

Performance assessment of time-domain optical brain imagers, part 1: basic instrumental performance protocol

Heidrun Wabnitz
Dieter
Richard Taubert
Mikhail Mazurenka
Oliver Steinkellner
Alexander Jelzow
Rainer Macdonald
Daniel Milej
Piotr Sawosz
Michał Kacprzak
Adam Liebert
Robert Cooper
Jeremy Hebden
Antonio Pifferi

Andrea Farina
Ilaria Bargigia
Davide Contini
Matteo Caffini
Lucia Zucchelli
Lorenzo Spinelli
Rinaldo Cubeddu
Alessandro Torricelli

Performance assessment of time-domain optical brain imagers, part 1: basic instrumental performance protocol

Heidrun Wabnitz,^{a,*} Dieter Richard Taubert,^a Mikhail Mazurenka,^a Oliver Steinkellner,^a Alexander Jelzow,^a Rainer Macdonald,^a Daniel Milej,^b Piotr Sawosz,^b Michał Kacprzak,^b Adam Liebert,^b Robert Cooper,^c Jeremy Hebden,^c Antonio Pifferi,^{d,e} Andrea Farina,^e Ilaria Bargigia,^f Davide Contini,^d Matteo Caffini,^d Lucia Zucchelli,^d Lorenzo Spinelli,^e Rinaldo Cubeddu,^{d,e} and Alessandro Torricelli^d

^aPhysikalisch-Technische Bundesanstalt (PTB), Abbestraße 2-12, 10587 Berlin, Germany

^bNałęcz Institute of Biocybernetics and Biomedical Engineering, Polish Academy of Sciences, ul. Trojdena 4, 02-109 Warsaw, Poland

^cUniversity College London, Department of Medical Physics and Biomedical Engineering, Gower Street, London WC1E 6BT, United Kingdom

^dPolitecnico di Milano, Dipartimento di Fisica, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

^eIstituto di Fotonica e Nanotecnologie, CNR, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

^fIstituto Italiano di Tecnologia, Center for Nano Science and Technology @ PoliMi, via Pascoli 70/3, 20133 Milano, Italy

Abstract. Performance assessment of instruments devised for clinical applications is of key importance for validation and quality assurance. Two new protocols were developed and applied to facilitate the design and optimization of instruments for time-domain optical brain imaging within the European project nEUROpt. Here, we present the “Basic Instrumental Performance” protocol for direct measurement of relevant characteristics. Two tests are discussed in detail. First, the responsivity of the detection system is a measure of the overall efficiency to detect light emerging from tissue. For the related test, dedicated solid slab phantoms were developed and quantitatively spectrally characterized to provide sources of known radiance with nearly Lambertian angular characteristics. The responsivity of four time-domain optical brain imagers was found to be of the order of 0.1 m² sr. The relevance of the responsivity measure is demonstrated by simulations of diffuse reflectance as a function of source-detector separation and optical properties. Second, the temporal instrument response function (IRF) is a critically important factor in determining the performance of time-domain systems. Measurements of the IRF for various instruments were combined with simulations to illustrate the impact of the width and shape of the IRF on contrast for a deep absorption change mimicking brain activation. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.19.8.086010](https://doi.org/10.1117/1.JBO.19.8.086010)]

Keywords: optical brain imaging; time-resolved measurements; instrumentation; performance assessment; tissue-like phantoms.

Paper 140246PR received Apr. 17, 2014; revised manuscript received Jul. 17, 2014; accepted for publication Jul. 21, 2014; published online Aug. 14, 2014.

1 Introduction

In recent years, instruments and methodologies to noninvasively probe human tissues by diffuse light in the near-infrared spectral range have undergone translation from laboratory bench to bedside. For example, clinical trials have been performed targeting the female breast. Optical mammography has been developed and applied to detect and characterize breast lesions,^{1,2} as well as to monitor the outcome of neoadjuvant chemotherapy of the breast.^{3,4} Diffuse optical spectroscopy is a promising technique for cancer risk assessment.^{5,6} In parallel, many clinical studies have been conducted targeting the brain. Optical imaging has been used to observe the hemodynamic response to cognitive, sensory, and motor stimuli,⁷ and to assess changes associated with traumatic brain injury,⁸ stroke,⁹ and epilepsy,¹⁰ with the aim of providing a less expensive and portable alternative to the existing neuroimaging techniques such as magnetic resonance imaging or x-ray computed tomography. Meanwhile, optical techniques have also been applied to other tissues such as muscle, joints, bone, and skin as a means of providing diagnostic information.

One of the principal impediments to a widespread adoption of optical imaging techniques in the clinic has been the difficulty in providing absolute physiological parameters from optical measurements. Even when the individual systems demonstrate excellent repeatability, the derived parameters are often system specific. Progress has been hampered by the lack of standardization and quality assessment tools, which would enable measurements acquired with one system to be compared with those acquired with another.¹¹ Appropriate protocols are clearly required in order to assess the performance of one instrument against another or to assess subsequent upgrades of the same system. The availability of such protocols may also be beneficial during research and development of methods and devices—particularly at an industrial level—toward an improvement in the measurement of those parameters that are relevant to specific clinical applications. Standardized procedures for characterization of instruments and techniques would also facilitate a more reliable comparison between data acquired on a given system for different clinical studies. The large variety of measurement techniques (e.g., continuous wave, frequency, and time domain), instrument configurations (e.g., multichannel imagers and single-channel monitor),¹² and clinical applications makes it even more desirable that a simple, well-defined procedure exists for testing and validating the relevant optical systems.

*Address all correspondence to: Heidrun Wabnitz, E-mail: heidrun.wabnitz@ptb.de

In the field of diffuse optical imaging and spectroscopy, the first systematic study of various instruments was performed within a European Thematic Network on the basis of the MEDPHOT protocol.¹³ Recently, the development and testing of phantoms have been reported for quality assurance in a multicenter clinical trial to measure the response of breast tumors to neoadjuvant chemotherapy by means of diffuse optical imaging.¹⁴

The work presented here was carried out within the European project nEUROpt (FP7-HEALTH-F5-2008-201076) that concerned the development of time-domain systems for optical imaging of the brain based on novel technological advances and new methodological approaches. The new methodologies were evaluated with respect to their potential to improve the spatial resolution, sensitivity, and quantification of optical properties. Two new protocols were developed to provide guidelines for the comparison of various instruments and to estimate the impact of new technological and methodological advances on performance, namely the BIP protocol for “Basic Instrumental Performance” assessment and the nEUROpt protocol for performance assessment of time-domain brain imagers. These protocols were employed for a series of common measurements performed by contributing project partners within the nEUROpt project and were also used for quality assurance purposes, particularly during clinical studies.

Overall, the three BIP, MEDPHOT, and nEUROpt protocols constitute a consolidated and integrated framework for the performance assessment of time-domain instruments in diffuse optics, with already broad consensus and experimental use among many EU laboratories. The key features of all three protocols are summarized in Table 1. The BIP protocol is focused on the key properties of the systems on a basic level, independent of the sample properties to be measured and of the related data analysis methods. In particular, it deals with the detection efficiency of the system, with its temporal performance characterized by the instrument response function (IRF), and a number

of other key parameters to be collected. The MEDPHOT protocol is a subsequent “high-level” protocol intended to grade the capability of the system taken as a whole—including the analysis software—to retrieve the optical properties of a homogeneous medium, with emphasis on the features that are most relevant for the clinical application. In particular, it is configured to characterize the measurement of the absorption (μ_a) and reduced scattering (μ'_s) coefficients of a turbid medium in terms of accuracy, linearity, noise, stability, and reproducibility. Finally, the nEUROpt protocol represents another “high-level” protocol, based on specific inhomogeneous phantoms, to test the performance of an imaging device. It addresses the ability of optical brain imaging systems to detect, localize, and quantify absorption changes in the brain. The BIP tests play an important role in the interpretation of the results of these high-level protocols, particularly when combined with appropriate simulations.

The present work is devoted to the BIP protocol, while the nEUROpt protocol will be introduced in a companion paper.¹⁵ After a description of the individual tests and their implementation in the following section, the application of the protocol is illustrated by presenting and discussing results obtained using several instruments developed by four different laboratories.

2 BIP Protocol

This protocol concerns the recording of basic characteristics of time-domain instruments, which influence the quality and accuracy of measurements in clinical applications. It is applicable to instruments based on pulsed laser sources with repetition rates of the order of several tens of MHz, fast single-photon detectors, and time-correlated single photon counting (TCSPC). The primary measurand is the histogram of photon flight times $N(t)$, usually denoted as the temporal point spread function or the distribution of times of flight (DTOF). The tests of this protocol aim to assess factors that influence the quality of the measured

Table 1 Overview of protocols for performance characterization of instruments in diffuse optics (cw—continuous wave, fd—frequency domain, and td—time domain).

Protocol		BIP	MEDPHOT ¹³	nEUROpt ¹⁵
Key features	Purpose	Characterization of an instrument without a measuring object	Capability of an instrument to measure homogeneous optical properties	Capability of an instrument to detect, localize, and quantify absorption changes
	Target quantities	Characteristics of source and detection component and temporal IRF	Absolute homogeneous μ'_s and μ_a	$\Delta\mu_a$ of an inhomogeneity
	Models for data analysis	—	Model-based retrieval of homogeneous μ'_s and μ_a	Model-based retrieval of $\Delta\mu_a$ or semi-empirical data analysis
	Spatial aspects	—	—	Tests for spatial localization of $\Delta\mu_a$
Reference implementation	Phantoms	Responsivity phantom (as diffuse light source)	Set of solid homogeneous phantoms with different μ'_s and μ_a	Inhomogeneous liquid phantoms with (a) inclusions and (b) two layers
Application	Primary fields	td photon migration instruments	Photon migration instruments of any kind	td optical brain imagers
	Possible extension	cw and fd photon migration instruments; other td instruments		cw and fd optical brain imagers; optical mammographs

distribution $N(t)$, in particular its shape, signal-to-noise ratio, and stability.

The instruments under test can be subdivided into a source component and a detection component. The source component consists of one or more lasers and the means to deliver light to the tissue, e.g., an optical fiber. The detection component comprises a fast single-photon detector, TCSPC electronics, and the optics (including e.g., a fiber bundle) to transport light from the tissue to the detector. The tests of this protocol address:

1. Relevant parameters of the source component of the instrument,
2. The responsivity of the detection system,
3. The differential nonlinearity (DNL) of the timing electronics, and
4. The temporal IRF, its shape, its background, as well as its stability.

Test (1) covers the separate characterization of the source component, while tests (2) and (3) are related only to the detection component. The IRF (test 4) depends on the properties of both components. Tests (1) and (2) address the time-integrated efficiency of the instrument, while tests (3) and (4) concern the characteristics of the time-resolved measurement. All four tests are briefly presented below. In Sec. 4, we will focus on the results related to tests (2) and (4) and discuss their implications in the context of *in vivo* time-domain optical brain imaging.

2.1 Source Parameters

The source component of the instrument usually comprises one or more pulsed lasers (possibly of tunable wavelength), attenuators, coupling free-space optics, and/or fiber optics (e.g., one or more source fibers, fiber splitters/combiners, or fiber switches). The following parameters are most relevant for characterizing this component: (1) average laser output power P_{laser} (at given repetition rate); (2) center wavelength and width of the laser output spectrum; (3) average laser power delivered to the sample at the source optode(s) P_{source} ; and (4) illuminated area A_{source} on the surface of the sample. A high P_{source} is desirable to increase the signal-to-noise ratio, although safety considerations limit the (total of all wavelengths) source power per unit area. Increasing A_{source} enables P_{source} to be increased, but a too large A_{source} may compromise (lateral) spatial resolution.

2.2 Responsivity of the Detection System

The efficiency of detecting low light levels emerging from tissue is crucial for any *in vivo* photon migration measurement. In general, the responsivity of a detector is the ratio between the measured signal and the magnitude of the input illumination. In the context of the instruments considered here, this is the ratio of the photon count rate and the amount of light emitted by the tissue directly beneath the detector optode. Light emerging from tissue after diffusive propagation can (under certain conditions) be modeled by a uniform Lambertian source^{16,17} with photon radiance L_p (in $\text{s}^{-1} \text{m}^{-2} \text{sr}^{-1}$). The $\cos \Theta$ angular distribution (Θ —angle with the surface normal) is a reasonable approximation in the diffusion regime and for a ratio >1 of the refractive index inside and outside the turbid medium. The lateral uniformity is fulfilled if the detection area is sufficiently far from the source.

The spectral responsivity of the detection system with respect to photon radiance is obtained from the count rate divided by the input photon radiance

$$s_{\text{det}}^L(\lambda) = \dot{N}_{\text{det}}/L_p(\lambda), \quad (1)$$

where the detected count rate \dot{N}_{det} is equal to the time-integrated total photon count N_{tot} divided by the measurement time t_{meas} . It should be noted that s_{det}^L addresses the overall sensitivity, irrespective of time resolution.

To facilitate the understanding of the responsivity of the detection system, the following equation illustrates its major components:

$$\begin{aligned} s_{\text{det}}^L(\lambda) &= A\eta_{\text{det}}(\lambda)\eta_{\text{TCSPC}} \int_0^{2\pi} d\Phi \int_0^{\Theta_{\text{max}}} \cos \Theta T_{\text{optics}}(\Theta, A, \lambda) \sin \Theta d\Theta, \end{aligned} \quad (2)$$

where A is the tissue area from which light is collected, η_{det} is the quantum efficiency of the detector, and η_{TCSPC} is the efficiency of recording single-photon pulses generated by the detector. The integration over the solid angle is performed up to a maximal polar angle Θ_{max} , usually restricted by the numerical aperture (NA) of the detection fiber or fiber bundle. The transmittance T_{optics} of the whole relay optics between tissue and detector includes any losses in the detection fibers or fiber bundles and the (in part angle-dependent) efficiency of light transfer to the detector. The dependence of T_{optics} on A symbolizes the limitations introduced, e.g., by the finite area of the detector or vignetting in the optical path. The efficiency η_{TCSPC} may be influenced by clipping of the single-photon pulse-height distribution by the discriminator threshold, by incomplete recording of the photon histogram or, at very high count rates, by dead-time effects. Many of these factors are not precisely known for a given system, and thus calculating the responsivity according to Eq. (2) would involve significant uncertainty. Therefore, we proposed to directly measure the responsivity.

The measurement of s_{det}^L requires an approximately uniform Lambertian source of known photon radiance. To realize such sources, integrating spheres are often used. However, they are not easy to handle when comparing different instruments in different locations. Therefore, in the present work, dedicated thick-slab phantoms with known diffuse transmittance were used as working standards. These turbid slab phantoms transform a given input power P_{in} (continuous wave or time averaged) from a collimated pencil laser beam into a known wavelength-dependent photon radiance at the exit surface opposite to the laser beam (in a transmission geometry). Input power and output photon radiance are related via the equation

$$L_p(\lambda) = \kappa_p(\lambda)P_{\text{in}}(\lambda), \quad (3)$$

where $\kappa_p(\lambda)$ is a phantom-specific photon transmittance factor (in units of $\text{W}^{-1} \text{s}^{-1} \text{m}^{-2} \text{sr}^{-1}$). This factor includes a transformation of a power (W) into a photon-related quantity, which makes it easy to apply. Combining Eqs. (1) and (3), the responsivity of the detection system can be measured as

$$s_{\text{det}}^L(\lambda) = N_{\text{tot}}/[t_{\text{meas}}\kappa_p(\lambda)P_{\text{in}}(\lambda)]. \quad (4)$$

The unit of s_{det}^L is $\text{m}^2 \text{sr}$.

It is worth noting that there is a straightforward relationship between $\kappa_p(\lambda)$ and the time-integrated diffuse transmittance T_{tot} of the slab phantom opposite to the source position

$$T_{\text{tot}}(\lambda) = \pi\kappa_p(\lambda)E_{\text{phot}}(\lambda), \quad (5)$$

where $E_{\text{phot}} = (hc/\lambda)$ is the photon energy. The factor π results from integration over the full hemisphere assuming a Lambertian distribution.

2.3 DNL of the Timing Electronics

The DNL is the nonuniformity of the time channel width in a TCSPC system.¹⁸ Since the number of photons collected in a given channel is proportional to the channel width, any nonuniformity appears as a modulation in the recorded photon distribution $N(t)$. It can result in systematic errors in TCSPC measurements. In most cases, DNL problems are due to the electronic crosstalk between start and stop pulse signals or the spurious pickup of start- or stop-related signals, such as from laser drivers, and can often be avoided by adequate shielding of detectors and cables. It is important to identify such problems with a DNL measurement. Since they may change when changing the configuration of the instrument or moving it to another location, a routine DNL test is mandatory for quality assurance.

The DNL is recorded as a response to a continuous signal. A battery-powered light source is preferable to avoid any electrical interference. To obtain the DNL with a good signal-to-noise ratio, each time channel should contain $\geq 10^5$ counts. Ideally, the photon counts in all time channels are expected to be equal, i.e., $N_{\text{DNL}}(t)$ should be constant in this case. The deviation from this situation is characterized by the peak-to-peak difference normalized to the mean value

$$\varepsilon_{\text{DNL}} = \frac{N_{\text{DNL,max}} - N_{\text{DNL,min}}}{\bar{N}_{\text{DNL}}}. \quad (6)$$

A correction of measured $N(t)$ can be performed by numerical equalization of the width of time channels based on $N_{\text{DNL}}(t)$ as a measure of their nonuniformity. It is only required if ε_{DNL} exceeds a few percent.

2.4 Temporal IRF

Whereas the tests described above are related to the source or detection component alone, the IRF characterizes the time resolution of the instrument as a whole. Exact knowledge of the IRF is crucial for any model-based reconstruction of tissue optical properties. The IRF is measured by inserting a reference sample between source and detection components of the instrument which contributes a negligible additional temporal dispersion. The IRF depends on the laser pulse shape, the temporal response of the detector and electronics, and pulse broadening due to temporal dispersion in the fiber-optics, and is represented as a convolution of these individual effects. Regarding fiber dispersion, it is essential that the IRF measurement adequately duplicates the conditions of the measurements on tissue, i.e., the angular distribution of the collected light must be similar, which normally implies filling the acceptance angle of the detection fiber or fiber bundles.¹⁹ This is achieved by employing a reference sample consisting of a layer of highly scattering material, e.g., sheets of paper or Teflon tape, small and thin enough to avoid pulse broadening by multiple

scattering within the layer itself.²⁰ To obtain an IRF shape over a wide dynamic range with a good signal-to-noise ratio, a total photon count of at least 10^6 is recommended. Data overflow in the peak region can be avoided by summing up repeated measurements, e.g., of 1-s duration.

It is common to characterize the IRF by its full-width-at-half-maximum (FWHM). However, since long flight-time photons are often of particular interest in time-domain brain imaging, the shape of the trailing edge of the IRF is also important. Signals produced by photomultipliers (PMTs) sometimes exhibit afterpeaks, caused by afterpulses with a short delay²¹ occurring up to a few nanoseconds after the main peak. Meanwhile, single-photon avalanche diodes typically exhibit an exponentially decaying diffusion tail. To study the influence of such features on a measurement, the full profile of the IRF has to be taken into account rather than a single or a few parameters.

Another important characteristic of the IRF is the background that influences the dynamic range. A signal-independent component of the background due to dark counts and residual ambient light can be obtained from a “dark” measurement with the laser source removed. In addition, PMTs often have a signal-dependent background due to afterpulsing caused by positive ions.¹⁸ These afterpulses occur on the time scale of a few microseconds after a single-photon pulse. At laser repetition rates of several 10 MHz, these afterpulses accumulate and lead to a virtually temporally constant, signal-dependent background component. A comprehensive characterization of afterpulsing employs an autocorrelation measurement of detected pulses.¹⁸ In the present work, we propose another, more simplified approach, i.e., the calculation of an “afterpulsing ratio” directly from an IRF and a “dark” measurement. This ratio relates the total counts in the background, after subtraction of the dark background and rescaling to the full laser period T_{laser} (reciprocal repetition rate), to the total counts in the IRF signal $N_{\text{tot,IRF}}$ (after background subtraction):

$$R_{\text{AP}} = \frac{(N_{\text{mean,bkg}} - N_{\text{mean,dark}}) T_{\text{laser}}}{N_{\text{tot,IRF}} \Delta t}, \quad (7)$$

where $N_{\text{mean,bkg}}$ and $N_{\text{mean,dark}}$ are the mean values of the background in the IRF measurement and in the “dark” measurement, respectively, and Δt is the time channel width (typically of the order of 10 ps). $N_{\text{mean,bkg}}$ is determined within an IRF interval of constant background level. It should be noted that the method described here cannot distinguish between the detector afterpulsing background and possible time-independent laser background radiation. However, the types of lasers which are relevant here do not exhibit such a background.

Finally, the stability of the IRF is another relevant factor for the capability of a time-domain brain imager to measure small physiologically induced changes in photon flight time (e.g., due to a functional stimulation of the brain). It is essential to know how quickly the entire system (including the sources, the detectors, and the electronics) reaches thermal equilibrium. We characterized the stability by continuously recording the IRF for at least 1 h after switching on the instrument and by analyzing the total intensity, the temporal position, and the shape of the IRF. Such recordings also provide a measure of the fluctuations in these parameters once the initial stabilization phase is completed.

2.5 Summary of the Protocol and Its Implementation

Table 2 provides an overview of the tests described above, together with the recommended count rates and measuring times for those involving TCSPC measurements. Note that the recommended measuring time typically consists of a number of repeated measurements with a collection time of 1 s, short enough to avoid a memory overflow. All tests should be performed when initially characterizing an instrument, and then repeated following any relevant modification of its components or settings. Some of the measurements, highlighted by (*) in Table 2, are recommended for daily quality control testing, particularly during clinical studies.

Together with the results of the tests and the related specific parameters, it is essential to record the configuration of the instrument with its specific components and all relevant settings of lasers, detectors, and electronics. For our study, we developed and employed a set of spreadsheets to keep track of all this information and to facilitate the comparison of different instruments.

3 Experiment

3.1 Instruments

The various instruments (including, in some cases, several different versions of them), whose performances were assessed according to the BIP protocol, are listed in Table 3. The major goals of the present work were to (1) characterize and compare the time-domain brain imagers of the various participating groups and (2) study the effect of modifying components and

configurations during the process of development and optimization.

Five time-domain optical brain imagers were assessed including their modified versions and a laboratory system for broadband time-domain diffuse spectroscopy. The four brain imagers of PTB, IBIB, and POLIMI (for abbreviations, see caption of Table 3) were developed mainly for depth-selective functional near-infrared spectroscopy (fNIRS) on adults. In each case, the optical sources consisted of picosecond diode lasers at either two or three wavelengths: one around 690 nm and one around 830 nm, and a third optional wavelength in between. The spectral width was typically less than 5 nm. The repetition rates were chosen between 40 and 80 MHz. The laser power exiting the source optodes ranged from several tenths of milliwatts to a few milliwatts. The detection components of these instruments consisted of compact fast PMTs, preamplifiers, and multiboard TCSPC systems. The various brain imagers mainly differed in the particular type of PMTs, the optical systems, and fiberoptic components. These characteristics are summarized in Table 3. The setups PTB_2 and IBIB_2 were based on the original brain imagers, but equipped with alternative detectors. The UCL monstir-II system has been developed to perform three-dimensional optical tomography of the entire newborn infant brain, with special focus on imaging slow hemodynamic activity associated with seizures. It was based on 32 photon counting PMTs, a TCSPC module (SPC-630, Becker & Hickl GmbH, Berlin, Germany), and a tunable supercontinuum pulsed laser (Fianium Ltd., Southampton, United Kingdom). The laser, combined with a pair of acoustic-optic tunable filters (AOTFs), enabled data to be simultaneously

Table 2 Test measurements and recommended parameters; (*)—recommended as daily routine tests. Measurements with picosecond time resolution are indicated by t , while T represents the time scale of minutes to hours.

Part of instrument	Test quantity	Measurements	Count rate	Measuring time	Analysis	Remarks
Source component	Laser power	P_{laser}				Measure source parameters for all source optodes and wavelengths
	Source power (*)	P_{source}				
	Area	A_{source}				
	Wavelength	$\lambda, \Delta\lambda$				
Detection component	Responsivity s_{det}^l	P_{in} N_{tot}	$<0.5 \times 10^6 \text{ s}^{-1}$	$20 \times 1 \text{ s}$	Eq. (4)	Use slab phantom with known $\kappa_p(\lambda)$ (see Sec. 3.2). Attenuate source power by filters to adjust the count rate
	Differential nonlinearity (*)	$N_{\text{DNL}}(t)$	$1 \times 10^6 \text{ s}^{-1}$	$1 \times 100 \text{ s}$	Eq. (6)	Use continuous light source (battery powered). Same TCSPC settings as in main experiment
Entire system (IRF)	IRF profile (*)	$N_{\text{IRF}}(t)$	$1 \times 10^6 \text{ s}^{-1}$	$20 \times 1 \text{ s}$	E.g., FWHM; full profile for simulations	Use thin diffuser to fill the entire acceptance angle of detection fiber or bundle
	Afterpulsing ratio R_{AP}	$N_{\text{IRF}}(t)$ $N_{\text{dark}}(t)$	$1 \times 10^6 \text{ s}^{-1}$	$20 \times 1 \text{ s}$ $20 \times 1 \text{ s}$	Eq. (7)	
	IRF stability	$N_{\text{IRF}}(t, T)$	$1 \times 10^6 \text{ s}^{-1}$	$3600 \times 1 \text{ s}$	E.g., $N_{\text{tot}}(T)$, $m_1(T)$	Extend measuring time necessary to reach stability. Record room temperature in parallel

Table 3 Characteristics of the detection part of the instruments and their modified versions participating in the comparison. Codes (acronyms of institutions; for complete information, see author affiliations): PTB—Physikalisch-Technische Bundesanstalt, IBIB—Nalecz Institute of Biocybernetics and Biomedical Engineering, and POLIMI—Politecnico di Milano, UCL—University College London; the codes of the instruments are consistent with those used in the companion paper¹⁵ parameters of collection fibers or fiber bundles: D —diameter, L —length, and NA—numerical aperture.

Code	Instrument	Detector type (photocathode, diameter)	Detection fiber or bundle			References
			D (mm)	L (m)	NA	
PTB_1	Brain imager “BI-1”	R7400U-02 (multialkali, 8 mm)	4	1.5	0.54	9, 22, 23
PTB_2	– with alternative detector modules	H7422-50 (GaAs, 5 mm)	4	1.5	0.54	24
IBIB_1	Brain imager	R7400U-02 (multialkali, 8 mm)	4	1.5	0.54	25, 26
IBIB_2	– with direct detector	R7400U-02 (multialkali, 8 mm)	–	–	–	27
POLIMI_1	Brain imager “fOXY”	R5900U-20-M4 (multialkali, 9 mm square)	3	1.5	0.57	28, 29
POLIMI_2	Brain imager “fOXY2”	HPM-100-50 (GaAs, 3 mm)	3	1.5	0.57	30
POLIMI_3	Diffuse spectrometer	MCP-PMT R1564U (S1)	1	1.5	0.39	31
UCL	Neonatal brain imager “monstir-II”	H8224P-50 (GaAs, 5 mm)	3	2.5	0.21	32

recorded at any combination of four wavelengths, selected within the range from 600 to 880 nm. The laboratory system POLIMI_3 was devised for high time resolution and coverage of a wide spectral range. It was based on a Ti:sapphire laser and a microchannel-plate PMT (MCP-PMT). All PMTs were manufactured by Hamamatsu Photonics, Hamamatsu, Japan; HPM-100-50 is a hybrid detector module (Becker & Hickl GmbH) based on a Hamamatsu R10467 tube.³³

3.2 “Responsivity” Phantoms

These phantoms were not intended to represent a standard for certain scattering and absorption properties, but rather for (time-integrated) diffuse transmittance. A set of solid, virtually identical phantoms was prepared, characterized, and distributed to the project partners. The five cylindrical solid phantoms were made of epoxy resin with TiO_2 particles added as a scattering medium and black toner as the absorbing medium, following the recipe published by Swartling et al.³⁴ The reduced scattering and absorption coefficients were of the order of 1 and 0.01 mm^{-1} , respectively. The phantoms were machined into a set of 10 cylindrical slices with a 2-cm thickness, 10.5-cm diameter, and smooth surfaces. For the following characterization, 5 out of the 10 phantoms were selected for best homogeneity in the central region. Black PVC housings were manufactured

with central openings on both sides to attach any specific optodes. Figure 1(a) shows one of the phantoms.

The diffuse transmittance factor $\kappa_p(\lambda)$ of these phantoms was measured using the arrangement illustrated in Fig. 1(b). An SC500-6-custom supercontinuum laser with AOTF (Fianium Ltd., Southampton, United Kingdom) was used as a tunable light source. The measurement was repeated at two wavelengths (686 and 808 nm) with more stable picosecond diode lasers (Sepia II, PicoQuant GmbH, Berlin, Germany). The collimated laser beam (diameter $< 5 \text{ mm}$) was directed onto the center of the surface of the phantom via an aperture in the housing of diameter 8.5 mm. Light diffusely transmitted through the phantom was able to pass through another housing aperture, of diameter $2r_1 = (5.00 \pm 0.05) \text{ mm}$, immediately opposite to the entrance aperture. A Si-photodiode detector (S1338-1010BQ, Hamamatsu Photonics) was located at a distance $d = (92.0 \pm 0.5) \text{ mm}$ from the exit surface of the phantom. A diaphragm immediately in front of the photodiode, slightly tilted at 5 deg, had an area determined radiometrically as $A_2 = (51.43 \pm 0.17) \text{ mm}^2$. The average power of the main beam $P_{in,0}(\lambda)$ (several milliwatts) was measured before and after the set of phantom measurements (i.e., all phantoms at all wavelengths) using a calibrated thermopile powermeter (LabMax with head PS19Q, Coherent Inc., Santa Clara, California). The reference power measurements recorded

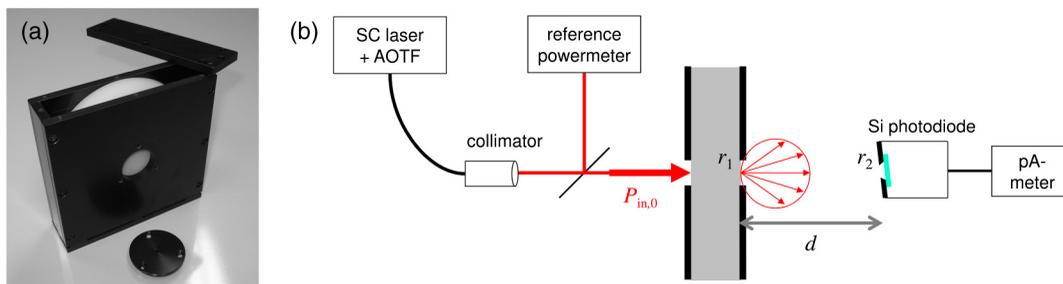


Fig. 1 (a) Responsivity phantom. (b) Setup to measure the diffuse transmittance factor $\kappa_p(\lambda)$.

throughout the phantom measurements were used to correct for changes in power during the experiment. The photocurrent (several nanoampere) was recorded using a Keithley 6485 picoamperemeter. The spectral irradiance responsivity $s_E(\lambda)$ ($\text{A W}^{-1} \text{m}^2$) of the Si photodiode had been calibrated before in the Department of Detector Radiometry and Radiation Thermometry of PTB. The radiation was measured in a small solid angle around the optical axis.

The photon radiance emerging from the phantom was derived as

$$L_{p,0}(\lambda) = \frac{I_{\text{photo}}(\lambda) \lambda}{s_E(\lambda) G hc}, \quad (8)$$

where $G = \pi r_1^2/d^2$ is the geometry factor. The diffuse transmittance factor is calculated from

$$\kappa_p(\lambda) = L_{p,0}(\lambda)/P_{\text{in},0}(\lambda). \quad (9)$$

Figure 2 shows the result of the characterization of the phantoms. The individual phantoms had rather different transmittances, although this is irrelevant for their application since their actual transmittance factors were measured. The phantom slabs 2o, 2u and 4o, 4u, respectively, which were cut from the same epoxy resin blocks, exhibited nearly the same transmittance factors. Equation (5) allows the time-integrated diffuse transmittance of the phantom opposite to the source position to be derived. A typical value of $\kappa_p = 1 \times 10^{20}/(\text{W s m}^2 \text{ sr})$ at $\lambda = 750 \text{ nm}$ corresponds to $T_{\text{tot}} \approx 0.83 \times 10^{-4} \text{ mm}^{-2}$.

The wavelength dependence $\kappa_p(\lambda)$ could be reasonably well approximated by a linear relationship. This feature facilitates the application for any wavelength within the range investigated. The overall relative uncertainty of $\kappa_p(\lambda)$, including the application of the linear approximation, was estimated to be $<10\%$. The comparability of responsivity measurements (at the same wavelength) using different phantoms from the set was much better. The corresponding relative uncertainty due to instrumental noise was $<1\%$ (standard deviation).

The thickness of the responsivity phantom is a parameter that not only is important to achieve a good spatial uniformity of the radiance within the acceptance area of the detector optode, but is

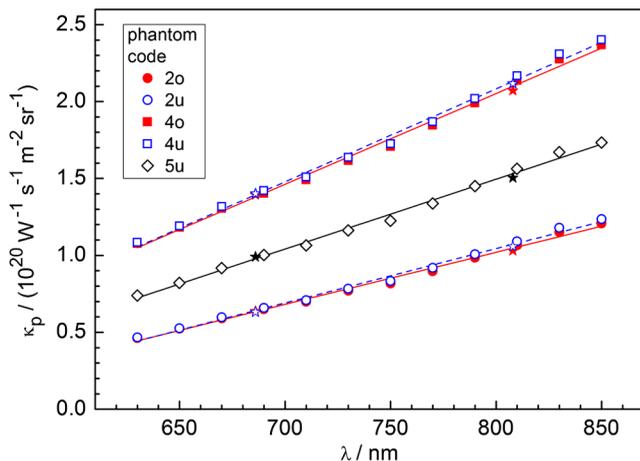


Fig. 2 Wavelength-dependent diffuse transmittance factor $\kappa_p(\lambda)$ of the five phantoms, together with linear approximation (star symbols—additional measurements with picosecond diode lasers as sources). For uncertainties see text.

also essential for the choice of the method to measure $\kappa_p(\lambda)$. For the given optical properties, a thickness of 2 cm turned out to be a reasonable compromise. For a much thicker slab, the straightforward measurement of a photocurrent as described above would no longer be feasible. The assumption of a Lambertian characteristic was checked for the angular range that is typically accepted by fiber bundles. The maximum deviation from a $\cos \Theta$ distribution in the range $0 < \Theta < 30 \text{ deg}$ was found to be less than 3%.

4 Results and Discussion

In this section, we focus on the results related to the responsivity of the detection system and the IRF. These measurements are most relevant when assessing the performance of instruments for clinical time-domain brain imaging. Exemplar simulations illustrate the use of the results of these tests to predict specific aspects of the performance in *in vivo* measurements.

4.1 Responsivity of the Detection System

4.1.1 Comparison of results for various instruments

The responsivity was measured for the various instruments as described in Sec. 3.1 and Table 3. The results obtained for several wavelengths are shown in Fig. 3. The data points pertaining to the instruments POLIMI_1, POLIMI_2, IBIB_1, and IBIB_2 are each connected by dotted lines for clarity. The data for the instruments PTB_1 and PTB_2 are given by symbols connected by vertical solid lines to indicate ranges of responsivity. The detection modules of both types (for PTB_1, see Ref. 22 and for PTB_2, see Ref. 24) contained motor-driven variable attenuators as well as iris diaphragms for independent control of the effective NA. Both options were routinely used in *in vivo* measurements to achieve an optimum count rate at a minimum effective NA, reducing the broadening effect of fiber dispersion on the IRF as far as possible. The symbols marking the upper limits of the ranges correspond to completely open diaphragms and minimum attenuation. The values at the lower limits were obtained for an effective NA of 0.15 and an attenuator transmittance of 0.2. These lower settings were typical for measurements on relatively “transparent” subjects, e.g., elderly patients.

It is interesting to note that the main brain imagers POLIMI_1, IBIB_1, and PTB_1, devised for clinical studies,

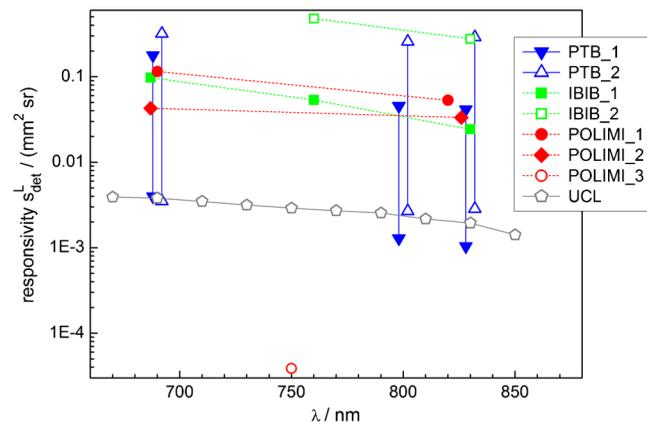


Fig. 3 Responsivity of the detection system of various instruments as a function of wavelength measured for the instruments listed in Table 3. Codes (acronyms of institutions; for complete information, see author affiliations): PTB_1 and PTB_2 was added for clarity.

yielded very similar responsivity values, differing by a factor of 2 at most, in spite of employing different PMTs, fiber bundles, and optical systems. This finding is one factor that enables the comparability of the results of common *in vivo* studies. For these systems, the responsivity clearly decreases with increasing wavelength, which corresponds to the wavelength dependence of the cathode sensitivity of the PMTs. For the detector module with a gallium arsenide (GaAs) PMT (PTB_2), generally higher s_{det}^L values were measured with a flat wavelength dependence. This behavior can be explained by the comparably high and rather constant sensitivity of this photocathode type up to about 850 nm. The brain imager POLIMI_2 that has a detector with the same type of photocathode yet has a smaller diameter (3 mm versus 5 mm for the H7422-50 PMT), shows a similar wavelength dependence, but responsivity values closer to the three main systems mentioned above. The neonatal brain imager of UCL exhibited responsivity values about an order of magnitude lower. This is due to the longer detection bundle with a comparably low NA and other aspects of the optics design.

The highest responsivity values were found for IBIB_2, i.e., a detector directly attached to the surface of the turbid medium without any optics in between. Here, the full size of the photocathode (diameter 8 mm) was effective. On the other hand, for the laboratory system POLIMI_3, the responsivity at 750 nm was considerably smaller. It should be noted that this system was optimized for high time resolution and measurements over a wide spectral range, employing, in particular, a detector with an S1 photocathode of comparably low quantum yield and a fiber of 1-mm diameter rather than a fiber bundle.

To facilitate the interpretation of the results, the following considerations regarding maximum possible values are helpful. Starting from Eq. (2) and assuming that η_{det} , η_{TCSPC} , and T_{optics} are all equal to 1, s_{det}^L is determined by the area and the solid angle over which light is collected. Theoretically, this product cannot exceed $A \times \pi$. The area A represents the smaller value of the optode and detector areas, a realistic value being 10 mm². The factor π (for a Lambertian source) results from integration over the full hemisphere. A maximum realistic aperture half-angle of a fiber bundle is about 30 deg, which results in a factor $\pi/4$ instead and $s_{\text{det,max}}^L \sim 8$ mm² sr. The values found for the brain imagers remain about 1 to 2 orders of magnitude lower, which can be explained by realistic quantum efficiencies and losses in the optical systems.

4.1.2 Simulations: Implications for In Vivo Measurements

The responsivity of the detection system and the laser power applied are the major instrumental determinants for the signal-to-noise ratio that can be achieved in an *in vivo* measurement at a certain source-detector separation. However, knowledge of both values and the photon count rate allows the diffuse reflectance R_{tot} of the tissue under investigation to be determined. This relationship can be expressed as follows:

$$\dot{N}_{\text{det}}/\dot{N}_{\text{in}} = s_{\text{det}}^L R_{\text{tot}}/\pi, \quad (10)$$

where $\dot{N}_{\text{in}} = P_{\text{in}}/E_{\text{photon}}$ is the input photon flux corresponding to the input power, and \dot{N}_{det} is the count rate recorded by the detection system. Equation (10) was derived from Eq. (1) and uses the relation $L_p = \dot{N}_{\text{in}} R_{\text{tot}}/\pi$ that is valid for a Lambertian angular distribution.

Figure 4 displays the results of a simulation to estimate the influence of the responsivity on the maximum source-detector

separation at which a good signal-to-noise ratio can be achieved. According to our experience with *in vivo* time-domain fNIRS measurements, the signal-to-noise ratio was usually sufficient as soon as the count rate \dot{N}_{det} was at least 1×10^6 s⁻¹. The input power (1 mW at 800 nm) was chosen according to the typical values for the brain imagers described in Sec. 3.1. The simulation of the time-integrated diffuse reflectance R_{tot} at separation r was based on the analytical solution of the diffusion equation for a homogeneous semi-infinite medium with extrapolated boundary conditions.³⁵ Considering the case where $\mu_s' = 1$ mm⁻¹ and $\mu_a = 0.01$ mm⁻¹, the given ratio of the detected count rate and input power would be reached at a source-detector separation $r = 5$ cm for a responsivity value in the range of those determined for the clinical brain imagers POLIMI_1, IBIB_1, and PTB_1. A decrease in responsivity of 1 order of magnitude results in a decrease of the maximum r by about 1 cm. The optical properties naturally have a substantial influence on the maximum r . This fact is illustrated in Fig. 4 by the curves for a factor of 2 higher and for lower values of μ_a and μ_s' . If μ_a is increased from 0.01 mm⁻¹ to 0.02 mm⁻¹, then the maximum r drops below 40 mm. Likewise, higher scattering would prevent a measurement at a larger separation. It is known that the intersubject variability of optical properties of the head is substantial. It should be noted that the simulation is based on assumed bulk optical properties while absorption at the tissue surface due to skin color or hair beneath the optodes additionally reduces the amount of light available for detection in *in vivo* measurements.

An association with *in vivo* measurements can be established from a recent study on stroke patients⁹ (brain imager PTB_1, 785 nm). The measurements were performed in the region of the motor cortex, with a source-detector separation of 3 cm. A count rate of $\sim 1 \times 10^6$ s⁻¹ was typically achieved with settings 0.15 for effective NA and 0.20 for attenuation, corresponding to the lower limit of the range indicated for this instrument in Fig. 3. With $s_{\text{det}}^L \sim 1.3 \times 10^{-3}$ mm² sr and an input

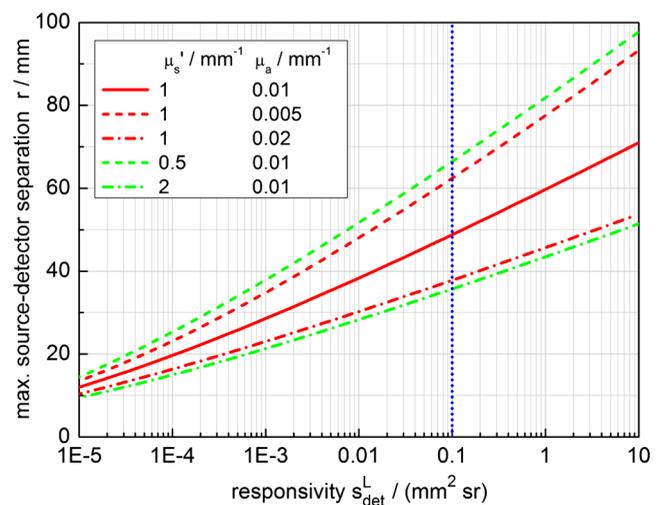


Fig. 4 Maximum source-detector separation r at which a time-resolved diffuse reflectance can be measured with good signal-to-noise ratio, as a function of responsivity of the detection system. The simulation was performed assuming an input power of 1 mW at 800 nm, a detected photon count rate of 1×10^6 s⁻¹, absorption and reduced scattering coefficients as indicated in the legend, and refractive index of the medium $n = 1.4$. The vertical dotted line illustrates a typical responsivity value found for the clinical brain imagers.

power of 5 mW, a time-integrated diffuse reflectance of $R_{\text{tot}} \sim 4 \times 10^{-3} \text{ mm}^{-2}$ can be inferred, a value close to the $R_{\text{tot}} = 5.5 \times 10^{-3} \text{ mm}^{-2}$ obtained for the homogeneous optical properties $\mu'_s = 1 \text{ mm}^{-1}$, $\mu_a = 0.01 \text{ mm}^{-1}$, and $n = 1.4$. However, for measurements on young adult subjects, particularly those with dark hair, similar count rates could only be achieved with a considerably increased responsivity.

4.2 IRF

4.2.1 Temporal profile

In this section, we present examples of measured IRFs for various instruments and discuss the expected impact of their shape on time-domain measurements of brain activation. Figure 5 shows the IRFs measured with three of the clinical brain imagers and a laboratory system (see Table 3). After subtraction of a constant background, all curves were normalized to their maxima and shifted to peak at $t = 0$.

The system POLIMI_3 with an MCP-PMT and supercontinuum laser has the best time resolution as characterized by the FWHM, but also has a fast-decaying trailing edge. The IRFs of the brain imagers are much wider with FWHMs of about

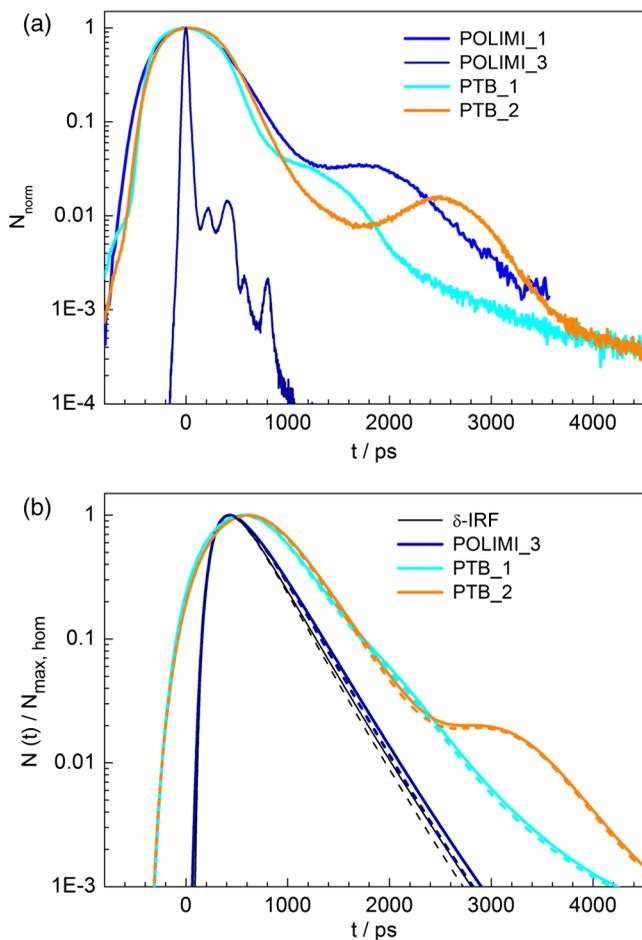


Fig. 5 (a) Instrument response functions (IRFs) of several instruments (see Table 3). (b) Simulated distributions of times of flight (DTOFs) for a homogeneous semi-infinite medium without (solid lines) and with (dashed lines) deep absorption perturbation (see text) for a delta-pulse IRF and after convolution with three of the IRFs shown in (a). Each curve is normalized to the maximum of the corresponding unperturbed curve.

750 ps. This fact can be explained by the use of picosecond diode lasers operated close to their maximum output power and fiber dispersion effects in high-aperture fiber bundles. Another feature of these IRFs is the occurrence of distinct shoulders and afterpeaks on the trailing edge. They obviously depend on the type of photocathode, which differs for the instruments shown.

The influence of the shape of the IRF on time-domain fNIRS measurements can be studied by combining simulations of light propagation to mimic brain activation with measured IRFs. This approach facilitates the interpretation of measurements on inhomogeneous phantoms as proposed in the nEUROPt protocol.¹⁵ The influence of the IRF shape on the contrast measured for a deep absorbing perturbation (mimicking an absorption change in the cortex) can be illustrated by means of the following example using a simple simulation. An unperturbed DTOF was first derived from the solution of the diffusion equation for a semi-infinite homogeneous medium with a transport scattering coefficient $\mu'_s = 1 \text{ mm}^{-1}$, absorption coefficient $\mu_a = 0.01 \text{ mm}^{-1}$, refractive indices inside and outside the medium of 1.4 and 1, respectively, and employing extrapolated boundary conditions. A source-detector separation $r = 2 \text{ cm}$ was assumed. A perturbation was then modeled as a point-like absorber according to Ref. 36. The inhomogeneity was buried 1.5-cm deep in the midplane between the source and detector. To define the magnitude of the perturbation, the product of the change in the absorption coefficient $\Delta\mu_a$ and the volume of the inhomogeneity was assumed to be 0.05 cm^2 . Figure 5(b) shows the normalized DTOFs obtained for the pure simulation, corresponding to a delta-pulse IRF, and after convolution with three noticeably different cases of real IRFs that were selected from those plotted in Fig. 5(a). The broad IRFs with afterpeaks cause a smaller difference between the unperturbed and the perturbed curves compared to the ideal case (δ -IRF). A decrease in the contrast, i.e., the relative difference with and without perturbation, is particularly evident in the region influenced by a strong afterpeak such as that for the system PTB_2. The presence of an afterpeak in the system IRF causes a small fraction of shorter-flight-time photons to be detected within the tail of the DTOF, thus contaminating a measurement of later arriving photons and reducing the contrast. On the contrary, a narrow IRF with a fast decaying tail (POLIMI_3) only has a minor influence on the contrast. It should be noted that this influence of the IRF on contrast could, in principle, be eliminated by deconvolving the IRF from a measured DTOFs prior to calculating the contrast.

Figure 6 displays the relative contrast obtained from a time-window analysis for the simulated DTOFs based on all IRFs from Fig. 5(a). For the ideal case (δ -IRF), the absolute value of the contrast for the deep inclusion is low for short photon flight times, but increases monotonically at longer flight times. The narrow IRF of POLIMI_3 leads to contrast values that are only slightly worse than those of the δ -IRF case. The finite FWHM only has a marginal influence, since it is small compared with the width of the DTOF. However, the three broad IRFs with afterpeaks lead to substantially lower absolute contrast values. In particular, an imprint of the shape of the shoulders and afterpeaks on the IRFs can be discerned in the shape of the contrast curve as a function of the position of the time window. The temporal position of these imprints not only depends on the IRF, but also on the amount of broadening in the medium which is, e.g., influenced by the source-detector separation (data not shown).

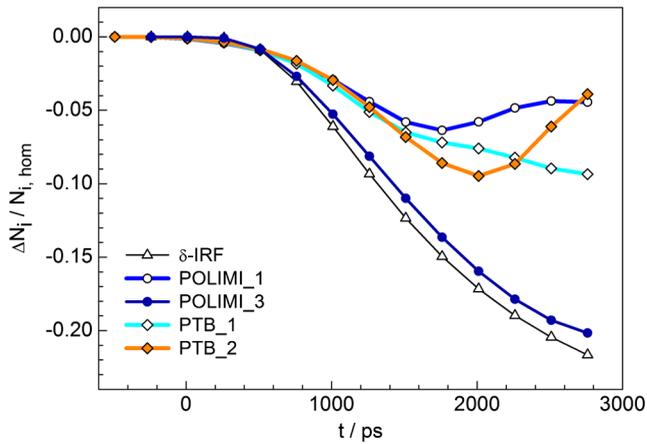


Fig. 6 Relative contrast due to a deep absorption perturbation (see text) for photon counts in time windows of 250-ps width derived from the DTOFs after convolution with the IRFs shown in Fig. 5(a). For comparison, the case of a delta-pulse IRF is plotted. The position of the symbols corresponds to the lower bounds of the respective time windows.

Note that the time-window analysis (Fig. 6) was performed in only a part of the full time range of Fig. 5. At much later times, the decrease in signal will be too significant to maintain a reasonable signal-to-noise ratio. A contrast-to-noise analysis for the time-window approach would be a logical continuation of this analysis but is beyond the scope of the present paper.

4.2.2 Stability

Measurements of the IRF stability are relevant to (1) determine the duration of the necessary warm-up phase prior to *in vivo* measurements and (2) assess fluctuations after completion of the warm-up. From a time series of IRF measurements, two characteristics of the temporal profile are derived, i.e., the integral N_{tot} and the first moment m_1 . The first moment is calculated as $m_1 = \sum_{i=a}^b iN_i/N_{\text{tot}} \cdot \Delta t$ from the counts N_i in the time channels (width Δt) of the histogram memory between limits a and b . The integral reflects the changes in laser power and also in detector sensitivity, e.g., due to changes in the high voltage supplied to the PMTs. The first moment provides information about timing drifts and jitter.

Figure 7 shows exemplar stability results for the total photon count and the first moment of the IRF. Both integration limits (a , b) were set at 1% of the maximum. The results for the POLIMI_2 brain imager are related to the laser operating at 830 nm and were measured every 1 min with a collection time of 1 s at a count rate of about 10^6 s^{-1} . Similar results were obtained for the second wavelength and for the other detectors of this instrument. The integral varied by about 5% over 7 h with a rather monotonic decrease within the first 2 h after system switch-on. After about 100 min, subsequent intensity variation was less than 1%. The first moment varied by about 250 ps in total, but after 300 min, subsequent variation was within ± 5 ps. The major reason for the observed drift is the (thermal) stabilization of the picosecond diode laser. A second example is the configuration of PTB_2 with its optional diode laser modules BHL P-700 (Becker & Hickl GmbH). These laser modules operating at 785 nm were employed in a clinical study of bedside monitoring of cerebral perfusion.⁹ Their fast warm-up time (about 10 min) and stable behavior were essential for this

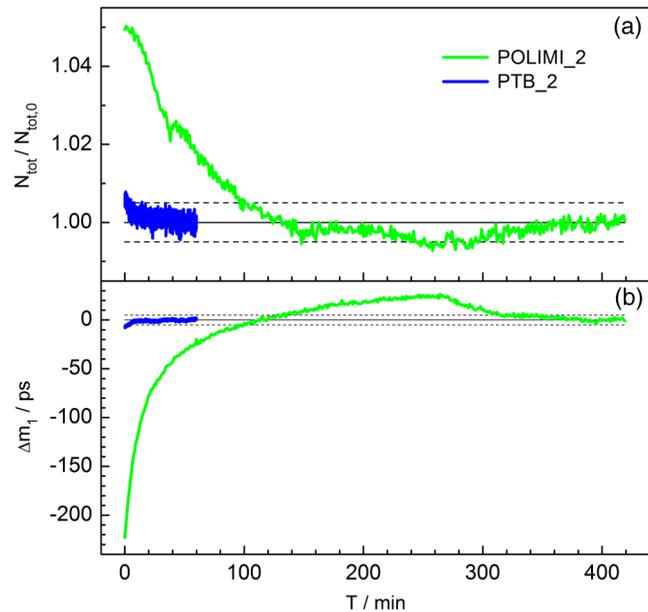


Fig. 7 An example of stability test for (a) total photon count N_{tot} and (b) first moment m_1 of the IRF for one detector of POLIMI_2, recorded for 7 h every minute, and of PTB_2 (operated at 785 nm), recorded for 1 h every second, in both cases with a collection time of 1 s and a count rate of about 10^6 s^{-1} . The thin dashed lines indicate the ranges of $\pm 0.5\%$ for N_{tot} and ± 5 ps for m_1 .

application. The high timing stability is achieved by temperature stabilization of the laser diode and by deriving the synchronization signal for the TCSPC timing directly from the pins of the laser diode.

Similar to the approach pursued in the previous section, such stability measurements can be used to study the impact of drifts and fluctuations on *in vivo* measurements. For example, uncertainty in the time origin t_0 (the time at which photons enter the medium) is relevant when fitting models to DTOFs which assume a known value of t_0 . The impact of instrumental fluctuations on fNIRS signals can be investigated by applying the same procedures used for *in vivo* measurement to IRF stability measurements, including, for example, block averaging as is typically performed in functional activation studies. This allows an estimation to be made of the contribution of the instrumental fluctuations to measured fNIRS signals.

5 Conclusions

The BIP protocol for time-domain optical brain imagers comprises a set of tests addressing the essential characteristics of all components of the system and most especially of the source and detector. The present paper introduces the individual tests and provides guidance for their implementation. While some of the tests, including the characterization of differential nonlinearity and IRF, are already commonly employed, the responsivity test has been specifically developed for this protocol. The design and characterization of the dedicated phantoms produced for this test were described in detail.

The results presented in this work were focused on the two particularly relevant measures: responsivity and IRF. The quantification and comparison of the responsivity of several instruments and laboratory setups of four different partner institutions of the nEUROpt project provided new insights. Since the efficiency of photon detection is essential for achieving a good

signal-to-noise ratio in *in vivo* studies (particularly for dynamic measurements), the assessment of responsivity is particularly useful in the process of instrumental development. Subsequent tests allow possible degradation of detectors or of alignment of optics in the detection path to be assessed. Furthermore, the difference between actual responsivity values and values estimated based on Eq. (2) can facilitate the identification of deficiencies in the detection system of a diffuse optical instrument. Moreover, quantitative knowledge of the responsivity of the instrument could be utilized to derive additional information from *in vivo* measurements by quantifying diffuse reflectance or transmittance of the tissue under investigation for a given source-detector geometry.

It should be noted that the assessment and interpretation of the responsivity of the detection system to some extent rely on the assumption of a Lambertian angular profile for the light exiting the phantom or tissue. This assumption is often invalid,¹⁷ particularly when measuring over small (approximately few millimeter) source-detector separations. In the diffuse regime, relevant to brain or breast imaging applications, the deviation from a Lambertian profile depends on the ratio of the refractive index inside and outside the turbid medium.³⁷ However, for biological tissue and solid phantoms in air where the ratio is around 1.4 to 1.6, this deviation is no larger than a few percent. The degree of roughness of the surface of the tissue or phantom is another factor that may limit the agreement between measurements and simulations of light propagation, in particular the angular distribution of outgoing radiation.^{17,37}

The measurement of the IRF, the second focus of the present paper, is highly relevant when assessing the performance of any time-domain instrument. The width (FWHM) of the IRF is often used to specify the time resolution. However, this single parameter is not sufficient to characterize the entire shape of the IRF which is highly dependent on the type of detector. In time-domain brain imaging, longer flight-time photons play an important role, and thus any afterpeaks and slowly decaying tails of the IRF are particularly relevant. Hence, the full temporal profile of the IRF needs to be considered. We demonstrated the utilization of measured IRFs in simulations to identify the impact of their shape on the contrast achieved for brain activation measurements. Apart from its temporal shape, the stability of the IRF is important, as was illustrated with exemplar measurements.

The BIP tests can be performed as a stand-alone protocol, but they also complement the more application-oriented MEDPHOT and nEUROpt protocols,^{13,15} and are relevant when interpreting the results of these protocols. Depending on the specific algorithm employed for data analysis, the properties of the instrument (e.g., width of the IRF, drift of laser power, or timing) may significantly influence the results of the measurements. Robustness against instrumental artifacts is an important requirement which can be assessed by the combination of basic and “high-level” tests, complemented by simulations taking into account the actual instrumental properties.

The BIP protocol was developed with the focus on time-domain optical brain imagers. Nevertheless, the complete set of basic instrumental tests can also be applied to the assessment of other time-domain photon-migration instruments based on TCSPC technology, such as time-domain optical mammography systems and diffuse spectrometers. The range of applicability of some of the individual tests extends even further. The assessment of responsivity is not restricted to time-domain instruments

and can, in principle, be adapted for any instrument that measures diffuse reflectance or transmittance of tissues. Meanwhile, the tests which characterize time-resolved measurements (e.g., IRF, temporal stability, and differential nonlinearity) and the use of simulations to assess the influence of nonideal behavior on measurements could be applied to any TCSPC-based systems, including instrumentation for recording fluorescence lifetimes.

The BIP protocol was part of a more comprehensive assessment and comparison of time-domain optical brain imagers. The nEUROpt protocol, a topic of the companion paper,¹⁵ is based on measurements on inhomogeneous phantoms and addresses the specific capability of these instruments to detect, localize, and quantify brain activation.

Acknowledgments

The research leading to these results has received funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreements No. FP7-HEALTH-F5-2008-201076 and LASERLAB EUROPE No. 284464. The authors wish to thank Alwin Kienle (Institut fuer Lasertechnologien in der Medizin und Messtechnik, Ulm, Germany) as well as Fabrizio Martelli and Giovanni Zaccanti (Università degli Studi di Firenze, Florence, Italy) for helpful discussions.

References

1. D. R. Leff et al., “Diffuse optical imaging of the healthy and diseased breast: a systematic review,” *Breast Cancer Res. Treat.* **108**(1), 9–22 (2008).
2. A. Poellinger et al., “Breast cancer: early- and late-fluorescence near-infrared imaging with indocyanine green—a preliminary study,” *Radiology* **258**(2), 409–416 (2011).
3. R. Choe et al., “Diffuse optical tomography of breast cancer during neoadjuvant chemotherapy: a case study with comparison to MRI,” *Med. Phys.* **32**(4), 1128–1139 (2005).
4. L. Enfield et al., “Monitoring the response to neoadjuvant hormone therapy for locally advanced breast cancer using three-dimensional time-resolved optical mammography,” *J. Biomed. Opt.* **18**(5), 056012 (2013).
5. P. Taroni et al., “Noninvasive assessment of breast cancer risk using time-resolved diffuse optical spectroscopy,” *J. Biomed. Opt.* **15**(6), 060501 (2010).
6. P. Taroni et al., “Optical identification of subjects at high risk for developing breast cancer,” *J. Biomed. Opt.* **18**(6), 060507 (2013).
7. M. Ferrari and V. Quaresima, “A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application,” *NeuroImage* **63**(2), 921–935 (2012).
8. W. Weigl et al., “Assessment of cerebral perfusion in post-traumatic brain injury patients with the use of ICG-bolus tracking method,” *NeuroImage* **85**(Pt 1), 555–565 (2014).
9. O. Steinkellner et al., “Optical bedside monitoring of cerebral perfusion: technological and methodological advances applied in a study on acute ischemic stroke,” *J. Biomed. Opt.* **15**(6), 061708 (2010).
10. H. Obrig, “NIRS in clinical neurology—a ‘promising’ tool?,” *NeuroImage* **85** (Pt 1), 535–546 (2014).
11. J. Hwang, J. C. Ramella-Roman, and R. Nordstrom, “Introduction: feature issue on phantoms for the performance evaluation and validation of optical medical imaging devices,” *Biomed. Opt. Express* **3**(6), 1399–1403 (2012).
12. T. Durdurán et al., “Diffuse optics for tissue monitoring and tomography,” *Rep. Prog. Phys.* **73**(7), 076701 (2010).
13. A. Pifferi et al., “Performance assessment of photon migration instruments: the MEDPHOT protocol,” *Appl. Opt.* **44**(11), 2104–2114 (2005).
14. A. E. Cerussi et al., “Tissue phantoms in multicenter clinical trials for diffuse optical technologies,” *Biomed. Opt. Express* **3**(5), 966–971 (2012).

15. H. Wabnitz et al., "Performance assessment of time-domain optical brain imagers, part 2: nEUROpt protocol," *J. Biomed. Opt.* **19**(8), 086012 (2014).
16. A. Kienle et al., "Spatially resolved absolute diffuse reflectance measurements for noninvasive determination of the optical scattering and absorption coefficients of biological tissue," *Appl. Opt.* **35**(13), 2304–2314 (1996).
17. A. Kienle and F. Foschum, "250 years Lambert surface: does it really exist?," *Opt. Express* **19**(5), 3881–3889 (2011).
18. W. Becker, *Advanced Time-Correlated Single Photon Counting Techniques*, Springer Series in Chemical Physics, Vol. 81, Springer Berlin Heidelberg (2005).
19. A. Liebert et al., "Fiber dispersion in time domain measurements compromising the accuracy of determination of optical properties of strongly scattering media," *J. Biomed. Opt.* **8**(3), 512–516 (2003).
20. J. Carlsson et al., "Time-resolved studies of light propagation in paper," *Appl. Opt.* **34**(9), 1528–1535 (1995).
21. Hamamatsu Photonics K.K. (editorial committee), *Photomultiplier Tubes—Basics and Applications*, 3rd ed., Hamamatsu Photonics K.K., Hamamatsu City (2007).
22. H. Wabnitz et al., "A time-domain NIR brain imager applied in functional stimulation experiments," *Proc. SPIE* **5859**, 58590H (2005).
23. H. Wabnitz et al., "Time-resolved near-infrared spectroscopy and imaging of the adult human brain," *Adv. Exp. Med. Biol.* **662**, 143–148 (2010).
24. A. Hagen et al., "Development of a multi-channel time-domain fluorescence mammograph," *Proc. SPIE* **6434**, 64340Z (2007).
25. M. Kacprzak et al., "Time-resolved optical imager for assessment of cerebral oxygenation," *J. Biomed. Opt.* **12**(3), 034019 (2007).
26. D. Milej et al., "Time-resolved multi-channel optical system for assessment of brain oxygenation and perfusion by monitoring of diffuse reflectance and fluorescence," *Opt.-Electron. Rev.* **22**(1), 55–67 (2014).
27. A. Liebert et al., "Assessment of inflow and washout of indocyanine green in the adult human brain by monitoring of diffuse reflectance at large source-detector separation," *J. Biomed. Opt.* **16**(4), 046011 (2011).
28. D. Contini et al., "Multi-channel time-resolved system for functional near infrared spectroscopy," *Opt. Express* **14**(12), 5418–5432 (2006).
29. D. Contini et al., "A multichannel time-domain brain oximeter for clinical studies," *Proc. SPIE* **7369**, 73691D (2009).
30. R. Re et al., "Multi-channel medical device for time domain functional near infrared spectroscopy based on wavelength space multiplexing," *Biomed. Opt. Express* **4**(10), 2231–2246 (2013).
31. P. Taroni et al., "Diffuse optical spectroscopy of breast tissue extended to 1100 nm," *J. Biomed. Opt.* **14**(5), 054030 (2009).
32. R. J. Cooper et al., "MONSTIR II: a 32-channel, multispectral, time-resolved optical tomography system for neonatal brain imaging," *Rev. Sci. Instrum.* **85**(5), 053105 (2014).
33. W. Becker et al., "FLIM and FCS detection in laser-scanning microscopes: increased efficiency by GaAsP hybrid detectors," *Microsc. Res. Tech.* **74**(9), 804–811 (2011).
34. J. Swartling, J. S. Dam, and S. Andersson-Engels, "Comparison of spatially and temporally resolved diffuse-reflectance measurement systems for determination of biomedical optical properties," *Appl. Opt.* **42**(22), 4612–4620 (2003).
35. D. Contini, F. Martelli, and G. Zaccanti, "Photon migration through a turbid slab described by a model based on diffusion approximation. I. Theory," *Appl. Opt.* **36**(19), 4587–4599 (1997).
36. S. Carraresi et al., "Accuracy of a perturbation model to predict the effect of scattering and absorbing inhomogeneities on photon migration," *Appl. Opt.* **40**(25), 4622–4632 (2001).
37. F. Martelli et al., *Light Propagation through Biological Tissue and Other Diffusive Media*, Chapter 4.8, SPIE, Bellingham, Washington (2010).

Biographies of the authors are not available.