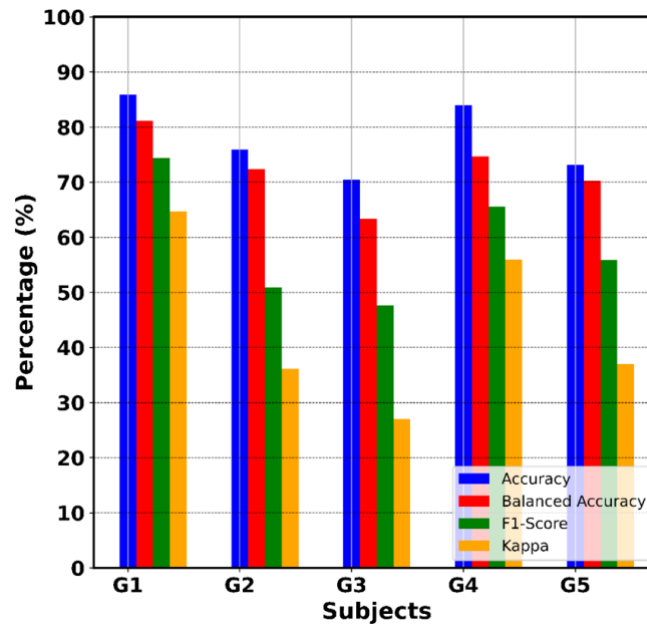
**Table 6:** Average and standard deviation of quantitative metrics employed to assess classifier performance using the LTSO cross-validation approach, calculated across all measurements.

Classifier	Accuracy (%)	Balanced Accuracy (%)	F1-score (%)	Kappa (%)
CNN <sup>1</sup>	78 ± 6	72 ± 6	59 ± 10	44 ± 14
SVM <sup>2</sup>	72 ± 5	69 ± 7	53 ± 11	35 ± 14
AdaB <sup>3</sup>	72 ± 2	61 ± 10	39 ± 22	22 ± 21
KNN <sup>4</sup>	69 ± 9	63 ± 8	46 ± 10	25 ± 16
NB <sup>5</sup>	68 ± 11	70 ± 8	55 ± 6	34 ± 14
RF <sup>6</sup>	78 ± 5	67 ± 11	45 ± 23	35 ± 21
XGB <sup>7</sup>	78 ± 5	69 ± 10	50 ± 21	38 ± 21

<sup>1</sup> Convolutional Neural Network. <sup>2</sup> Support Vector Machine. <sup>3</sup> AdaBoost. <sup>4</sup> K-Nearest Neighbors. <sup>5</sup> Naive Bayes. <sup>6</sup> Random Forest. <sup>7</sup> XGBoost.

Figure 6 illustrates a detailed performance analysis of the proposed CNN classifier, conducted within the framework of the LTSO cross-validation methodology. This assessment encompasses a suite of performance metrics, including accuracy, balanced accuracy, F1-score, and Kappa. Notably, the classifier's performance exhibited variations across different subject groups (see section 2.5.2). Specifically, when tested on Group 1 (G1), the classifier achieved superior results in terms of accuracy, balanced accuracy, F1-score, and Kappa compared to other groups. In contrast, the classifier's performance was notably lower when tested on Group 3 (G3), registering the lowest values across all performance metrics. Looking at the variability of the performance metrics among groups, we observe that the Kappa had more variability from below 30% to above 60% than other metrics. Additionally, a consistent

trend emerges across all groups, with accuracy consistently exceeding balanced accuracy, which, in turn, surpasses the F1-score, while Kappa consistently lags behind as the metric with the lowest values.



**Figure 6:** Bar chart of the quantitative metrics (i.e., accuracy, balanced accuracy, F1-score, and Kappa) used for assessing the performance of the proposed CNN classifier in each subject group, obtained using the LTSO cross-validation approach.

Table 7 summarizes numerically the accuracy, balanced accuracy, F1-score, and Kappa for the proposed CNN classifier per group when using the LTSO cross-validation approach. Among the subject groups, the classifier evaluated on Group 1 (G1) demonstrated the highest levels of performance across multiple metrics, including accuracy (86%), balanced accuracy (81%), F1-score (74%), and Kappa (65%). In contrast, the classifier assessed on Group 3 (G3) displayed the lowest performance, recording accuracy (70%), balanced accuracy (63%), F1-score (48%), and Kappa (27%) at the lowest levels. The most pronounced variation among the groups was observed in the Kappa performance metric, with a range of 38%, while the smallest degree of variation was found in the accuracy metric, with a range of 16%.

**Table 7:** Quantitative metrics employed to assess the performance of the proposed CNN classifier in each subject group, obtained using the LTSO cross-validation approach.

Group NO.	Accuracy (%)	Balanced Accuracy (%)	F1-score (%)	Kappa (%)
1	86	81	74	65
2	76	72	51	36
3	70	63	48	27
4	84	75	66	56
5	73	70	56	37



## 4. Discussion

In this study, we harnessed the potential of multi-modal feature extraction from NIRS signals acquired at an enhanced sampling rate of 100 Hz, evaluating its feasibility of characterizing sleep stages among preterm infants using only a single NIRS sensor. Leveraging the CNN classifier, we differentiated AS from QS by extracting and fusing different features from high sampling rate NIRS, including physiological signals such as HR and RR, motion-related information, and raw NIRS signals. Our approach allowed for the assessment of NIRS signal quality and the extraction of diverse physiological and non-physiological features, including motion artifacts and brain activation, providing a comprehensive platform for sleep assessment. This amalgamation of features, not extensively explored in previous studies, enabled us to construct a more holistic framework for assessing sleep patterns in preterm infants, even those as young as 24 weeks of gestation or with birth weights as low as 500 grams, solely from data acquired from a single sensor. This classifier distinguishing between AS which aids neural processing and learning consolidation [40–42], and QS that promotes physical growth and recovery [43,44], applied on data with minimal contact to the preterm, can be of pivotal benefit to both clinical and research settings.

The performance of the proposed classifier was evaluated through a quantitative analysis involving key performance metrics, including accuracy, balanced accuracy, F1-score, and Kappa. The assessment of classifier performance was undertaken via two distinct cross-validation strategies, namely, pooling cross-validation and LTSO cross-validation. In the context of pooling cross-validation employing a 10-fold scheme, the proposed classifier achieved an accuracy of  $88 \pm 12\%$ , a balanced accuracy of  $94 \pm 7\%$ , an F1-score of  $91 \pm 9$ , and a Kappa of  $95 \pm 5\%$ . Moreover, we also assessed the outcomes of the classifier within the LTSO cross-validation paradigm, with a reported accuracy of  $78 \pm 6\%$ , a balanced accuracy of  $72 \pm 6$ , an F1-score of  $59 \pm 10$ , and a Kappa of  $44 \pm 14$  (as summarized in Tables 5, 6, and 7).

Furthermore, a comparative analysis was conducted to benchmark the performance of the proposed CNN classifier against six established classifiers, including SVM, AdaB, KNN, NB, RF, and XGB. Across both cross-validation methodologies, the CNN classifier consistently exhibited superior performance in terms of balanced accuracy, F1-score, and Kappa, although it is worth noting that RF and XGB classifiers displayed competitive results in terms of overall accuracy, as detailed in Tables 5 and 6. Notably, when considering metrics sensitive to imbalanced classification in the pooling cross-validation framework (i.e., balanced accuracy, F1-score, and Kappa), the CNN classifier demonstrated a substantial advantage over the benchmark classifiers, exceeding their performance by more than 10%. In contrast, the margin of superiority was somewhat reduced in the LTSO cross-validation approach, where the CNN classifier outperformed the benchmark classifiers by 2%, 4%, and 6% in terms of balanced accuracy, F1-score, and Kappa, respectively.

Although direct comparisons with previous PSG studies are challenging due to differences in demographics, sample sizes, and classification criteria, our CNN classifier for NIRS performs well in distinguishing AS from QS. In this regard, some related previous research can be mentioned and compared. For instance, Sentner et al. [16] integrated physiological parameters such as HR, RR, and SpO2 derived from PSG recordings to delineate sleep stages among 39 preterm infants, achieving an area under the Receiver Operating

Characteristic curve (AUC-ROC) ranging from 0.61 to 0.78. Similarly, Werth et al. [17] leveraged features extracted from electrocardiography to attain a Kappa score of 0.43 in the classification of AS versus QS within a cohort of 34 preterm infants. Several studies exclusively concentrated on EEG and amalgamated diverse EEG-derived features for sleep stage classification. For instance, Hermans et al. [18] devised a CNN classifier that amalgamated multiple EEG-derived features to distinguish between sleep stages, achieving a Kappa score spanning from 0.33 to 0.44 among fifteen term infants. Similarly, Wang et al. [19] exclusively utilized EEG features, achieving an AUC-ROC of 0.75 in their study involving 17 preterm infants. Koolen et al. [20] introduced a SVM classifier trained on EEG features, demonstrating an accuracy and balanced accuracy of 85% across a cohort of 67 infants, while Ansari et al. [21] reported average accuracy, balanced accuracy, and AUC-ROC values of 79%, 79.5%, and 87%, respectively, in their utilization of exclusive EEG features to discriminate between QS and non-QS states among 97 preterm infants. Comparing these outcomes with the aforementioned results of our study, we conclude that it either outperforms or shows comparable results to existing methods.

It is essential to acknowledge several limitations of the current study, which warrant attention in future research endeavors. First, the extraction of HR and RR in this study relied upon algorithms outlined in [28,29], with measurements obtained at 7.5-second intervals. Notably, if sleep stages exert an influence on the short-term dynamics of HR and RR, the 7.5-second interval for computation could potentially lead to the omission of critical sleep-related information, thereby potentially impinging upon the classifier's efficacy. Nevertheless, the promising results obtained in this study confirm that the computation times for HR and RR have been appropriate for achieving precise sleep classification (AS vs. QS) in the population studied. Second, despite our assessment involving 723 one-minute trials derived from a cohort of 9 subjects in a highly challenging intensive care setting, validation in a larger and more diverse patient population needs to be performed. Third, the current research was centered on offline feature extraction and classification procedures, utilizing data sourced from preterm infants in a NICU. However, for practical clinical implementation, an online bedside version for feature extraction and classification is needed. This crucial step would serve as a litmus test to evaluate the applicability and performance of our proposed sleep classification approach within a genuine clinical environment. To progress in this direction, a key focus would involve refining the existing RR and HR algorithms to adapt them for real-time, online processing [28,29]. By incorporating the preprocessing and feature extraction procedures outlined in the paper into the online system, the foundation for accurate data analysis can be established. Once the online algorithms are in place, integrating the trained classifier into this setup and then developing practical software would provide clinicians with a powerful tool for real-time sleep pattern analysis in clinical environments.

Exploring potential future avenues for extending this study can open up a multitude of potential applications across diverse settings. Foremost, it could be used in hospital settings. Healthcare providers stand to benefit from enhanced sleep monitoring capabilities, enabling more precise and efficient diagnosis of sleep disorders. The simplicity of using a single head-mounted sensor in our portable system has the additional advantage of reducing potential discomfort and distress that can arise from the placement of multiple sensors typically required for PSG in preterm infants [45]. This would facilitate the planning of elective care and thereby improve sleep (quality). Furthermore, NIRS-based sleep assessment could be used for more precise and comprehensive home-based sleep stage classification [46–48]. Given the increasing popularity of home-based sleep monitoring, while considering the limitations of

current sensor modalities [49], our NIRS-based method offers a potent alternative. Sleep analysis within the comfort of one's own home holds great advantages, not only in terms of logistics (transportation to and from the hospital) but also in the assessment of sleep in a patient's daily environment. This could facilitate proactive health management [50,51]. In addition, the portability and user-friendliness would enable sleep assessment in environments with limited (monitoring) resources, which could significantly augment the quality of care in such environments. Lastly, our approach introduces exciting prospects for prospective research studies in the domain of neurodevelopmental assessment, complemented by concurrent sleep monitoring [52,53]. This innovative approach holds the potential for shedding light on the intricate relationship between sleep patterns and neurodevelopmental outcomes, thus enriching our understanding of this critical aspect of neonatal health and (brain)development.

## 5. Conclusion

Our study demonstrates the viability of accurately categorizing active sleep and quiet sleep in preterm infants within an intensive care setting using a single high sampling rate NIRS sensor and a sophisticated feature extraction approach and CNN classifier. Through rigorous quantitative analyses and cross-validation, our approach consistently outperforms in terms of balanced accuracy, F1-score, and Kappa when compared to the benchmark classifiers. Our findings offer a simplified, non-invasive, and user-friendly alternative to traditional PSG, reducing discomfort for preterm infants and facilitating more precise and efficient diagnosis of sleep disorders.

## References

1. Moutquin, J.-M. Classification and Heterogeneity of Preterm Birth. *BJOG* **2003**, *110*, 30–33, doi:10.1046/J.1471-0528.2003.00021.X.
2. Kostović, I.; Sedmak, G.; Judaš, M. Neural Histology and Neurogenesis of the Human Fetal and Infant Brain. *Neuroimage* **2019**, *188*, 743–773, doi:10.1016/J.NEUROIMAGE.2018.12.043.
3. Del Rio-Bermudez, C.; Kim, J.; Sokoloff, G.; Blumberg, M.S. Active Sleep Promotes Coherent Oscillatory Activity in the Cortico-Hippocampal System of Infant Rats. *Cereb Cortex* **2020**, *30*, 2070–2082, doi:10.1093/CERCOR/BHZ223.
4. Graven, S.N.; Browne, J. V. Sleep and Brain Development: The Critical Role of Sleep in Fetal and Early Neonatal Brain Development. *Newborn and Infant Nursing Reviews* **2008**, *8*, 173–179, doi:10.1053/J.NAINR.2008.10.008.
5. Tham, E.K.H.; Schneider, N.; Broekman, B.F.P. Infant Sleep and Its Relation with Cognition and Growth: A Narrative Review. *Nat Sci Sleep* **2017**, *9*, 135, doi:10.2147/NSS.S125992.
6. Yang, F.N.; Xie, W.; Wang, Z. Effects of Sleep Duration on Neurocognitive Development in Early Adolescents in the USA: A Propensity Score Matched, Longitudinal, Observational Study. *Lancet Child Adolesc Health* **2022**, *6*, 705–712, doi:10.1016/S2352-4642(22)00188-2.
7. Petty, J.; van den Hoogen, A. Brain Development, Promoting Sleep and Well-Being in the Context of Neonatal Developmental Care. In *Neonatal Nursing: A Global Perspective*; Springer International Publishing: Cham, 2022; pp. 135–149.
8. Besedovsky, L.; Lange, T.; Born, J. Sleep and Immune Function. *Pflugers Arch* **2012**, *463*, 121–137, doi:10.1007/S00424-011-1044-0/FIGURES/4.
9. Scher, M.S.; Loparo, K.A. Neonatal EEG/Sleep State Analyses: A Complex Phenotype of Developmental Neural Plasticity. *Dev Neurosci* **2009**, *31*, 259–275, doi:10.1159/000216537.
10. Allen, K.A. Promoting and Protecting Infant Sleep. *Adv Neonatal Care* **2012**, *12*, 288, doi:10.1097/ANC.0B013E3182653899.
11. Colombo, G.; De Bon, G. Strategies to Protect Sleep. *J Matern Fetal Neonatal Med* **2011**, *24 Suppl 1*, 30–31, doi:10.3109/14767058.2011.607571.
12. de Groot, E.R.; Bik, A.; Sam, C.; Wang, X.; Shellhaas, R.A.; Austin, T.; Tataranno, M.L.; Benders, M.J.N.L.; van den Hoogen, A.; Dudink, J. Creating an Optimal Observational Sleep Stage Classification System for Very and Extremely Preterm Infants. *Sleep Med* **2022**, *90*, 167–175, doi:10.1016/J.SLEEP.2022.01.020.
13. Knoop, M.S.; de Groot, E.R.; Dudink, J. Current Ideas about the Roles of Rapid Eye Movement and Non-Rapid Eye Movement Sleep in Brain Development. *Acta Paediatr* **2021**, *110*, 36–44, doi:10.1111/APA.15485.

14. Werth, J.; Atallah, L.; Andriessen, P.; Long, X.; Zwartkruis-Pelgrim, E.; Aarts, R.M. Unobtrusive Sleep State Measurements in Preterm Infants – A Review. *Sleep Med Rev* **2017**, *32*, 109–122, doi:10.1016/J.SMRV.2016.03.005.
15. de Groot, E.R.; Knoop, M.S.; van den Hoogen, A.; Wang, X.; Long, X.; Pillen, S.; Benders, M.; Dudink, J. The Value of Cardiorespiratory Parameters for Sleep State Classification in Preterm Infants: A Systematic Review. *Sleep Med Rev* **2021**, *58*, doi:10.1016/J.SMRV.2021.101462.
16. Sentner, T.; Wang, X.; De Groot, E.R.; Van Schaijk, L.; Tataranno, M.L.; Vijlbrief, D.C.; Benders, M.J.N.L.; Bartels, R.; Dudink, J. The Sleep Well Baby Project: An Automated Real-Time Sleep-Wake State Prediction Algorithm in Preterm Infants. *Sleep* **2022**, *45*, doi:10.1093/SLEEP/ZSAC143.
17. Werth, J.; Radha, M.; Andriessen, P.; Aarts, R.M.; Long, X. Deep Learning Approach for ECG-Based Automatic Sleep State Classification in Preterm Infants. *Biomed Signal Process Control* **2020**, *56*, 101663, doi:10.1016/J.BSPC.2019.101663.
18. Hermans, T.; Thewissen, L.; Gewillig, M.; Cools, B.; Jansen, K.; Pillay, K.; De Vos, M.; Van Huffel, S.; Naulaers, G.; Dereymaeker, A. Functional Brain Maturation and Sleep Organisation in Neonates with Congenital Heart Disease. *European Journal of Paediatric Neurology* **2022**, *36*, 115–122, doi:10.1016/j.ejpn.2021.12.008.
19. Wang, X.; Bik, A.; de Groot, E.R.; Tataranno, M.L.; Benders, M.J.N.L.; Dudink, J. Feasibility of Automated Early Postnatal Sleep Staging in Extremely and Very Preterm Neonates Using Dual-Channel EEG. *Clinical Neurophysiology* **2023**, *146*, 55–64, doi:10.1016/j.clinph.2022.11.018.
20. Koolen, N.; Oberdorfer, L.; Rona, Z.; Giordano, V.; Werther, T.; Klebermass-Schrehof, K.; Stevenson, N.; Vanhatalo, S. Automated Classification of Neonatal Sleep States Using EEG. *Clin Neurophysiol* **2017**, *128*, 1100–1108, doi:10.1016/J.CLINPH.2017.02.025.
21. Ansari, A.H.; De Wel, O.; Lavanga, M.; Caicedo, A.; Dereymaeker, A.; Jansen, K.; Vervisch, J.; De Vos, M.; Naulaers, G.; Van Huffel, S. Quiet Sleep Detection in Preterm Infants Using Deep Convolutional Neural Networks. *J Neural Eng* **2018**, *15*, doi:10.1088/1741-2552/AADC1F.
22. Dereymaeker, A.; Pillay, K.; Vervisch, J.; Van Huffel, S.; Naulaers, G.; Jansen, K.; De Vos, M. An Automated Quiet Sleep Detection Approach in Preterm Infants as a Gateway to Assess Brain Maturation. *Int J Neural Syst* **2017**, *27*, doi:10.1142/S012906571750023X.
23. Cordeiro, M.; Peinado, H.; Montes, M.T.; Valverde, E. Evaluation of the Suitability and Clinical Applicability of Different Electrodes for AEEG/CEEG Monitoring in the Extremely Premature Infant. *An Pediatr (Engl Ed)* **2021**, *95*, 423–430, doi:10.1016/J.ANPEDE.2020.09.010.
24. Costa, F.G.; Hakimi, N.; Van Bel, F. Neuroprotection of the Perinatal Brain by Early Information of Cerebral Oxygenation and Perfusion Patterns. *Int J Mol Sci* **2021**, *22*, doi:10.3390/IJMS22105389.
25. Hakimi, N.; Kamaledin Setarehdan, S. Stress Assessment by Means of Heart Rate Derived from Functional Near-Infrared Spectroscopy.

- <https://doi.org/10.1117/1.JBO.23.11.115001> **2018**, *23*, 115001, doi:10.1117/1.JBO.23.11.115001.
26. Hakimi, N.; Shahbakhti, M.; Sappia, S.; Horschig, J.M.; Bronkhorst, M.; Floor-Westerdijk, M.; Valenza, G.; Dudink, J.; Colier, W.N.J.M. Estimation of Respiratory Rate from Functional Near-Infrared Spectroscopy (FNIRS): A New Perspective on Respiratory Interference. *Biosensors (Basel)* **2022**, *12*, doi:10.3390/BIOS12121170.
  27. Shahbakhti, M.; Hakimi, N.; Horschig, J.M.; Floor-Westerdijk, M.; Claassen, J.; Colier, W.N.J.M. Estimation of Respiratory Rate during Biking with a Single Sensor Functional Near-Infrared Spectroscopy (FNIRS) System. *Sensors* **2023**, *Vol. 23*, Page 3632 **2023**, *23*, 3632, doi:10.3390/S23073632.
  28. Hakimi, N.; Shahbakhti, M.; Horschig, J.M.; Alderliesten, T.; Van Bel, F.; Colier, W.N.J.M.; Dudink, J. Respiratory Rate Extraction from Neonatal Near-Infrared Spectroscopy Signals. *Sensors* **2023**, *Vol. 23*, Page 4487 **2023**, *23*, 4487, doi:10.3390/S23094487.
  29. Hakimi, N.; Horschig, J.M.; Alderliesten, T.; Bronkhorst, M.; Floor-Westerdijk, M.J.; Bel, F. V.; Colier, W.N.J.M.; Dudink, J. Heart Rate Extraction from Neonatal Near-Infrared Spectroscopy Signals. *IEEE Trans Instrum Meas* **2023**, doi:10.1109/TIM.2023.3295015.
  30. Papile, L.A.; Burstein, J.; Burstein, R.; Koffler, H. Incidence and Evolution of Subependymal and Intraventricular Hemorrhage: A Study of Infants with Birth Weights Less than 1,500 Gm. *J Pediatr* **1978**, *92*, 529–534, doi:10.1016/S0022-3476(78)80282-0.
  31. Delpy, D.T.; Cope, M.; Van Der Zee, P.; Arridge, S.; Wray, S.; Wyatt, J. Estimation of Optical Pathlength through Tissue from Direct Time of Flight Measurement. *Phys Med Biol* **1988**, *33*, 1433, doi:10.1088/0031-9155/33/12/008.
  32. Sappia, M.S.; Hakimi, N.; Colier, W.N.J.M.; Horschig, J.M. Signal Quality Index: An Algorithm for Quantitative Assessment of Functional near Infrared Spectroscopy Signal Quality. *Biomedical Optics Express*, *Vol. 11*, Issue 11, pp. 6732-6754 **2020**, *11*, 6732–6754, doi:10.1364/BOE.409317.
  33. Jahani, S.; Setarehdan, S.K.; Boas, D.A.; Yücel, M.A. Motion Artifact Detection and Correction in Functional Near-Infrared Spectroscopy: A New Hybrid Method Based on Spline Interpolation Method and Savitzky–Golay Filtering. <https://doi.org/10.1117/1.NPh.5.1.015003> **2018**, *5*, 015003, doi:10.1117/1.NPH.5.1.015003.
  34. Piper, S.K.; Krueger, A.; Koch, S.P.; Mehnert, J.; Habermehl, C.; Steinbrink, J.; Obrig, H.; Schmitz, C.H. A Wearable Multi-Channel FNIRS System for Brain Imaging in Freely Moving Subjects. *Neuroimage* **2014**, *85*, 64–71, doi:10.1016/J.NEUROIMAGE.2013.06.062.
  35. Scholkmann, F.; Spichtig, S.; Muehlemann, T.; Wolf, M. How to Detect and Reduce Movement Artifacts in Near-Infrared Imaging Using Moving Standard Deviation and Spline Interpolation. *Physiol Meas* **2010**, *31*, 649, doi:10.1088/0967-3334/31/5/004.
  36. Morais, G.A.Z.; Scholkmann, F.; Balardin, J.B.; Furucho, R.A.; Paula, R.C.V. de; Claudinei Eduardo Biazoli, Jr.; Sato, J.R. Non-Neuronal Evoked and Spontaneous

- Hemodynamic Changes in the Anterior Temporal Region of the Human Head May Lead to Misinterpretations of Functional near-Infrared Spectroscopy Signals. *Neurophotonics* **2018**, *5*, 1, doi:10.1117/1.NPH.5.1.011002.
37. Cui, X.; Bray, S.; Reiss, A.L. Functional near Infrared Spectroscopy (NIRS) Signal Improvement Based on Negative Correlation between Oxygenated and Deoxygenated Hemoglobin Dynamics. *Neuroimage* **2010**, *49*, 3039–3046, doi:10.1016/J.NEUROIMAGE.2009.11.050.
  38. Li, B.; Liang, H.; Sun, L.; Wei, Q.; He, L.; Xu, P. Review on Studies of Machine Learning Algorithms. *J Phys Conf Ser* **2019**, *1187*, 052103, doi:10.1088/1742-6596/1187/5/052103.
  39. Pedregosa FABIANPEDREGOSA, F.; Michel, V.; Grisel OLIVIERGRISEL, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Vanderplas, J.; Cournapeau, D.; Pedregosa, F.; Varoquaux, G.; et al. Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.
  40. Denenberg, V.H.; Thoman, E.B. Evidence for a Functional Role for Active (REM) Sleep in Infancy. *Sleep* **1981**, *4*, 185–191, doi:10.1093/SLEEP/4.2.185.
  41. Roffwarg, H.P.; Muzio, J.N.; Dement, W.C. Ontogenetic Development of the Human Sleep-Dream Cycle. *Science (1979)* **1966**, *152*, 604–619, doi:10.1126/SCIENCE.152.3722.604/ASSET/A7B90230-BE4A-472F-A59A-A8DCFB61DD69/ASSETS/SCIENCE.152.3722.604.FP.PNG.
  42. Tarullo, A.R.; Balsam, P.D.; Fifer, W.P. Sleep and Infant Learning. *Infant Child Dev* **2011**, *20*, 35–46, doi:10.1002/ICD.685/FULL.
  43. Altimier, L.; Phillips, R. The Neonatal Integrative Developmental Care Model: Advanced Clinical Applications of the Seven Core Measures for Neuroprotective Family-Centered Developmental Care. *Newborn and Infant Nursing Reviews* **2016**, *16*, 230–244, doi:10.1053/J.NAINR.2016.09.030.
  44. Holditch-Davis, D.; Scher, M.; Schwartz, T.; Hudson-Barr, D. Sleeping and Waking State Development in Preterm Infants. *Early Hum Dev* **2004**, *80*, 43–64, doi:10.1016/J.EARLHUMDEV.2004.05.006.
  45. Afsar, F.S. Skin Care for Preterm and Term Neonates. *Clin Exp Dermatol* **2009**, *34*, 855–858, doi:10.1111/J.1365-2230.2009.03424.X.
  46. Didden, R.; Curfs, L.M.G.; Van Driel, S.; De Moor, J.M.H. Sleep Problems in Children and Young Adults with Developmental Disabilities: Home-Based Functional Assessment and Treatment. *J Behav Ther Exp Psychiatry* **2002**, *33*, 49–58, doi:10.1016/S0005-7916(02)00012-5.
  47. Patriquin, M.A.; Mellman, T.A.; Glaze, D.G.; Alfano, C.A. Polysomnographic Sleep Characteristics of Generally-Anxious and Healthy Children Assessed in the Home Environment. *J Affect Disord* **2014**, *161*, 79–83, doi:10.1016/J.JAD.2014.02.037.
  48. McCabe, S.M.; Blackmore, A.M.; Abbiss, C.R.; Langdon, K.; Elliott, C. Sleep Concerns in Children and Young People with Cerebral Palsy in Their Home Setting. *J Paediatr Child Health* **2015**, *51*, 1188–1194, doi:10.1111/JPC.12933.



49. Dereymaeker, A.; Pillay, K.; Vervisch, J.; De Vos, M.; Van Huffel, S.; Jansen, K.; Naulaers, G. Review of Sleep-EEG in Preterm and Term Neonates. *Early Hum Dev* **2017**, *113*, 87–103, doi:10.1016/J.EARLHUMDEV.2017.07.003.
50. Kahn, A.; Sottiaux, M.; Appelboom-Fondu, J.; Blum, D.; Rebuffat, E.; Levitt, J. Long-Term Development of Children Monitored as Infants for an Apparent Life-Threatening Event During Sleep: A 10-Year Follow-up Study. *Pediatrics* **1989**, *83*, 668–673, doi:10.1542/PEDS.83.5.668.
51. Thoman, E.B.; Whitney, M.P. Sleep States of Infants Monitored in the Home: Individual Differences, Developmental Trends, and Origins of Diurnal Cyclicity. *Infant Behav Dev* **1989**, *12*, 59–75, doi:10.1016/0163-6383(89)90053-2.
52. El-Dib, M.; Massaro, A.N.; Glass, P.; Aly, H. Neurodevelopmental Assessment of the Newborn: An Opportunity for Prediction of Outcome. *Brain Dev* **2011**, *33*, 95–105, doi:10.1016/J.BRAINDEV.2010.04.004.
53. Blackmer, A.B.; Feinstein, J.A. Management of Sleep Disorders in Children With Neurodevelopmental Disorders: A Review. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy* **2016**, *36*, 84–98, doi:10.1002/PHAR.1686.



# CHAPTER 10

## Discussion and Conclusion

General discussion

Conclusions

Future directions





# 1. General discussion

Using near-infrared spectroscopy (NIRS) to measure brain and tissue oxygen levels has become an important surveillance tool in neonatal perioperative and intensive care settings [1–5]. Although there are some challenges when using NIRS, such as systemic artifacts [6], these obstacles can also provide opportunities for improvement [7]. In particular, physiological artifacts, which are often seen as hindrances, can actually be helpful by providing a way to evaluate the quality of the signal [8,9]. Utilizing this approach can aid in identifying optimal sensor placement for achieving the highest signal quality, thereby enhancing the precision and reliability of data collection. Moreover, signals resulting from cardiac and respiratory activity can be analyzed to extract heart rate (HR) and respiratory rate (RR) measurements [7,10–14]. This approach, which involves analyzing these physiological signals to derive vital measurements, effectively transforms potential disturbances into crucial health data.

The noninvasive nature of NIRS stands out particularly for its gentle approach on the delicate skin of neonates, potentially eliminating the need for adhesive sensors [15–17]. This method not only minimizes discomfort and stress associated with the routine application and removal of glue based sensors but also underscores the strategy of utilizing inherent challenges to refine the accuracy of signal quality assessments, along with HR and RR measurements. This thesis endeavors to shed new light on the methodologies of NIRS data collection within clinical environments, by harnessing physiological artifacts for a dual purpose of improving data quality and extracting vital signs.

The research initially focused on adult subjects in Part I, as adult data is more accessible and has a higher signal-to-noise ratio than neonatal data, especially in clinical scenarios. The insights gained from adult data were instrumental in developing algorithms that were then adapted and applied to neonatal data. Part II of the thesis presents the application of these algorithms to data collected from neonates in clinical settings, demonstrating the adaptability and relevance of the initial findings to the more sensitive neonatal group.

## **Part I: physiological information in adult NIRS**

In the first part of the thesis, **chapter two** focuses on the development of an algorithm for assessing signal quality in NIRS data. The newly developed Signal Quality Index (SQI) algorithm is capable of quantitatively evaluating NIRS signals on a numeric scale from 1, indicating very low quality, to 5, representing very high quality. This is the first algorithm of its kind to provide more than two levels of signal quality assessment, unlike existing algorithms such as Scalp Coupling Index (SCI) [8] and PHOEBE (placing headgear optodes efficiently before experimentation) [9], which only offer a binary good/bad evaluation. The SQI algorithm has demonstrated significantly better performance in quantitatively rating NIRS signal quality compared to PHOEBE and SCI, with a higher positive correlation to expert ratings and improved binary classification in terms of accuracy, sensitivity, and F1-score. Unlike SCI, which uses a single metric, or PHOEBE, which combines two metrics, SQI employs a three-stage rating process. This process includes identifying very low and very high-quality signals, followed by a comprehensive signal quality rating, using selected features that translate visual assessments into numeric ratings. The algorithm's efficacy is evident in its ability to identify signal quality with high precision, and its design allows for both offline and online application over 10-second sliding windows. This enables the assessment of signal quality variations over time and the exclusion of poor quality trials. The use of SQI ensures that the quality of NIRS

data is not dependent on the researcher's experience, thereby facilitating the collection of reliable hemodynamic information and reducing artifacts.

The results from the SQI algorithm show promise, but improvements are possible. Our study didn't address motion-affected signal segments, limiting our analysis to motion-artifact-free data. Motion artifacts, primarily caused by head and body movements, introduce unpredictable, non-stationary noise, making them challenging to filter out without losing critical information. Furthermore, the algorithm's training and validation were only performed on data from healthy young adults, which overlooks significant hemodynamic variations related to age and systemic conditions. For instance, aging could alter cerebral blood flow in association and limbic cortices during rest [18,19]. Studies using functional NIRS have shown that elderly individuals have a reduced hemodynamic response functional activation [20,21]. Additionally, cardiac abnormalities such as congenital heart disease could lead to altered cerebral hemodynamics, including lower cerebral blood flow and increased oxygen extraction in neonates [22]. These age- and condition-related changes can significantly affect NIRS signals. Future studies should expand the dataset to include infants, elderly, and individuals with cardiac abnormalities to enhance the algorithm's applicability and accuracy.

Continuing on signal quality assessment, in **chapter three**, we leveraged machine learning to refine the assessment of signal quality in NIRS. Our developed algorithm, Machine Learning-based SQI (MLSQI), quantitatively evaluates NIRS signal quality on a scale ranging from 1 for very low quality to 5 for very high quality, similar to SQI algorithm. It operates by analyzing 14 temporal and morphological features from the NIRS signals using a medium Gaussian SVM regressor. When compared to three advanced algorithms—SCI, PHOEBE, and SQI—MLSQI demonstrated superior performance in both quantitative and binary evaluations of NIRS signal quality. The study also revisited the extraction of various features from NIRS signals that translate the heuristic visual assessment of signal quality into structured analysis. The cardiac pulses' morphological and temporal characteristics within the NIRS signals provide insights into their quality. Although summarizing this information into a single metric would be ideal, our study found that using a greater number of features significantly improves algorithm performance. Additionally, we proposed the MLSQI algorithm to assess shorter, 4-second signal segments, offering quicker feedback compared to the previously proposed SQI algorithm that assessed 10-second segments.

In **chapter four**, we explored the potential of utilizing respiratory-related physiological artifacts in NIRS data for adults. Our investigation centered on the feasibility of estimating RR from NIRS. This approach represents a novel perspective on the interpretation of respiratory interference in NIRS data, which is traditionally viewed as an unwanted disturbance [23]. The significance of our study lies in its ability to extract an additional metric, namely RR, from cerebral activity analysis without the need for a reference signal.

Our study's promising findings are tempered by notable limitations that necessitate further exploration. Our data collection, restricted to a simple seated breathing task, didn't capture a broad spectrum of RRs, limiting our analysis. NIRS and respiratory signals were concurrently collected from subjects during controlled breathing tasks at a constant rate from 0.1 Hz to 0.4 Hz. This controlled setting, while useful for consistency, does not reflect the variability encountered in more dynamic, real-world scenarios. Future research should incorporate dynamic activities to obtain a more representative range of RRs, as discussed in the subsequent chapter. Additionally, the study's reliability was also only verified in young, healthy subjects, underscoring the need for validation across diverse groups, such as the elderly,

neonates, and patients. This is particularly crucial because age-related changes in physiological mechanisms, like the reduction in respiratory sinus arrhythmia and chest wall expansion, can affect respiratory signal quality and consequently the performance of respiratory rate estimation algorithms [24].

Progressing with RR estimation from NIRS in adults, in **chapter five**, we addressed the challenge of data corrupted by motion artifacts. We introduced a robust method for RR estimation during physical activities, such as bicycling, using NIRS signals captured with portable devices. This method leverages the reduced sensitivity of NIRS to subject movement, aiming to overcome the practical constraints identified in contemporary studies of RR estimation during exercise.

Despite the promising outcomes of our method, several limitations highlight areas for future improvement. Testing on subjects engaging in more vigorous exercises beyond the current study's scope could also validate its effectiveness under diverse physical conditions. Additionally, exploring advanced fusion strategies, such as the Smart Fusion method proposed in [25], might improve performance, despite minimal observed differences in estimated RRs across modalities. This underscores the need for broader and more intensive testing to refine the RR estimation method.

## **Part II: physiological information in neonatal NIRS**

Transitioning from adult to neonatal data, in **chapter seven**, we applied our acquired knowledge and developed algorithms to create a new method for HR extraction from NIRS signals of neonates admitted to the neonatal intensive care unit (NICU). To our knowledge, our NHR (NIRS HR) algorithm, developed here, is the first HR extraction method tailored to neonatal NIRS data collected in clinical settings. In contrast, the existing automatic multiscale-based peak detection (AMPD) algorithm, introduced by Scholkmann et al. [26], was validated on signals from healthy adults in controlled environments. Similarly, the Perdue algorithm was designed for signals from infants over three months old, also in controlled settings [11]. Our algorithm, however, is specifically developed for signals from (pre)term neonates in NICUs. Our findings demonstrate that the NHR algorithm significantly outperforms ( $p < 0.05$ ) the Perdue ( $r = 71.8\%$ ) and AMPD ( $r = 73.2\%$ ) algorithms in HR extraction from neonatal NIRS signals, as evidenced by a higher agreement with a greater correlation coefficient ( $r = 99.5\%$ ). Despite these advancements, the NHR algorithm is constrained by a 12.5-second delay in updates, which may limit its application in scenarios requiring analysis of rapid physiological changes.

Building upon the use of physiological information in neonatal NIRS, **chapter eight** introduces the development of the NRR (NIRS RR) algorithm. This innovative algorithm is designed to extract RR from neonatal NIRS signals obtained within the NICU. The NRR algorithm is specifically tailored for clinical neonatal data. Unlike previous methods validated on adult data during resting state in controlled settings [14], the NRR algorithm has been validated on neonatal data in the active and variable conditions of a NICU. The algorithm's performance was evaluated, showing a high agreement and strong linear correlation ( $r = 84.5\%$ ) between the reference and extracted RRs. However, it similarly suffers from a limitation of a 7.5-second delay in updates, restricting its effectiveness in analyzing quick, transient physiological shifts.

The process of data acquisition in clinical environments presents unique challenges that necessitated innovative modifications to the NHR and NRR algorithms. Unlike controlled settings, clinical data often contain a higher incidence of motion artifacts due to factors such

as patient movement, sensor adjustments, and ventilator operations. To address this, we implemented specific steps to detect and correct motion artifacts on a segment-by-segment basis. The algorithm selectively excludes signal segments where motion artifacts are present in the majority of the signal segment. On average, this process excluded approximately 6% of the signal segments (see Table 1 in Chapter 8). Moreover, the inherent low quality of clinical data, typically resulting from electromagnetic interference and limitations in data acquisition in such settings, posed another obstacle. To mitigate this, we introduced a processing step that involves selecting the highest quality NIRS channel, evaluating the signal quality for each segment, and incorporating only those signals deemed of sufficient quality into our analysis. Variability in the signal source, primarily due to potentially greater physiological changes in clinical settings compared to controlled environments, is yet another challenge encountered in clinical data. To effectively handle this variability, the NHR and NRR algorithms adaptively adjust the HR and RR frequency bandwidths for each measurement.

While the NHR and NRR algorithms currently exhibit delays in updates—12.5 seconds for HR and 7.5 seconds for RR—these limitations open avenues for future enhancements. The existing framework of both algorithms provides a strong foundation for adaptation towards real-time, online usage. For the NHR algorithm, real-time adaptation can be achieved by computing SQI from the beginning of the measurement up to the current moment at several-minute intervals and storing segment SQI values for reuse. Additionally, the adaptive HR frequency bandwidth can be recalculated using recent O<sub>2</sub>Hb signals. For the NRR algorithm, the signal quality could be assessed in each 30-s window, and the HR frequency bandwidth could be updated during the measurement every 10 min. Such advancements could significantly increase their clinical utility by enabling dynamic signal quality assessments and adaptive frequency adjustments based on real-time data.

In **chapter nine**, a relevant application of signal quality assessment, along with HR and RR extraction from neonatal NIRS was demonstrated in the context of sleep stage assessment in NICU-admitted neonates. Utilizing multi-modal feature extraction from NIRS signals at an enhanced sampling rate of 100 Hz, we assessed the capability of these signals to characterize sleep stages in preterm neonates using a solitary NIRS sensor. Employing a convolutional neural network classifier, we distinguished between active sleep and quiet sleep by integrating various features derived from the high-frequency NIRS data. These features encompass physiological signals like HR and RR, motion-related information, and the raw NIRS signals themselves. Our comprehensive approach facilitated the evaluation of NIRS signal quality and the extraction of a broad spectrum of physiological and non-physiological features. This novel combination of features, previously underexplored, allowed for the creation of an inclusive framework for analyzing sleep patterns in preterm neonates as young as 24 weeks gestational age or with birth weights as minimal as 500 grams, using only data from a single sensor.

This study's limitations include the reliance on algorithms that compute HR and RR at 7.5-second intervals, which may overlook short-term dynamics affected by sleep stages, potentially affecting classifier performance for sleep classification [27–29]. Additionally, while the study's validation involved trials from a limited number of (pre)term neonates in an intensive care setting, broader validation and the development of an online system for real-time feature extraction and classification are essential steps for practical clinical application.



## 2. Conclusions

- The SQI algorithm, developed for evaluation of NIRS signal quality, significantly enhances the ability to manage data integrity by allowing researchers to discard poor quality signals without losing valuable channels. Additionally, there is potential for adapting this algorithm for online use, as evidenced by similar functionalities in Artinis software.
- Implementation of the SQI algorithm could standardize the evaluation of NIRS signal quality, reducing dependency on the user's expertise and allowing for broader application across various research settings.
- The feasibility of extracting HR and RR from neonatal NIRS signals in NICU settings has been demonstrated. NIRS offers a non-invasive monitoring alternative that could potentially supplement traditional methods like electrocardiography and photoplethysmography, especially given its reduced susceptibility to motion artifacts [28].
- This thesis highlights the potential of NIRS technology to develop standalone systems capable of simultaneously monitoring multiple physiological parameters (cerebral oxygen saturation, HR, and RR) using a single device. Such systems that provide a more comprehensive view of a patient's physiological state with minimal discomfort and reduced risk of skin damage could revolutionize patient monitoring in clinical settings.
- Initial investigations into the use of NIRS for neonatal sleep monitoring suggest the possibility of tracking sleep stages by analyzing integrated physiological data from NIRS signals (e.g., HR, RR, and cerebral oxygen saturation). This method might offer a non-invasive and continuous way to monitor neonatal sleep patterns, potentially providing valuable insights into neurodevelopment and overall health [29–32], although further research is necessary to confirm its effectiveness.

## 3. Future directions

Future advancements in NIRS technology, by integrating SQI for signal quality evaluation, HR and RR extraction capabilities for neonatal monitoring, and methods for assessing sleep stages, present a significant opportunity to enhance the precision, reliability, and scope of use in both clinical and research settings. As these technologies are further refined, NIRS could potentially streamline the use of multiple monitoring devices. Moreover, integrating NIRS-derived measurements of HR, RR, and cerebral oxygenation can significantly improve multimodal monitoring in NICUs, enhancing care responsiveness and minimizing neonate stress.

### 3.1. Signal quality assessment

Improving signal quality assessment is paramount for advancing NIRS technology. As mentioned, our study excluded major motion-affected segments, limiting analysis to motion-artifact-free data. Future research should prioritize the development of algorithms that can accurately identify and mitigate major motion artifacts, which are a significant source of noise in NIRS signals [35,36]. This is especially important for applications involving infants [37], patients undergoing rehabilitation therapy [38,39], and sports science [40] where subject movement is inevitable. Integrating machine learning techniques to dynamically adapt to the presence of artifacts and improve signal fidelity in real-time could revolutionize signal quality assessment in NIRS systems. Additionally, the exploration of novel sensor designs and placement strategies to minimize the impact of motion and enhance optical coupling between the optodes and the scalp will further augment signal quality [41].

### 3.2. HR extraction

To improve HR extraction, achieving real-time beat-by-beat HR computation is a critical objective. In our study, we used a frequency domain-based approach to extract HR due to the low quality of our data, at the expense of reduced time resolution. However, with high-quality data, a time domain-based approach, such as peak detection (e.g., using AMPD), could provide beat-by-beat HR information. Reducing the delay in HR computation and increasing the algorithm's sensitivity to short-term HR dynamics are essential for applications in acute care settings and for monitoring conditions that require immediate response [42,43]. Future developments could include the implementation of adaptive algorithms that can adjust to the signal's quality in real-time and employ advanced signal processing techniques to enhance accuracy and reduce latency in HR detection.

### 3.3. RR extraction

Expanding the robustness and accuracy of RR extraction algorithms is another key direction. Future studies should aim at validating the performance of RR extraction algorithms in more dynamic and realistic scenarios, such as during intense physical activities [44] or in diverse clinical settings [45]. Additionally, the development of algorithms that can adapt to different cohorts, taking into account factors like age, race, skin color, and body mass index (BMI), will enhance the universality and reliability of RR measurement using NIRS technology.

### 3.4. Sleep assessment

In sleep assessment, refining algorithms for the detection of sleep stages and their influence on HR and RR dynamics is crucial. Future research should focus on reducing computation intervals to capture more granular sleep-related changes in physiological signals. Moreover, validating these algorithms in a broader range of subjects, including diverse patient populations in clinical settings, will enhance their applicability. Developing online versions of these algorithms for real-time monitoring and classification of sleep stages, both in clinical environments and for home-based healthcare, would provide valuable tools for healthcare professionals and enhance patient care in diverse settings.

Our approach, which utilizes NIRS technology to derive HR and RR, could potentially be employed as a noninvasive, real-time bedside automated sleep classification system for preterm infants, similar to the Sleep Well Baby algorithm developed by Sentner et al. [46]. This

innovative approach could eliminate the need for the extra electrodes that are standard in traditional polysomnography, thus simplifying routine sleep assessment and enabling continuous monitoring of sleep patterns in the NICU.

The practical implementation of NIRS systems equipped with advanced SQI, HR, RR, and sleep assessment features necessitates the development of user-friendly, robust, and portable devices that can deliver real-time feedback and monitoring in clinical and everyday settings. Such development requires interdisciplinary collaboration among engineers, clinicians, and end-users.

### 3.5. Integrated multimodal monitoring

Advancements in algorithms for extracting HR and RR from neonatal NIRS signals could potentially enhance multimodal monitoring capabilities in NICUs. By integrating HR and RR measurements with cerebral oxygenation data concurrently, NIRS technology can potentially reduce the number of sensors required. Thus it could potentially minimize stress on neonates [15,16], and improve the accuracy of managing conditions like hypoxic-ischemic encephalopathy [47–49], respiratory distress [50,51], and monitor cerebral vascular autoregulation. To assess cerebral vascular autoregulation, one method involves analyzing the relationship between blood pressure and NIRS-derived cerebral oxygen saturation [52], which serves as a surrogate for cerebral blood flow [53]. Integrating HR—a surrogate for blood pressure in preterm infants [54]—with cerebral oxygen saturation could allow for the assessment of cerebral vascular autoregulation using a standalone NIRS system in preterm infants.

This integrated approach using NIRS could allow for real-time, tailored adjustments to treatment protocols, enhancing the responsiveness of neonatal care. Evidence from the SafeBoosC Phase II trial shows that monitoring cerebral oxygen saturation (StO<sub>2</sub>) alongside cardiovascular and respiratory status—guided by specific treatment guidelines—effectively reduces the risks of hypo- and hyperoxia [55]. This approach could allow for immediate interventions and reduce neonate stress with fewer sensors. Additionally, it could provide valuable diagnostic information without relying on absolute set-off values, addressing a challenge in current StO<sub>2</sub> implementation in clinics.

Furthermore, integrating NIRS-derived data on HR, RR, and cerebral oxygenation—with its proven role as a surrogate for cerebral perfusion [53]—with other monitoring technologies like electroencephalography [56] and continuous transcranial Doppler (the NeoDoppler device) [57], can provide a more comprehensive evaluation of a neonate's cerebral health. This integration allows for real-time tracking of the balance between cerebral oxygen supply and demand, needed for detecting early signs of cerebral dysregulation [40]. Such an integrated approach underscores the potential of continuous, multimodal monitoring systems to enhance the responsiveness and effectiveness of individualized neonatal critical care.

## References

1. Yu, Y.; Zhang, K.; Zhang, L.; Zong, H.; Meng, L.; Han, R. Cerebral Near-infrared Spectroscopy (NIRS) for Perioperative Monitoring of Brain Oxygenation in Children and Adults. *Cochrane Database Syst Rev* **2018**, 2018, doi:10.1002/14651858.CD010947.PUB2.
2. Petrova, A.; Mehta, R. Near-Infrared Spectroscopy in the Detection of Regional Tissue Oxygenation during Hypoxic Events in Preterm Infants Undergoing Critical Care. *Pediatric Critical Care Medicine* **2006**, 7, 449–454, doi:10.1097/01.PCC.0000235248.70482.14.
3. Scheeren, T.W.L.; Schober, P.; Schwarte, L.A. Monitoring Tissue Oxygenation by near Infrared Spectroscopy (NIRS): Background and Current Applications. *J Clin Monit Comput* **2012**, 26, 279, doi:10.1007/S10877-012-9348-Y.
4. Van Bel, F.; Lemmers, P.; Naulaers, G. Monitoring Neonatal Regional Cerebral Oxygen Saturation in Clinical Practice: Value and Pitfalls. *Neonatology* **2008**, 94, 237–244, doi:10.1159/000151642.
5. Chock, V.Y.; Vesoulis, Z.A.; El-Dib, M.; Austin, T.; van Bel, F. The Future of Neonatal Cerebral Oxygenation Monitoring: Directions After the SafeBoosC-III Trial. *J Pediatr* **2024**, 270, 114016, doi:10.1016/J.JPEDI.2024.114016.
6. Tachtsidis, I.; Scholkmann, F. False Positives and False Negatives in Functional Near-Infrared Spectroscopy: Issues, Challenges, and the Way Forward. <https://doi.org/10.1117/1.NPh.3.3.031405> **2016**, 3, 031405, doi:10.1117/1.NPH.3.3.031405.
7. Hakimi, N.; Kamaledin Setarehdan, S. Stress Assessment by Means of Heart Rate Derived from Functional Near-Infrared Spectroscopy. <https://doi.org/10.1117/1.JBO.23.11.115001> **2018**, 23, 115001, doi:10.1117/1.JBO.23.11.115001.
8. Pollonini, L.; Olds, C.; Abaya, H.; Bortfeld, H.; Beauchamp, M.S.; Oghalai, J.S. Auditory Cortex Activation to Natural Speech and Simulated Cochlear Implant Speech Measured with Functional Near-Infrared Spectroscopy. *Hear Res* **2014**, 309, 84–93, doi:10.1016/J.HEARES.2013.11.007.
9. Pollonini, L.; Bortfeld, H.; Oghalai, J.S. PHOEBE: A Method for Real Time Mapping of Optodes-Scalp Coupling in Functional near-Infrared Spectroscopy. *Biomedical Optics Express*, Vol. 7, Issue 12, pp. 5104-5119 **2016**, 7, 5104–5119, doi:10.1364/BOE.7.005104.
10. Hakimi, N.; Jodeiri, A.; Mirbagheri, M.; Setarehdan, S.K. Proposing a Convolutional Neural Network for Stress Assessment by Means of Derived Heart Rate from Functional near Infrared Spectroscopy. *Comput Biol Med* **2020**, 121, 103810, doi:10.1016/J.COMPBIOMED.2020.103810.
11. Perdue, K.L.; Westerlund, A.; McCormick, S.A.; Nelson, C.A.; Iii, C.A.N. Extraction of Heart Rate from Functional Near-Infrared Spectroscopy in Infants.

- <https://doi.org/10.1117/1.JBO.19.6.067010>      **2014**,      19,      067010,  
doi:10.1117/1.JBO.19.6.067010.
12. Perdue, K.L.; Edwards, L.A.; Tager-Flusberg, H.; Nelson, C.A. Differing Developmental Trajectories in Heart Rate Responses to Speech Stimuli in Infants at High and Low Risk for Autism Spectrum Disorder. *J Autism Dev Disord* **2017**, *47*, 2434–2442, doi:10.1007/S10803-017-3167-4/FIGURES/4.
  13. Holper, L.; Seifritz, E.; Scholkmann, F. Short-Term Pulse Rate Variability Is Better Characterized by Functional near-Infrared Spectroscopy than by Photoplethysmography. *J Biomed Opt* **2016**, *21*, 091308, doi:10.1117/1.JBO.21.9.091308.
  14. Izzetoglu, M.; Holtzer, R. Effects of Processing Methods on FNIRS Signals Assessed during Active Walking Tasks in Older Adults. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2020**, *28*, 699–709, doi:10.1109/TNSRE.2020.2970407.
  15. Kuller, J.M.M. Skin Breakdown: Risk Factors, Prevention, and Treatment. *Newborn and Infant Nursing Reviews* **2001**, *1*, 35–42, doi:10.1053/NBIN.2001.22874.
  16. Afsar, F.S. Skin Care for Preterm and Term Neonates. *Clin Exp Dermatol* **2009**, *34*, 855–858, doi:10.1111/J.1365-2230.2009.03424.X.
  17. Tarar, A.A.; Mohammad, U.; Srivastava, S.K. Wearable Skin Sensors and Their Challenges: A Review of Transdermal, Optical, and Mechanical Sensors. *Biosensors 2020, Vol. 10, Page 56* **2020**, *10*, 56, doi:10.3390/BIOS10060056.
  18. Martin, A.J.; Friston, K.J.; Colebatch, J.G.; Frackowiak, R.S.J. Decreases in Regional Cerebral Blood Flow with Normal Aging. <http://dx.doi.org/10.1038/jcbfm.1991.121> **1991**, *11*, 684–689, doi:10.1038/JCBFM.1991.121.
  19. Nakano, S.; Asada, T.; Matsuda, H.; Uno, M.; Takasaki, M. [Effects of Healthy Aging on the Regional Cerebral Blood Flow Measurements Using 99mTc-ECD SPECT Assessed with Statistical Parametric Mapping]. *Nihon Ronen Igakkai Zasshi* **2000**, *37*, 49–55, doi:10.3143/GERIATRICALS.37.49.
  20. Hock, C.; Müller-Spahn, F.; Schuh-Hofer, S.; Hofmann, M.; Dirnagl, U.; Villringer, A. Age Dependency of Changes in Cerebral Hemoglobin Oxygenation during Brain Activation: A Near-Infrared Spectroscopy Study. <http://dx.doi.org/10.1038/jcbfm.1995.137> **1995**, *15*, 1103–1108, doi:10.1038/JCBFM.1995.137.
  21. Schroeter, M.L.; Zysset, S.; Kruggel, F.; Von Cramon, D.Y. Age Dependency of the Hemodynamic Response as Measured by Functional Near-Infrared Spectroscopy. *Neuroimage* **2003**, *19*, 555–564, doi:10.1016/S1053-8119(03)00155-1.
  22. Cheng, H.H.; Ferradal, S.L.; Vyas, R.; Wigmore, D.; McDavitt, E.; Soul, J.S.; Franceschini, M.A.; Newburger, J.W.; Grant, P.E. Abnormalities in Cerebral Hemodynamics and Changes with Surgical Intervention in Neonates with Congenital Heart Disease. *J Thorac Cardiovasc Surg* **2020**, *159*, 2012, doi:10.1016/J.JTCVS.2019.08.045.

23. Nguyen, H.D.; Yoo, S.H.; Bhutta, M.R.; Hong, K.S. Adaptive Filtering of Physiological Noises in FNIRS Data. *Biomed Eng Online* **2018**, *17*, 1–23, doi:10.1186/S12938-018-0613-2/FIGURES/17.
24. Charlton, P.H.; Bonnici, T.; Tarassenko, L.; Alastruey, J.; Clifton, D.A.; Beale, R.; Watkinson, P.J. Extraction of Respiratory Signals from the Electrocardiogram and Photoplethysmogram: Technical and Physiological Determinants. *Physiol Meas* **2017**, *38*, 669, doi:10.1088/1361-6579/AA670E.
25. Karlen, W.; Raman, S.; Ansermino, J.M.; Dumont, G.A. Multiparameter Respiratory Rate Estimation from the Photoplethysmogram. *IEEE Trans Biomed Eng* **2013**, *60*, 1946–1953, doi:10.1109/TBME.2013.2246160.
26. Scholkmann, F.; Boss, J.; Wolf, M. An Efficient Algorithm for Automatic Peak Detection in Noisy Periodic and Quasi-Periodic Signals. *Algorithms* **2012**, *Vol. 5*, Pages 588–603 **2012**, *5*, 588–603, doi:10.3390/A5040588.
27. Yan, C.; Li, P.; Yang, M.; Li, Y.; Li, J.; Zhang, H.; Liu, C. Entropy Analysis of Heart Rate Variability in Different Sleep Stages. *Entropy* **2022**, *24*, doi:10.3390/E24030379.
28. Latremouille, S.; Lam, J.; Shalish, W.; Sant’Anna, G. Neonatal Heart Rate Variability: A Contemporary Scoping Review of Analysis Methods and Clinical Applications. *BMJ Open* **2021**, *11*, e055209, doi:10.1136/BMJOPEN-2021-055209.
29. de Groot, E.R.; Knoop, M.S.; van den Hoogen, A.; Wang, X.; Long, X.; Pillen, S.; Benders, M.; Dudink, J. The Value of Cardiorespiratory Parameters for Sleep State Classification in Preterm Infants: A Systematic Review. *Sleep Med Rev* **2021**, *58*, doi:10.1016/J.SMRV.2021.101462.
30. Hartmann, V.; Liu, H.; Chen, F.; Qiu, Q.; Hughes, S.; Zheng, D. Quantitative Comparison of Photoplethysmographic Waveform Characteristics: Effect of Measurement Site. *Front Physiol* **2019**, *10*, 431697, doi:10.3389/FPHYS.2019.00198/BIBTEX.
31. Spaeder, M.C.; Klugman, D.; Skurow-Todd, K.; Glass, P.; Jonas, R.A.; Donofrio, M.T. Perioperative Near-Infrared Spectroscopy Monitoring in Neonates with Congenital Heart Disease: Relationship of Cerebral Tissue Oxygenation Index Variability with Neurodevelopmental Outcome. *Pediatric Critical Care Medicine* **2017**, *18*, 213–218, doi:10.1097/PCC.0000000000001056.
32. Su, W.C.; Colacot, R.; Ahmed, N.; Nguyen, T.; George, T.; Gandjbakhche, A. The Use of Functional Near-Infrared Spectroscopy in Tracking Neurodevelopmental Trajectories in Infants and Children with or without Developmental Disorders: A Systematic Review. *Front Psychiatry* **2023**, *14*, 1210000, doi:10.3389/FPSYT.2023.1210000/BIBTEX.
33. Chock, V.Y.; Kwon, S.H.; Ambalavanan, N.; Batton, B.; Nelin, L.D.; Chalak, L.F.; Tian, L.; Van Meurs, K.P. Cerebral Oxygenation and Autoregulation in Preterm Infants (Early NIRS Study). *Journal of Pediatrics* **2020**, *227*, 94–100.e1, doi:10.1016/j.jpeds.2020.08.036.
34. Ren, H.; Jiang, X.; Xu, K.; Chen, C.; Yuan, Y.; Dai, C.; Chen, W. A Review of Cerebral Hemodynamics During Sleep Using Near-Infrared Spectroscopy. *Front Neurol* **2020**, *11*, 524009, doi:10.3389/FNEUR.2020.524009/BIBTEX.

35. Robertson, F.C.; Douglas, T.S.; Meintjes, E.M. Motion Artifact Removal for Functional near Infrared Spectroscopy: A Comparison of Methods. *IEEE Trans Biomed Eng* **2010**, *57*, 1377–1387, doi:10.1109/TBME.2009.2038667.
36. Cui, X.; Baker, J.M.; Liu, N.; Reiss, A.L. Sensitivity of FNIRS Measurement to Head Motion: An Applied Use of Smartphones in the Lab. *J Neurosci Methods* **2015**, *245*, 37–43, doi:10.1016/J.JNEUMETH.2015.02.006.
37. Di Lorenzo, R.; Pirazzoli, L.; Blasi, A.; Bulgarelli, C.; Hakuno, Y.; Minagawa, Y.; Brigadoi, S. Recommendations for Motion Correction of Infant FNIRS Data Applicable to Multiple Data Sets and Acquisition Systems. *Neuroimage* **2019**, *200*, 511–527, doi:10.1016/J.NEUROIMAGE.2019.06.056.
38. Li, Q.; Feng, J.; Guo, J.; Wang, Z.; Li, P.; Liu, H.; Fan, Z. Effects of the Multisensory Rehabilitation Product for Home-Based Hand Training after Stroke on Cortical Activation by Using NIRS Methods. *Neurosci Lett* **2020**, *717*, 134682, doi:10.1016/J.NEULET.2019.134682.
39. Strangman, G.; Goldstein, R.; Rauch, S.L.; Stein, J. Near-Infrared Spectroscopy and Imaging for Investigating Stroke Rehabilitation: Test-Retest Reliability and Review of the Literature. *Arch Phys Med Rehabil* **2006**, *87*, 12–19, doi:10.1016/J.APMR.2006.07.269.
40. De Fazio, R.; Mastronardi, V.M.; De Vittorio, M.; Visconti, P. Wearable Sensors and Smart Devices to Monitor Rehabilitation Parameters and Sports Performance: An Overview. *Sensors* **2023**, Vol. 23, Page 1856 **2023**, *23*, 1856, doi:10.3390/S23041856.
41. Yücel, M.A.; Selb, J.; Boas, D.A.; Cash, S.S.; Cooper, R.J. Reducing Motion Artifacts for Long-Term Clinical NIRS Monitoring Using Collodion-Fixed Prism-Based Optical Fibers. *Neuroimage* **2014**, *85*, 192–201, doi:10.1016/J.NEUROIMAGE.2013.06.054.
42. Johnson, P.A.; Schmölzer, G.M. Heart Rate Assessment during Neonatal Resuscitation. *Healthcare* **2020**, *8*, doi:10.3390/HEALTHCARE8010043.
43. Anton, O.; Fernandez, R.; Rendon-Morales, E.; Aviles-Espinosa, R.; Jordan, H.; Rabe, H. Heart Rate Monitoring in Newborn Babies: A Systematic Review. *Neonatology* **2019**, *116*, 199–210, doi:10.1159/000499675.
44. Hamaoka, T.; McCully, K.K.; Niwayama, M.; Chance, B. The Use of Muscle Near-Infrared Spectroscopy in Sport, Health and Medical Sciences: Recent Developments. *Philos Trans A Math Phys Eng Sci* **2011**, *369*, 4591–4604, doi:10.1098/RSTA.2011.0298.
45. Peng, Z.; Kommers, D.; Liang, R.H.; Long, X.; Cottaar, W.; Niemarkt, H.; Andriessen, P.; van Pul, C. Continuous Sensing and Quantification of Body Motion in Infants: A Systematic Review. *Heliyon* **2023**, *9*, e18234, doi:10.1016/J.HELİYON.2023.E18234.
46. Sentner, T.; Wang, X.; De Groot, E.R.; Van Schaijk, L.; Tataranno, M.L.; Vijlbrief, D.C.; Benders, M.J.N.L.; Bartels, R.; Dudink, J. The Sleep Well Baby Project: An Automated Real-Time Sleep-Wake State Prediction Algorithm in Preterm Infants. *Sleep* **2022**, *45*, doi:10.1093/SLEEP/ZSAC143.
47. Massaro, A.N.; Govindan, R.B.; Vezina, G.; Chang, T.; Andescavage, N.N.; Wang, Y.; Al-Shargabi, T.; Metzler, M.; Harris, K.; du Plessis, A.J. Impaired Cerebral Autoregulation and Brain Injury in Newborns with Hypoxic-Ischemic Encephalopathy



- Treated with Hypothermia. *J Neurophysiol* **2015**, *114*, 818–824, doi:10.1152/JN.00353.2015/ASSET/IMAGES/LARGE/Z9K0131532000002.JPEG.
48. Kurth, C.D.; McCann, J.C.; Wu, J.; Miles, L.; Loepke, A.W. Cerebral Oxygen Saturation-Time Threshold for Hypoxic-Ischemic Injury in Piglets. *Anesth Analg* **2009**, *108*, 1268–1277, doi:10.1213/ANE.0B013E318196AC8E.
  49. Mitra, S.; Bale, G.; Meek, J.; Tachtsidis, I.; Robertson, N.J. Cerebral Near Infrared Spectroscopy Monitoring in Term Infants With Hypoxic Ischemic Encephalopathy—A Systematic Review. *Front Neurol* **2020**, *11*, 1–17, doi:10.3389/FNEUR.2020.00393.
  50. Liu, J.; Sorantin, E. Neonatal Respiratory Distress Syndrome. *Neonatal Lung Ultrasonography* **2019**, 17–39, doi:10.1007/978-94-024-1549-0\_3/COVER.
  51. Reuter, S.; Moser, C.; Baack, M. Respiratory Distress in the Newborn. *Pediatr Rev* **2014**, *35*, 417–428, doi:10.1542/PIR.35-10-417.
  52. Kooi, E.M.W.; Verhagen, E.A.; Elting, J.W.J.; Czosnyka, M.; Austin, T.; Wong, F.Y.; Aries, M.J.H. Measuring Cerebrovascular Autoregulation in Preterm Infants Using Near-Infrared Spectroscopy: An Overview of the Literature. *Expert Rev Neurother* **2017**, *17*, 801–818, doi:10.1080/14737175.2017.1346472.
  53. Kooi, E.M.W.; Richter, A.E. Cerebral Autoregulation in Sick Infants: Current Insights. *Clin Perinatol* **2020**, *47*, 449–467, doi:10.1016/J.CLP.2020.05.003.
  54. Mitra, S.; Czosnyka, M.; Smielewski, P.; O'Reilly, H.; Brady, K.; Austin, T. Heart Rate Passivity of Cerebral Tissue Oxygenation Is Associated with Predictors of Poor Outcome in Preterm Infants. *Acta Paediatr* **2014**, *103*, e374–e382, doi:10.1111/APA.12696.
  55. Pellicer, A.; Greisen, G.; Benders, M.; Claris, O.; Dempsey, E.; Fumagalli, M.; Gluud, C.; Hagmann, C.; Hellström-Westas, L.; Hyttel-Sorensen, S.; et al. The SafeBoosC Phase II Randomised Clinical Trial: A Treatment Guideline for Targeted Near-Infrared-Derived Cerebral Tissue Oxygenation versus Standard Treatment in Extremely Preterm Infants. *Neonatology* **2013**, *104*, 171–178, doi:10.1159/000351346.
  56. Thewissen, L.; Caicedo, A.; Lemmers, P.; Van Bel, F. V.; Van Huffel, S. V.; Naulaers, G. Measuring Near-Infrared Spectroscopy Derived Cerebral Autoregulation in Neonates: From Research Tool toward Bedside Multimodal Monitoring. *Front Pediatr* **2018**, *6*, 341349, doi:10.3389/FPED.2018.00117/BIBTEX.
  57. Lovett, M.E.; MacDonald, J.M.; Mir, M.; Ghosh, S.; O'Brien, N.F.; LaRovere, K.L. Noninvasive Neuromonitoring Modalities in Children Part I: Pupillometry, Near-Infrared Spectroscopy, and Transcranial Doppler Ultrasonography. *Neurocrit Care* **2024**, *40*, 130–146, doi:10.1007/S12028-023-01730-4/TABLES/2.



# CHAPTER 11

Nederlandse samenvatting  
(summary in Dutch)





Near-Infrared Spectroscopy (NIRS) is een essentieel hulpmiddel geworden in de neonatale intensieve care unit (NICU) voor effectieve monitoring van zuurstofniveaus in de hersenen en andere weefsels zoals de nier [1–5]. Hoewel artefacten in NIRS-gegevens uitdagingen vormen tijdens de analyse, kunnen deze voordelig worden gebruikt [6]. Zo kunnen systemische fysiologische artefacten niet alleen worden gebruikt voor het beoordelen van de signaalkwaliteit [7,8], maar ook voor het extraheren van bijvoorbeeld de hartslag (HR) [6,9–13]. Daarnaast kunnen ademhalingspatronen in NIRS-signalen worden geanalyseerd om ademhalingsfrequenties (RR) te extraheren, vergelijkbaar met technieken die worden gebruikt in photoplethysmography [14,15], waardoor potentiële verstoringen worden omgezet in bruikbare gezondheidsgegevens.

Deze strategie om fysiologische informatie via NIRS te benutten is vooral waardevol bij neonaten. Het kan potentieel het ongemak en de stress verminderen die samenhangen met het frequent plaatsen en verwijderen van meerdere sensoren. Dit wordt mogelijk gemaakt door gelijktijdige analyse van HR, RR en cerebrale zuurstofverzadiging, terwijl ook de kwaliteit van de gegevens wordt gewaarborgd voor een nauwkeurige analyse. In deze scriptie onderzoeken we innovatieve methoden voor het verzamelen van NIRS-gegevens in klinische settings, met als dubbel doel het verbeteren van de kwaliteit van de gegevens en het extraheren van vitale functies zoals HR en RR.

Aanvankelijk richtte het onderzoek zich in Deel I op volwassenen, aangezien gegevens van volwassen proefpersonen in een gecontroleerde omgeving toegankelijker zijn dan de neonatale gegevens in klinische scenario's. De inzichten die uit deze studies met volwassenen zijn verkregen, waren cruciaal bij de ontwikkeling van algoritmen die later werden aangepast voor gebruik bij neonaten. Deel II van de scriptie toont de toepassing van deze algoritmen op neonatale gegevens.

### **Deel I: fysiologische informatie in volwassen NIRS**

De eerste hoofdstukken in dit deel introduceren twee nieuwe algoritmen—de Signaal Kwaliteitsindex (SQI) en de op Machine Learning gebaseerde SQI (MLSQI)—voor de beoordeling van de kwaliteit van NIRS-signalen. **Hoofdstuk 2** beschrijft de ontwikkeling van SQI, het eerste algoritme dat de kwaliteit van NIRS-signalen kwantitatief beoordeelt op een meerledige schaal, wat de traditionele binaire goede/slechte beoordelingen door bestaande algoritmen overtreft [7,8]. Dit algoritme blijkt sterk te correleren met visuele beoordelingen van experts, waardoor gegarandeerd wordt dat NIRS-gegevens betrouwbaar worden geïnterpreteerd zonder afhankelijk te zijn van subjectief oordeel. In **Hoofdstuk 3** verschuift de focus naar MLSQI, dat voortbouwt op de fundamenteën van SQI door machine learning technieken te incorporeren om een breder scala aan signaal kenmerken te analyseren. Dit algoritme toont superieure prestaties in het beoordelen van signaalkwaliteit, essentieel voor het aanpassen van NIRS-technologie voor realtime besluitvorming.

In **hoofdstukken 4 en 5** verkennen we innovatieve benaderingen voor het schatten van RR uit NIRS-gegevens bij volwassenen, waarmee de bruikbaarheid van NIRS verder wordt uitgebreid dan traditionele toepassingen. **Hoofdstuk 4** introduceert een nieuwe methode voor het extraheren van RR zonder dat een referentiesignaal nodig is, waardoor de manier waarop ademhalingsgerelateerde fysiologische artefacten in NIRS-gegevens worden gebruikt, typisch beschouwd als verstoringen, wordt getransformeerd [16]. Voortbouwend op deze basis adresseert **hoofdstuk 5** de uitdaging van door beweging gecorrumpeteerde gegevens tijdens fysieke activiteiten zoals fietsen. We ontwikkelden een robuuste methode voor RR-schatting

die gebruikmaakt van de inherente lage gevoeligheid van NIRS voor bewegingen van het onderwerp, wat de betrouwbaarheid van RR-metingen onder dynamische omstandigheden effectief verhoogt. Deze vooruitgang markeert een significante vooruitgang in het verfijnen van NIRS-toepassingen, waardoor het een veelzijdiger hulpmiddel wordt voor het monitoren van fysiologische reacties in real-world scenario's.

## **Deel II: fysiologische informatie in neonatale NIRS**

De overgang naar neonatale toepassingen, in **hoofdstukken 7 en 8** van deze scriptie, benadrukken we significante vooruitgang in neonatale monitoring binnen NICU-instellingen door de ontwikkeling van nieuwe algoritmen voor het extraheren van HR en RR uit NIRS-signalen. Het nieuw geïntroduceerde NHR (NIRS HR) algoritme is specifiek ontworpen voor HR-extractie uit (pre)term neonaten in NICU's, wat een baanbrekende stap markeert in de aanpassing van NIRS-technologie voor real-world klinische toepassingen. Deze methode toont een opmerkelijke verbetering ten opzichte van bestaande algoritmen, bereikt een hoge correlatiecoëfficiënt ( $r = 99,5\%$ ) en presteert aanzienlijk beter dan de automatische multischaal gebaseerde piekdetectie (AMPD) [17] en Perdue-algoritmen [10], die werden gevalideerd onder minder uitdagende omstandigheden. Voortbouwend hierop werd het NRR (NIRS RR) algoritme ontwikkeld om RR nauwkeurig te extraheren uit (pre)term neonaten in NICU's. Dit algoritme is het eerste dat specifiek de complexiteiten van neonatale RR-meting adresseert, en toont sterke prestaties met een robuuste correlatie ( $r = 84,5\%$ ) tussen de geëxtraheerde en referentie RR's. Samen vertegenwoordigen deze algoritmen transformerende verbeteringen in het vermogen van NIRS om cruciale fysiologische gegevens te leveren voor de meest kwetsbare patiënten.

**Hoofdstuk 9** van de scriptie toont een toepassing van HR, RR en signaal kwaliteitsbeoordeling met behulp van een enkel NIRS-apparaat voor het beoordelen van slaapstadia bij neonaten opgenomen in NICU's. Door gebruik te maken van een hoge frequentie (100 Hz) bemonstering, werden kenmerken van NIRS-signalen geanalyseerd met behulp van een convolutioneel neurale netwerk classifier om onderscheid te maken tussen actieve en stille slaap bij (pre)term neonaten. Deze aanpak integreerde fysiologische en niet-fysiologische gegevens, en demonstreerde het nut van een enkele sensor in gedetailleerde slaappatroon analyse voor neonaten zo jong als 24 weken zwangerschap leeftijd of met geboortegewichten zo laag als 500 gram.

Deze scriptie benadrukt het transformerende potentieel van NIRS in klinische instellingen, met name door de verfijnde toepassingen in neonatale zorg. Door algoritmen te ontwikkelen voor het extraheren van vitale fysiologische metingen zoals HR en RR, en deze te integreren met cerebrale zuurstof gegevens, zou NIRS-technologie niet alleen de behoefte aan meerdere sensoren kunnen minimaliseren—waardoor neonatale stress wordt verminderd [18,19]—maar ook het beheer van aandoeningen zoals hypoxisch-ischemische encefalopathie [20–22] en cerebrale vasculaire autoregulatie [23] kunnen verbeteren. Deze integratie, gecombineerd met technologieën zoals electroencephalography [24] en transcraniële Doppler [25], zou ook een uitgebreide beoordeling van de cerebrale gezondheid van een patiënt kunnen bieden, waardoor continue, real-time monitoring cruciaal wordt voor de vroege detectie van cerebrale dysregulatie [4]. Deze ontwikkelingen zouden een nieuwe standaard kunnen stellen voor niet-invasieve patiëntmonitoring in neonatale zorg, waarbij NIRS's capaciteit wordt gedemonstreerd om patiëntenzorg te transformeren door gedetailleerde en real-time fysiologische beoordelingen.

## Referenties

1. Yu, Y.; Zhang, K.; Zhang, L.; Zong, H.; Meng, L.; Han, R. Cerebral Near-infrared Spectroscopy (NIRS) for Perioperative Monitoring of Brain Oxygenation in Children and Adults. *Cochrane Database Syst Rev* **2018**, 2018, doi:10.1002/14651858.CD010947.PUB2.
2. Petrova, A.; Mehta, R. Near-Infrared Spectroscopy in the Detection of Regional Tissue Oxygenation during Hypoxic Events in Preterm Infants Undergoing Critical Care. *Pediatric Critical Care Medicine* **2006**, 7, 449–454, doi:10.1097/01.PCC.0000235248.70482.14.
3. Scheeren, T.W.L.; Schober, P.; Schwarte, L.A. Monitoring Tissue Oxygenation by near Infrared Spectroscopy (NIRS): Background and Current Applications. *J Clin Monit Comput* **2012**, 26, 279, doi:10.1007/S10877-012-9348-Y.
4. Van Bel, F.; Lemmers, P.; Naulaers, G. Monitoring Neonatal Regional Cerebral Oxygen Saturation in Clinical Practice: Value and Pitfalls. *Neonatology* **2008**, 94, 237–244, doi:10.1159/000151642.
5. Chock, V.Y.; Vesoulis, Z.A.; El-Dib, M.; Austin, T.; van Bel, F. The Future of Neonatal Cerebral Oxygenation Monitoring: Directions After the SafeBoosC-III Trial. *J Pediatr* **2024**, 270, 114016, doi:10.1016/J.JPEDI.2024.114016.
6. Hakimi, N.; Kamaledin Setarehdan, S. Stress Assessment by Means of Heart Rate Derived from Functional Near-Infrared Spectroscopy. <https://doi.org/10.1117/1.JBO.23.11.115001> **2018**, 23, 115001, doi:10.1117/1.JBO.23.11.115001.
7. Pollonini, L.; Olds, C.; Abaya, H.; Bortfeld, H.; Beauchamp, M.S.; Oghalai, J.S. Auditory Cortex Activation to Natural Speech and Simulated Cochlear Implant Speech Measured with Functional Near-Infrared Spectroscopy. *Hear Res* **2014**, 309, 84–93, doi:10.1016/J.HEARES.2013.11.007.
8. Pollonini, L.; Bortfeld, H.; Oghalai, J.S. PHOEBE: A Method for Real Time Mapping of Optodes-Scalp Coupling in Functional near-Infrared Spectroscopy. *Biomedical Optics Express*, Vol. 7, Issue 12, pp. 5104-5119 **2016**, 7, 5104–5119, doi:10.1364/BOE.7.005104.
9. Hakimi, N.; Jodeiri, A.; Mirbagheri, M.; Setarehdan, S.K. Proposing a Convolutional Neural Network for Stress Assessment by Means of Derived Heart Rate from Functional near Infrared Spectroscopy. *Comput Biol Med* **2020**, 121, 103810, doi:10.1016/J.COMPBIOMED.2020.103810.
10. Perdue, K.L.; Westerlund, A.; McCormick, S.A.; Nelson, C.A.; Iii, C.A.N. Extraction of Heart Rate from Functional Near-Infrared Spectroscopy in Infants. <https://doi.org/10.1117/1.JBO.19.6.067010> **2014**, 19, 067010, doi:10.1117/1.JBO.19.6.067010.
11. Perdue, K.L.; Edwards, L.A.; Tager-Flusberg, H.; Nelson, C.A. Differing Developmental Trajectories in Heart Rate Responses to Speech Stimuli in Infants at High and Low Risk

- for Autism Spectrum Disorder. *J Autism Dev Disord* **2017**, *47*, 2434–2442, doi:10.1007/S10803-017-3167-4/FIGURES/4.
12. Holper, L.; Seifritz, E.; Scholkmann, F. Short-Term Pulse Rate Variability Is Better Characterized by Functional near-Infrared Spectroscopy than by Photoplethysmography. *J Biomed Opt* **2016**, *21*, 091308, doi:10.1117/1.JBO.21.9.091308.
  13. Izzetoglu, M.; Holtzer, R. Effects of Processing Methods on FNIRS Signals Assessed during Active Walking Tasks in Older Adults. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2020**, *28*, 699–709, doi:10.1109/TNSRE.2020.2970407.
  14. Hernando, A.; Peláez-Coca, M.D.; Lozano, M.T.; Lázaro, J.; Gil, E. Finger and Forehead PPG Signal Comparison for Respiratory Rate Estimation. *Physiol Meas* **2019**, *40*, 095007, doi:10.1088/1361-6579/AB3BE0.
  15. Madhav, K.V.; Ram, M.R.; Krishna, E.H.; Komalla, N.R.; Reddy, K.A. Robust Extraction of Respiratory Activity from PPG Signals Using Modified MSPCA. *IEEE Trans Instrum Meas* **2013**, *62*, 1094–1106, doi:10.1109/TIM.2012.2232393.
  16. Nguyen, H.D.; Yoo, S.H.; Bhutta, M.R.; Hong, K.S. Adaptive Filtering of Physiological Noises in FNIRS Data. *Biomed Eng Online* **2018**, *17*, 1–23, doi:10.1186/S12938-018-0613-2/FIGURES/17.
  17. Scholkmann, F.; Boss, J.; Wolf, M. An Efficient Algorithm for Automatic Peak Detection in Noisy Periodic and Quasi-Periodic Signals. *Algorithms 2012, Vol. 5, Pages 588-603* **2012**, *5*, 588–603, doi:10.3390/A5040588.
  18. Kuller, J.M.M. Skin Breakdown: Risk Factors, Prevention, and Treatment. *Newborn and Infant Nursing Reviews* **2001**, *1*, 35–42, doi:10.1053/NBIN.2001.22874.
  19. Afsar, F.S. Skin Care for Preterm and Term Neonates. *Clin Exp Dermatol* **2009**, *34*, 855–858, doi:10.1111/J.1365-2230.2009.03424.X.
  20. Massaro, A.N.; Govindan, R.B.; Vezina, G.; Chang, T.; Andescavage, N.N.; Wang, Y.; Al-Shargabi, T.; Metzler, M.; Harris, K.; du Plessis, A.J. Impaired Cerebral Autoregulation and Brain Injury in Newborns with Hypoxic-Ischemic Encephalopathy Treated with Hypothermia. *J Neurophysiol* **2015**, *114*, 818–824, doi:10.1152/JN.00353.2015/ASSET/IMAGES/LARGE/Z9K0131532000002.JPEG.
  21. Kurth, C.D.; McCann, J.C.; Wu, J.; Miles, L.; Loepke, A.W. Cerebral Oxygen Saturation-Time Threshold for Hypoxic-Ischemic Injury in Piglets. *Anesth Analg* **2009**, *108*, 1268–1277, doi:10.1213/ANE.0B013E318196AC8E.
  22. Mitra, S.; Bale, G.; Meek, J.; Tachtsidis, I.; Robertson, N.J. Cerebral Near Infrared Spectroscopy Monitoring in Term Infants With Hypoxic Ischemic Encephalopathy—A Systematic Review. *Front Neurol* **2020**, *11*, 1–17, doi:10.3389/FNEUR.2020.00393.
  23. Kooi, E.M.W.; Verhagen, E.A.; Elting, J.W.J.; Czosnyka, M.; Austin, T.; Wong, F.Y.; Aries, M.J.H. Measuring Cerebrovascular Autoregulation in Preterm Infants Using Near-Infrared Spectroscopy: An Overview of the Literature. *Expert Rev Neurother* **2017**, *17*, 801–818, doi:10.1080/14737175.2017.1346472.



24. Thewissen, L.; Caicedo, A.; Lemmers, P.; Van Bel, F. V.; Van Huffel, S. V.; Naulaers, G. Measuring Near-Infrared Spectroscopy Derived Cerebral Autoregulation in Neonates: From Research Tool toward Bedside Multimodal Monitoring. *Front Pediatr* **2018**, *6*, 341349, doi:10.3389/FPED.2018.001117/BIBTEX.
25. Lovett, M.E.; MacDonald, J.M.; Mir, M.; Ghosh, S.; O'Brien, N.F.; LaRovere, K.L. Noninvasive Neuromonitoring Modalities in Children Part I: Pupillometry, Near-Infrared Spectroscopy, and Transcranial Doppler Ultrasonography. *Neurocrit Care* **2024**, *40*, 130–146, doi:10.1007/S12028-023-01730-4/TABLES/2.

## CHAPTER 12

List of abbreviations

List of coauthors

List of publications

Curriculum vitae

Acknowledgements



## LIST OF ABBREVIATIONS

AHRFB	=	adaptive heart rate frequency bandwidth
AE	=	absolute error
AdaB	=	AdaBoost
AMPD	=	automatic multiscale-based peak detection
AUC-ROC	=	area under the receiver operating characteristic curve
BAR	=	bland-altman ratio
BeSSPI	=	behavioral sleep stage classification for preterm infants
BLL	=	Beer-Lambert law
BLA	=	Bland-Altman limits of agreement
BMI	=	body mass index
BPM	=	breaths per minute
bpm	=	beats per minute
BP	=	blood pressure
BW	=	baseline wandering
CAR	=	cerebral autoregulation
CNN	=	convolutional neural network
CO <sub>2</sub>	=	carbon dioxide
CSI	=	critical success index
CBF	=	cerebral blood flow
CVAR	=	cerebral vascular autoregulation
dWMI	=	diffuse white matter injury
DPF	=	differential pathlength factor
EEG	=	electroencephalography
ECG	=	electrocardiogram
EHR	=	extracted heart rate
FDA	=	food and drug administration
FFT	=	fast fourier transform
FIR	=	finite impulse response
fMRI	=	functional magnetic resonance imaging
fNIRS	=	functional near-infrared spectroscopy
GA	=	gestational age
GB	=	gigabyte
HHb	=	deoxygenated hemoglobin
HIE	=	hypoxic-ischemic encephalopathy
HR	=	heart rate
hsPDA	=	hemodynamically significant patent ductus arteriosus
IQR	=	interquartile range
IRDS	=	infant respiratory distress syndrome
IS	=	included segments
LOA	=	Bland-Altman limits of agreement
LTSO	=	leave-two-subjects-out
MA	=	moving average

---

MBLL	=	modified Beer-Lambert law
ME	=	mean of error
MLSQI	=	machine learning-based signal quality index
MRI	=	magnetic resonance imaging
ms	=	milliseconds
NB	=	naive bayes
NHR	=	NIRS heart rate
NICU	=	neonatal intensive care unit
NIRS	=	near-infrared spectroscopy
NRR	=	NIRS respiratory rate
OD	=	optical density
O <sub>2</sub> Hb	=	oxygenated hemoglobin
PAIS	=	perinatal arterial ischemic stroke
PIVH	=	periventricular-intraventricular hemorrhage
PHOEBE	=	placing headgear optodes efficiently before experimentation
PPG	=	photoplethysmography
PSG	=	polysomnography
PI	=	performance index
pCO <sub>2</sub>	=	partial pressure of carbon dioxide
QS	=	quiet sleep
RF	=	random forest
RHR	=	reference heart rate
RMSE	=	root mean square error
RR	=	respiratory rate
rSO <sub>2</sub>	=	regional oxygen saturation
rScO <sub>2</sub>	=	regional cerebral oxygen saturation
R <sup>2</sup>	=	coefficient of determination
SaO <sub>2</sub>	=	arterial oxygen saturation
SD	=	standard deviation
SNR	=	signal-to-noise ratio
SpO <sub>2</sub>	=	peripheral oxygen saturation
SQI	=	signal quality index
StdE	=	standard deviation of error
SVM	=	support vector machine
StO <sub>2</sub>	=	tissue oxygen saturation
TF	=	transfer function
tHb	=	total hemoglobin
TSI	=	tissue saturation index
WMI	=	white matter injury
XGB	=	extreme gradient boosting or XGBoost



## LIST OF COAUTHORS

### **Jeroen Dudink**

Department of Neonatology, Wilhelmina Children's Hospital, University Medical Center Utrecht, Lundlaan 6, Utrecht 3584 EA, the Netherlands

### **Thomas Alderliesten**

Department of Neonatology, Wilhelmina Children's Hospital, University Medical Center Utrecht, Lundlaan 6, Utrecht 3584 EA, the Netherlands

### **Jörn M. Horschig**

Artinis Medical Systems, B.V., Einsteinweg 17, 6662 PW Elst, The Netherlands

### **Frank Van Bel**

Department of Neonatology, Wilhelmina Children's Hospital, University Medical Center Utrecht, Lundlaan 6, Utrecht 3584 EA, the Netherlands

### **M. Sofia Sappia**

Artinis Medical Systems, B.V., Einsteinweg 17, 6662 PW Elst, The Netherlands

Radboud University Nijmegen, Donders Institute for Brain, Behaviour and Cognition, 6525 EN Nijmegen, the Netherlands

### **Mohammad Shahbakhti**

Artinis Medical Systems, B.V., Einsteinweg 17, 6662 PW Elst, The Netherlands

Biomedical Engineering Institute, Kaunas University of Technology, K. Barsausko 59, LT-51423 Kaunas, Lithuania

### **Willy N.J.M. Colier**

Artinis Medical Systems, B.V., Einsteinweg 17, 6662 PW Elst, The Netherlands

### **Liucija Svinkunaite**

Artinis Medical Systems, B.V., Einsteinweg 17, 6662 PW Elst, The Netherlands

**Mathijs Bronkhorst**

Artinis Medical Systems, B.V., Einsteinweg 17, 6662 PW Elst, The Netherlands

**Marianne Floor-Westerdijk**

Artinis Medical Systems, B.V., Einsteinweg 17, 6662 PW Elst, The Netherlands

**Gaetano Valenza**

Bioengineering and Robotics Research Center E. Piaggio and the Department of Information Engineering, School of Engineering, University of Pisa, Via G. Caruso 16, 56122 Pisa, Italy

**Jurgen Claassen**

Donders Institute for Brain, Cognition and Behaviour, Radboud University, Houtlaan 4, 6525 XZ Nijmegen, The Netherlands

**Filipe Gonçalves Costa**

Department of Neonatology, Wilhelmina Children's Hospital, University Medical Center Utrecht, Lundlaan 6, Utrecht 3584 EA, the Netherlands

**Emad Arasteh**

Department of Neonatology, Wilhelmina Children's Hospital, University Medical Center Utrecht, Lundlaan 6, Utrecht 3584 EA, the Netherlands

**Maren Zahn**

Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

Artinis Medical Systems, B.V., Einsteinweg 17, 6662 PW Elst, The Netherlands

## LIST OF PUBLICATIONS

Sappia, M.S.; **Hakimi, N.**; Colier, W.N.J.M.; Horschig, J.M. Signal Quality Index: An Algorithm for Quantitative Assessment of Functional near Infrared Spectroscopy Signal Quality. *Biomedical Optics Express*, Vol. 11, Issue 11, pp. 6732–6754 2020, 11, 6732–6754, doi:10.1364/BOE.409317.

Sappia, M.S.; **Hakimi, N.**; Svinkunaite, L.; Alderliesten, T.; Horschig, J.M.; Colier, W.N.J.M. FNIRS Signal Quality Estimation by Means of a Machine Learning Algorithm Trained on Morphological and Temporal Features. <https://doi.org/10.1117/12.2587188> 2021, 11638, 29–39, doi:10.1117/12.2587188.

Costa, F.G.; **Hakimi, N.**; Van Bel, F. Neuroprotection of the Perinatal Brain by Early Information of Cerebral Oxygenation and Perfusion Patterns. *Int J Mol Sci* 2021, 22, doi:10.3390/IJMS22105389.

**Hakimi, N.**; Shahbakhti, M.; Sappia, S.; Horschig, J.M.; Bronkhorst, M.; Floor-Westerdijk, M.; Valenza, G.; Dudink, J.; Colier, W.N.J.M. Estimation of Respiratory Rate from Functional Near-Infrared Spectroscopy (FNIRS): A New Perspective on Respiratory Interference. *Biosensors (Basel)* 2022, 12, doi:10.3390/BIOS12121170.

Shahbakhti, M.; **Hakimi, N.**; Horschig, J.M.; Floor-Westerdijk, M.; Claassen, J.; Colier, W.N.J.M. Estimation of Respiratory Rate during Biking with a Single Sensor Functional Near-Infrared Spectroscopy (FNIRS) System. *Sensors* 2023, Vol. 23, Page 3632 2023, 23, 3632, doi:10.3390/S23073632.

**Hakimi, N.**; Horschig, J.M.; Alderliesten, T.; Bronkhorst, M.; Floor-Westerdijk, M.J.; Bel, F. V.; Colier, W.N.J.M.; Dudink, J. Heart Rate Extraction from Neonatal Near-Infrared Spectroscopy Signals. *IEEE Trans Instrum Meas* 2023, doi:10.1109/TIM.2023.3295015.

**Hakimi, N.**; Shahbakhti, M.; Horschig, J.M.; Alderliesten, T.; Van Bel, F.; Colier, W.N.J.M.; Dudink, J. Respiratory Rate Extraction from Neonatal Near-Infrared Spectroscopy Signals. *Sensors* 2023, Vol. 23, Page 4487 2023, 23, 4487, doi:10.3390/S23094487.



## CURRICULUM VITAE

Naser Hakimi graduated with a B.Sc. in Electrical Engineering from Shahid Beheshti University, Tehran, Iran, in 2015, and earned his M.Sc. in Biomedical Engineering from the University of Tehran, Tehran, Iran in 2018. Following his master's studies, he moved to the Netherlands in 2019 to advance his research career. Naser joined the INFANS (INtegrating Functional Assessment measures for Neonatal Safeguard) project as an early-stage researcher, supported by the Marie Skłodowska-Curie Actions programme. He conducted his research at Artinis Medical Systems B.V., Elst, the Netherlands, and simultaneously served as a PhD candidate in the Neonatology department at Wilhelmina Children's Hospital, part of the University Medical Center Utrecht, Utrecht, the Netherlands. His primary research focus has been on analyzing near-infrared spectroscopy (NIRS) signals in adults and infants. He has developed methods for extracting physiological information, such as heart rate and respiratory rate, from NIRS signals—data typically overlooked as artifacts. Additionally, he has worked on NIRS signal quality assessment and the application of NIRS physiological information in stress and sleep analysis. Currently, he is a Clinical Research Engineer at the Centre for Human Drug Research (CHDR) in Leiden, where he continues to contribute to advancements in medical technology and research.



## ACKNOWLEDGEMENTS

First, I would like to express my deepest gratitude to my promotor, Jeroen, and my co-promotor, Thomas. Your unwavering support and guidance have been instrumental in the completion of this PhD. There were times when I felt disheartened and unsure about continuing, but your openness to my concerns and your efforts to assist me provided the motivation I needed. Jeroen, your motivational personality has been a source of great encouragement for me. You taught me the challenges of data collection in a neonatal intensive care unit, and your support was crucial in helping me meet the tight schedule for my defense.

I would also like to extend my thanks to Jörn, my other co-supervisor. Starting at Artinis and moving to the Netherlands was a significant transition, and your assistance with various settlements was invaluable to me. At a time when I had no other support, your kindness and willingness to listen to my thoughts made a profound difference. Although circumstances at Artinis changed, I will always cherish the support and guidance you provided.

In addition to my supervisors, I would like to thank Frank, who was always open for discussion about my questions and ideas and very helpful with giving valuable feedback. I wish I could have had more opportunities to take advantage of your great knowledge.

I would also like to extend my sincere thanks to the participants in my studies and the parents who agreed to involve their children. Your willingness to contribute, even though the study did not have a direct effect on your children's health, has been invaluable and deeply appreciated.

I would like to thank my wife, my love, Sheyda. Every year in my life, I am grateful that my journey, which started in 2019 with moving abroad and was filled with hardship and challenges, led me to you. If I had known that journey would lead to you, I would have followed it with passion and love. Meeting you has been a turning point in my life. Without your support, I definitely could not have finished this PhD. You cared more than I did about what was most suitable for me.

My father, you did your best and spent your life helping us reach our goals. You prepared an environment for me to grow, even though it was not your preference to be far from your son. It was not my preference either, but sometimes we need to pay for success, and this has been our payment. Getting this PhD is not only my success; it is also yours. It is also my mother's success. I will never forget the last promise I made to her to continue my studies no matter what happens. Perhaps that promise has been my driving force in finishing this journey. I am always grateful for all your support.

I would like to thank my parents-in-law, Akbar and Mamak. You are truly the most logical people I have ever met. You have always been a great support for us, and I hope I can give back your kindness.

I would also like to thank other members of my family: Narges, Leila, Sajad, Yaser, Yasin, Amir, Fatemeh, and Faezeh. You supported me in moving abroad and did your best to make it easy for me. Yasin, my oldest brother, you were my role model during school, and it was you who first showed me how much progress we can make in education. Narges, my only sister, you have been someone I could talk to comfortably without feeling judged, always supporting and sympathizing with me. Yaser, as time passes, I understand more what a great brother you are. I always cherish the memories we have. Sajad, my little brother, through you

I learned how sweet it is to have a younger brother. Please remember that I will always be one of your supporters. Leila, you have played a crucial role in keeping our family together with your great personality, always prioritizing others first. Without you, we would not have come this far.

I would like to thank my friends who made living in the Netherlands easier for me. Sadeh and Sara, your supportive personalities have made our friendship a great honor for me. I will never forget your efforts in forming an Iranian community in Nijmegen. Through your efforts, I met Sheyda and wonderful friends Khalil, Mahsa, Hadi, Maryam, Milad, Mona, Arezoo, Alireza, Fatemeh, Vahid, and Mitra. Thank you all for making living abroad a more comfortable experience.

I also want to express my gratitude to the friends I met while working at Artinis. My paranymphs, Sofia and Lara, you are my first non-Iranian friends. You taught me that people from different countries, even different continents, can have remarkably similar personalities. You are kind, empathetic, and responsible individuals. I wish we could have collaborated more as you are brilliant researchers. Thank you for being my paranymphs; I couldn't imagine anyone else who started this PhD journey with me.

I would also like to thank Mohammad, Somayeh, and Parivash. Mohammad, you prioritize friendship over competition, and your help and support have been invaluable. I wish we could have continued working together and going to the gym, as I also deeply value our friendship. I always think of you whenever I hear "Kabab!"

Last but not least, I extend my sincere thanks to the members of my committee, Prof. dr. M.N. Bekker, Prof. dr. D. van Dijk, Dr. F. Groenendaal, Prof. dr. G. Naulaers, and Prof. dr. A.A. Salah. Your invaluable feedback and insightful questions have significantly contributed to the depth and rigor of my research. I am deeply grateful for your time, expertise, and guidance throughout this process. Your support was also crucial in helping me adhere to the tight schedule for my defense.