Published in final edited form as:

J Med Eng Technol. 2011 ; 35(0): 370–376. doi:10.3109/03091902.2011.595533.

Variability comparison of simultaneous brain near-infrared spectroscopy (NIRS) and functional MRI (fMRI) during visual stimulation

Ludovico Minati, MSc CEng^{1,2}, Elisa Visani, MSc³, Nick G Dowell, PhD⁴, Nick Medford, MD^{1,5}, and Hugo D Critchley, PhD^{1,5,6}

¹Department of Psychiatry, Brighton & Sussex Medical School (BSMS), Falmer, UK

²Scientific Department, Fondazione IRCCS Istituto, Neurologico Carlo Besta, Milano, IT

³Department of Neurophysiology, Fondazione IRCCS Istituto, Neurologico Carlo Besta, Milano, IT

⁴Clinical Imaging Sciences Centre (CISC), Brighton & Sussex Medical School (BSMS), Falmer, UK

⁵Sackler Centre for Consciousness Science, University of Sussex, Falmer, UK

⁶Sussex Partnership NHS Foundation Trust, Brighton, UK

Abstract

Brain near-infrared spectroscopy (NIRS) is emerging as a potential alternative to functional MRI (fMRI). To date, no study has explicitly compared the two techniques in terms of measurement variability, a key parameter dictating attainable statistical power. Here, NIRS and fMRI were simultaneously recorded during event-related visual stimulation. Inter-subject coefficients of variation (CVs) for peak response amplitude were considerably larger for NIRS than fMRI, but inter-subject CVs for response latency and intra-subject CVs for response amplitude were overall comparable. Our results may represent an optimistic estimate of the CVs of NIRS measurements, as optode positioning was guided by structural MRI, which is normally unavailable. We conclude that fMRI may be preferable to NIRS for group comparisons, but NIRS is equally powerful when comparing conditions within participants. The discrepancy between inter- and intra-subject CVs is likely related to variability in head anatomy and tissue properties which may be better accounted for by emerging NIRS technology.

Keywords

Near-infrared spectroscopy (NIRS); Functional Magnetic Resonance Imaging (fMRI); Measurement variability; Coefficients of variation (CVs); Visual stimulation

1. Introduction

Near-infrared spectroscopy (NIRS), a non-invasive optical technique measuring tissue oxygenation, has received increasing interest as a functional neuroscientific modality complementing electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) [1,2]. Typical NIRS devices emit light at two wavelengths, tuned to the absorption

Corresponding author currently at: Scientific Department, Fondazione IRCCS Istituto Neurologico Carlo Besta, Via Celoria 11, 20133 Milano MI, Italy, Tel. +39 02 2394 2158; +44 784 960 3287, lminati@iee.org.

Conflicts of interest: All authors declare that they do not have any real or perceived conflicts of interest pertaining to the present study.

peaks of oxy- and deoxyhaemoglobin (respectively, O₂Hb and HHb) and provide continuous quantitative measurements of changes in their regional concentration. In brain NIRS, light follows a ‘banana-shaped’ wavepath through skin, subcutaneous fat, scalp muscles and skull, eventually passing through superficial regions of cerebral cortex. Back-scattered radiation therefore originates from extra- and intra-cranial regions, making measurements potentially sensitive to systemic factors in addition to cortical haemodynamics [1,2].

Brain NIRS has recognized advantages over fMRI in terms of equipment cost and portability, and compatibility with limited patient tolerance for the potentially claustrophobic MRI environment. Brain NIRS is proving its usefulness in studies of motor and language function, executive control and appraisal of affective stimuli [3]. Like fMRI, brain NIRS infers neural activation from neurally-coupled haemodynamic responses occurring over 4-6 s. Both techniques are therefore intrinsically limited in temporal resolution, however brain NIRS has the advantage of a higher temporal sampling density than fMRI, i.e. >10 Hz vs. <1 Hz. Several NIRS systems are available, ranging from inexpensive wearable devices based on light-emitting diodes and photodiodes, to large multichannel setups employing pulsed laser sources and photomultipliers [4].

Despite their common physiological basis, NIRS and fMRI signals can have different statistical properties due to the physics of the measurement process, e.g. in relation to sensitivity to anatomical variability and systemic factors. For example, this is realized in studies where their complementary high spatial and temporal sampling features are used to isolate specific aspects of neural activity, or to disentangle the systemic vs. neural components of observed fluctuations [5,6]. A key concern is the inter- and intra-subject measurement variability, which dictates the power attainable for a given experimental design. Limited data are available to date. A recent NIRS investigation found large inter-individual variability during finger-tapping, with detectable O₂Hb increases found in 90% of participants and HHb decreases only in 76%; paradoxical responses were also occasionally observed [7]. Two investigations in which NIRS and fMRI were recorded simultaneously during visual stimulation [8] and cognitive tasks [3] obtained much weaker activation significance for NIRS at individual and group level. However another study using motor stimulation suggested that the sensitivity of NIRS and fMRI is similar and that variance is dominated by cross-modality, physiological factors [9]. There is a substantial body of literature comparing NIRS and fMRI (e.g., [10] for a review), however none of the existing studies explicitly calculated the intra- and inter-individual coefficients of variation (CVs), which are required to compare the statistical power. Extensive characterization of measurement variability is a pre-requisite for more widespread adoption of brain NIRS as an alternative to fMRI, and CVs represent a compact way to compare the relative statistical power attainable with different modalities in the context of a given experimental design. CVs can represent the effect of measurement noise but also of other sources of error such as unaccounted inter-individual variance.

Here, we directly compared the intra- and inter-individual CVs for NIRS- and fMRI-derived response amplitude and latency measurements during visual stimulation with reversing chequerboards of variable duration. We also provide data on the correlation between BOLD and O₂Hb/HHb signals during an event-related task.

2. Methods

2.1 Participants

Eight participants (7 F, 1 M), age 22±5 (range 18-31) years, were enrolled after providing written informed consent. The study was approved by the Brighton and Sussex Medical School Research Ethics and Governance Committee (BSMS REGC). All participants had

normal vision and were right handed. None was taking psychoactive or cardiovascular medication.

2.2 Data acquisition procedure

Functional Magnetic Resonance Imaging (fMRI) was performed on a Magnetom Avanto 1.5 Tesla system (Siemens AG, Erlangen, DE), in two steps: i) iterative cap adjustment and structural scanning in supine position using a 4-channel head coil, and ii) functional scanning in sideways position using a 10 cm flexible loop coil.

Upon arrival, the optode holder was applied, shifting the hair sideways from the occipital region. The holder consisted of a modified silicone swimming cap, to which optode receptacles had been secured corresponding to the four intended optode positions, i.e. approximately 2.6 cm above, below, left and right of the Oz site (defined according to the international 10/20 system). The measured distance between transmit and receive optodes across subjects was 3.6 ± 0.5 cm. This is in line with previous studies with a similar setup [11,12], and with recent work demonstrating that, under optimal conditions, optode spacings up to 5 cm are usable with commercial NIRS equipment [13]. According to a 'banana-shaped' model, which assumes that the penetration is similar for all wavelengths, the resulting wavepath had a maximum penetration depth of about 1.8 cm [1-3,14]. The optode receptacles were designed to hold the fibres approximately perpendicular to the scalp, and consisted of a nylon screw insulator and two nylon washers (RS no. 178-800 and 280-521). The washers were positioned on either side of the cap to achieve a secure grip on the silicone, and the screw insulator was positioned inside them to provide support for the fibre bundle. The washers, insulator and cap were bound together with hot melt glue. A 1 mm-thick tube was glued around each washer (diameter 9 mm) and injected with fish oil for MRI visibility.

Participants were imaged using a coronal T₂-weighted fast low-angle shot (FLASH) sequence, having 1 mm in-plane resolution, 256×192 matrix, 4 mm thickness, 21 slices, 10% gap, TR=4,500 ms and TE=120 ms. The position of the four washers was inspected ensuring that: 1) top and bottom holders appeared aligned with the mid-sagittal plane, and 2) left and right holders appeared aligned with the level of the occipital pole. When these conditions were not met within 4 mm, the cap was re-adjusted. For most participants, one re-adjustment was sufficient. Across participants, the final optode positions in MNI space were $(-7 \pm 3, -128 \pm 4, -26 \pm 14)$ mm and $(-6 \pm 4, -129 \pm 4, 5 \pm 19)$ mm for the receivers and $(29 \pm 5, -124 \pm 4, -9 \pm 23)$ mm and $(-37 \pm 5, -125 \pm 6, -7 \pm 19)$ mm for the transmitters.

Anatomical images were thereafter acquired with a magnetization-prepared gradient-echo volumetric T₁-weighted sequence (MP-RAGE, 1 mm³ isotropic voxels, 192 slices, no gap, TR=1,640 ms, TE=2 ms).

The head coil was subsequently removed and the optodes were inserted in the holders. The optodes consisted of bundles of high-flexibility multimode glass fibres, housed in cylindrical plastic holders. The two transmit optodes had a diameter of 3 mm and were positioned left and right of Oz. The two receive optodes had a diameter of 4 mm and were positioned above and below it. The fibres were strapped to the left shoulder to minimize the risk of cap movement due to weight loading. To further stabilize the optodes, a 5 cm-thick polyurethane slab was applied onto the occipital region and secured in place with Velcro. Participants laid on their right side with the head resting on a memory-foam pillow, and the loop coil was applied over the occipital region and aligned with the isocentre. With the table in scanning position, a mirror mounted on a Plexiglas bracket was installed enabling the participant to view a back-projection screen. The mirror was positioned 3 cm away from the eyes, centred

with respect to the nose, and provided an image covering about half the visual field. The scanner and ambient lights were switched off.

The visual stimulation was delivered in three blocks, each lasting about 8 min, separated by pauses of 2 min. Participants passively attended to radial black/white chequerboards reversing at a rate of 5 Hz, which occupied 80% of the screen. The centre was marked by a red cross, which was always visible and which participants were requested to fixate constantly. The reversing chequerboards were presented in epochs of 1,000 ms, 1,500 ms, 2,000 ms, 2,500 ms, 3,000 ms and 3,500 ms durations. Each simulation epoch was delivered 20 times, yielding a total of 120 stimuli. The resting condition consisted of fixating the central cross superimposed on a black screen. The inter-stimulus interval was jittered at 12.0 ± 1.7 s (range 9–15 s).

NIRS measurements were acquired using a continuous-wave device emitting at 764 ± 20 nm and 859 ± 20 nm, and recording at 50 Hz (Oxymon Mk III, Artinis BV, Zetten, NL). The recording apparatus was located outside the scanner room and connected through 10 m-long amagnetic fibres to the optodes. Changes in O₂Hb and HHb concentration were calculated through the modified Beer–Lambert law, incorporating a differential path length factor as given by Duncan's formula [1].

Two further FLASH acquisitions were performed in sagittal and coronal planes to position the slice-packet for functional imaging. T₂*-weighted volumes were obtained through a gradient-echo echo-planar sequence, having 3.3 mm in-plane resolution, 80×80 matrix, 3.3 mm thickness, 20 slices, no gap, TR=2,500 ms, TE=55 ms, 0.9 kHz bandwidth, which was aligned with the floor of the posterior ventricle horns and centred on the occipital pole. A pre-scan was also performed using the body coil to compute the sensitivity map of the loop coil and normalize signal intensity, which was large for the posterior half of the brain and decayed substantially only in frontal regions.

An MR-compatible pulse oximeter (model 8600FO, Nonin Inc., Plymouth MN, USA) was also applied to monitor heart rate. Timing signals for volume acquisition, stimulus presentation and heart beats were recorded using a power1401 ADC (CED Ltd., Cambridge, UK).

2.3 Data analysis

NIRS signals were low-pass filtered at 5 Hz, then detrended by fitting the baseline with a 3rd-degree polynomial and subtracting it. For each channel, a model of the heart pulsation-related artefact was obtained by averaging over cardiac cycles and the corresponding signal was removed. Low-pass filtering was then repeated at 1 Hz. The signals from the four channels were thereafter averaged together. This operation was performed because, as discussed below, considering that the low number of NIRS channels prevents us from establishing topographical correlations, it was decided to focus the experimental hypotheses on the CVs of the overall response amplitude and latency, averaged over the whole cortical area covered. The four channels were equally weighted in the averaging process, and no channel had to be excluded for any participant.

After realignment and slice-timing correction, functional images were co-registered with the structural scans and normalized into MNI space using SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK). A general linear model was set up for each participant, to visualize which brain regions had a BOLD time-course significantly correlated with the NIRS signals, with two regressors corresponding to O₂Hb and HHb recorded throughout the whole session and averaged within each TR. The realignment parameters were inserted as nuisance covariates. Individual statistical maps of the F-value for the combined effect of

O₂Hb and HHb concentration changes were entered into a one-sample t-test, to identify the areas which were significantly related to these signals over all participants.

We investigated the relationship between NIRS and BOLD responses, using a custom program running under Matlab 7 (MathWorks Inc., Natick MA, USA) to automatically segment the optode holder rings on coronal FLASH images acquired immediately prior to functional scanning, and project their positions on the scalp. This yielded a binary brain mask corresponding to each NIRS channel, defined as all voxels within 25% of the inter-optode distance from an assumed wavepath, corresponding to an arc connecting the scalp positions of the transmit and receive optodes with a maximum penetration of 50% of the inter-optode distance [15]. The masks of the four channels were combined and the corresponding average fMRI BOLD time-series was extracted. After removal of movement-related variance by linear regression, the signal was spline-interpolated to 50 Hz, temporally realigned to the NIRS signals and filtered with the same settings.

NIRS and BOLD time-courses were then epoched between -0.5 and 12 s with respect to visual stimulus onsets, and the average pre-stimulus (0.5 s) baseline was subtracted. After rejecting outlier epochs (using a ± 3 SD criterion), the remaining data were averaged according to the six levels of stimulus duration. Response peak amplitude and latency were derived by picking the maximum (for O₂Hb and BOLD) or minimum (for HHb) value in the 0-12 s post-stimulus period and analyzed with within-subject ANOVAs. Peak picking was performed on the average responses for each stimulus duration level, and there was never more than one maximum/minimum which would have resulted in ambiguous peak-picking.

The signal-to-noise ratio (SNR) of the evoked responses was calculated on responses averaged for each participant and duration level, as peak amplitude divided by the standard deviation of the pre-stimulus period, and compared among O₂Hb, HHb and BOLD with within-subject ANOVA. Inter-subject coefficients of variation (CVs), defined as SD over participants divided by the grand-average, were calculated for response amplitude and latency and compared across the three signals using within-subject ANOVAs. Intra-subject CVs were calculated for amplitude only, on the basis of single-trial values at the latency of the average peak identified for each condition and participant; this was necessary because the response SNR of individual trials was insufficient for reliable peak extraction. These were analyzed using within-subject ANOVA with factors for stimulus duration (1,000 ... 3,500 ms) and signal type (O₂Hb, HHb or BOLD). Further, for all parameters we tested for outliers using a ± 3 SD criterion, and verified distribution normality by means of the Kolmogorov-Smirnov test.

Additionally, to exclude effects related to the detrending procedure, we repeated the calculation of the NIRS inter-subject amplitude CVs excluding the polynomial detrending step, and subtracting the average of the 1 s pre-stimulus and 12-13 s post-stimulus windows. Further, to evaluate the relationship between the three signals, linear, quadratic and cubic regressions were performed between O₂Hb and HHb, O₂Hb and BOLD, and HHb and BOLD responses.

2.4 Image SNR comparison

In order to improve the general validity of the study, the raw image SNR of the functional echo-planar images obtained with the loop coil described above was measured and compared with that of similar images obtained on an identical scanner using the same sequence, timings and geometry with a standard four-channel matrix head coil. The SNR of raw (i.e., unfiltered and not normalized) images was measured using planar regions-of-interest drawn on four axial slices corresponding to the occipital region shown in Fig. 1 (see results), and quantified as mean signal intensity divided by the standard deviation of the

background outside the head. For the loop coil, the raw image SNR was 73 ± 19 . For eight participants imaged with the multichannel coil, the SNR was 122 ± 13 . The difference was strongly significant ($p < 0.001$).

3. Results

The superficial portion of the striate cortex (area V1) and the surrounding extrastriate cortex (principally area V2) exhibited a BOLD time-course significantly correlated to O_2Hb and HHb concentration changes (Fig. 1). These regions overlapped substantially with the anatomical masks modelled using a 'banana-shaped' light wavepath of the NIRS signal. No correlation was observed for any other region.

Due to the presence of large slow-trends in the NIRS signals, polynomial detrending removed $29 \pm 13\%$ of the initial variance for O_2Hb and $45 \pm 17\%$ for HHb.

The stimulus-evoked responses are charted in Fig. 2a, and corresponding peak amplitudes and latencies are given in Fig. 2c. Response amplitude increased monotonically with stimulus duration for O_2Hb ($F(5,35)=3.9$, $p=0.007$, $\eta_p^2=0.36$), HHb $F(5,35)=8.4$, $p<0.001$, $\eta_p^2=0.55$) and BOLD ($F(5,35)=16.8$, $p<0.001$, $\eta_p^2=0.71$). Similarly, response latency increased for O_2Hb ($F(5,35)=7.0$, $p<0.001$, $\eta_p^2=0.50$), HHb ($F(5,35)=16.4$, $p<0.001$, $\eta_p^2=0.70$) and BOLD ($F(5,35)=13.0$, $p<0.001$, $\eta_p^2=0.65$). There were no outliers, and the distribution was normal across all design cells.

The evoked response SNR (given as average over all stimulus durations, SD across participants) was 93.2 ± 56.6 for O_2Hb , 25.4 ± 18.7 for HHb and 122.1 ± 107.9 for BOLD; this was significantly different across signals ($F(2,14)=3.9$, $p=0.04$, $\eta_p^2=0.36$), i.e. larger for O_2Hb than HHb ($p=0.02$), larger for BOLD than HHb ($p=0.03$) and comparable between O_2Hb and BOLD ($p=0.5$).

Inter-subject amplitude CVs (obtained processing the NIRS signals through polynomial detrending and reported as mean $\pm$ SD over stimulus durations) were 0.52 ± 0.14 , median 0.48, range 0.31-0.72 for O_2Hb , 0.61 ± 0.11 , median 0.55, range 0.53-0.79 for HHb and 0.34 ± 0.08 , median 0.34, range 0.21-0.43 for BOLD; this was significantly different across signals ($F(2,15)=9.0$, $p=0.003$, $\eta_p^2=0.54$), i.e. larger for O_2Hb than BOLD ($p=0.02$), larger for HHb than BOLD ($p=0.001$) and comparable between O_2Hb and HHb ($p=0.2$). Repeating this step excluding the polynomial detrending step yielded analogous findings, i.e. the inter-subject amplitude CVs were 0.50 ± 0.05 , median 0.48, range 0.46-0.59 for O_2Hb , 0.62 ± 0.03 , median 0.63, range 0.56-0.67 for HHb; this was significantly different across signals ($F(2,15)=37.4$, $p<0.001$, $\eta_p^2=0.83$), i.e. larger for O_2Hb than BOLD ($p<0.001$), larger for HHb than BOLD ($p<0.001$) and comparable between O_2Hb and HHb ($p=0.002$). There were no outliers, and the distribution was normal across all design cells.

Inter-subject latency CVs were 0.14 ± 0.02 , median 0.14, range 0.10-0.17 for O_2Hb , 0.12 ± 0.01 , median 0.11, range 0.10-0.13 for HHb and 0.12 ± 0.02 , median 0.11, range 0.09-0.16 for BOLD; there were no significant differences ($p=0.2$).

Intra-subject amplitude CVs, plotted in Fig. 2d, significantly decreased with stimulus duration ($F(5,35)=4.1$, $p=0.005$, $\eta_p^2=0.37$). There was a weak main effect of signal type (1.13 ± 0.71 , median 0.94, range 0.38-2.8 for O_2Hb , 0.76 ± 0.45 , median 0.62, range 0.27-2.5 for HHb and 0.70 ± 0.42 , median 0.56, range 0.22-1.9 for BOLD; $F(2,14)=3.8$, $p=0.045$, $\eta_p^2=0.35$), i.e. CVs were larger for O_2Hb than HHb ($p=0.02$), and comparable between BOLD and O_2Hb /HHb. There were no outliers, and the distribution was normal across all design cells.

Signal correlation plots are given in Fig. 2b. Curve fitting for O₂Hb vs. HHb responses yielded $F(1,1870)=52,000$ for linear, $F(1,1869)=26,200$ for quadratic and $F(1,1868)=17,500$ for cubic models ($p<0.001$ for all). Curve fitting for BOLD vs. O₂Hb yielded $F(1,1870)=77,200$ for linear, $F(1,1869)=38,900$ for quadratic and $F(1,1868)=28,400$ for cubic models ($p<0.001$). Curve fitting for BOLD vs. HHb yielded $F(1,1870)=70,400$ for linear, $F(1,1869)=35,200$ for quadratic and $F(1,1868)=24,900$ for cubic models ($p<0.001$). Across participants, the linear correlation r -value was 0.76 ± 0.17 for BOLD vs. O₂Hb, and 0.85 ± 0.06 for BOLD vs. HHb responses; the r -values were not significantly different (paired-samples t -test, $p=0.1$).

4. Discussion

Our results indicate that the inter-subject CVs for response amplitude during a typical visual stimulation task are larger for NIRS than fMRI, by a factor of about 1.5 for O₂Hb and 1.8 for HHb. In practice, this would translate into a need to recruit more than twice as many participants to obtain equal power for a group comparison using parametric statistics (e.g., [16]). One needs to consider that the optodes were carefully positioned using an ad-hoc setup under direct guidance from structural MRI, which is normally unavailable. In real-life applications, inter-subject CVs may be even larger due to unaccounted anatomical variability resulting in less precise positioning. By contrast, inter-subject CVs for peak response latency and intra-subject CVs for amplitude were comparable. Importantly, with implementation of an effective jittered design, fMRI was able to represent haemodynamic latency effects similar to NIRS, despite the much lower sample rate.

Previous investigations on the correlation between fMRI and NIRS signals have extracted BOLD signals using spherical ROIs at fixed positions [17], manually positioned ROIs [9] and by identifying the shallowest activated voxels [18]. Here, we assumed a ‘banana-shaped’ wavepath, which had previously been found to yield better correlation than other models [3]. This model is limited with respect to the Monte Carlo techniques used in diffuse optical tomography [19,20] in that it makes several simplifying assumptions, such as neglecting different penetration between wavelengths, however it remains very widely used in brain NIRS. The validity of the masks utilized for BOLD signal extraction was confirmed by substantial overlap observed with the areas displaying a BOLD time-course directly correlated to the O₂Hb/HHb time courses.

Increasing response amplitude and latency with stimulus duration is in agreement with the notion that the cortical haemodynamic response represents a time-integral of postsynaptic activity, and is in line with previous observations [17,21]. Here, consistently shorter latencies were observed for O₂Hb than HHb. This is in line with previous observations (e.g., [10]) and likely a consequence of the fact that, while elevated blood flow leads to an immediate increase in O₂Hb concentration, its dilution effect on HHb is initially counteracted by increasing oxygen extraction [22,23]. The intermediate peak latency values observed for BOLD may be due to its different sensitivity at the microscopic scale, i.e. to the fact that BOLD signal changes are dominated by large draining veins, which are associated with a blood propagation delay with respect to the microvasculature to which NIRS is mostly sensitive [10,18,24].

We observed empirically that a linear model is sufficient (i.e., yields the highest F -values) to approximate the relationship between O₂Hb/HHb and BOLD responses, in spite of the complex differences in the underlying physics. The average correlation coefficient was larger for HHb ($r=0.85$) than O₂Hb ($r=0.76$), though this did not reach statistical significance. Previous studies have yielded highly inconsistent results, i.e. some reported stronger correlation with O₂Hb [8,18], some stronger correlation with HHb [9,17] and some

mixed findings [3]. Since O₂Hb is essentially magnetically inert and HHb is strongly paramagnetic, the concentration of the latter should dominate effects on the BOLD signal, through T₂* relaxivity changes. However, there are practical factors biasing the observed correlations. First, stimulus-evoked concentration changes are much smaller for HHb than O₂Hb, resulting in a much lower response SNR. In our data, however, the correspondence was stronger for HHb despite the fact that its SNR was almost four times smaller. Second, for as-yet unclear reasons there is a known topographical incongruence between O₂Hb and HHb effects, the former being significantly more diffuse [10,25,26]. This can introduce bias through partial voluming effects, which are more severe for NIRS than for fMRI. Our data are uninformative in this regard, because due to the small number of available optodes we did not attempt to characterize the topographical distribution of the NIRS responses. Third, there are substantial sensitivity differences at the microscopic scale: due to photon diffusion properties, NIRS is sensitive to changes within micro-vessels smaller than the typical absorption length (i.e., 1 mm), whereas gradient-echo fMRI emphasizes the effect of wide, superficial veins [18,24]. Importantly, one needs to consider that the relationship observed between NIRS and fMRI may not be generalizable to patients with cerebrovascular abnormalities, for example due to tumours or stroke, in which paradoxical effects may even be observed [27].

One needs to consider the limitations of the NIRS and fMRI methods employed in the present study. As regards to NIRS, we have utilized a relatively small number of channels, which were centred around the Oz site of the occipital pole. In this region, the skull is usually thickest, and the photons may only marginally diffuse into the cortex, limiting signal quality and increasing relative sensitivity to non-brain signal sources. Further research should be conducted considering, for example, the motor cortex, which represents a more advantageous site for NIRS, in light of the thinner skull and lower density of hair. Another limitation of our NIRS setup is the limited number of channels: high-density systems (e.g., 20 channels or more) enable more focal comparison of local responses, potentially improving specificity [25,26]. As a consequence of this limitation, it was decided to compare overall responses, averaging over the whole area covered. As regards to fMRI, imaging was performed with a loop coil that yielded approximately half of the raw image SNR obtained with a standard head matrix coil. The use of a 3 T scanner would have provided a further increase in SNR. Our results therefore likely represent a substantial overestimation of the CVs for fMRI with respect to optimal settings, even though one needs to consider that the inter- and intra-individual CVs do not necessarily show a strong correlation with raw image SNR due to the presence of sources of noise unrelated to the measurement process. Hence, the effective disparity between fNIRS and fMRI in terms of inter-individual CVs may in practice be even larger than we have observed, and it is possible that with higher SNR fMRI may also be preferable in terms of intra-subject variability.

Our data are likely to be generalizable to other continuous-wave NIRS setups, which remain the most common form of NIRS employed in psychophysiological studies, but may not be representative of time-resolved NIRS, which allows for greater anatomical specificity and enhanced immunity to physiological noise by excluding extra-cranial signals [4].

Especially, the findings of the present study are not valid for diffuse optic tomography, which can arguably offer substantial improvements in measurement reproducibility with respect to NIRS, thanks to much denser sampling and more advanced modelling of photon diffusion across tissue types and within the brain [19,20]. However, our estimates of the inter-individual CVs are unlikely to be excessive with respect to a typical continuous-wave setup currently used in psychophysiological research, also considering that the effects of inter-individual anatomical variability were mitigated by the use of MRI to guide positioning. The finding of larger inter-individual CVs needs confirmation in other

experimental settings, e.g. during motor and cognitive tasks. If confirmed, it would imply that, where available and viable for the population and task under study, fMRI is preferable to NIRS for group comparisons. Inter-individual CVs for response latency were similar, but this parameter alone is usually not sufficient to capture differences in neural activation. However, we obtained comparable intra-individual CVs, in spite of markedly different SNR between BOLD and NIRS signals. This is reassuring, indicating the viability of NIRS to study within-participants effects, and is in line with earlier findings [28]. The discrepancy between inter- and intra-individual CVs is likely to be related, at least in part, to differences in head anatomy and skull tissue properties between participants. Currently, no consensus or comparative studies exist regarding the optode holders used for brain NIRS. Further work is needed to develop and standardize holders which, alongside the use of denser arrays, more advanced models of photon diffusion and time-resolved measurements, will mitigate the effect of these differences improving the consistency of NIRS measurements across subjects [1,2,4].

Acknowledgments

The study was supported by a programme grant from the Wellcome Trust to HDC (R127). All data were acquired at the Clinical Imaging Sciences Centre (CISC) of the Brighton & Sussex Medical School (BSMS). The authors would like to thank the radiographic staff at CISC for operational assistance during data acquisition, and an anonymous reviewer for insightful feedback on an earlier version of the manuscript.

References

- [1]. Strangman G, Boas DA, Sutton JP. Non-invasive neuroimaging using near-infrared light. *Biol Psychiatry*. 2002; 52:679–693. [PubMed: 12372658]
- [2]. Wolf M, Ferrari M, Quaresima V. Progress of near-infrared spectroscopy and topography for brain and muscle clinical applications. *J Biomed Opt*. 2007; 12:062104. [PubMed: 18163807]
- [3]. Cui X, Bray S, Bryant DM, Glover GH, Reiss AL. A quantitative comparison of NIRS and fMRI across multiple cognitive tasks. *Neuroimage*. 2010 in press.
- [4]. Torricelli A, Contini S, Pifferi A, Spinelli L, Cubeddu R. Functional brain imaging by multi-wavelength time-resolved near infrared spectroscopy. *Opto-Elect Rev*. 2008; 16:131–135.
- [5]. Kashou NH, Xu R, Roberts CJ, Leguire LE. Using FMRI and FNIRS for localization and monitoring of visual cortex activities. *Conf Proc IEEE Eng Med Biol Soc*. 2007;2634–8. [PubMed: 18002536]
- [6]. Tong Y, Frederick BD. Time lag dependent multimodal processing of concurrent fMRI and near-infrared spectroscopy (NIRS) data suggests a global circulatory origin for low-frequency oscillation signals in human brain. *Neuroimage*. 2010; 53:553–64. [PubMed: 20600975]
- [7]. Sato H, Fuchino Y, Kiguchi M, Katura T, Maki A, Yoro T, Koizumi H. Intersubject variability of near-infrared spectroscopy signals during sensorimotor cortex activation. *J Biomed Opt*. 2005; 10:44001. [PubMed: 16178635]
- [8]. Toronov VY, Zhang X, Webb AG. A spatial and temporal comparison of hemodynamic signals measured using optical and functional magnetic resonance imaging during activation in the human primary visual cortex. *Neuroimage*. 2007; 34:1136–48. [PubMed: 17134913]
- [9]. Huppert TJ, Hoge RD, Diamond SG, Franceschini MA, Boas DA. A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. *Neuroimage*. 2006; 29:368–82. [PubMed: 16303317]
- [10]. Steinbrink J, Villringer A, Kempf F, Haux D, Boden S, Obrig H. Illuminating the BOLD signal: combined fMRI-fNIRS studies. *Magn Reson Imaging*. 2006; 24:495–505. [PubMed: 16677956]
- [11]. Minati L, Jones CL, Gray MA, Medford N, Harrison NA, Critchley HD. Emotional modulation of visual cortex activity: a functional near-infrared spectroscopy study. *Neuroreport*. 2009; 20:1344–50. [PubMed: 19738501]

- [12]. Minati L, Kress IU, Visani E, Medford N, Critchley HD. Intra- and extra-cranial effects of transient blood pressure changes on brain near-infrared spectroscopy (NIRS) measurements. *J Neurosci Methods*. 2011 in press.
- [13]. Tachtsidis I, Tisdall M, Delpy DT, Smith M, Elwell CE. Measurement of cerebral tissue oxygenation in young healthy volunteers during acetazolamide provocation: a transcranial Doppler and near-infrared spectroscopy investigation. *Adv Exp Med Biol*. 2008; 614:389–96. [PubMed: 18290350]
- [14]. Cui W, Kumar C, Chance B. Experimental study of migration depth for the photons measured at sample surface. *SPIE*. 1991; 1431:180–91.
- [15]. Firbank M, Okada E, Delpy DT. A theoretical study of the signal contribution of regions of the adult head to near-infrared spectroscopy studies of visual evoked responses. *Neuroimage*. 1998; 8:69–78. [PubMed: 9698577]
- [16]. Ellis, PD. *The Essential Guide to Effect Sizes: An Introduction to Statistical Power, Meta-Analysis and the Interpretation of Research Results*. Cambridge University Press; UK: 2010.
- [17]. Toyoda H, Kashikura K, Okada T, Nakashita S, Honda M, Yonekura Y, Kawaguchi H, Maki A, Sadato N. Source of nonlinearity of the BOLD response revealed by simultaneous fMRI and NIRS. *Neuroimage*. 2008; 39:997–1013. [PubMed: 18054252]
- [18]. Strangman G, Culver JP, Thompson JH, Boas DA. A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation. *Neuroimage*. 2002; 17:719–31. [PubMed: 12377147]
- [19]. Gregg NM, White BR, Zeff BW, Berger AJ, Culver JP. Brain specificity of diffuse optical imaging: improvements from superficial signal regression and tomography. *Front Neuroenergetics*. 2010; 2:14. [PubMed: 20725524]
- [20]. White BR, Culver JP. Quantitative evaluation of high-density diffuse optical tomography: in vivo resolution and mapping performance. *J Biomed Opt*. 2010; 15:026006. [PubMed: 20459251]
- [21]. Soltysik DA, Peck KK, White KD, Crosson B, Briggs RW. Comparison of hemodynamic response nonlinearity across primary cortical areas. *Neuroimage*. 2004; 22:1117–27. [PubMed: 15219583]
- [22]. Buxton RB, Wong EC, Frank LR. Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magn Reson Med*. 1998; 39:855–64. [PubMed: 9621908]
- [23]. Karen T, Morren G, Haensse D, Bauschatz AS, Bucher HU, Wolf M. Hemodynamic response to visual stimulation in newborn infants using functional near-infrared spectroscopy. *Hum Brain Mapp*. 2008; 29:453–60. [PubMed: 17525986]
- [24]. Liu H, Chance B, Hielscher AH, Jacques SL, Tittel FK. Influence of blood vessels on the measurement of hemoglobin oxygenation as determined by time-resolved reflectance spectroscopy. *Med Phys*. 1995; 22:1209–17. [PubMed: 7476706]
- [25]. Franceschini MA, Toronov V, Filiaci M, Gratton E, Fantini S. On-line optical imaging of the human brain with 160-ms temporal resolution. *Opt Express*. 2000; 6:49–57. [PubMed: 19401744]
- [26]. Boas DA, Gaudette T, Strangman G, Cheng X, Marota JJ, Mandeville JB. The accuracy of near infrared spectroscopy and imaging during focal changes in cerebral hemodynamics. *Neuroimage*. 2001; 13:76–90. [PubMed: 11133311]
- [27]. Sakatani K, Murata Y, Fujiwara N, Hoshino T, Nakamura S, Kano T, Katayama Y. Comparison of blood-oxygen-level-dependent functional magnetic resonance imaging and near-infrared spectroscopy recording during functional brain activation in patients with stroke and brain tumors. *J Biomed Opt*. 2007; 12:062110. [PubMed: 18163813]
- [28]. Sato H, Kiguchi M, Maki A, Fuchino Y, Obata A, Yoro T, Koizumi H. Within-subject reproducibility of near-infrared spectroscopy signals in sensorimotor activation after 6 months. *J Biomed Opt*. 2006; 11:014021. [PubMed: 16526898]

Figure 1.

a) Cumulative maps of the binary masks used for BOLD signal averaging, derived from measured optode positions and computed assuming 'banana-shaped' wavepaths. Percentages represent the proportion of participants for whom a voxel was included. b) Group-level statistical maps for the effect of O₂Hb and HHb on the BOLD signal. Visualization threshold: voxel-wise $p < 0.001$ un-corrected.

Figure 2.

a) Average O_2Hb , HHb and BOLD responses evoked by the reversing chequerboards, indicating overall increasing response amplitude and latency with stimulus duration. b) Corresponding relationships between the signals, indicating strong linear correlations. c) Peak amplitudes and latencies. d) Intra-subject CVs for response amplitude, which were comparable between BOLD and O_2Hb/HHb . All error bars denote 1 SD.