

# Tissue oximetry: a comparison of mean values of regional tissue saturation, reproducibility and dynamic range of four NIRS-instruments on the human forearm

Simon Hyttel-Sorensen,<sup>1,\*</sup> Line C. Sorensen,<sup>2</sup> Joan Riera,<sup>3,4</sup> and Gorm Greisen<sup>1</sup>

<sup>1</sup>Department of Neonatology, National University Hospital, Rigshospitalet, Blegdamsvej 9, Copenhagen, DK-2100, Denmark

<sup>2</sup>Department of Paediatrics, National University Hospital, Hvidovre Hospital, Kettegard Alle 30, Hvidovre, DK-2650, Denmark

<sup>3</sup>Department of Neonatology, La Paz University Hospital, Autónoma University of Madrid, Paseo de la Castellana 261, E-28046 Madrid, Spain

<sup>4</sup>Center for Biomedical Technology (Technical University of Madrid and Complutense University of Madrid), 28040 Madrid, Spain

\*Simonhyttelsoren@me.com

**Abstract:** We compared absolute values of regional tissue hemoglobin saturation (StO<sub>2</sub>), reproducibility, and dynamic range of four different instruments on the forearm of adults. The sensors were repositioned 10 times on each subject. Dynamic range was estimated by exercise with subsequent arterial occlusion. Mean StO<sub>2</sub> was 70.1% ± 6.7 with INVOS 5100, 69.4% ± 5.0 with NIRO 200 NX, 63.4% ± 4.5 with NIRO 300, and 60.8% ± 3.6 with OxyPrem. The corresponding reproducibility S<sub>w</sub> was 5.4% (CI 4.4–6.9), 4.4% (CI 3.5–5.2), 4.1% (CI 3.3–4.9), and 2.7% (CI 2.2–3.2), respectively. The dynamic ranges ΔStO<sub>2</sub> were 45.0%, 46.8%, 44.8%, and 27.8%, respectively. In conclusion, the three commercial NIRS instruments showed different absolute values, whereas reproducibility and dynamic range were quite similar.

© 2011 Optical Society of America

OCIS codes: (170.1470) Blood or tissue constituent monitoring; (300.6190) Spectrometers.

---

## References and links

1. M. Heringlake, C. Garbers, J. H. Käbler, I. Anderson, H. Heinze, J. Schön, K.-U. Berger, L. Dibbelt, H.-H. Sievers, and T. Hanke, "Preoperative cerebral oxygen saturation and clinical outcomes in cardiac surgery," *Anesthesiology* **114**(1), 58–69 (2011).
2. M. C. Toet, P. M. A. Lemmers, L. J. van Schelven, and F. van Bel, "Cerebral oxygenation and electrical activity after birth asphyxia: their relation to outcome," *Pediatrics* **117**(2), 333–339 (2006).
3. J. C. Hirsch, J. R. Charpie, R. G. Ohye, and J. G. Gurney, "Near infrared spectroscopy (NIRS) should not be standard of care for postoperative management," *Semin. Thorac. Cardiovasc. Surg. Pediatr. Card. Surg. Annu.* **13**(1), 51–54 (2010).
4. G. Greisen, "Is near-infrared spectroscopy living up to its promises?" *Semin. Fetal Neonatal Med.* **11**(6), 498–502 (2006).
5. R. E. Gagnon, A. J. Macnab, F. A. Gagnon, D. Blackstock, and J. G. LeBlanc, "Comparison of two spatially resolved NIRS oxygenation indices," *J. Clin. Monit. Comput.* **17**(7/8), 385–391 (2002).
6. A. Dullenkopf, B. Frey, O. Baenziger, A. Gerber, and M. Weiss, "Measurement of cerebral oxygenation state in anaesthetized children using the INVOS 5100 cerebral oximeter," *Paediatr. Anaesth.* **13**(5), 384–391 (2003).
7. L. C. Sorensen and G. Greisen, "Precision of measurement of cerebral tissue oxygenation index using near-infrared spectroscopy in preterm neonates," *J. Biomed. Opt.* **11**(5), 054005 (2006).
8. J. Menke, U. Voss, G. Möller, and G. Jorch, "Reproducibility of cerebral near infrared spectroscopy in neonates," *Biol. Neonate* **83**(1), 6–11 (2003).
9. M. Wolf, M. Keel, V. Dietz, K. von Siebenthal, H. U. Bucher, and O. Baenziger, "The influence of a clear layer on near-infrared spectrophotometry measurements using a liquid neonatal head phantom," *Phys. Med. Biol.* **44**(7), 1743–1753 (1999).

10. L. C. Sorensen, T. S. Leung, and G. Greisen, "Comparison of cerebral oxygen saturation in premature infants by near-infrared spatially resolved spectroscopy: observations on probe-dependent bias," *J. Biomed. Opt.* **13**(6), 064013 (2008).
11. F. Y. Wong, N. B. Witcombe, S. R. Yiallourou, S. Yorkston, A. R. Dymowski, L. Krishnan, A. M. Walker, and R. S. C. Horne, "Cerebral oxygenation is depressed during sleep in healthy term infants when they sleep prone," *Pediatrics* **127**(3), e558–e565 (2011).
12. P. M. A. Lemmers, M. C. Toet, and F. van Bel, "Impact of patent ductus arteriosus and subsequent therapy with indomethacin on cerebral oxygenation in preterm infants," *Pediatrics* **121**(1), 142–147 (2008).
13. S. R. Arridge, M. Cope, and D. T. Delpy, "The theoretical basis for the determination of optical pathlengths in tissue: temporal and frequency analysis," *Phys. Med. Biol.* **37**(7), 1531–1560 (1992).
14. P. van der Zee, M. Cope, S. R. Arridge, M. Essenpreis, L. A. Potter, A. D. Edwards, J. S. Wyatt, D. C. McCormick, S. C. Roth, E. O. Reynolds, and D. T. Delpy, "Experimentally measured optical pathlengths for the adult head, calf and forearm and the head of the newborn infant as a function of inter optode spacing," *Adv. Exp. Med. Biol.* **316**, 143–153 (1992).
15. S. J. Matcher, P. Kirkpatrick, K. Nahid, M. Cope, and D. T. Delpy, "Absolute quantification methods in tissue near infrared spectroscopy," *Proc. SPIE* **2389**, 486–495 (1995).
16. "NIRS Technology," [www.somanetics.com/our-technology/nirs-technology#trend](http://www.somanetics.com/our-technology/nirs-technology#trend).
17. D. M. Hueber, S. Fantini, A. E. Cerussi, and B. Barbieri, "New optical probe designs for absolute (self-calibrating) nir tissue hemoglobin measurements," *Proc. SPIE* **3597**, 618–631 (1999).
18. J. M. Bland and D. G. Altman, "Measurement error," *BMJ* **313**(7059), 744 (1996).
19. K. Yoshitani, M. Kawaguchi, K. Tatsumi, K. Kitaguchi, and H. Furuya, "A comparison of the INVOS 4100 and the NIRO 300 near-infrared spectrophotometers," *Anesth. Analg.* **94**(3), 586–590 (2002).
20. E. G. McKeating, J. R. Monjardino, D. F. Signorini, M. J. Souter, and P. J. Andrews, "A comparison of the Invos 3100 and the Critikon 2020 near-infrared spectrophotometers as monitors of cerebral oxygenation," *Anaesthesia* **52**(2), 136–140 (1997).
21. M. Thavasoathy, M. Broadhead, C. Elwell, M. Peters, and M. Smith, "A comparison of cerebral oxygenation as measured by the NIRO 300 and the INVOS 5100 Near-Infrared Spectrophotometers," *Anaesthesia* **57**(10), 999–1006 (2002).
22. N. Nagdyman, P. Ewert, B. Peters, O. Miera, T. Fleck, and F. Berger, "Comparison of different near-infrared spectroscopic cerebral oxygenation indices with central venous and jugular venous oxygenation saturation in children," *Paediatr. Anaesth.* **18**(2), 160–166 (2008).
23. T. Komiya, V. Quaresima, H. Shigematsu, and M. Ferrari, "Comparison of two spatially resolved near-infrared photometers in the detection of tissue oxygen saturation: poor reliability at very low oxygen saturation," *Clin. Sci.* **101**(6), 715–718 (2001).
24. K. Shiroishi, R. Kime, T. Osada, N. Murase, K. Shimomura, and T. Katsumura, "Decreased muscle oxygenation and increased arterial blood flow in the non-exercising limb during leg exercise," *Adv. Exp. Med. Biol.* **662**, 379–384 (2010).
25. C. E. Cooper, S.-M. Penfold, C. E. Elwell, and C. Angus, "Comparison of local adipose tissue content and SRS-derived NIRS muscle oxygenation measurements in 90 individuals," *Adv. Exp. Med. Biol.* **662**, 177–181 (2010).
26. T. J. Germon, A. E. Young, A. R. Manara, and R. J. Nelson, "Extracerebral absorption of near infrared light influences the detection of increased cerebral oxygenation monitored by near infrared spectroscopy," *J. Neurol. Neurosurg. Psychiatry* **58**(4), 477–479 (1995).
27. G. Grubhofer, A. Lassnigg, F. Manlik, E. Marx, W. Trubel, and M. Hiesmayr, "The contribution of extracranial blood oxygenation on near-infrared spectroscopy during carotid thrombendarterectomy," *Anaesthesia* **52**(2), 116–120 (1997).
28. P. G. Al-Rawi, P. Smielewski, and P. J. Kirkpatrick, "Evaluation of a near-infrared spectrometer (NIRO 300) for the detection of intracranial oxygenation changes in the adult head," *Stroke* **32**(11), 2492–2500 (2001).
29. C. Jenny, M. Biallas, I. Trajkovic, J.-C. Fauchère, H.-U. Bucher, and M. Wolf, "Reproducibility of cerebral tissue oxygen saturation measurements by near-infrared spectroscopy in newborn infants," *J. Biomed. Opt.* **16**(9), 097004 (2011).
30. A. Dullenkopf, A. Kolarova, G. Schulz, B. Frey, O. Baenziger, and M. Weiss, "Reproducibility of cerebral oxygenation measurement in neonates and infants in the clinical setting using the NIRO 300 oximeter," *Pediatr. Crit. Care Med.* **6**(3), 344–347 (2005).
31. J. M. Bland and D. G. Altman, "Measuring agreement in method comparison studies," *Stat. Methods Med. Res.* **8**(2), 135–160 (1999).
32. M. Pocivalnik, G. Pichler, H. Zotter, N. Tax, W. Müller, and B. Urlesberger, "Regional tissue oxygen saturation: comparability and reproducibility of different devices," *J. Biomed. Opt.* **16**(5), 057004 (2011).
33. H. M. Watzman, C. D. Kurth, L. M. Montenegro, J. Rome, J. M. Steven, and S. C. Nicolson, "Arterial and venous contributions to near-infrared cerebral oximetry," *Anesthesiology* **93**(4), 947–953 (2000).
34. F. Y. Wong, T. Alexiou, T. Samarasinghe, V. Brodecky, and A. M. Walker, "Cerebral arterial and venous contributions to tissue oxygenation index measured using spatially resolved spectroscopy in newborn lambs," *Anesthesiology* **113**(6), 1385–1391 (2010).
35. F. Felici, V. Quaresima, L. Fattorini, P. Sbriccoli, G. C. Filligoi, and M. Ferrari, "Biceps brachii myoelectric and oxygenation changes during static and sinusoidal isometric exercises," *J. Electromyogr. Kinesiol.* **19**(2), e1–e11 (2009).

36. K. J. Kek, R. Kibe, M. Niwayama, N. Kudo, and K. Yamamoto, "Optical imaging instrument for muscle oxygenation based on spatially resolved spectroscopy," *Opt. Express* **16**(22), 18173–18187 (2008).
37. N. Nagdyman, T. Fleck, S. Schubert, P. Ewert, B. Peters, P. E. Lange, and H. Abdul-Khaliq, "Comparison between cerebral tissue oxygenation index measured by near-infrared spectroscopy and venous jugular bulb saturation in children," *Intensive Care Med.* **31**(6), 846–850 (2005).
38. G. Buunk, J. G. van der Hoeven, and A. E. Meinders, "A comparison of near-infrared spectroscopy and jugular bulb oximetry in comatose patients resuscitated from a cardiac arrest," *Anaesthesia* **53**(1), 13–19 (1998).
39. M. S. Ali, M. Harmer, R. S. Vaughan, J. A. Dunne, and I. P. Latta, "Spatially resolved spectroscopy (NIRO-300) does not agree with jugular bulb oxygen saturation in patients undergoing warm bypass surgery," *Can. J. Anaesth.* **48**(5), 497–501 (2001).
40. C. D. Kurth, J. C. McCann, J. Wu, L. Miles, and A. W. Loepke, "Cerebral oxygen saturation-time threshold for hypoxic-ischemic injury in piglets," *Anesth. Analg.* **108**(4), 1268–1277 (2009).
41. X. Hou, H. Ding, Y. Teng, C. Zhou, X. Tang, S. Li, and H. Ding, "Research on the relationship between brain anoxia at different regional oxygen saturations and brain damage using near-infrared spectroscopy," *Physiol. Meas.* **28**(10), 1251–1265 (2007).

---

## 1. Introduction

Cerebral tissue oximetry by near-infrared spectroscopy (NIRS) is a growing monitoring modality during surgery and intensive care. Accumulating evidence suggest that cerebral NIRS could have prognostic value [1,2], however clinical utility remains to be proven. Concerns have been raised about the validity and precision of the technology [3,4]. Here we introduce a practical method to compare the most important characteristics of tissue oximetry and apply it to 3 commercial NIRS instruments and a prototype.

The preterm infant is born in a state of hemodynamic instability hence reliable absolute values are needed. The values of regional tissue hemoglobin saturation ( $StO_2$ ) from different cerebral oximeters are correlated [5], but the absolute values differ systematically. Dullenkopf et al. found that there was a mean difference of  $11.3\% \pm 5.4$  and  $13.8 \pm 7.9$ , when the pediatric sensor of INVOS 5100 was compared to the adult sensor and the NIRO 300, respectively [6]. Furthermore, poor reproducibility is an important concern when applying NIRS in a clinical setting [4]. Using NIRO 300, the within-infant standard deviation ( $S_w$ ) was 5.2% with repeated cerebral measurements in preterm newborns [7]. Using a different instrument, Critikon 2020, the  $S_w$  was only 1.7% on the neonatal head [8], but this instrument also underestimated changes in concentration of hemoglobin [9]. Thus in order to assess a instrument for clinical use the absolute  $StO_2$  values, reproducibility, and dynamic range need to be tested.

This study presents the first step in a series of studies of NIRS oximeters. While clinical research in neonatology is difficult and test of dynamic range impossible other methods of testing are needed. Our overall goal is to facilitate the process of collecting good evidence of the benefits and harms of cerebral oximetry. Knowledge of the instruments is important. We chose three different continuous wave instruments, that have previously been used in neonatal studies: INVOS 5100 (Somanetics, Troy, Michigan, USA), NIRO 300 and NIRO 200 NX (Hamamatsu Phototonics, Hamamatsu City, Japan) [10–12] and the prototype OxyPrem (Biomedical Optic Research Laboratory, Zurich, Switzerland)—a newly developed instrument for neonates.

We propose a simple and practical setup to test and compare the absolute values, reproducibility, and dynamic range of  $StO_2$  by different instruments. This has to our knowledge not been done before. As these instruments use different source-detector separations from 2 to 4 cm, we also studied the effect different source-detector separations with the NIRO 200 NX. All measurements were done on the forearm of healthy adults.

## 2. Methods

This project was approved by the local Ethics Committee (Journal no. H-1-2010-055) and the National Board of Health (Sundhedsstyrelsen) (Journal no. 8313-75). Six subjects participated and written informed consent was obtained.

Inclusion criteria were a skinfold thickness of less than 10 mm on the lower arm as measured by the Harpenden caliper. Exclusion criteria were pregnancy and chronic disease.

The mean  $\text{StO}_2$  values and reproducibility of measurements under steady state were tested with subjects placed in upright position with the lower arm at heart level. The sensors were placed at ten appropriate, but slightly different sites on the upper part of the flexor muscles of the lower arm (Fig. 1), and ten measurements were done with each instrument before the next instrument was applied. After 30 seconds of stabilization a recording of one minute was made. The sensors were held in place by hand and great care was taken not to induce pressure on the tissue. During the resiting of each instrument the NIRO 200 NX measured continuously in order to assess the spontaneous fluctuations in  $\text{StO}_2$ .

Testing of dynamic range was done inflating a cuff around the upper arm at heart level combined with exercise. The sensors of all four instruments were simultaneously placed on four discrete positions of the lower arm (Fig. 1). We made sure that light source-light detector distance between adjacent instruments was at least 6 cm. All sensors were positioned so that the light path was parallel to the longitudinal direction of the muscles below. Sensors were fixated with self-adhesive wrap (Coban LP). Then 30 seconds of maximal force was made alternating isometric flexion and extension of the wrist against a resistance provided manually by SHS, followed by arterial occlusion by cuff inflation to 250 mmHg. The cuff was deflated after 90 seconds. This routine was done three times after which the sensors were rotated and measurements were repeated. We aimed for the same positions throughout when rotating the sensors. The result was 12 dynamic measurements for each instrument on each subject.

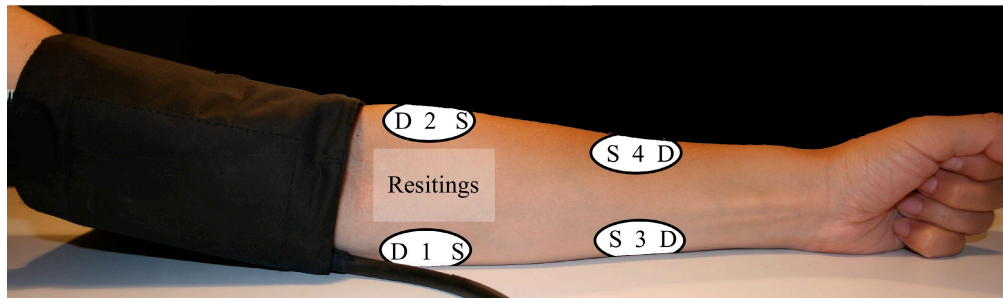


Fig. 1. Illustration of sensor positions during resittings (translucent white rectangle) and dynamic measurements (four opaque white ovals with position number). S, source. D, detector.

In a setup similar to the dynamic range study, the distance between the sensor and light source was altered so that 3 times 3 measurements were done with distances of 2.2 cm, 3.0 cm and 4.0 cm, respectively. These measurements were all done on the flexor muscles corresponding to position one on Fig. 1. Only NIRO 200 NX was used, because the design allowed putting the light source and detector closer than intended by the manufacturer. The distance 2.2 cm is not recommended by Hamamatsu. Empirical data suggest that the differential path length factor is not constant below distances of 3 cm [13,14].

### 3. NIRS instruments

The NIRO 300 and 200 NX (Hamamatsu) use spatially resolved spectroscopy. It is based on solution of the diffusion approximation for a highly scattering semi-infinite homogeneous media. By measuring the decrease in reflected light as a function of distance the effective light attenuation coefficient can be estimated. Then by assuming wavelength dependence of the reduced scattering coefficient the spectral shape of the absorption coefficient can be calculated and tissue saturation estimated [15]. The distance between the two sensors is 0.8 cm in both instruments. NIRO 200 NX uses LED (735, 810, and 850 nm) while NIRO 300 uses laser diodes (775, 810, 847, and 919 nm). In this study we used a source-detector

separation of 4 cm with the NIRO 300 and 3 cm with NIRO 200 NX unless specified otherwise.

INVOS 5100 gives a regional tissue oxygenation ( $SO_2$ ). It uses LED (730 and 810 nm). It applies two source-detector distances, 3 and 4 cm. The algorithm is not known to the public, but supposedly it “subtracts” the short distance signal from the longer distance in order to diminish the contribution of the skin and scalp [16]. In this study we used the adult sensor.

The OxyPrem was designed and constructed in the Biomedical Optic Research Laboratory in Zurich, Switzerland. The design is dedicated to preterm brain. It uses two wavelengths, 760 nm and 870 nm, four sources and two detectors. The pair-wise source-detector separation is 1.5 and 2.5 cm. The algorithm involves an auto-calibration that benefits from the multiple light paths [17]. In this prototype two simultaneous values of  $StO_2$  are calculated – one for each detector. The mean of the two is presented here.

We will refer to the regional tissue oxygenation as  $StO_2$  irrespective of instrument.

## 4. Statistical analysis

### 4.1. Reproducibility

The distribution of the data was explored by Shapiro-Wilk’s test for normality. The data was normally distributed. Overall mean  $StO_2$  and standard deviation for each subject and each instrument was determined. A univariate ANOVA with  $StO_2$  as the dependent and instrument as fixed factor and subject as random factor was done, if significant variation was revealed pairwise comparisons of the means for each instrument with post hoc Tukey HSD test would be done. Reproducibility for each instrument was determined by one-way ANOVA with subject as the factor. The within-subject standard deviation,  $S_w$ , was then estimated from the square root of the residual mean square. The 95% confidence interval was calculated as suggested by Bland and Altman [18]. The within-subject standard deviations for each instrument were pair-wise compared with the variance ratio test (F-test). The Bonferroni method was applied due to multiple testing. For spontaneous variation in  $StO_2$  ten randomly chosen, not overlapping, one minute timeframes from the continuous NIRO 200 NX reading would be analyzed.

### 4.2. Dynamic range

The dynamic range was calculated as the difference between the pre-exercise steady state oxygenation and the deoxygenated steady state. In a Matlab® (Mathworks) script a period of 10 seconds just before exercise would be chosen by visual assessment. The script would then calculate the mean  $StO_2$  of those 10 seconds and for another 10 seconds exactly 80 seconds afterwards, which would typically, be at the lowest  $StO_2$  levels. This method was chosen because the  $StO_2$  would quite often not find a steady state at the lowest values of  $StO_2$ , but be sloping slightly downwards or upwards.

The  $StO_2$  was not normally distributed for subject, position, or instrument. The median  $\Delta StO_2$  was calculated for each instrument and the nonparametric independent samples Kruskal-Wallis test was applied with  $\Delta StO_2$  as the dependent and instrument as the independent factor. If showing significant variation post-hoc pairwise comparisons were done by the SPSS nonparametric test and p-values adjusted according to the Bonferroni method. Similar statistics were applied to the study of changing source-detector separations.

All statistics were calculated in SPSS 18.0;  $p < 0.05$  is considered significant.

## 5. Results

### 5.1. Population

Six subjects, five men and one woman participated. They had a double skinfold mean of 4.1 mm and were from 19 to 30 years of age. The study of source-detector separation was done with another 6 male subjects, age range 30–59 years, with double skinfold mean of 3.9 mm.

## 5.2. Mean StO<sub>2</sub>

The mean StO<sub>2</sub> for each instrument of the ten one minute measurements in the six subjects is shown in Table 1. INVOS 5100 had the highest mean StO<sub>2</sub> of 70.1 ± 6.7%, OxyPrem the lowest StO<sub>2</sub> of 60.8 ± 3.6% (Fig. 2). The univariate ANOVA showed significant difference between means of instruments. In pairwise comparison significant difference was found between all instruments (p < 0.00001), except NIRO 200 NX and INVOS 5100. The mean StO<sub>2</sub> averaging all four instruments on each subject (n = 40) varied significantly (p < 0.003), range 61.8 ± 6.0% to 68.9 ± 6.3%.

**Table 1. Mean regional tissue oxygenation StO<sub>2</sub> of each instrument<sup>a</sup>**

Instrument	INVOS 5100	NIRO 200 NX	NIRO 300	OxyPrem
StO <sub>2</sub> ± sd (%)	70.2 ± 6.7	69.4 ± 5.0	63.4 ± 4.5*	60.8 ± 3.6*

<sup>a</sup>The mean of ten measurements in each of six subjects: n = 60.

\*Significantly different from the other three devices (p < 0.00001)

## 5.3. Reproducibility

The within-subject standard deviation for each instrument is presented in Table 2. Ranging from 5.4% (CI 4.4–6.9) in INVOS 5100 to 2.7% (CI 2.2–3.2) in OxyPrem. In pairwise comparison after correction for multiple testing the two NIROs and the INVOS had similar S<sub>w</sub>, while OxyPrem had a significantly lower S<sub>w</sub> than the other three instruments (p < 0.01). Of the total variance S<sub>w</sub> accounted for 61%, 82%, 71%, and 53% with INVOS, NIRO 200 NX, NIRO 300, and OxyPrem, respectively. The spontaneous fluctuations represented by S<sub>w</sub> of randomly selected 1-minute periods measured by NIRO 200 NX was 1.5% (CI 1.2–1.8)

## 5.4. Dynamic Range

The median ΔStO<sub>2</sub> of INVOS 5100, NIRO 200 NX, and NIRO 300 was 45.0%, 46.8%, 44.8%, respectively. The median ΔStO<sub>2</sub> of the OxyPrem was 27.8%, which was significantly lower than the other three devices (p < 0.001) (Table 2)(Fig. 2). A typical tracing of StO<sub>2</sub> of the NIRO 200 NX is given in Fig. 3. There was missing data from all 4 devices—ranging from 6 measurements with NIRO 300 to 1 measurement with INVOS 5100. The ΔStO<sub>2</sub> was dependent on sensor position. Position one and two on Fig. 1 gave overall ΔStO<sub>2</sub> mean values of 49.4% (CI 44.9–54.0) and 49.0% (CI 45.6–52.3), respectively, while position three and four gave mean values of 26.3% (CI 22.9–29.6) and 33.1% (CI 29.6–36.7).

## 5.5. Source-detector separation

The median ΔStO<sub>2</sub> was 68.9% (quartile range (IQR) 9.6) with 2.2 cm, 63.1% (IQR 8.5) with 3.0 cm, and 45.9% (IQR 6.6) with 4 cm source-detector separation, respectively. The median ΔStO<sub>2</sub> varied significantly across the different source-detector separations (p < 0.001). It was not possible to measure StO<sub>2</sub> during de-oxygenation with a distance of 4 cm in three of the subjects due to “signal underflow.”

**Table 2. Reproducibility S<sub>w</sub> and dynamic range ΔStO<sub>2</sub> for each instrument<sup>a</sup>**

Instrument	S <sub>w</sub> (%)	SE (%)	95% CI (%)	ΔStO <sub>2</sub> (%)	IQR (%)
INVOS 5100	5.4	0.5	4.4–6.5	45.0	29.7
NIRO 200 NX	4.4	0.4	3.5–5.2	46.8	29.4
NIRO 300	4.1	0.4	3.3–4.9	44.8	40.1
OxyPrem	2.7*	0.3	2.2–3.2	27.8**	10.3

<sup>a</sup>S<sub>w</sub>, Ten resittings (m) in six subjects (n): mn = 60. SE, Standard error. ΔStO<sub>2</sub>: Three repeated measurements on four different sites in six subjects: n = 72.

\*Significantly different from the other three instruments (p < 0.01).

\*\*Significantly different from the other three instruments (p < 0.001).

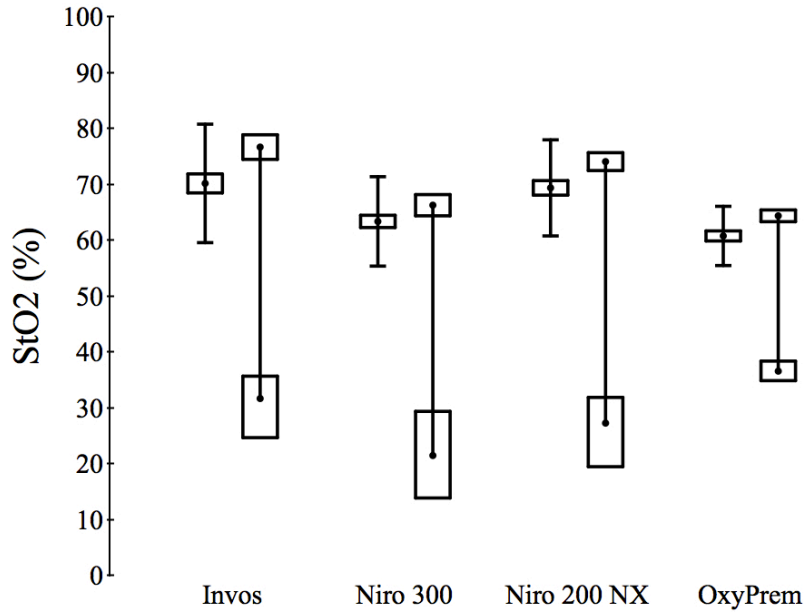


Fig. 2. For each instrument: On the left side the mean StO<sub>2</sub> with 95% CI (box) and 95% reference range. On the right side dynamic range with mean StO<sub>2</sub> pre- and post-exercise with 95% CI (boxes)

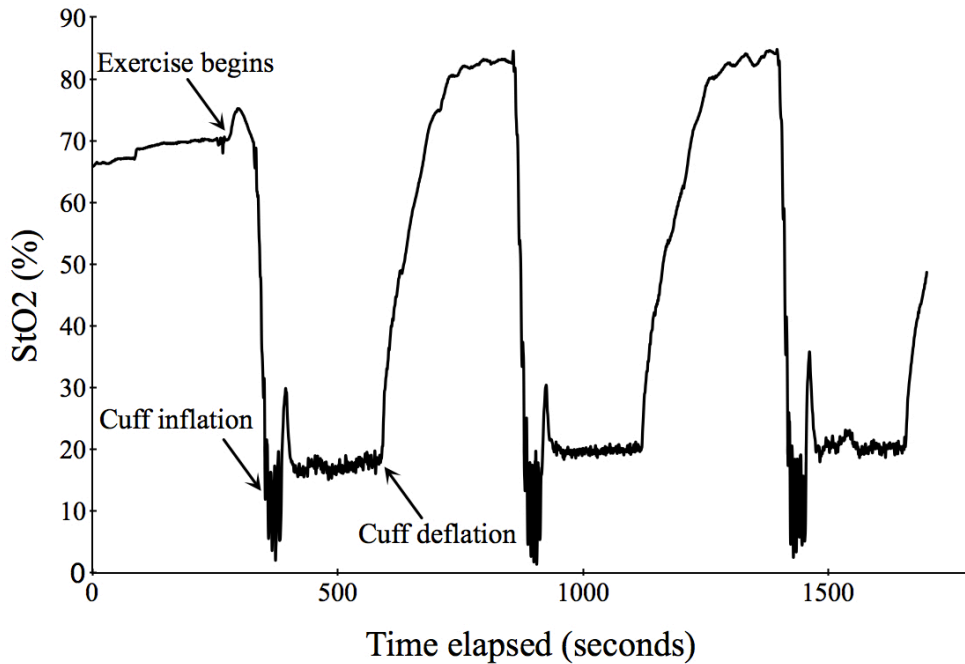


Fig. 3. A tracing of StO<sub>2</sub> with the NIRO 200 NX during exercise and arterial occlusion

## 6. Discussion

### 6.1. Statement of principal findings

We found that the steady state mean StO<sub>2</sub> was similar between NIRO 200 NX and INVOS 5100, while the mean StO<sub>2</sub> of NIRO 300 was significant less. The NIRO 200 NX, NIRO 300,

and INVOS 5100 had similar reproducibility and dynamic range on the human forearm. The OxyPrem had better reproducibility, the lowest steady state mean, and a dynamic range of only 60% of the other instruments.

### 6.2. Appraisal of methods

We used a simple, feasible setup to compare NIRS instruments. Numerous comparisons between different instruments have been reported [5,6,19–22]. The novel aspect of our approach is that we combine repeated resittings and dynamic measurements and estimate the reproducibility of several instruments in same setting. We propose that forearm muscle can serve as a model for the brain of the preterm infants, because the light has to travel through multiple layers offering comparable optical heterogeneity. We chose lean subjects to mimic the thin scalp and skull in preterm infants. It is likely that both during fluctuations in muscle and brain oxygenation, the oxygenation of the overlying skin and bone will vary less. Therefore the challenge of discriminating between superficial and more profound tissue is the same. Oxygen-hemoglobin saturation can be controlled more precisely in an in-vitro phantom, but it is difficult to create the optical heterogeneity of tissue, particular at the small scale, such as caused by hair, pigmentation, and blood vessels. Furthermore, instruments using different wavelengths will respond differently to a phantom using non-hemoglobin chromophores. We were able to achieve minimum values of  $\text{StO}_2$  below 10% by combining maximal exercise and arterial occlusion fast enough to allow multiple measurements. This provides a good signal to noise ratio and thus presents an effective setup for comparisons between different instruments. A similar dynamic range can only be achieved in-vivo for the brain by the use of animal models. Animal models are anatomically different and they are less easily available. For comparison, using the human adult forearm entails no risk and minimal inconvenience.

Some limitations of our approach deserve notice. 1) It is uncertain how absolute values on the arm agree with values on the head. Important differences in geometry, optical properties, and water content between arm and brain are likely to be processed differently by different instruments with different algorithms. 2) The testing of dynamic range had an inherent source of error. After the first de-oxygenation the subsequent hyper-perfusion would result in larger  $\Delta\text{StO}_2$  during the next 11 measurements. We compensated for this by varying the order of the instruments in each subject. 3) It is possible that the instruments probe bone especially at position 3 and 4 on Fig. 1. How this influences the dynamic measurements is uncertain, but it could introduce considerably bias depending on algorithm and source-detector distance. 4) The  $\Delta\text{StO}_2$  was highly dependent on sensor position, i.e., ranging from a mean  $\Delta\text{StO}_2$  of 49.4% to 26.3%. 5) The  $\Delta\text{StO}_2$  would occasionally differ more than 40 percentage points between repeats. This can probably be explained by minor displacement of the sensors during the exercise.

Despite these limitations it seems reasonable that a better reproducibility on the forearm translates to a better reproducibility on the head. And if the OxyPrem is significantly less sensitive to changes in oxygenation on the lower arm it is likely that this translates into less sensitivity to changes in the brain. Thus we suggest that clinical relevant information can be obtained from this test model.

### 6.3. Comparison with previous work

The mean resting  $\text{StO}_2$  in this study compares variably with previous studies. Komiyama et al. found a mean of  $65.6 \pm 5.6\%$  ( $n = 18$ ) with NIRO 300 in the flexor digitorum muscle on the lower arm [23]. This is similar to our results with NIRO 300, i.e. mean difference (MD) 2.2 (Confidence interval (CI)  $-0.4$ – $4.8$ ). Rodriguez et al. used the INVOS 5100 on the medial forearm and measured a mean  $\text{StO}_2$  of  $67.3 \pm 7.1\%$  ( $n = 50$ ). Compared with our result with the INVOS there is a significant difference with a MD of 2.9 (CI  $0.2$ – $5.6$ ). Shiroishi et al. used the NIRO 200 on the flexor muscles of the forearm and found a resting  $\text{StO}_2$  of  $61.4 \pm$



4.5% (n = 10) [24]. This is significant different from our results with the NIRO 200 NX (MD 8.0 CI 4.6 –11.4), but similar to our results with the NIRO 300 (MD 2.0 CI –1.1–5.1). A possible explanation for this varied agreement with literature could be different populations. For example the thickness of the skin, that may influence the mean values, is not documented in most of the studies [25].

As previously reported, our study confirms, that instruments and sensors differ in StO<sub>2</sub> values [6,20]. The difference in source-detector separation could be a possible explanation. It is still debated to what extent the more superficial tissues influence the measurement, e.g. Germon et al found a significant effect of varying the extra cranial perfusion [26] while others report negligible contributions from the extra cranial tissues [27,28]. The skin is believed to have a lower oxygen extraction than muscle. In spatially resolved spectroscopy the depth of measurement is correlated with the source-detector separation. The wider the distance is, the deeper the measurement. However all our study subjects had very thin subcutaneous layer with double skinfolds less than 5 mm, which obviously diminishes the skin's contribution to the signal. The possible probing of bone are less likely during these measurements since the mean values were collected in the area of resitings in Fig. 1. Thus it seems reasonable to assume that the StO<sub>2</sub> differences between the different instruments are primarily caused by other factors, most likely the different algorithms, although surprisingly there was significant difference between the two NIROs. NIRO 200 NX with the shortest source-detector separation giving the highest mean value.

To our knowledge the reproducibility of NIRS StO<sub>2</sub> measurements has not been evaluated on the lower arm before. The OxyPrem showed lowest S<sub>w</sub>, which contributed least to the total variance. However the estimates of between subject variance are imprecise due to the limited sample size. In neonates a very recent study with an earlier version of the OxyPrem found a similar S<sub>w</sub> of 2.76% [29]. In addition Sorensen et al. found a S<sub>w</sub> of 5.2% (CI 4.7–5.7) with NIRO 300 [7]. Comparable to this we measured a S<sub>w</sub> of 4.1% (CI 3.3–4.9) using NIRO 300. Similarly Dullenkopf et al. found that measurements with NIRO 300 on the same location of the scalp by two identical sensors 1 minute apart showed limits of agreement from –17.8% to 17.6% in a mixed population of neonates and pediatric patients corresponding to a S<sub>w</sub> of 6.4% (CI 4.4–8.4), again comparable to our results [30,31]. In a recent study by Pocivalnik et al. the NIRO 300 was compared with INVOS 5100 with 5 repeated measurement on the head of neonates, predominately preterms. They found comparable reproducibility, but a significantly higher mean StO<sub>2</sub> with the INVOS neonatal sensor 84.1% ± 6.4 vs. 73.1% ± 6.9 with NIRO 300 [32]. Higher values of StO<sub>2</sub> when using the neonatal INVOS sensor has previously been shown [6]. Pocivalnik et al. estimated the reproducibility by the mean standard deviation, which is not comparable to the within-subject standard deviation. In conclusion our results compares well with studies of reproducibility of cerebral StO<sub>2</sub> in neonates.

Does the lack of agreement and reproducibility represent optical heterogeneity, physiological fluctuations, and/or regional differences in oxygenation? Watzman et al. found that the arterial-venous ratio differed between subjects ranging from 0:100 to 40:60 [33]. However the study did not take the possible lack of precision of the used prototype NIRS oximeter into consideration and the variation could be the result of optical heterogeneity. Wong et al. suggests that there could be a within-subject variation in arterial-venous ratio. In lambs, they found that during hypoxaemia the arteries contribute more to the signal indicating possible arterial vasodilatation [34]. This makes sense, but is based on the assumption that the accuracy of NIRS is independent of level of oxygenation. It, however, cannot explain the steady state imprecision/variation we found. Moreover studies where sensors were repositioned at different sites [7] and studies where they were repositioned at the same site [30] have shown similar reproducibility. During our reproducibility study we had a continuous StO<sub>2</sub> measurement with the NIRO 200 NX. Ten randomly chosen 1-minute periods representing pseudo-resisting and thereby quantifying spontaneous fluctuations in oxygenation gave a S<sub>w</sub> of 1.5% (CI 1.2–1.8). This indicate that most of the variation, that we

found were not due to spontaneous fluctuations. We therefore believe that the lack of reproducibility of cerebral oximetry is primarily caused by optical heterogeneity.

Our test of dynamic range is a measure of sensitivity to changes in tissue oxygenation. To our knowledge this way of comparing instruments has not been used before. The changes in oxygenation achieved compare reasonably well with other dynamic muscle studies using NIRS oximeters [35,36]. The three commercially available instruments showed similar performance, indicating equivalent sensitivity to changes in muscle oxygenation. The OxyPrem was far less sensitive to oxygenation changes and  $StO_2$  rarely went below 30%. We altered the source-detector separation on the NIRO 200 NX to test whether the shorter distance on the OxyPrem could explain our results. Our data does not support this hypothesis.

Our results indicate that NIRO 300, NIRO 200 NX, and INVOS 5100 are equally suited for simple dynamic muscle measurements, if only changes in  $StO_2$  are needed. The steady state reproducibility and absolute  $StO_2$  values are not that important in studies of muscle physiology, but matters highly in clinical care. The OxyPrem did show better reproducibility, but perhaps at the expense of good sensitivity. However the intention of testing cerebral NIRS oximeters for neonatal clinical care on the adult arm should be validated in a suitable animal model.

#### 6.4. Validity

The cerebral  $StO_2$  correlates with venous jugular bulb saturation, but the agreement is poor with a mean bias about -4% and limits of agreement from -20% to +12% [37-39]. These studies are done without analysis of the within-subject agreement. It could be that the lack of agreement is primarily caused by optical heterogeneity. Perfect agreement is however unlikely, since jugular bulb saturation is flow-weighted, while cerebral oximetry is volume weighted. During states of microcirculatory heterogeneity, like sepsis, a lot of vessels in the brain will not be perfused while others will be hyperperfused. This will theoretically affect jugular bulb saturation less than NIRS oximetry that also will measure the deoxygenated pool of blood in the underperfused vessels. It remains to be determined what modality is the best predictor of cerebral outcome, but NIRS has the great advantage of non-invasiveness.

#### 6.5. Clinical implications

Two problems with cerebral oximetry are highlighted by our results: the difference in absolute  $StO_2$  values between instruments and the lack of reproducibility. The varying absolute values between instruments causes the determination of normal ranges and thresholds for hypoxic-ischemic brain injury to be problematic. From studies in piglets  $StO_2$  levels below 35-40% has been damaging to the brain [40,41], but since different non-commercial instruments are used in these studies, it is not simple to translate this into clinical care, where other instruments are available. During surgery it has become standard to use changes from pre-operative  $StO_2$ , but in neonatal care where infants are born in a state of cardio-pulmonary instability, no such set point exist. The reliability of the absolute values is therefore important. The OxyPrem data shows that such reproducibility can be achieved. However since the OxyPrem was less sensitive to changes in oxygenation the ability to distinguish between deleterious low and safe normal  $StO_2$  might not be better than with the commercial instruments. Estimating the signal/noise ratio ( $range/S_w$ ) gave 8.2, 10.8, 10.9 and 10.2 with INVOS 5100, NIRO 200 NX, NIRO 300 and OxyPrem, respectively. This suggests, with the caveat that this is a less than perfect model for the neonatal brain, that the instruments are equally suited for clinical purposes.

Optical heterogeneity is probably the primary cause of the poor reproducibility of the commercial instruments and a possible solution could be to use 2 channels at a time and work with the mean of these. The alternative, to do repositioning manually is too cumbersome outside a trial setting. The clinical potential of cerebral NIRS is still unclear, but further

development of the technology to improve the precision and reproducibility, while maintaining good sensitivity seems paramount.

#### *6.6. Conclusion*

We compared three commercial NIRS instruments on the forearm of adults and found that the absolute values of StO<sub>2</sub> differ between instruments, while reproducibility and dynamic range were similar. The results with the prototype OxyPrem suggest that the reproducibility of cerebral NIRS can be improved.

#### **Acknowledgments**

Preparation of this manuscript was supported by grants from the Ludvig and Sara Elsass Foundation. We thank Hamamatsu for the loan of NIRO 200 NX.