

Reproducibility, interchangeability of measures, time to measure stabilization, and reference values of two tissue oximeters in healthy volunteers

Alexis Cournoyer
André Denault
Sylvie Cossette
Annik Fortier
Raoul Daoust
Massimiliano Iseppon
Jean-Marc Chauny
Eric Notebaert

Reproducibility, interchangeability of measures, time to measure stabilization, and reference values of two tissue oximeters in healthy volunteers

Alexis Cournoyer,^{a,*} André Denault,^b Sylvie Cossette,^c Annik Fortier,^d Raoul Daoust,^a Massimiliano Iseppon,^a Jean-Marc Chauny,^a and Eric Notebaert^a

^aUniversité de Montréal, Faculty of Medicine, Department of Family Medicine and Emergency Medicine, 5400 Gouin Ouest, Montréal H4J 1C5, Canada

^bUniversité de Montréal, Faculty of Medicine, Department of Anesthesia and Critical Care Division, 5000 Bélanger, Montréal H1T 1C8, Canada

^cUniversité de Montréal, Faculty of Nursing, 2375 Côte-Ste-Catherine, Montréal H3T 1A8, Canada

^dMontréal Health Innovations Coordinating Center, 4100 Molson #400, Montréal H1Y 3N1, Canada

Abstract. This study aimed to compare two tissue oximeters, the INVOS 5100c and the Equanox 7600, in terms of their reproducibility and the interchangeability of their measures. In a randomized order, three measurements were taken at six different sites on both sides of the body in 53 healthy volunteers. Intraclass correlation coefficients (ICC) and within-subject standard deviation (S_w) were calculated for each device. The ICCs were compared using Fisher r-to-z transformation and the S_w were compared using paired-sample t -tests. We found no difference between the reproducibility of the INVOS [ICC = 0.92 [95% confidence interval (CI) 0.90 to 0.93]] and Equanox [ICC = 0.90 (95% CI 0.88 to 0.93)] in terms of ICCs ($p = 0.06$). However, the Equanox [$S_w = 1.96$ (95% CI 1.91 to 2.02)] showed a better S_w than the INVOS [$S_w = 2.11$ (95% CI 2.05 to 2.17)] ($p = 0.019$). Also, when compared directly to stable condition, the readings produced by the two oximeters varied considerably [ICC 0.43 (95% CI 0.36 to 0.49)]. When taken individually, both tissue oximeters displayed good reproducibility, the Equanox being slightly better than the INVOS in terms of absolute reproducibility. However, when compared, the oximeters showed poor interdevices agreement. Reference values were also described. © 2016 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JBO.21.9.097003]

Keywords: near-infrared spectroscopy; tissue oximetry; reproducibility; interdevice agreement; reference values.

Paper 160384R received Jun. 8, 2016; accepted for publication Aug. 30, 2016; published online Sep. 16, 2016.

1 Introduction

Patients presenting hemodynamic instability should be aggressively managed without delay.^{1,2} This often involves the use of invasive monitoring tools, such as arterial lines or central venous catheters, in order to monitor tissue perfusion. Although they provide useful information, these tools are often unavailable and require time to install. During resuscitation, the utility of laboratory values is blunted by the fact that they provide static information on a continuously evolving clinical process. The time required to access these values can result in outdated information that provides little real-time feedback on the efficacy of the resuscitation. Near-infrared spectroscopy (NIRS) is a non-invasive, continuous and painless method of monitoring the oxygen saturation of hemoglobin in any given superficial tissue (tissue saturation) and, as such, can be used as a way of detecting tissue hypoperfusion. Unlike pulse oximetry, which only measures the arterial saturation of hemoglobin, NIRS measures the saturation in all vessels smaller than 1 mm (arterioles, capillaries, and venules).³ NIRS devices can also provide saturation measures in situations of circulatory arrest, as they do not depend on pulsation to provide its readings. Tissue perfusion (and saturation) is affected in situations of hemodynamic instability.⁴⁻⁶ Thus, the quantification of tissue saturation by NIRS oximetry could prove to be an interesting surrogate

measure for tissue perfusion and a useful tool in guiding resuscitation in settings where more invasive tools are unavailable such as when out of the hospital or in the emergency room. They can also prove to be an interesting addition to other specialties of critical care, such as anaesthesiology and the intensive care.

NIRS technology was developed after an experiment by Jöbsis, which showed that light spectrum could be used as a way to monitor patients. This is based on two unique properties. The first is that near-infrared photons (i.e., photons with a 700 to 1300 nm wavelength) have the capacity to go through organic tissues without being significantly absorbed. The second is that oxyhemoglobin and deoxyhemoglobin have different absorptions properties along this wavelength spectrum, allowing them to be differentiated.⁷ Using the Beer-Lambert law and an algorithm, NIRS devices calculate the total concentration of both these molecules, thus providing a value representing tissue saturation in oxygen (StO_2).

Although promising, there are still areas of uncertainty with this technology. One of these areas is the reproducibility of tissue oximeters, which has not been thoroughly studied. Current data has shown that the difference between two successive measurements can be quite large, although some oximeters have shown better reproducibility than others.⁸⁻¹² Also, newer tissue oximeters have only been studied with forehead and forearm

*Address all correspondence to: Alexis Cournoyer, E-mail: alexiscournoyer@gmail.com

readings in adults, other sites have been included in past clinical research.^{11,12}

The main objective of this study was to compare the reproducibility of two commonly used tissue oximeters, the INVOS 5100c (Covidien, Mansfield, Massachusetts) and the Equanox 7600 (Nonin Medical, Plymouth, Minnesota). For the primary outcome, we hypothesized that the Equanox would display a better reproducibility than the INVOS owing to its more recent development. The effects of sensor switching, measurement site and side of body (left versus right) on reproducibility were evaluated as secondary outcomes. The second objective was to describe the interchangeability of the oximeters and sides of the body to provide similar raw tissue saturation measurements. The third objective was to assess the time necessary for each of the oximeters to provide a stable value for tissue oximetry. Finally, the last objective was to describe and compare reference values for NIRS saturation for all sites evaluated.

2 Methods

2.1 Design and Setting

This prospective cohort study was conducted at the Hôpital du Sacré-Coeur de Montréal, in association with the Université de Montréal, Canada. It was submitted and approved by its Ethics Committee prior to enrolment.

2.2 Population and Sample

A convenience sample was comprised of healthy adult volunteers. Recruitment was conducted via snowball sampling, mainly among hospital employees between March and April 2015. Exclusion criteria included pregnancy, active or chronic systemic illness, the use of medication other than oral contraceptives, skin disease, very coarse hair overlying measurement sites, active smoking, substance use, weight less than 40 kg, or a body mass index lower than 18 or higher than 35. All recruited volunteers signed informed consent waivers prior to their inclusion in the study.

2.3 Material

The two tissue oximeters we used in this study both use the multi-distance spectroscopy technology. By using multiple light-emitting diodes, sensors at varying distances, and multiple wavelengths, they are able to measure variations in the absorption of the near-infrared photons.^{13,14} This allows them to blunt superficial tissue's contribution to their readings, better representing the target tissue (brain, muscle, and so on) saturation.

The INVOS 5100C with adult SomaSensor has two light-emitting diodes (730 and 810 nm) at one position and two receptors at distances of 30 and 40 mm. The EQUANOX 7600 with Equanox Advance 8004CA sensors uses eight light-emitting diodes split in groups of four (730, 760, 810, and 880 nm) at two positions separated by 60 mm. It uses two receptors placed between the emitting groups. Each receptor is 20 mm proximal to an emitting group and 40 mm distal to the other. The INVOS estimates that the mixed arterio-venous oxygen saturation it measures is fixed at 25:75, whereas the Equanox estimates it is 30:70.

2.4 Measures of Cerebral and Somatic Oximetry

Prior to measures being taken, participants were asked to provide their basic sociodemographic characteristics and their phototype, using the Fitzpatrick phototyping scale.¹⁵ At the beginning of the experiment, participants were asked to lie on their backs while their vital signs and skinfold thickness (using a Harpenden Skinfold Caliper) were measured. Six sites (forehead, deltoid, forearm, knee, calf, and foot) on both sides of the body were then marked with either a pen or tape to ensure that the same exact location was used for each sensor throughout the measure collection (cf., Fig. 1).

The order in which each tissue oximeter was used was randomized. The sensors were placed on each of the measurement sites until the reading was stable. A stable reading was defined

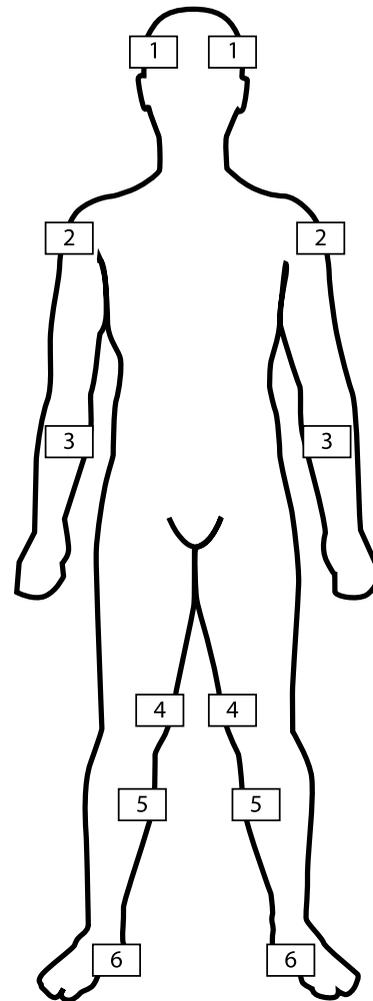


Fig. 1 Sites of tissue saturation measurements. 1. Forehead: above the eyebrow, long side perpendicular to the body axis and following the median line of the face. 2. Deltoid: at the height of the head of the humerus, long side perpendicular to the body axis and centered on the coronal plane of the body. 3. Forearm: distally to the elbow crease, centered on the palmar face of the forearm and long side perpendicular to the body axis. 4. Knee: on the medial side of the thigh at the base of the junction between the patella and the vastus medialis of the quadriceps, long side parallel to the leg axis and just medial to the patella. 5. Calf: on the medial line of the calf, 3 cm below the articular line of the knee and long side parallel to the leg axis. 6. Foot: on the medial line of the plantar face, long side parallel to the leg axis and proximal side just distally to the heel.

as an identical value in two consecutive measurements, measures alternating between two contiguous values, or a maximum delay of 20 s from the first reading. In the latter case, the last recorded value was retained. The time delay to stabilization was noted. If a measure was not provided within 30 s of sensor apposition, the sensor, cables, and connections were verified. Should the oximeter persist in its failure to provide a measure, its functionality was tested on the experimenter. In instances where the oximeter provided a measure on the experimenter but not on the subject, the data were considered as missing, and the experimenter proceeded to the next site with the same sensor. Values from various measurement sites were taken in a predetermined order (forehead, deltoid, forearm, knee, calf then foot) with a first sensor (s1). Once all sites had been evaluated on both sides of the body (t1), the process was repeated at 5 min (t2) a second time with the same sensor (s1) to evaluate the intrasensor reproducibility. After the second run, the procedure was repeated a third time at 10 min (t3) with a different sensor of the same oximeter model (s2) to assess the intersensor reproducibility. The same procedure was then used for the second oximeter. Because up to 24 measures would ultimately need to be taken with the same sensor, it was decided that the sensors would be held in place manually by the experimenter instead of using the adhesive, in order to avoid skin irritation. That being said, great care was taken neither to compress the sensors nor to expose them to ambient light during measurements, so as to avoid impacting the values measured. Every measure was either collected or supervised by the primary author (AC). This process was repeated for each of the two oximeters. A total of 72 measures were recorded for each subject in the study (six sites per side of the body and three values per site for each of the two oximeters).

2.5 Sample Size Determination

The measurement reproducibility for each oximeter was assessed using the intraclass correlation coefficient (ICC). Given the absence of literature on the power required for a test for the comparison of dependant ICCs based on two-way ANOVAs, we based our sample size calculation on the precision of a unidirectional confidence interval (CI) for a parametric ICC. An ICC is a correlation coefficient whose value varies between 0 and 1. It is used to assess reproducibility and depends not only on the differences observed between the repeated measurements, but also on the variation between pairs of the sample. An ICC over 0.75 is considered an indicator of good reproducibility while an ICC under 0.40 represents poor reproducibility. For an expected ICC value of 0.75, we concluded that 53 participants needed to be recruited in order to obtain a one-sided 95% CI (computed using the large sample normal approximation) for an ICC based on two measurements, given that the CI will extend about 0.1 toward the low values from the observed ICC. A good precision for the CI of the ICCs was assumed to be associated with good statistical power for the comparisons of ICCs.

2.6 Statistical Analysis

Collected data was held and managed at Hôpital du Sacré-Coeur de Montréal. Statistical analyses were done both at the aforementioned hospital and the Montréal Health Innovations Coordinating Center, using SAS 9.4 (SAS Institute Inc., Cary, North Carolina).

2.6.1 Main objective

Both intrasensor and intersensor ICCs and their respective 95% CI were calculated for each individual measurement site and for each side of the body, thus providing 24 ICC values (95% CI) for each oximeter. Overall intrasensor and intersensor ICCs (95% CI) were also calculated for each oximeter. To complement the analysis of reproducibility we calculated global intrasensor and intersensor within-subject standard deviation (S_w) (% and 95% CI), mean bias (%), and limits of agreement (%) for each oximeter using the Bland-Altman method.¹⁶ To compare our results with the currently available literature on the forearm site, we also calculated both oximeters forearm intrasensor S_w . Plots of Identity and Bland-Altman plots were used to illustrate the data. S_w are also a measure a reproducibility and represent the standard deviation of repeated measurements in an individual. Whereas variations within the sample influence the ICC (making it a marker of relative reproducibility), they do not influence the S_w , making it a marker of absolute reproducibility. A lower S_w represents a better reproducibility than a higher S_w . Mean bias represent the average difference between two repeated measurements. A mean bias close to zero is a sign of good concordance. Limits of agreement represent the range of values in which the difference between repeated measurements will be comprised 95% of the time. A smaller range of limits of agreement represent better reproducibility than a wider range of limits of agreement.

The primary outcome of the main objective was the comparison of the global intersensor ICCs between the INVOS and Equanox. We retained the intersensor ICC instead of the intrasensor ICC because it represents the reproducibility of the oximeter and its sensor combined. Comparisons of ICCs were done using the Fisher's r-to-z transformation method. All other ICC comparisons (INVOS versus Equanox intrasensor and intrasensor versus intersensor for both oximeters) were secondary outcomes. Also as secondary outcomes, S_w comparisons (INVOS versus Equanox intersensors, INVOS versus Equanox intrasensor and intrasensor versus intersensor for both oximeters) were done using paired sample *t*-tests.

Last, for the main objective, global ICCs, S_w , mean bias and limits of agreement were also calculated for each site of sensor application and side of the body. Comparisons of ICCs between sites of sensor application and between sides of the body were done using the Fisher's r-to-z transformation method while comparisons of S_w between sites of sensor application and between sides of the body were done using paired sample *t*-tests.

2.6.2 Second to fourth objectives

For these objectives, the average of the three values for both tissue saturation and time to measure stabilization obtained (c1t1, c1t2, and c2t3) on each of the 12 sites were used. For our second objective, we describe the interchangeability of the devices and sides of the body by comparing individual raw tissue saturation measurements. This comparison was done by calculating ICCs, S_w , mean bias, and limits of agreement presented as descriptive statistics.

For our third objective, we collected data on the time necessary for both oximeters to give a stable value and, for our fourth objective, we described the reference values for each site and side of the body for each oximeter. For comparisons of time to measure stabilization and reference values between the two oximeters, three-way repeated measure ANOVA were used.

3 Results

3.1 Baseline Characteristics

Fifty-three healthy volunteers were included in this study, with ages ranging between 20 and 81 years old. With the exception of oral contraceptives ($n = 14$), none took any medication on a regular basis. No patients enrolled had a chronic or active illness. Participant characteristics are presented in Table 1.

On 23 occasions [21 (1.1%) with the INVOS and two (0.1%) with the Equanox], despite otherwise functioning normally, the oximeters failed to provide a saturation value. No adverse reactions were observed during the course of the trial.

Table 1 Characteristics of the 53 healthy volunteers.

Variables	Mean (SD) or N (%)
Age, years	31.5 (9.6)
Gender, men	27 (51.9)
Height, m	1.69 (0.10)
Weight, kg	69.4 (12.4)
BMI, kg/m ²	24.2 (3.1)
Blood pressure	120/75 (11/9)
Pulse	72 (13)
Pulse oximetry	98 (1)
Fitzpatrick phototyping scale (15)	
I	4 (7.6)
II	24 (45.3)
III	20 (37.7)
IV	1 (1.9)
V	0 (0)
VI	4 (7.6)
Skinfold measurement location, mm	
Left deltoid	17.9 (7.6)
Right deltoid	17.9 (8.0)
Left forearm	6.8 (3.8)
Right forearm	6.9 (3.0)
Left thigh	16.4 (9.0)
Right thigh	16.8 (10.3)
Left calf	12.8 (6.9)
Right calf	12.1 (6.9)

Note: SD: standard deviation; BMI: body mass index.

3.2 Measures of Reproducibility

3.2.1 INVOS

Intersensor ICCs for the INVOS oximeter ranged from 0.78 to 0.91 while the intrasensor ICCs ranged from 0.76 to 0.96 (cf., Table 2). Global ICCs for the INVOS oximeter, as well as S_w , mean bias and limits of agreement are presented in Table 3, Figs. 2 and 3. The INVOS forearm intrasensor S_w was 2.97% (95% CI 2.77 to 3.18) (data not shown).

3.2.2 Equanox

Intersensor ICCs for the Equanox oximeter ranged from 0.73 to 0.94 while intrasensor ICCs ranged from 0.77 to 0.96 (cf., Table 4). Global ICCs for the Equanox oximeter, as well as S_w , mean bias and limits of agreement are presented in Table 5, Figs. 4 and 5. The Equanox forearm intrasensor S_w was 2.68% (95% CI 2.50 to 2.87) (data not shown).

3.2.3 Comparisons of reproducibility between the two oximeters

For the primary outcome of our main objective, we found no significant difference when comparing both oximeters in terms of global intersensor ICC using the fisher r-to-z transformation ($p = 0.06$) (cf., Tables 3 and 5). For the secondary outcomes, using the same method, we also found no significant difference between the two oximeters in terms of global intrasensor ICCs ($p = 0.78$). However, the intrasensor ICC of the Equanox was better than its intersensor ICC ($p = 0.003$), which was not the case for the INVOS ($p = 0.42$) (cf., Tables 3 and 5). Using the Bland-Altman method, the Equanox displayed a significantly better inter and intrasensor S_w than did the INVOS ($p = 0.019$ and $p < 0.001$, respectively). Again, the intrasensor S_w of the Equanox was better than its intersensor S_w ($p < 0.001$) while no such difference was found for the INVOS ($p = 0.28$) (cf., Tables 3 and 5 and Figs. 2–5). Both oximeters displayed adequate consistency across the range of observed results, as shown by the similarities displayed in dot dispersion around the reference line of the Bland-Altman plots (Figs. 3 and 5).

3.2.4 Comparisons of reproducibility between sides of the body and sites of measurement

Last, for the main objective, ICCs and S_w were similar when comparing both sides of the body ($p = 0.58$ and $p = 0.74$, respectively). When comparing ICCs, among all possible comparisons between sites, the knee site displayed better ICCs than all the other sites (p ranging from 0.005 to <0.001). Also, the deltoid and calf sites were superior to the forearm site ($p = 0.009$ and $p < 0.001$, respectively). Using the Bland-Altman method, measures from the knee and calf sites had better S_w than did all other sites (p -values ranging from 0.005 to <0.001). Furthermore, the forehead, deltoid, and foot had better S_w than the forearm site (p -values ranging from 0.002 to <0.001) (cf., Table 6).

3.3 Interchangeability of the Oximeters and Sides of the Body to Provide Similar Raw Tissue Saturation Measurements

For the second objective, tissue saturation measures provided by the two oximeters varied considerably [ICC 0.43 (95% CI 0.36

Table 2 Calculated measures of reproducibility of the INVOS oximeter (ICC and 95% CI).

	Intersensor		Intrasensor	
	Left	Right	Left	Right
Forehead	0.91 (0.84 to 0.95)	0.83 (0.72 to 0.90)	0.91 (0.86 to 0.95)	0.88 (0.79 to 0.93)
Deltoid	0.84 (0.73 to 0.90)	0.84 (0.69 to 0.91)	0.86 (0.76 to 0.92)	0.90 (0.83 to 0.94)
Forearm	0.78 (0.65 to 0.87)	0.86 (0.76 to 0.92)	0.81 (0.68 to 0.88)	0.78 (0.65 to 0.87)
Knee	0.92 (0.87 to 0.96)	0.94 (0.91 to 0.97)	0.92 (0.87 to 0.95)	0.92 (0.86 to 0.95)
Calf	0.89 (0.82 to 0.93)	0.91 (0.86 to 0.95)	0.93 (0.88 to 0.96)	0.82 (0.71 to 0.89)
Foot	0.88 (0.78 to 0.94)	0.88 (0.80 to 0.93)	0.96 (0.93 to 0.98)	0.92 (0.86 to 0.95)

Note: ICC: intraclass correlation coefficient; 95% CI: 95% confidence interval.

Table 3 Global measures of reproducibility of the INVOS oximeter.

		ICC and 95% CI	S_w and 95% CI (%)	Mean bias and limits of agreement (%)
INVOS	Intersensor	0.92 (0.90 to 0.93)	2.11 (2.05 to 2.17)	-0.85 (-9.11 to 7.41)
	Intrasensor	0.92 (0.91 to 0.93)	2.22 (2.16 to 2.29)	-0.35 (-8.42 to 7.72)

Note: ICC: intraclass correlation coefficient; 95% CI: 95% confidence interval; S_w : within-subject standard deviation.

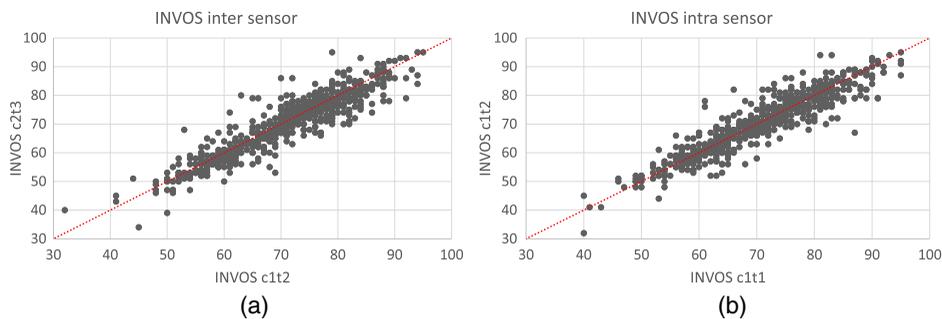


Fig. 2 Plot of identity describing the (a) intersensor and (b) intrasensor reproducibility of the INVOS oximeter.

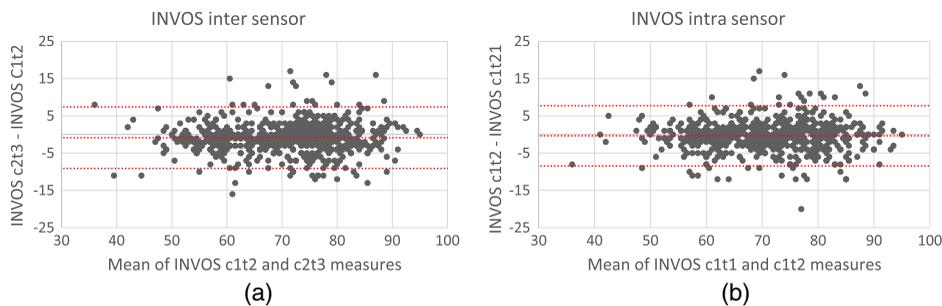


Fig. 3 Bland-Altman plot describing the (a) intersensor and (b) intrasensor reproducibility of the INVOS oximeter.

Table 4 Calculated measures of reproducibility of the Equanox oximeter (ICC and 95% CI).

	Intersensor		Intrasensor	
	Left	Right	Left	Right
Forehead	0.76 (0.62 to 0.85)	0.73 (0.57 to 0.84)	0.82 (0.71 to 0.89)	0.77 (0.63 to 0.86)
Deltoid	0.89 (0.91 to 0.93)	0.92 (0.86 to 0.95)	0.95 (0.92 to 0.97)	0.87 (0.78 to 0.92)
Forearm	0.78 (0.65 to 0.87)	0.77 (0.63 to 0.86)	0.82 (0.72 to 0.89)	0.79 (0.66 to 0.97)
Knee	0.94 (0.90 to 0.97)	0.93 (0.89 to 0.96)	0.96 (0.93 to 0.97)	0.95 (0.92 to 0.97)
Calf	0.92 (0.87 to 0.95)	0.90 (0.83 to 0.94)	0.95 (0.92 to 0.97)	0.93 (0.88 to 0.96)
Foot	0.86 (0.76 to 0.91)	0.82 (0.71 to 0.89)	0.91 (0.86 to 0.95)	0.88 (0.81 to 0.93)

Note: ICC: intraclass correlation coefficient; 95% CI: 95% confidence interval.

Table 5 Global measures of reproducibility of the Equanox oximeter.

		ICC and 95% CI	S_w and 95% CI (%)	Mean bias and limits of agreement (%)
EQUANOX	Intersensor	0.90 (0.88 to 0.91)	1.96 (1.91 to 2.02)	-0.39 (-7.78 to 7.00)
	Intrasensor	0.92 (0.91 to 0.93)	1.67 (1.63 to 1.72)	0.32 (-6.07 to 6.72)

Note: ICC: intraclass correlation coefficient; 95% CI: 95% confidence interval; S_w : within-subject standard deviation.

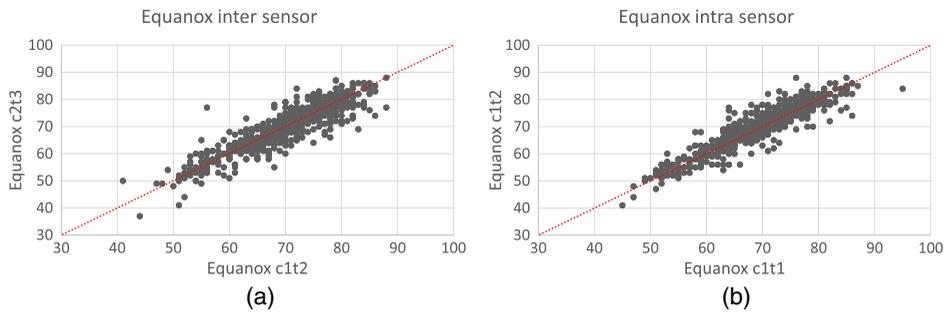


Fig. 4 Plot of identity describing the (a) intersensor and (b) intrasensor reproducibility of the Equanox oximeter.

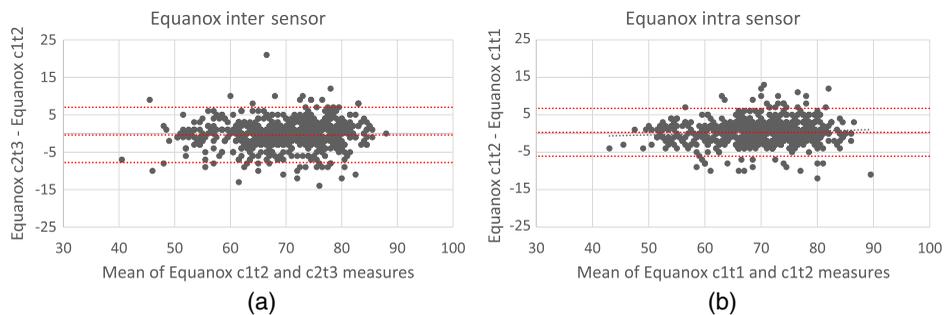


Fig. 5 Bland-Altman plot describing the (a) intersensor and (b) intrasensor reproducibility of the Equanox oximeter.

Table 6 Comparison of measures of reproducibility of two tissue oximeters between sides and sites.

		ICC and 95% CI	S_w and 95% CI (%)	Mean bias and limits of agreement (%)
Left	Intersensor	0.91 (0.90 to 0.92)	2.10 (2.05 to 2.16)	-0.49 (-8.43 to 7.45)
Right	Intersensor	0.91 (0.89 to 0.92)	2.08 (2.02 to 2.13)	-0.75 (-8.49 to 6.99)
Forehead	Intersensor	0.86 (0.83 to 0.89)	2.16 (2.05 to 2.26)	-0.81 (-9.06 to 7.43)
Deltoid	Intersensor	0.88 (0.84 to 0.91)	2.07 (1.97 to 2.17)	-0.84 (-8.41 to 6.73)
Forearm	Intersensor	0.81 (0.76 to 0.86)	2.86 (2.72 to 3.00)	-1.23 (-11.35 to 8.89)
Knee	Intersensor	0.94 (0.92 to 0.95)	1.54 (1.47 to 1.62)	-0.17 (-5.86 to 5.52)
Calf	Intersensor	0.91 (0.88 to 0.93)	1.72 (1.63 to 1.80)	0.27 (-6.36 to 6.89)
Foot	Intersensor	0.87 (0.83 to 0.90)	2.20 (2.09 to 2.30)	-0.94 (-8.72 to 6.83)

Note: ICC: intraclass correlation coefficient; 95% CI: 95% confidence interval; S_w : within-subject standard deviation.

to 0.49), Bias -0.77 (Limits of agreement -19.88 to 18.34]). When evaluated together, all tissue saturation measurements from the Equanox showed less variation (SD 8.0) than the ones from the INVOS (SD 10.2). Thus, in comparison to the INVOS' tissue saturation measurements, the Equanox generally displayed higher tissue saturations when in the lower saturations range (e.g., around 40% to 60%) and lower tissue saturations in the higher saturation range (e.g., around 75% to 90%). This can be appreciated in Fig. 6 in the Bland-Altman plot by noticing that the dot cloud seems to follow a negative linear regression. On the other hand, the tissue saturations varied much less when we compared the readings from the left and right sides of the body [ICC 0.91 (95% CI 0.89 to 0.92), Bias -0.34 (limits of agreement -7.94 to 7.25)] (cf., Fig. 7). Those results are similar to the ones previously described for inter- and intrasensor reproducibility (cf., Tables 3 and 5 and Figs. 2–5).

3.4 Time Necessary for Measure Stabilization

For the third objective, the time necessary for stabilization of the oximetry measures were calculated using mean values for each site in each of the study participants. The Equanox displayed a quicker time to stabilization than the INVOS [4.7 s (SD 1.4) versus 9.5 s (SD 3.3), $p < 0.0001$]. All other comparisons done (between measurement site and side) were not considered clinically significant (data not shown).

3.5 Reference Values

Four the last objective, reference values for each tissue oximeters calculated using the mean values of each site for each volunteer are presented in Table 7. They ranged from 58.6 (SD 8.8) to 77.0 (SD 7.4). Using a three-way repeated measures ANOVA (oximeter * side * site), small to no differences were observed between sides of the body. However, the differences were substantial when comparing sites and oximeters [e.g., INVOS left deltoid (77.0%) versus INVOS left foot (58.6%), $p < 0.0001$].

4 Discussion

The present study sought to compare the reproducibility of two commonly used tissue oximeters. Both oximeters exhibited good reproducibility. Interestingly, no differences were observed when ICCs were used to assess this metrological characteristic, but the Equanox displayed better reproducibility when the Bland-Altman method was used, displaying smaller S_w and limits of agreement. The reproducibility also varied among measurement sites, but not between each side of the body. Despite the good intraoximeter reproducibility, the interchangeability of raw saturation value between both oximeter was deficient, reflecting poor interoximeter agreement. To our knowledge, this is the first time that that both of these characteristics (reproducibility and interchangeability) were evaluated for tissue saturation at sites other than the forehead and

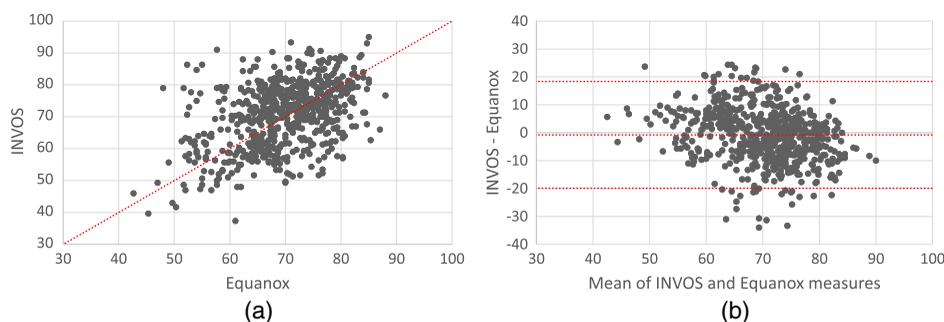


Fig. 6 (a) Plot of identity and (b) Bland-Altman plot describing the interchangeability among two tissue oximeters.

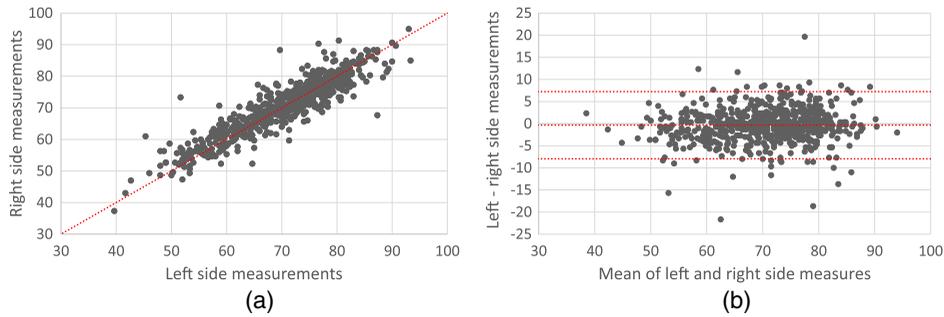


Fig. 7 (a) Plot of identity and (b) Bland-Altman plot describing the interchangeability between the oximetry readings of the left and right side.

the forearm, or for multiple sites of the body. The wide range of observed values in our experiment allowed a good appraisal of these characteristics. In addition, we showed the Equanox was faster at providing a stable saturation measurement than the INVOS was, a characteristic that had never been evaluated in these two devices. Finally, reference values for the two oximeters used were also described for the first time at sites most often currently used in clinical practice.

The Equanox showed better absolute reproducibility, as reflected by better S_w values than the INVOS, but similar relative reproducibility, as reflected by the similar ICC values. This is likely explained by the fact that, as explained earlier, the ICC is a correlation coefficient and therefore takes into account the variability and range of the observed measures, thus providing a relative measure while the S_w and limits of agreement are absolute measures of reproducibility. The Equanox also provided a smaller range of observed values across the measurement sites and sides of the body than did the INVOS. Based on these results, the Equanox could be favoured over the INVOS, especially when NIRS monitoring is used intermittently

or in unstable conditions since good reproducibility of a device is more important in these conditions than in continuous or stable monitoring. That being said, the magnitude of the differences observed in terms of reproducibility between the two oximeters is quite small and should not be the sole determinant of the type of oximeter used. Other issues that should be considered include price, versatility, convenience, or institutional preferences. The Equanox’s intrasensor reproducibility was better than its intersensor reproducibility. This should encourage clinicians using this device to use a single sensor per patient whenever possible.

These results contrast in some aspects with prior published studies. First, studies evaluating NIRS reproducibility only presented S_w values, and not ICCs, making ICC comparisons impossible. These studies reported S_w ranging from 3% to 7%.^{8,9,11,12,17–20} Notably, in three different studies, Hyttel-Sorensen observed S_w of 5.4%,¹¹ 2.9%,¹² and 4.0%¹⁹ using value measured on the forearms of healthy adults using the INVOS. We, on the other hand, observed slightly lower S_w of 3.0% for the same measurement site. As for the Equanox, a S_w of 4.6% was observed in the sole study it was used in, as compared to the 2.7% S_w we observed.¹² Therefore, the reproducibility observed in our study is among the better values previously mentioned in the available published data. This is possibly because we standardized the data collection procedures and technique. In our experiment, all measurement sites were marked, participant body position was stable and all measurements were taken or directly supervised by the same investigator (AC). This may have contributed to our better measures of reproducibility, which are similar to the best values observed by Hyttel-Sorensen.¹² This emphasizes the necessity to be both familiar and rigorous with the procedure used when NIRS monitoring is utilized, especially in case where intermittent measures are taken. The Fore-Sight (Casmed, Branford, Connecticut), another oximeter less commonly used that we did not evaluate, compared favourably to both the INVOS and Equanox in terms of reproducibility in other experiments. In light of our results, this oximeter likely merits more rigorous evaluation in further studies of reproducibility.^{12,18,21}

We identified variations in the reproducibility between measurement sites, but not between sides of the body. Variations in the reproducibility between measurement sites had already been observed in a paediatric study, which only included measures from the forehead and arm.¹⁹ This should influence the way intermittent monitoring with this technology is done. Indeed, when intermittent peripheral NIRS monitoring is used to assess body perfusion, such as in the intensive care unit, the knee and

Table 7 Absolute values of two tissue oximeters by site and side of the body (%) [mean (SD)].

		INVOS	EQUANOX
Forehead	Left	66.1 (9.4)	72.8 (4.38)
	Right	66.6 (9.1)	71.4 (4.63)
Deltoid	Left	77.0 (7.4)	71.5 (7.96)
	Right	76.3 (7.4)	71.9 (6.75)
Forearm	Left	74.1 (7.7)	69.3 (7.17)
	Right	74.8 (8.5)	69.8 (7.19)
Knee	Left	74.2 (8.0)	71.0 (8.46)
	Right	74.2 (8.1)	72.0 (7.69)
Calf	Left	73.2 (8.3)	72.50 (7.25)
	Right	73.4 (7.4)	73.52 (7.24)
Foot	Left	58.6 (8.8)	60.84 (6.99)
	Right	59.6 (8.3)	61.69 (6.35)

calf sites should be favoured over other sites because of their better reproducibility. Of course, this does not apply when upper extremity specific saturations are warranted, such as in situations where upper extremity ischemia is suspected.

Despite similar mean values in both samples, interoximeter agreement was poor on an individual basis, with limits of agreement of $\pm 19\%$. This is even larger than what had already been observed in multiple settings and between other brands of oximeters.^{22–26} This is likely explained by the measurements over sites—such as the deltoid, knee or foot—for which this had never been evaluated. Moreover, it is probable that ischemia worsen disagreements between oximeters' readings.²⁴ Thus, our results seem to reinforce the fact that readings from different brands of oximeters are not interchangeable and that absolute thresholds of clinically significant ischemia are oximeter-specific and need to be studied individually. This high variability is likely caused by the fact that each type of NIRS oximeter has unique characteristics (number of light-emitting diodes and receptors, distances between diodes, wavelengths used). This will influence the size and depth at which the underlying tissue is evaluated (wide distances between emitters and receptors increase the volume of tissue evaluated).^{27,28} Moreover, tissue oximeters use a specific algorithm to calculate the overall concentration of oxyhemoglobin and deoxyhemoglobin.²⁹ This is explained by the absence of a recognized gold-standard for tissue oxygenation, which led to lack of standardization for the NIRS technology.

On the other hand, interside agreement was much better with limits of agreement of $\pm 8\%$, a value that is somewhat smaller than what has been previously observed in the cardiac and vascular surgery population.³⁰ This difference is probably explained by the assumed absence of significant vascular disease in our sample of healthy volunteers, a characteristic that could potentially create asymmetry limbs perfusion. This description could help guide clinicians when caring for a healthier population when searching for perfusion asymmetry between limbs.

The Equanox was quicker to give a stable reading than the INVOS. While the difference observed (~ 5 s) might not be clinically relevant in a hospital setting, the Equanox could be favored if this technology is used as a triage assessment tool in mass casualty events, helping to identify patients who require emergent, life-saving interventions.^{6,31}

Finally, we described the reference values for the INVOS and Equanox at most sites used in clinical practice. The variations observed between sites can be explained by the perfusion status of these sites, but also of the nature of the tissue evaluated (brain, muscle, connective tissue, fat, and so on). These results will likely help clinicians when using this technology at less conventional sites. The results observed in the present study were similar to most results described previously in healthy volunteers.^{11,12,23} Davie and Fellahi^{25,32} observed higher tissue saturations in their experiment on the forehead and calf, respectively. This could be explained by different sensor positioning and a smaller sample size ($n = 12$ and 20). Interestingly, patients with stable severe systemic disease about to undergo cardiac surgery seem to have similar cerebral saturation as healthy volunteers.^{30,33} This suggests that only critical systemic insults will affect baseline cerebral saturations.³³ The possible determinants of raw tissue saturations across all measurements sites (e.g., age, comorbidity, skin color, body mass index, or skinfold thickness) will need to be better studied and defined in future studies.²¹

4.1 Limitations

The majority of the patients recruited were Caucasian and, since melanin concentration has been shown to affect NIRS values, our results might not be applicable to patients with darker skin pigmentation, especially in situations of low tissue saturations ($<40\%$).²¹ Also, despite a wide range of tissue saturation being observed among the different sites in our study, the results may not be generalizable in cases where extremely low levels of tissue saturation are measured ($<40\%$), such as during cardiac arrest, especially since the variation of the reproducibility of other oximetry devices, such as pulse oximetry, is uncertain in these situations.^{5,34} Furthermore, since the readings of the oximeters we used have been shown to be influenced by ambient light, sensor apposition was verified with great care, but it is always possible this could have affected some of the measures we took.³⁵ This, however, is a limitation that will also be present in real-life NIRS use. Even though we took great care in ensuring that the volunteers were comfortable during data collection, it is possible that some may have experienced changes in body temperature. Given the influence of superficial tissue on the readings of some oximeters, some of the measures could have been affected by this, especially with measures from the lower extremity.³² Since this study did not evaluate dynamic reproducibility, our results should not be used by clinicians using NIRS in such a way. Also, while many are currently available on the market, only two tissue oximeters were evaluated in the course of this study. Finally, since it is still unsure how tissue saturation of each peripheral site will vary in situations of in vivo low perfusion, peripheral monitoring of tissue saturation using NIRS should be incorporated in a multimodal approach to the evaluation of perfusion.

5 Conclusion

In summary, the Equanox could be favored over the INVOS, especially when NIRS monitoring is used intermittently or in unstable conditions, mainly because of its slightly better absolute reproducibility. The knee and calf sites displayed better reproducibility, and could potentially be favoured in situations where NIRS monitoring at these sites is clinically indicated. The description of reference values and normal range of asymmetry for each measurement site will help clinicians identify situations of abnormal perfusions. The differences in observed tissue saturation between devices suggests that clinicians should use the same device for each patient when possible. Also, tissue saturation thresholds are not interchangeable from one NIRS oximeter to another, given their poor interoximeter agreement. NIRS offer new avenues for assessing perfusion in an array of clinical situations in which noninvasive monitoring is required. However, an eventual standardization of NIRS oximeters will be essential to future development of this technology and the expansion of its use in critical care.

Acknowledgments

Conflict of interest: Dr. Denault is part of a speakers bureau for Covidien (Mansfield, MA, USA) and CAE Healthcare (Montréal, Quebec, Canada). Funding: This project received full funding from the “Fond des Urgentistes de l’Hôpital du Sacré-Coeur de Montréal,” the “Chaire du Docteur Sadok Besroux du Département de Médecine Familiale et de Médecine d’Urgence de l’Université de Montréal,” the “Richard I. Kauffman Endowment Fund in Anesthesia and Critical Care”

and the “Fondation de l’Institut de Cardiologie de Montréal.” No support of any kind was received from the industry.

References

- R. P. Dellinger et al., “Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012,” *Crit. Care Med.* **41**(2), 580–637 (2013).
- A. Holley et al., “Review article: part one: goal-directed resuscitation— which goals? Haemodynamic targets,” *Emergency Med. Australas.* **24**(1), 14–22 (2012).
- D. De Backer et al., “Monitoring the microcirculation in the critically ill patient: current methods and future approaches,” *Intensive Care Med.* **36**(11), 1813–1825 (2010).
- A. S. Neto et al., “Association between static and dynamic thenar near-infrared spectroscopy and mortality in patients with sepsis: a systematic review and meta-analysis,” *J. Trauma Acute Care Surg.* **76**(1), 226–233 (2014).
- A. Cournoyer et al., “Near-infrared spectroscopy monitoring during cardiac arrest: a systematic review and meta-analysis,” *Acad. Emergency Med.* **23**(8), 851–862 (2016).
- F. X. Guyette et al., “Prehospital dynamic tissue oxygen saturation response predicts in-hospital lifesaving interventions in trauma patients,” *J. Trauma Acute Care Surg.* **72**(4), 930–935 (2012).
- F. F. Jobsis, “Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters,” *Science* **198**(4323), 1264–1267 (1977).
- 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).
- M. Pocivalnik et al., “Regional tissue oxygen saturation: comparability and reproducibility of different devices,” *J. Biomed. Opt.* **16**(5), 057004 (2011).
- R. Kragelj, T. Jarm, and D. Miklavcic, “Reproducibility of parameters of postocclusive reactive hyperemia measured by near infrared spectroscopy and transcutaneous oximetry,” *Ann. Biomed. Eng.* **28**(2), 168–173 (2000).
- S. Hyttel-Sorensen et al., “Tissue oximetry: a comparison of mean values of regional tissue saturation, reproducibility and dynamic range of four NIRS-instruments on the human forearm,” *Biomed. Opt. Express* **2**(11), 3047–3057 (2011).
- S. Hyttel-Sorensen, T. W. Hessel, and G. Greisen, “Peripheral tissue oximetry: comparing three commercial near-infrared spectroscopy oximeters on the forearm,” *J. Clin. Monit. Comput.* **28**(2), 149–155 (2013).
- S. Fantini et al., “Frequency-domain multichannel optical detector for noninvasive tissue spectroscopy and oximetry,” *Opt. Eng.* **34**(1), 32–34 (1995).
- S. J. Matcher and C. E. Cooper, “Absolute quantification of deoxyhaemoglobin concentration in tissue near infrared spectroscopy,” *Phys. Med. Biol.* **39**(8), 1295–1312 (1994).
- T. B. Fitzpatrick, “The validity and practicality of sun-reactive skin types I through VI,” *Arch. Dermatol.* **124**(6), 869–871 (1988).
- J. M. Bland and D. G. Altman, “Measurement error,” *BMJ* **313**(7059), 744–744 (1996).
- R. G. Wijbenga, P. M. Lemmers, and F. van Bel, “Cerebral oxygenation during the first days of life in preterm and term neonates: differences between different brain regions,” *Pediatr. Res.* **70**(4), 389–394 (2011).
- T. W. Hessel, S. Hyttel-Sorensen, and G. Greisen, “Cerebral oxygenation after birth—a comparison of INVOS(R) and FORE-SIGHT near-infrared spectroscopy oximeters,” *Acta Paediatr.* **103**(5), 488–493 (2014).
- S. Hyttel-Sorensen et al., “A comparison between two NIRS oximeters (INVOS, OxyPrem) using measurement on the arm of adults and head of infants after caesarean section,” *Biomed. Opt. Express* **5**(10), 3671–3683 (2014).
- A. Dullenkopf et al., “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).
- P. E. Bickler, J. R. Feiner, and M. D. Rollins, “Factors affecting the performance of 5 cerebral oximeters during hypoxia in healthy volunteers,” *Anesthesia Analgesia* **117**(4), 813–823 (2013).
- L. M. Dix et al., “Comparing near-infrared spectroscopy devices and their sensors for monitoring regional cerebral oxygen saturation in the neonate,” *Pediatr. Res.* **74**(5), 557–563 (2013).
- M. Thavasothy et al., “A comparison of cerebral oxygenation as measured by the NIRO 300 and the INVOS 5100 Near-Infrared Spectrophotometers,” *Anaesthesia* **57**(10), 999–1006 (2002).
- A. Pisano et al., “Direct comparison between cerebral oximetry by INVOS(TM) and EQUANOX(TM) during cardiac surgery: a pilot study,” *Heart Lung Vessel* **6**(3), 197–203 (2014).
- J. L. Fellahi et al., “Dynamic evaluation of near-infrared peripheral oximetry in healthy volunteers: a comparison between INVOS and EQUANOX,” *J. Crit. Care* **28**(5), 881.e1–881.e6 (2013).
- J. H. Lee et al., “Comparison of two devices using near-infrared spectroscopy for the measurement of tissue oxygenation during a vascular occlusion test in healthy volunteers (INVOS(R) vs. InSpectra),” *J. Clin. Monit. Comput.* **29**(2), 271–278 (2015).
- G. E. Strangman, Z. Li, and Q. Zhang, “Depth sensitivity and source-detector separations for near infrared spectroscopy based on the Colin27 brain template,” *PLoS One* **8**(8), e66319 (2013).
- T. J. Germon et al., “Cerebral near infrared spectroscopy: emitter-detector separation must be increased,” *Br. J. Anaesthesia* **82**(6), 831–837 (1999).
- S. J. Matcher et al., “Performance comparison of several published tissue near-infrared spectroscopy algorithms,” *Anal. Biochem.* **227**(1), 54–68 (1995).
- N. G. Baikoussis et al., “Baseline cerebral oximetry values in cardiac and vascular surgery patients: a prospective observational study,” *J. Cardiothoracic Surg.* **5**, 41 (2010).
- A. C. Beekley et al., “Continuous noninvasive tissue oximetry in the early evaluation of the combat casualty: a prospective study,” *J. Trauma* **69**(Suppl 1), S14–S25 (2010).
- S. N. Davie and H. P. Grocott, “Impact of extracranial contamination on regional cerebral oxygen saturation: a comparison of three cerebral oximetry technologies,” *Anesthesiology* **116**(4), 834–840 (2012).
- M. Heringlake et al., “Preoperative cerebral oxygen saturation and clinical outcomes in cardiac surgery,” *Anesthesiology* **114**(1), 58–69 (2011).
- J. T. Moyle, “Uses and abuses of pulse oximetry,” *Arch. Dis. Child* **74**(1), 77–80 (1996).
- C. Zaouter and E. Arbeid, “Influence of ambient light on cerebral oximeters,” *Br. J. Anaesthesia* **105**(6), 873–874 (2010).

Alexis Cournoyer is both a resident in emergency medicine and a fellow in the clinician-scientist program at the Université de Montréal. He received a MD degree from the Université de Montréal in 2012 and the present work will be part of his master’s thesis.

Biographies for the other authors are not available.