Performance assessment of time-domain optical brain imagers, part 2: nEUROPt protocol

Heidrun Wabnitz
Alexander Jelzow
Mikhail Mazurenka
Oliver Steinkellner
Rainer Macdonald
Daniel Milej
Norbert Żołek
Michał Kacprzak
Piotr Sawosz
Roman Maniewski
Adam Liebert
Salavat Magazov
Jeremy Hebden
Fabrizio Martelli
Paola Di Ninni
Giovanni Zaccanti

Alessandro Torricelli
Davide Contini
Rebecca Re
Lucia Zucchelli
Lorenzo Spinelli
Rinaldo Cubeddu
Antonio Pifferi

SPIE.

Performance assessment of time-domain optical brain imagers, part 2: nEUROPt protocol

Heidrun Wabnitz,[a,][*] Alexander Jelzow,[a] Mikhail Mazurenka,[a] Oliver Steinkellner,[a] Rainer Macdonald,[a] Daniel Milej,[b] Norbert Zołek,[b] Michał Kacprzak,[b] Piotr Sawosz,[b] Roman Maniewski,[b] Adam Liebert,[b] Salavat Magazov,[c] Jeremy Hebden,[c] Fabrizio Martelli,[d] Paola Di Ninni,[d] Giovanni Zaccanti,[d] Alessandro Torricelli,[e] Davide Contini,[e] Rebecca Re,[e] Lucia Zucchelli,[e] Lorenzo Spinelli,[f] Rinaldo Cubeddu,[e,f] and Antonio Pifferi[e,f]

[a]Physikalisch-Technische Bundesanstalt (PTB), Abbestraße 2-12, 10587 Berlin, Germany
[b]Nałęcz Institute of Biocybernetics and Biomedical Engineering, Polish Academy of Sciences, ul. Trojdena 4, 02-109 Warsaw, Poland
[c]University College London, Department of Medical Physics and Biomedical Engineering, Gower Street, London WC1E 6BT, United Kingdom
[d]Dipartimento di Fisica e Astronomia dell'Università degli Studi di Firenze, Via G. Sansone 1, 50019 Sesto Fiorentino, Firenze, Italy
[e]Politecnico di Milano–Dipartimento di Fisica, Piazza Leonardo da Vinci 32, 20133 Milano, Italy
[f]Istituto di Fotonica e Nanotecnologie, CNR, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

Abstract. The nEUROPt protocol is one of two new protocols developed within the European project nEUROPt to characterize the performances of time-domain systems for optical imaging of the brain. It was applied in joint measurement campaigns to compare the various instruments and to assess the impact of technical improvements. This protocol addresses the characteristic of optical brain imaging to detect, localize, and quantify absorption changes in the brain. It was implemented with two types of inhomogeneous liquid phantoms based on Intralipid and India ink with well-defined optical properties. First, small black inclusions were used to mimic localized changes of the absorption coefficient. The position of the inclusions was varied in depth and lateral direction to investigate contrast and spatial resolution. Second, two-layered liquid phantoms with variable absorption coefficients were employed to study the quantification of layer-wide changes and, in particular, to determine depth selectivity, i.e., the ratio of sensitivities for deep and superficial absorption changes. We introduce the tests of the nEUROPt protocol and present examples of results obtained with different instruments and methods of data analysis. This protocol could be a useful step toward performance tests for future standards in diffuse optical imaging. © *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.19.8.086012]

1 Introduction

In biomedical optics, the characterization of instrumentation with the help of appropriate tissue-simulating phantoms[1] plays an important role, ranging from proof-of-principle tests at early stages of the development to quality assurance during routine clinical application.[2] Numerous performance studies have been carried out by individual research groups; however, there are only a few multilaboratory efforts to characterize and to compare multiple instruments. Such approaches are essential to provide a sound basis for achieving quantitatively comparable results in clinical studies. They require the development of standardized guidelines to perform the same measurements on phantoms with the same properties in different laboratories. In the field of photon migration instruments, the first systematic study of different instruments was performed on the basis of the MEDPHOT protocol.[3] Recently, the development and testing of phantoms has 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 spectroscopic imaging.[4] In the context of fluorescence imaging of tissues, phantoms were developed to

characterize and compare imaging systems and to train surgeons.[5]

The present work is related to the characterization of instruments for time-domain optical brain imaging. This effort was part of the tasks of the European project nEUROPt (FP7-HEALTH-F5-2008-201076) that aimed to develop new time-domain systems for optical imaging of the brain based on technological advances as well as novel methodological approaches. Common protocols were developed to provide guidelines for the assessment and comparison of instruments in coordinated measurement campaigns, for the estimation of the impact of technological and methodological advances on their performances, and for supporting quality assurance by routine tests, in particular, during clinical studies. This work represents a contribution to the ongoing efforts toward quality control[6] and standardization in the emerging field of functional near-infrared spectroscopy (fNIRS),[7,8] in general, and more specifically, for time-domain fNIRS imaging.[9]

The nEUROPt protocol was specifically developed to address the characteristic of instruments for optical brain imaging to detect, localize, and quantify changes in the optical properties of the brain and to strive to eliminate the influence of changes in extracerebral tissues on the measurement. It is focused on the assessment of sensitivity, spatial resolution, and quantification of an absorption change $\Delta \mu_a$ in the cerebral cortex as the most relevant physical quantity related to neurological

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

applications of diffuse optical imaging. The complete performance assessment of the time-domain optical brain imagers also included other relevant aspects covered by the following protocols: (1) the previously developed MEDPHOT protocol[3] that evaluates the capability of an instrument to measure the optical properties (absorption coefficient μ_a and reduced scattering coefficient μ_s') of a homogeneous turbid medium and (2) the new Basic Instrumental Performance (BIP) protocol that measures instrumental characteristics in a direct way as described in a companion paper.[10]

Given the nature of the problems concerned with neurological applications of diffuse optical imaging, inhomogeneous turbid media had to be considered and specific quantities (measurands) for the individual tests had to be defined. There are numerous accounts on the construction and application of phantoms to model various aspects of the tissue structure of the head and of brain activation, e.g., layered gel phantoms,[11] liquid phantoms with variable absorption in simple or more realistic geometries,[12,13] and dynamic solid phantoms with absorption changes induced by mechanical,[14] thermochromic,[15] or electrochromic[16] methods. The nEUROPt protocol was implemented with liquid phantoms based on aqueous dilutions of Intralipid and India ink, taking into account the ease of preparation and replication as well as the achievable accuracy and reproducibility of optical properties. To model localized absorption changes related to a localized brain activation, solid black inclusions were used. This choice was supported by the derivation of an equivalence relation between a realistic absorption change in a finite volume and the perturbation produced by a small black object.[17] Further, the liquid phantom also allowed a layered medium to be mimicked by separating two compartments by means of a thin Mylar foil.[18] We defined and applied tests and measurands, such as contrast, contrast-to-noise ratio, spatial resolution, and depth selectivity. The explicit specification of the operational parameters for the implementation of the protocol was another prerequisite to ensure the comparability of results across various laboratories.

This paper is organized as follows. First the measurands and individual tests are introduced, followed by the description of the liquid phantoms and the measurement conditions relevant for the implementation of the protocol. A number of instruments which were assessed during the joint measurement campaigns within the nEUROPt project are introduced, and finally, typical results are reported, which were obtained with the two types of phantoms and two different approaches for data analysis.

2 Definition of the Tests

The nEUROPt protocol is designed to characterize the performances of instruments for optical brain imaging by focusing on application-oriented figures while treating the whole instrument as a black box. Thus, the performance assessment is directed toward the final results of the measurement related to the clinical application, mimicked by appropriate phantom measurements. The results of these measurements are influenced by instrumental features and settings as well as by the analysis procedure. If the influence of hardware aspects of different instruments is to be compared, the experimental conditions and the analysis need to be kept the same. Similarly, the protocol can also be used to assess the performances of different algorithms of data analysis, based either on forward simulations or on a set of experimental data obtained in phantom measurements.

The protocol consists of a total of six tests which address three main features that are of direct relevance for the performance of optical brain imaging in clinical applications, namely (1) sensitivity, (2) spatial resolution, and (3) quantitation of absorption changes in the brain. Each feature is characterized by two tests. In particular, the sensitivity is described by contrast and contrast-to-noise ratio, the spatial resolution by depth sensitivity and lateral resolution, and the quantitation by accuracy and linearity.

The nEUROPt protocol was developed in the context of time-domain brain imaging, which relies on the measurement of time-resolved diffuse reflectance, most commonly by time-correlated single photon counting (TCSPC). The measurement provides the temporal profile $N(t)$ of the histogram of collected photons, also denoted as distribution of times of flight of photons (DTOF). Although the results presented in this paper refer to this specific kind of measurement, it is worth noting that the nEUROPt protocol is not restricted to time-resolved measurements, but can also be applied to continuous wave (cw) or frequency-domain instruments.

To ensure the applicability of the tests to any mode of measurement and data analysis, the tests were formulated in a generalized way. The tests of the nEUROPt protocol can be applied to any measurand M that reflects absorption changes in the brain. M can represent, in particular, (1) absorption changes $\Delta\mu_a$ retrieved by specific reconstruction algorithms and (2) semiempirical quantities as, for example, photon counts in parts of a measured DTOF (time windows) or moments of the DTOF (integral, mean time of flight, variance). These semi-empirical quantities can be regarded as primary steps in the analysis of the measured DTOFs that have been routinely used within the nEUROPt consortium. Here, they serve as examples of algorithms with different features.

In the following, the individual tests are introduced and the related measurands are first defined in an abstract manner. In order to apply the protocol in practice, a specific implementation identifying suitable phantoms and measurement conditions has to be specified. A possible implementation—the one actually used to compare different brain imagers within the nEUROPt project—will be presented in Sec. 3.

2.1 Sensitivity

The goal of this section is to define tests for the assessment of the detection capabilities of the techniques with respect to small absorption changes in the cortex. Two tests are identified, i.e., contrast and contrast-to-noise ratio.

2.1.1 Contrast

The primary assessment of the effect of a given absorption change in the compartment i on the measurand M is to consider the difference with respect to the baseline state (prior to the absorption change) M_0

$$\Delta M_i = M_i - M_0 \tag{1}$$

or the contrast, expressed as a relative difference

$$C_i = \Delta M_i / M_0. \tag{2}$$

Which of these options is more appropriate depends on the particular measurand M. For example, for the photon count N_k in a k'th time window, only the ratio $\Delta N_{k,i}/N_{k,0}$ is relevant,

whereas the attenuation related to a time window requires that the absolute difference $\Delta A_{k,i} = -\ln(N_{k,i}/N_{k,0})$ be considered. In other cases, e.g., for moments of DTOFs, both options can be applied.

A valid comparison of contrast is feasible only within measurands of the same kind, for example, within various ratios of photon counts in time windows. Comparison of measurands of different kinds is possible on the basis of the following tests: contrast-to-noise ratio, depth selectivity, lateral spatial resolution, and linearity.

2.1.2 Contrast-to-noise ratio

The detectability of a small absorption change $\Delta\mu_{a,i}$ in the compartment i depends on the contrast-to-noise ratio CNR_i, i.e., the ratio between the corresponding change in the measurand ΔM_i and the uncertainty of M due to random effects.

$$\text{CNR}_i = \frac{\Delta M_i}{\sigma(M_0)}. \tag{3}$$

$\sigma(M_0)$ is the standard deviation of a series of repeated measurements for the unperturbed (baseline) state. Typically, the major source of noise is photon noise. Hence, the contrast-to-noise ratio depends on the detected intensity and, in particular, on the injected laser power, on the source–detector separation, on the responsivity of the detection system,[10] as well as on the measuring time t_{meas} of the individual measurements. In addition, $\sigma(M_0)$ can be influenced by various random fluctuations due to instabilities in the measurement system, e.g., laser power.

2.2 Spatial Resolution

Depth discrimination and lateral localization of absorption changes are major tasks in optical brain imaging. Due to strong scattering in the tissue (typically $\mu_s' \sim 1 \text{ mm}^{-1}$), the spatial region in which an absorption change influences a measurement of diffuse reflectance has extensions on the order of 1 cm. In addition, the attenuation of light in tissue makes the detection of absorption changes in the cortex a challenging task. Depth selectivity is a practicable test related to depth localization and depth resolution. The spatial resolution parallel to the surface is addressed by the test lateral spatial resolution.

2.2.1 Depth selectivity

This test addresses the capability of instruments and/or methods of data analysis to distinguish between absorption changes occurring in different compartments of the head, in particular, in the cortex versus the overlying tissues. It compares the sensitivity of the measurand M to a small absorption change $\Delta\mu_a$ in the lower compartment (index 2, corresponding to the cortex) with the sensitivity to an absorption change in the upper compartment (index 1, corresponding to extracerebral, superficial tissue) by considering the ratio

$$\begin{aligned} S_{2,1} &= \frac{S_{M,2}}{S_{M,1}} \\ &= \frac{[M(\mu_{a,1},\mu_{a,2}+\Delta\mu_{a,2}) - M(\mu_{a,1},\mu_{a,2})]/\Delta\mu_{a,2}}{[M(\mu_{a,1}+\Delta\mu_{a,1},\mu_{a,2}) - M(\mu_{a,1},\mu_{a,2})]/\Delta\mu_{a,1}}. \end{aligned} \tag{4}$$

For ease of implementation, the reduced scattering coefficient μ_s' is kept constant and equal in both compartments.

The sensitivity ratio $S_{2,1}$ is derived from three measurements, $M(\mu_{a,1},\mu_{a,2}) = M_0$ (baseline state), $M(\mu_{a,1},\mu_{a,2}+\Delta\mu_{a,2})$ with a change in the lower compartment only, and $M(\mu_{a,1}+\Delta\mu_{a,1},\mu_{a,2})$ with a change in the upper compartment only. For $\Delta\mu_{a,1} = \Delta\mu_{a,2}$, the ratio $S_{2,1}$ equals the ratio of the corresponding contrasts, which is valid for both cases of contrast definitions, see Eqs. (1) and (2).

The ratio $S_{2,1}$ is dimensionless and can be used to compare measurands of any kind and dimension. Ideally, a measurand that is sensitive to absorption changes in the cortex should only exhibit an infinitely large value of $S_{2,1}$. The magnitude of $S_{2,1}$ critically depends on the thickness of the upper compartment.

2.2.2 Lateral spatial resolution

Lateral spatial resolution is determined as a spatial point spread function (PSF) by measuring the diffuse reflectance for a small absorber moved across this region. It is quantified by the full width at half maximum (FWHM) of the PSF related to a measurand M along two perpendicular directions parallel to the surface.

$$\Delta x_M = \text{FWHM}\big[M(x, y = 0; z, r)\big] \quad \text{and}$$
$$\Delta y_M = \text{FWHM}\big[M(x = r/2, y; z, r)\big] \tag{5}$$

at a predefined depth z below the surface and for a source–detector separation r. The source is located at $(0, 0, 0)$ and the detector at $(r, 0, 0)$. Since the resolution is most relevant with respect to changes in the cortex, z is preferably chosen to represent a typical depth of the cortex.

2.3 Quantification of Absorption Changes

The goal of quantification of concentration changes of oxy- and deoxyhemoglobin can only be achieved if the absorption changes in a certain compartment of the head can be accurately retrieved. In this case, the accuracy test is mandatory for performance assessment. In many cases, data analysis relies on a linear approximation that is only valid for small absorption changes. Therefore, the linearity test is important to determine the range of applicability of a linear method of analysis. This test can be applied to any measurand M.

2.3.1 Accuracy

The accuracy of a measured absorption change $\Delta\tilde{\mu}_{a,i}$ is characterized as a relative measurement error, i.e., the relative deviation from its (conventional) true value $\Delta\mu_{a,i}$.

$$\varepsilon_{\Delta\mu_{a,i}} = \frac{\Delta\tilde{\mu}_{a,i} - \Delta\mu_{a,i}}{\Delta\mu_{a,i}}. \tag{6}$$

The index i denotes the compartment in which a known absorption change $\Delta\tilde{\mu}_{a,i}$ was realized. The value of $\Delta\tilde{\mu}_{a,i}$ is available only if an absorption change can be derived by the application of a reconstruction algorithm to the measured DTOFs.

The accuracy depends on the measurement as well as on the reconstruction algorithm. Practically, the conventional true value $\Delta\mu_{a,i}$ realized in an inhomogeneous phantom can be obtained by prior accurate characterization of the individual turbid materials in homogeneous phantoms. The absolute (baseline) optical properties of the materials used should also be known and kept fixed when comparing the accuracy of $\Delta\mu_a$ measurements

performed with different instruments and/or reconstruction algorithms.

2.3.2 Linearity

To assess the linearity of the change in the measurand M with respect to an underlying absorption change $\Delta\mu_{a,i}$ in the compartment i, the values for ΔM (together with their uncertainty) are plotted as a function of $\Delta\mu_{a,i}$. The linearity range is defined as the maximum value of $\Delta\mu_{a,i}$ for which the deviation from proportionality does not exceed a certain percentage of the respective ΔM. This test requires several absorption changes to be realized, starting from small values and a sufficiently low uncertainty of ΔM due to random effects.

3 Implementation

This section describes the particular phantoms and measurement conditions to perform the tests of the nEUROPt protocol in a standardized and reproducible manner. The intention was to model the optical properties and geometry of the adult human head in a largely simplified, but as far as possible realistic, manner. The tests rely on two types of inhomogeneous turbid phantoms: (1) a localized inclusion in an otherwise homogeneous medium and (2) a two-layered medium.

3.1 Phantoms for the nEUROPt Protocol

3.1.1 Liquid phantoms

The decision to use well-characterized liquid phantoms for the performance characterization of instruments in the present study was based on the following considerations:

- Liquid phantoms offer a very high flexibility to (1) gradually change optical properties and (2) realize various inhomogeneous geometries, in particular, to vary the position of an inclusion.

- Liquid phantoms can be replicated much more easily than solid phantoms. Common measurement campaigns can be performed in parallel at several institutions.

- The optical properties can be adjusted with good reproducibility (a few percentage points) based on well-characterized Intralipid and prediluted ink as base materials. The optical properties of these base materials (1) remain stable for a long time, (2) are almost identical when taken from the same batch, and (3) are very similar for samples from different batches.[19,20]

A detailed procedure was developed at the University of Florence (UNIFI) to prepare liquid phantoms with known optical properties based on water dilutions of Intralipid as a diffusively scattering component and India ink as an absorber. The amounts of components to be mixed were solely determined by accurate weighing, which ensured that the concentrations were known with high accuracy. UNIFI provided the other partners with Intralipid and India ink from the same batches and with optical properties characterized with an uncertainty of <3%. The mixtures had to be freshly prepared on the day of the experiment. Accurate weighing and homogeneous mixing of the components is mandatory. It should be noted that these phantom materials were also characterized in an interlaboratory comparison of nine laboratories of the nEUROPt consortium and beyond. The intrinsic absorption coefficient of India ink

and the intrinsic reduced scattering coefficient of Intralipid-20% were determined with an uncertainty of ∼2% or better.[21]

Modular scattering cells made of black polyvinyl chloride (PVC) with small plexiglass windows were provided by UNIFI for measurements of time-resolved reflectance in homogeneous and layered geometry (see Refs. 18 and 22). For the implementation of the nEUROPt protocol, black front walls of 2 mm thickness were used, with three transparent windows (diameter 7 mm, adapted to the diameter of the fiber bundles used) for a source optode and a detector optode positioned 20 and 30 mm apart. For the cell, spacers of 10 and 30 mm thickness were provided. The effects of the finite lateral dimensions of the cell, the shape, and the refractive index mismatch of the inclusion were studied in detail by Monte Carlo (MC) simulations.[22] The inner dimensions of the cell were 120 mm width, 145 mm height, and 70 mm thickness (with three spacers). A Mylar foil of 30 μm thickness could optionally be inserted as a separator with minimal perturbation of light propagation to realize the two-layer geometry.[18] During data analysis of the measurement campaign, it turned out that it is not easy to maintain a constant and reproducible thickness of the upper layer (see Sec. 5.2). Therefore, a dedicated mounting plate for the Mylar foil was designed and manufactured by Physikalisch-Technische Bundesanstalt (PTB) for future use with a smaller area (60 mm × 80 mm) to be covered by the foil. This arrangement was used when characterizing the instrument POLIMI_2.

3.1.2 Absorbing objects

To mimic the perturbation due to small localized variations of the absorption coefficient for the nEUROPt protocol, UNIFI proposed the use of small black PVC cylinders immersed in the homogeneous liquid phantom.[22] They can, by far, more easily be manufactured and reproduced than small inhomogeneities made of a scattering and moderately absorbing material with well-known optical properties. The major idea behind this approach is the equivalence of the perturbation by a small black inclusion and by a certain moderate absorption change in a given volume. For details, see Martelli et al.[17] Thus, the small black cylinders can be used as a kind of universal inclusions provided their position is sufficiently far (>10 mm) from the source and detector and from the boundary of the medium. PVC cylinders of various sizes were provided by UNIFI, with dimensions (diameter equal to height) of 3.2, 4, 5, 6.8, and 8.6 mm, corresponding to volumes $V_{incl}/mm^3 = 25$, 50, 100, 250, and 500. They were held by thin, rigid metallic wires (0.5 mm music wire) that were painted white in order to reduce their influence on the measurement. The perturbation by the black cylinders can be directly matched to the equivalent finite $\Delta\mu_a$ changes over a certain larger volume (for instance, for a reference volume $V_0 = 1000$ mm^3, the equivalent values for the five cylinders were found to be $\Delta\mu_a/mm^{-1} = 0.0056$, 0.0087, 0.015, 0.037, and 0.094, within the restrictions regarding the position as mentioned above. Thus, linearity plots can be drawn against the equivalent $\Delta\mu_a$.

3.2 Specifications of Measurement Conditions

The measurement conditions were explicitly defined in order to ensure the comparability of the results of performance assessment across various devices and institutions. A consolidated table was prepared which specified all relevant parameters for the measurements pertaining to the BIP protocol,[10] the

MEDPHOT protocol,[3] as well as the nEUROPt protocol. These specifications included (1) the phantom configuration and measurement geometry as well as the baseline optical properties, (2) the geometrical parameters or optical properties to be changed, and (3) the parameters of data acquisition (measuring time, number of repetitions, count rates). In addition, a template lab report was prepared to facilitate the exchange of raw measured and preprocessed data between the institutions.

Table 1 represents a part of the complete implementation table. It contains the parameters of the phantoms to be changed in the measurements according to the nEUROPt protocol. For all tests, the target baseline optical properties were $\mu_s' = 1 \text{ mm}^{-1}$ and $\mu_a = 0.01 \text{ mm}^{-1}$ and the source–detector separation $r = 30 \text{ mm}$. The first column contains the tests addressed (see definitions in Sec. 2). A set of phantom measurements, i.e., a row in Table 1, can provide data for several tests. Meanwhile, the tests for accuracy and linearity can be realized by measurements on both types of phantoms. Section 4 gives examples for several tests on both phantoms.

4 Instruments and Data Analysis

4.1 Instruments

The instruments and configurations that have been characterized during the multilaboratory measurement campaign according to the nEUROPt protocol are listed together with their most relevant specifications in Table 2. In the discussion of the results in Sec. 5, they are referred to by their code. Table 2 specifies the laser, the detector types, and the parameters of the detection fiber bundles since these are the most relevant determinants of the instrument response function (IRF). The parameters of the source fibers (which are typically graded-index or low NA fibers and have less influence on the IRF) are omitted in this abbreviated table. The codes of the instruments in Table 2 are consistent with those of Table 3 in the companion paper,[10] but not all instruments and configurations underwent all tests.

The instruments denoted by "brain imager" are compact, portable systems for clinical applications on adults. All of these instruments contain lasers operating at two or more wavelengths (mostly between 690 and 830 nm), compact fast photomultipliers, preamplifiers, and multiboard TCSPC systems. PTB performed the experiments on the same phantom with two types of detector modules (PTB_1 and PTB_2) installed in the brain imager and, in addition, either with an external hybrid detector (PTB_3) or with a setup based on a supercontinuum laser source with an acousto-optic tunable filter and a hybrid detector (PTB_4). Figure 1 shows selected IRFs whose temporal profiles differ remarkably. Some of them exhibit afterpeaks at a level of several percent of the maximum at different temporal positions. The detector used with POLIMI_2 and PTB_4 has a narrower IRF without afterpeaks and an approximately exponential tail. The different widths of both IRFs result from the different pulse widths of the laser types used as well as from the different amounts of dispersion in the detection bundles.

4.2 Data Analysis

Two types of semiempirical algorithms have been routinely used within the nEUROPt consortium as the first steps for analyzing the measured DTOFs obtained in *in vivo* experiments. These algorithms are based on (1) moments and (2) time windows of the distributions. The moments considered are the integral N_{tot} (total photon count), the first moment m_1 (mean time of flight), and variance V (second central moment). They are derived from the counts N_i in the time channels (width Δt) of the histogram memory between limits a and b, according to

$$N_{\text{tot}} = \sum_{i=a}^{b} N_i,$$

$$m_1 = \sum_{i=a}^{b} iN_i/N_{\text{tot}} \cdot \Delta t,$$

$$V = \sum_{i=a}^{b} i^2 N_i/N_{\text{tot}} \cdot \Delta t^2 - m_1^2. \tag{7}$$

Photon counts in the k'th time window with limits $[a_k, b_k]$ are obtained as

$$N_k = \sum_{i=a_k}^{b_k} N_i. \tag{8}$$

To estimate the expected results of measurements and to check for deviations from ideal conditions, dedicated simulations

Table 1 Implementation of the tests of the nEUROPt protocol by phantom measurements (x, y, lateral; z, depth position coordinates; V_{cyl}, volume of black cylinder; d_1, thickness of upper layer).

Tests	Phantom with black cylinders			Two-layered phantom			Remarks
	x (mm)	z (mm)	V_{cyl} (mm³)	$\Delta\mu_{a,1}$ (mm⁻¹)	$\Delta\mu_{a,2}$ (mm⁻¹)	d_1 (mm)	
Contrast, (accuracy)[a], (linearity)[a]	0	6 to 30, step 2	25, 50, 100, 250, 500	—	—	—	Black polyvinyl chloride cylinders, midline parallel to wall, position refers to center of cylinder
Lateral spatial resolution	−40 to 40, step 5	10 (15)[b]	100	—	—	—	Repeat with y scan
Contrast, accuracy, linearity, depth selectivity	—	—	—	0 to 0.01, step 0.002	0 to 0.01, step 0.002	10 (15)[b]	Change in upper / lower layer separately

[a]Indirect method, requires prior application of equivalence relation for black absorbers.[17]
[b]Measurements reported here were performed with z and d_1 values of 10 mm. Values of 15 mm are suggested for future measurements, to better match the typical geometry of the adult human head.[23]

Table 2 Instruments and configurations characterized. Codes (acronyms of institutions, for complete information, see author affiliations): POLIMI, Politecnico di Milano; IBIB, Nałęcz Institute of Biocybernetics and Biomedical Engineering; UCL, University College London; PTB, Physikalisch-Technische Bundesanstalt. Superscript lowercase letters indicate manufacturers of components. Parameters of collection fibers or fiber bundles— D, diameter; L, length; NA, numerical aperture.

Code	Instrument	Laser type	Detector type (photocathode)	Detection fiber or bundle			References
				D (mm)	L (m)	NA	
POLIMI_1	Brain imager "fOXY"	Picosecond diode laser PDL 800[a]	R5900U-20-M4[c] (multialkali)	3	1.5	0.57	24, 25
POLIMI_2	Brain imager "fOXY2"	Picosecond diode laser Sepia II[a]	HPM-100-50[c,d] (GaAs)	3	1.5	0.57	26
IBIB_1	Brain imager	Picosecond diode laser Sepia II[a]	R7400U-02[c] (multialkali)	4	1.5	0.54	27, 28
UCL	Neonatal brain imager "monstir-II"	Supercontinuum laser SC400 with AOTF[b]	H8422-P50[c] (GaAs)	3	2.5	0.21	29
PTB_1	Brain imager "BI-1"	Picosecond diode laser Sepia[a]	R7400U-02[c] (multialkali)	4	1.5	0.54	30, 31
PTB_2	- with alternative detector module		H7422-50[c] (GaAs)	4	1.5	0.54	32
PTB_3	- combined with lab setup		HPM-100-50[c,d] (GaAs)	1	2	0.39	
PTB_4	Lab setup	Supercontinuum laser SC500-6 with AOTF[b]	HPM-100-50[c,d] (GaAs)	1	2	0.39	

Note: AOTF, acousto-optic tunable filter.
[a]PicoQuant GmbH, Berlin, Germany
[b]Fianium Ltd., Southampton, United Kingdom.
[c]Hamamatsu Photonics, Hamamatsu, Japan.
[d]Becker & Hickl GmbH, Berlin, Germany.

were performed. The perturbation of time-resolved diffuse reflectance due to black inclusions at varying positions was simulated by UNIFI using an MC code for photon migration through turbid media containing spherical absorbing objects.[22,33] Simulations for the two-layer geometry were carried out with the forward solver based on the solution of the diffusion equation for n-layered media.[34]
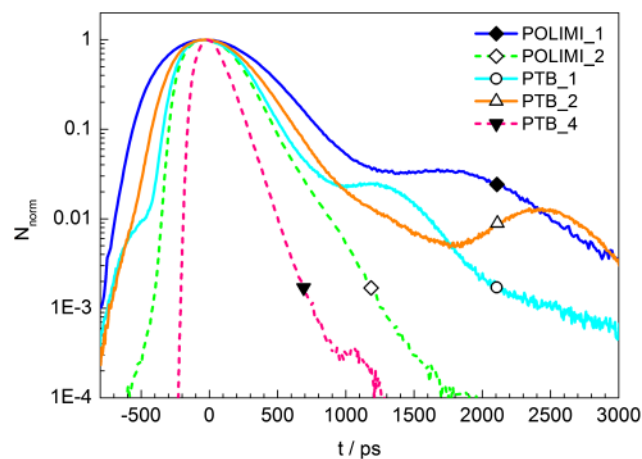

Fig. 1 Instrument response functions. The measuring time was 20 s in each case, constant background was subtracted. The position $t = 0$ was obtained as the baricenter of each instrument response function in the interval between the points at half maximum.

5 Results and Discussion

During the common measurement campaign, the tests according to the nEUROPt protocol were performed by four groups of the nEUROPt consortium, in part, with several instruments. The measurements were analyzed by the time window as well as the moments approach. A comprehensive set of meaningful data was obtained. Here, we cannot present examples of all individual tests. The results shown below were selected to give an overview of the kind of data obtained and to illustrate the influence of particular features of instruments and methods of data analysis. The results are arranged according to the type of phantom. Examples for both approaches of data analysis are given. In the case of depth selectivity, we show a direct comparison of results of several instruments. Additional examples of results of various other tests, for one particular instrument, can be found in the consolidated reporting sheet shown in Sec. 5.3, e.g., for lateral resolution and linearity.

5.1 Phantom with Localized Inhomogeneity

5.1.1 Depth-dependent contrast

Contrast and noise measured as a function of the depth of a black inclusion in the liquid phantom are presented in Fig. 2 for two instrumental configurations, for analysis of time windows. Figure 2(a) clearly illustrates the advantages of the time-resolved measurement to detect absorption changes in the brain. Photons with a short time of flight are only sensitive to shallow absorption changes. For late time windows, the highest contrast is found deeply below the surface, in this example at
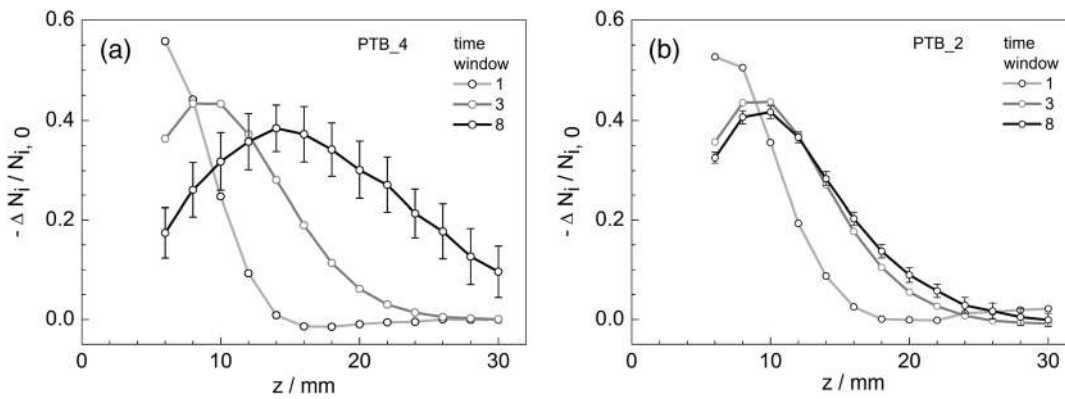
Fig. 2 Contrast of photon counts in time windows measured as a function of depth position z of a black polyvinyl chloride (PVC) cylinder of 100 mm³ volume, for instrumental configurations PTB_4 (a) and PTB_2 (b). The numbers indicate consecutive time windows of 500 ps width; error bars represent noise as obtained from the standard deviation of repeated measurements.

~15 mm. The error bars represent noise as obtained from the standard deviation of 100 repetitions of the measurement at a 1-s duration and a count rate of ~1 · 10⁶ s⁻¹. Only for the latest time window, where the photon count is low, do they exceed the size of the markers. The comparison of Figs. 2(a) and 2(b) highlights the influence of the IRF on the contrast, in particular, for late photons (see also related simulations in Ref. 10). The depth-dependent contrast for instrument PTB_2 for late times (eighth time window, ranging from 3500 to 4000 ps) is similar to that for the third time window. The reason is the existence of a substantial late afterpeak in the IRF of PTB_2 (see Fig. 1). Meanwhile, the hybrid detector employed in the setup PTB_4 together with a supercontinuum laser has a good time resolution and an IRF without afterpeaks. The contrast for a deep inclusion, e.g., at $z = 20$ mm, is considerably larger for PTB_4 compared to PTB_2.

Figure 3 provides another representation of the results of the same measurement. For the inclusion positioned at shallow depths, the contrast decreases with increasing time [Fig. 3(a)], whereas a deep inclusion becomes detectable only at later times [Fig. 3(b)]. For the configurations PTB_1 and PTB_2, a deviation from this behavior is observed at late times, i.e., too high a contrast for the shallow inclusion and too low a contrast for the deep inclusion. This finding can be explained by the fact that an afterpeak in the IRF (see Fig. 1) effectively transfers early photons to late times. At the same time, these falsely assigned photons lead to a decreased noise and deviations in the contrast-to-noise ratio [Fig. 3(c)]. A comparison of Figs. 3(c) and 3(d) reveals a considerably lower maximum CNR for the deep versus shallow position of the inclusion. For the inclusion at $z = 6$ mm, the maximum contrast and CNR are both found at early times. The maximum CNR for $z = 16$ mm, however, is found at intermediate times. At late times, where the contrast is highest, the low photon count at late times causes the CNR to drop. The results presented in Figs. 2 and 3 illustrate that contrast and CNR alone are not sufficient to characterize the discrimination between a superficial and a deep absorption change. For this characteristic, we refer to depth selectivity, see Sec. 5.2.

As an example of analysis of moments, the depth-dependent contrast for variance derived from DTOFs is shown in Fig. 4 for three of the black cylinders. Unlike in the case of a time-window analysis, the results for the three different experimental configurations applied in the same measurement almost exactly match.

This finding is explained by the fact that measurands based on differences in variance (as well as mean time of flight) are virtually independent of the IRF.[35] In addition, the measured variance contrast agrees very well with the results of MC simulations performed according to Ref. 33 for black spheres of equal volume (data not shown, see Ref. 22). Note the changing sign of variance contrast between 10- and 12-mm depths, a behavior that is typical for normalized moments.

5.2 Two-Layered Phantom

5.2.1 Contrast

The measurements on the two-layered phantom allow the contrast of laterally extended absorption changes in a superficial and a deep compartment to be compared. Results of a time-window analysis of the measurements on the two-layer phantom with an upper layer of nominal thickness of 10 mm are presented in Fig. 5, together with the results of two-layer simulations based on Ref. 34. The absorption is changed in each of the
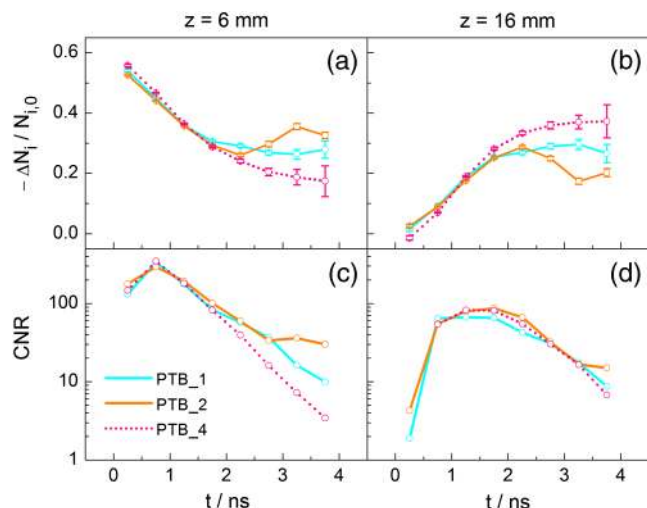

Fig. 3 Contrast [(a) and (b)] and contrast-to-noise ratio [(c) and (d)] for the black cylinder of 100 mm³ volume as a function of time, for depths $z = 6$ mm [(a) and (c)] and $z = 16$ mm [(b) and (d)]. The results are plotted for the configurations PTB_1, PTB_2, and PTB_4.
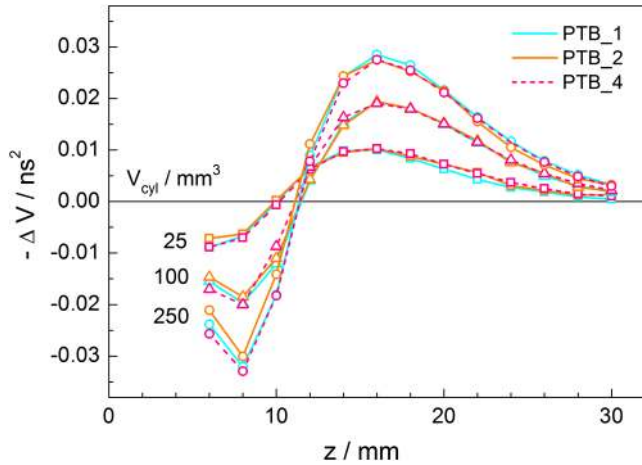
Fig. 4 Difference in variance for z scans of black PVC cylinders of various sizes (squares: 25 mm^3, triangles: 100 mm^3, circles: 250 mm^3). Line color indicates different experimental configurations.

compartments separately; the corresponding results are plotted in the upper and lower rows. Note that the range of the absorption change is rather large, μ_a is changed from 0.01 mm^{-1} (baseline) to 0.02 mm^{-1}. It is not surprising that the contrast is not linear in $\Delta\mu_a = \mu_a - 0.01$ mm^{-1} in the whole range.

The contrast for changes in the upper layer is comparable with the simulated data for all instruments, whereas the experimental results for changes in the lower layer show less contrast than expected from the simulations. A major reason is the finite width of the IRF that has not been accounted for in the simulations. In general, the sensitivity to absorption changes in the lower layer is higher for later photons, as observed in the case of

POLIMI_2. The experimental data for the instruments PTB_1 and PTB_2 show, however, that the maximum contrast for the lower layer is achieved for intermediate time windows. For the latest time windows, the contrast decreases again. This behavior can be explained by the influence of afterpeaks in the IRF (see Fig. 1 and Ref. 10).

The contrasts for moments obtained on the two-layered phantom with the instrument IBIB_1 is depicted in Fig. 6, together with simulations with the same nominal optical properties and thickness of the upper layer. Measured and simulated contrasts agree very well despite the fact that the IRF has not been taken into account in the simulations. As already discussed in Sec. 5.1, this behavior is expected since differences of moments are virtually independent of the IRF.[35]

In general, the contrast obtained from N_{tot} is larger in the upper layer (by about a factor of 3), and the contrast based on m_1 is approximately the same in both layers, whereas the contrast for V is larger (by about a factor of 2) for absorption changes in the lower layer. A quantitative comparison of these results is performed in the following section.

5.2.2 Depth selectivity

The definition of depth selectivity [Eq. (4)] based on a ratio is a way to compare different measurands even of different dimensions (see the different moments in Fig. 6) with respect to the discrimination between deep and superficial changes. This ratio is formed from data such as those displayed in Figs. 5 and 6 for the lower and upper compartments. In Fig. 7, the same quantity $S_{2,1}$ is plotted for analysis by time windows [Fig. 7(a)] and moments [Fig. 7(b)]. $S_{2,1}$ was obtained as the ratio of the slopes of the contrast curves for small absorption changes, i.e., for μ_a
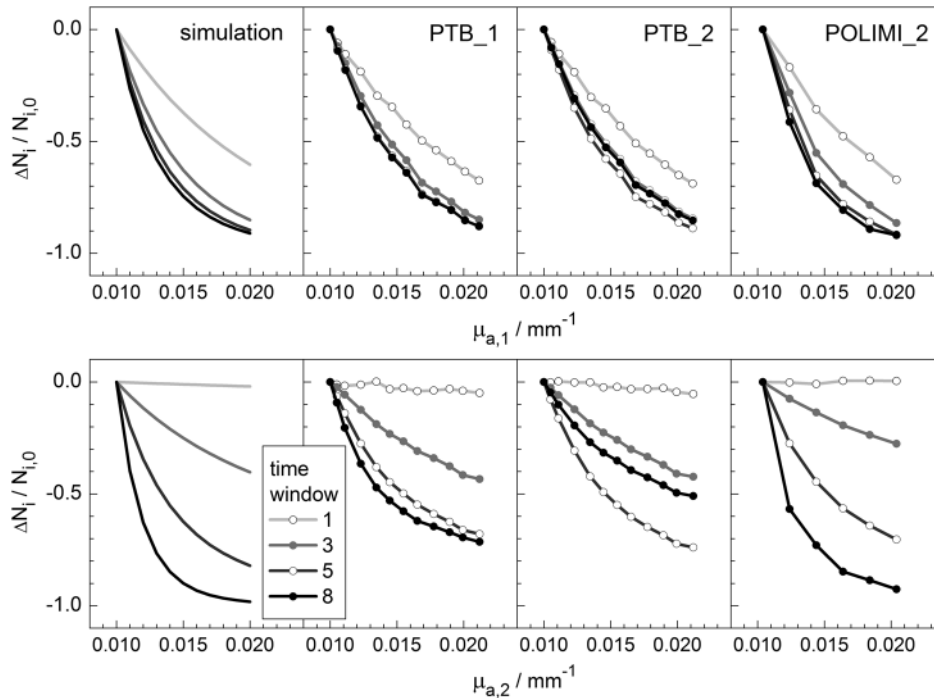

Fig. 5 Contrasts of selected time windows of 500 ps width (number given in the legend) as a function of a change in $\mu_{a,1}$ in the upper layer (upper row) and $\mu_{a,2}$ in the lower layer (lower row) of the two-layered phantom with nominal thickness of the upper layer $d = 10$ mm. First column: two-layer simulation based on Ref. 34, second column: PTB_1, third column: PTB_2, fourth column: POLIMI_2.
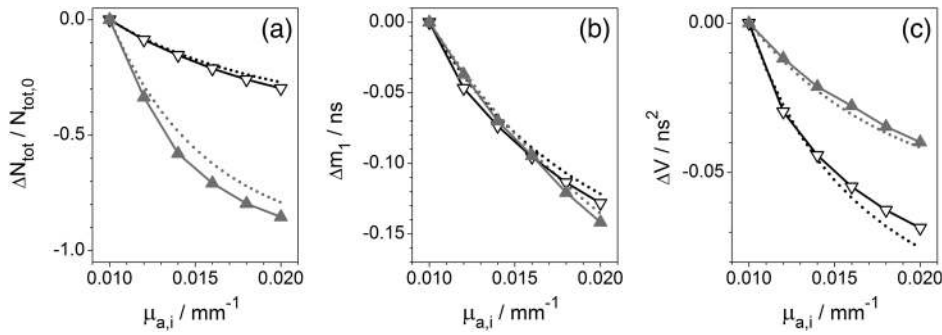
Fig. 6 Contrasts of moments, (a) integral, (b) mean time of flight, (c) variance, as a function of a change in $\mu_{a,1}$ in the upper layer (up triangles, gray) and $\mu_{a,2}$ in the lower layer (down triangles, black) for the instrument IBIB_1. Both integration limits for moments were set at 1% of the maximum. Dotted lines: simulations based on Ref. 34.

values in the interval [0.010 mm^{-1}, 0.012 mm^{-1}]. The consolidated plots in Fig. 7 include the results obtained for several different instruments and configurations together with results obtained from simulations. Comparing the plots for the instrumental configurations PTB_1, PTB_2, and PTB_3, which differ in the type of detector, reveals the different influence of the IRF with both approaches of analysis. For time windows [Fig. 7(a)], the maximum achievable depth selectivity for late photons is substantially degraded if the IRF exhibits remarkable afterpeaks as in case of the instruments PTB_1 and PTB_2. The existence of afterpeaks causes a fraction of early photons to be detected at late times. Such afterpeaks are not present in the case of the hybrid detectors in PTB_3 and POLIMI_2. These instruments provided the best depth selectivity at late times. Moreover, in these cases, the courses as a function of time essentially follow that for the simulated data where no IRF was taken into account.

For moments [Fig. 7(b)], the influence of the IRF is expected to vanish, as discussed above in Sec. 5.1. This can actually be observed when comparing the results for the three PTB detectors that were obtained together in the same phantom experiment. It should be noted that here the integration limits were set at 0.1% of the maximum (otherwise 1%) to avoid cutting within the region influenced by the afterpeak for PTB_2. Discrepancies between the measurements by different groups are most likely due to inaccuracies in the thickness of the upper layer caused by difficulty in flattening the inelastic Mylar foil that separates the layers. The comparison with corresponding two-layer simulations suggests that in the case of PTB, the thickness of the

upper layer was most likely ∼9 mm instead of 10 mm. A decreased thickness of the upper layer mainly leads to an increase in contrast for changes in the lower layer and, thus, to an overestimation of depth selectivity. The inaccuracy of the thickness hampers the comparison between different experiments. An improved mounting of the Mylar foil employing the mounting plate mentioned in Sec. 3.1 could solve this problem. This mounting plate was applied in the later experiment with the instrument POLIMI_2, which yielded depth selectivities for moments rather close to the simulated values.

Overall, the highest depth selectivity is obtained for variance. It should be noted that the depth selectivity of the time-window approach can be substantially improved by considering ratios of photon counts in late to early time windows.

5.3 Consolidated Presentation of Results

In order to facilitate the presentation of the results of all tests, the comparison of different instruments or different configurations of the same instrument, as well as the archival storage of system specifications, a consolidated reporting sheet was prepared. The report is automatically generated from the data inserted by the user in a tabular form. Figure 8 shows an example for a particular instrument and method of data analysis.

The first page of the report is specific for the instruments tested. It is divided into three sections: a first section (Section I) with some information on the system (e.g., name, institution, operating conditions); a second section (Section II)
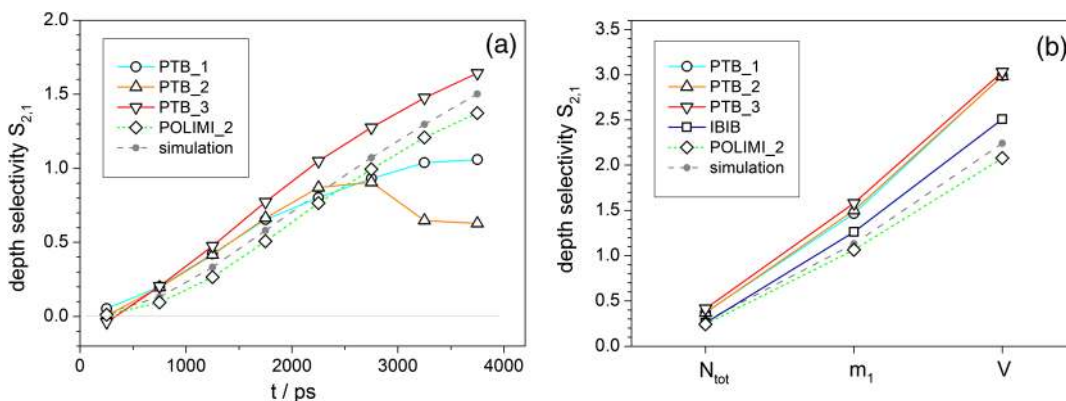

Fig. 7 Depth selectivity, i.e., the ratio of contrasts for small absorption changes in the lower layer and in the upper layer, for time windows (a) and moments (b). The positions of the symbols in (a) correspond to the center of the respective time windows.

with six figures reporting the results of the main tests (contrast, contrast-to-noise ratio, x and y lateral resolution, linearity, depth selectivity); and a third section (Section III) with several final synthetic descriptors derived automatically from the figures and used to summarize the key performances of the system in a short and quantitative way for easy grading of the instrument under test. The second page (not shown) describes the actual implementation

of the nEUROPt protocol and the conditions for analysis and reporting. In particular, the specific measurands reported on the first page are defined and details on the synthetic descriptors are provided.

In the example presented in Fig. 8, the measurands are photon counts in an early time window (N_E, 0 to 500 ps), containing information mainly on superficial regions, photon counts in
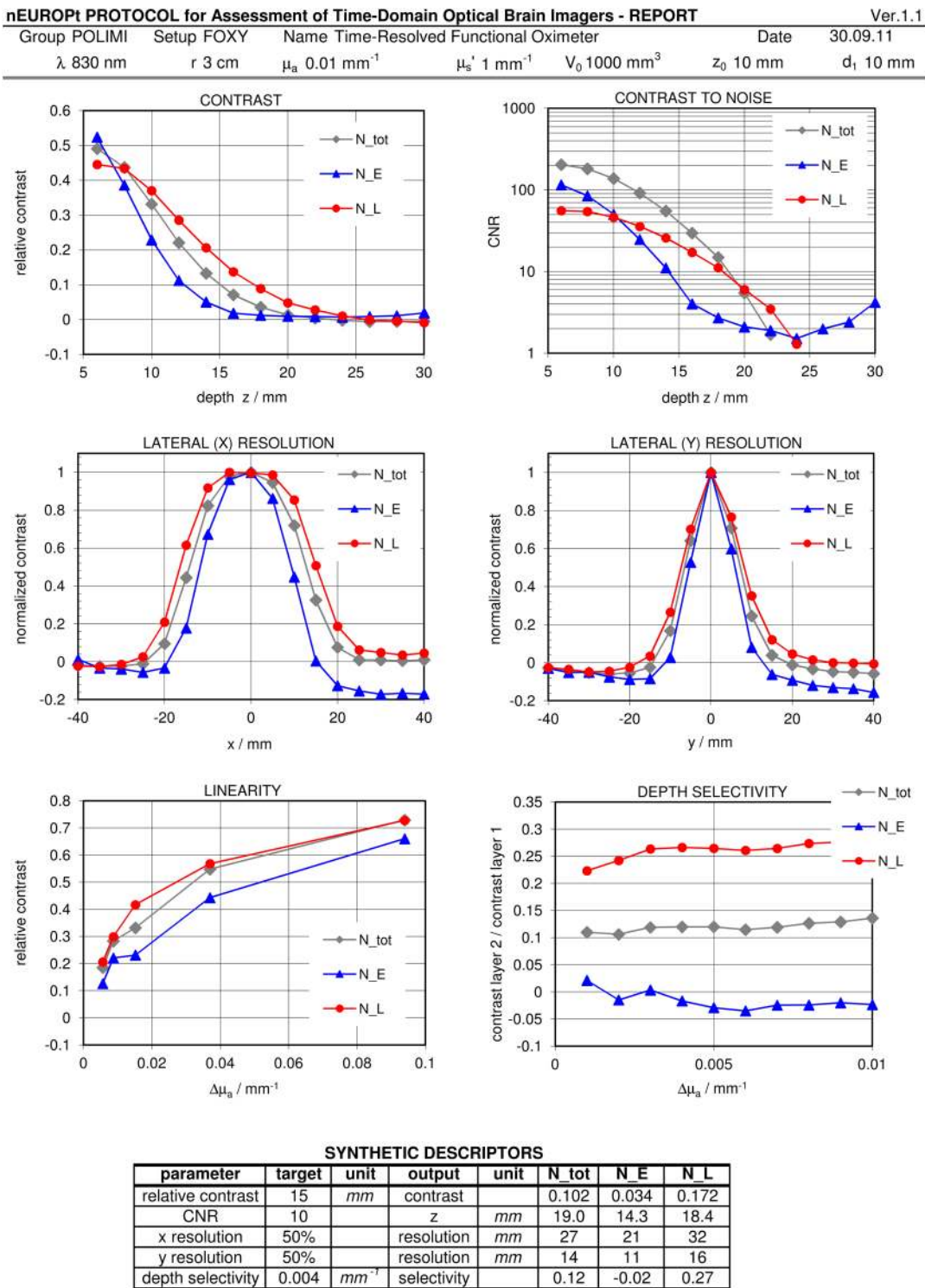

Fig. 8 Front page of reporting sheet for one of the instruments (POLIMI_1 at 830 nm) with data analysis by time windows (N_{tot}: total counts integrated from 0.0 to 4.0 ns, N_E: early time window integrated from 0 to 0.5 ns, N_L: late time window integrated from 2.5 to 3.0 ns).

a late time window (N_L, 2500 to 3000 ps), mainly probing deeper regions, and the time-integrated (N_tot) signal. The measurements were performed according to the specifications in Table 1. The figure related to relative contrast was obtained using a black PVC cylinder of volume $V_{cyl} = 100$ mm^3, equivalent to $\Delta\mu_a = 0.015$ mm^{-1} within a spherical volume $V_0 = 1000$ mm^3. The linearity plot was constructed from the change in the measurand produced by the black cylinders with increasing volumes at the specific depth $z_0 = 10$ mm. The contrast is plotted versus the equivalent increase in μ_a for $V_0 = 1000$ mm^3. For the transformation from V_{cyl} to $\Delta\mu_a$, we applied the nonlinear relation between the black object volume and a realistic equivalent absorption perturbation reported in a previous paper.[17] Finally, the depth selectivity assessed using the two-layer phantom is shown. Compared to the complete nEUROPt protocol, the accuracy test is missing here. The reason is that the specific measurands used in this example (photon counts in selected time windows) are not sufficient to retrieve absolute information on the absorption coefficient.

The choice of the synthetic descriptors aims at extracting several numbers for a quick and quantitative comparison of the instruments at the expense of loss of detail. Particular thresholds or conditions are specified to extract these values from the plots of Section II. In the case illustrated in Fig. 8, five descriptors were extracted, namely (1) the contrast obtained at $z = 15$ mm; (2) the CNR at $z = 10$ mm; (3) and (4) the FWHM of the x and y lateral resolution plots; and (5) the depth selectivity for $\Delta\mu_a = 0.004$ mm^{-1}.

6 Conclusions

As a joint effort of the nEUROPt consortium, a standardized protocol (nEUROPt protocol) has been developed to assess and compare the performances of time-domain brain imagers with respect to the detection, localization, and quantification of absorption changes in the brain. The protocol is composed of six tests, namely contrast and contrast-to-noise ratio to assess sensitivity of detection, depth selectivity and lateral resolution to address capabilities related to localization, and linearity and accuracy to grade quantification. The instrument as a whole, together with data analysis, is subjected to these tests. They aim at the determination of performance parameters that are directly relevant in the context of the *in vivo* application. The tests provide a snapshot of the capabilities of a particular system and allow various comparisons to be made.

A specific implementation of the protocol was proposed, exploiting well-characterized and reproducible inhomogeneous liquid phantoms[17,21,22] together with a compilation of experimental conditions for the test measurements. Following the protocol and the implementations described above, a total of eight instruments developed by four different laboratories were tested. The results of these measurements were represented with several examples comparing different instruments and methods of data analysis. These examples illustrate, in particular, the suitability of the protocol to evaluate the effect of technical improvements, e.g., the use of different detectors. The shape of the IRF turned out to be crucial for the performance of a time-domain brain imager. In particular, afterpeaks in the IRF caused a substantial decrease in the contrast for photons with long flight times, thus hampering detection of deep absorption changes. In this context, the link to the BIP protocol[10] is important. A thorough characterization of the instrument itself, complemented by simulations taking into account its characteristics, facilitates the understanding and interpretation of the results of the tests according to the nEUROPt protocol.

The results of the tests presented in this work can also serve as an example for comparing different algorithms of data analysis and assessing their robustness with respect to experimental factors. We focused on straightforward measurands that were routinely obtained in primary steps of the analysis of measurement data, including those recorded in clinical studies. Two approaches were considered here based on time windows and moments of DTOFs, respectively. Both types of measurands are suitable to study absorption changes in the cortex, but exhibit a substantially different sensitivity to the shape of the IRF. It should be noted that these basic steps of analysis do not directly yield the absolute value of $\Delta\mu_a$, which would require more sophisticated reconstruction methods. Hence, the accuracy test was not applied.

All test measurements started from a fixed and equal baseline photon count rate. This approach eliminates the influence of differing laser power, of the scheme of its distribution to various source optodes, and of the responsivity of the detection system. Thus, the overall sensitivity of the instruments is not reflected in the results. This decision was taken to avoid too much complexity. The BIP protocol[10] addresses these issues separately, specifically by means of the responsivity test together with recording the source parameters. The noise test of the MEDPHOT protocol,[3] i.e., evaluation of the noise level of optical properties as a function of input energy, illustrated the possibility of a combined approach.

The implementation of the nEUROPt protocol presented here was based on a fixed geometry (source–detector separation, layer thickness) and optical properties of the phantoms chosen to reflect realistic conditions of *in vivo* measurements. However, the tests could also be applied to simulated data to study the effect of these parameters. It is worth mentioning that three factors, i.e., (1) physics of light propagation in turbid media, (2) instrumental characteristics, and (3) features of data analysis methods, influence the various tests differently. Some tests are dominated by the physics of the problem, as in the case of the lateral resolution, which is only slightly affected by instrumental characteristics, e.g., the size of the optodes.

Although originally developed for time-domain brain imagers, the nEUROPt protocol is also applicable to instruments based on cw or frequency-domain technologies. The tests were formulated in a rather universal way in terms of an arbitrary measurand M which can also represent, in particular, the phase shift and demodulation obtained from frequency-domain measurements. When applying the protocol to cw or frequency-domain technologies, the protocol itself can remain valid. Even the implementation in terms of phantom configuration and measurement geometry as well as the baseline optical properties could be adopted. However, the conditions of data acquisition (e.g., measuring time, signal magnitude) need to be specified with respect to the particular technology. With some modifications, the nEUROPt protocol can also be extended to address other applications of diffuse optical imaging, in particular, optical mammography. While most of the tests can also be applied in transmission geometry, the depth selectivity test, as described above, is primarily relevant to measurements in reflection geometry.

The work presented here can be useful when designing performance tests in future documentary standards in these fields. Liquid phantoms are most likely impracticable for test

procedures in such a context, requiring easy and reliable means to be applied by the manufacturer under industrial conditions. In such a case, solid, easily accessible, and possibly commercially available phantoms need to be developed.

References

1. B. W. Pogue and M. S. Patterson, "Review of tissue simulating phantoms for optical spectroscopy, imaging and dosimetry," *J. Biomed. Opt.* **11**(4), 041102 (2006).
2. 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).
3. A. Pifferi et al., "Performance assessment of photon migration instruments: the MEDPHOT protocol," *Appl. Opt.* **44**(11), 2104–2114 (2005).
4. A. E. Cerussi et al., "Tissue phantoms in multicenter clinical trials for diffuse optical technologies," *Biomed. Opt. Express* **3**(5), 966–971 (2012).
5. A. M. De Grand et al., "Tissue-like phantoms for near-infrared fluorescence imaging system assessment and the training of surgeons," *J. Biomed. Opt.* **11**(1), 014007 (2006).
6. F. Orihuela-Espina et al., "Quality control and assurance in functional near infrared spectroscopy (fNIRS) experimentation," *Phys. Med. Biol.* **55**(13), 3701–3724 (2010).
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. D. A. Boas et al., "Twenty years of functional near-infrared spectroscopy: introduction for the special issue," *NeuroImage* **85**(Pt 1), 1–5 (2014).
9. A. Torricelli et al., "Time domain functional NIRS imaging for human brain mapping," *NeuroImage* **85**(Pt 1), 28–50 (2014).
10. H. Wabnitz et al., "Performance assessment of time domain optical brain imagers, part 1: basic instrumental performance protocol," *J. Biomed. Opt.* **19**(8), 086010 (2014).
11. M. A. Franceschini et al., "Influence of a superficial layer in the quantitative spectroscopic study of strongly scattering media," *Appl. Opt.* **37**(31), 7447–7458 (1998).
12. M. A. Franceschini et al., "Diffuse optical imaging of the whole head," *J. Biomed. Opt.* **11**(5), 054007 (2006).
13. X. Zhang and A. Webb, "Detectability of absorption and reduced scattering coefficients in frequency-domain measurements using a realistic head phantom," *Sensors* **13**(1), 152–164 (2013).
14. T. Funane et al., "Dynamic phantom with two stage-driven absorbers for mimicking hemoglobin changes in superficial and deep tissues," *J. Biomed. Opt.* **17**(4), 047001 (2012).
15. J. C. Hebden et al., "An electrically-activated dynamic tissue-equivalent phantom for assessment of diffuse optical imaging systems," *Phys. Med. Biol.* **53**(2), 329–337 (2008).
16. R. L. Barbour et al., "Validation of near infrared spectroscopic (NIRS) imaging using programmable phantoms," *Proc. SPIE* **6870**, 687002 (2008).
17. F. Martelli et al., "Phantoms for diffuse optical imaging based on totally absorbing objects. Part 1: basic concepts," *J. Biomed. Opt.* **18**(6), 066014 (2013).
18. S. Del Bianco et al., "Liquid phantom for investigating light propagation through layered diffusive media," *Opt. Express* **12**(10), 2102–2111 (2004).
19. P. Di Ninni, F. Martelli, and G. Zaccanti, "Intralipid: towards a diffusive reference standard for optical tissue phantoms," *Phys. Med. Biol.* **56**(2), N21–28 (2011).
20. P. Di Ninni, F. Martelli, and G. Zaccanti, "The use of India ink in tissue-simulating phantoms," *Opt. Express* **18**(26), 26854–26865 (2010).
21. L. Spinelli et al., "Determination of reference values for optical properties of liquid phantoms based on Intralipid and India ink," *Biomed. Opt. Express* **5**(7), 2037–2053 (2014).
22. F. Martelli et al., "Phantoms for diffuse optical imaging based on totally absorbing objects. Part 2: experimental implementation," *J. Biomed. Opt.* **19**(7), 076011 (2014).
23. M. Okamoto et al., "Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10-20 system oriented for transcranial functional brain mapping," *NeuroImage* **21**(1), 99–111 (2004).
24. D. Contini et al., "Multi-channel time-resolved system for functional near infrared spectroscopy," *Opt. Express* **14**(12), 5418–5432 (2006).
25. D. Contini et al., "A multichannel time-domain brain oximeter for clinical studies," *Proc. SPIE* **7369**, 73691D (2009).
26. 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).
27. M. Kacprzak et al., "Time-resolved optical imager for assessment of cerebral oxygenation," *J. Biomed. Opt.* **12**(3), 034019 (2007).
28. D. Milej et al., "Time-resolved multi-channel optical system for assessment of brain oxygenation and perfusion by monitoring of diffuse reflectance and fluorescence," *Opto-Electron. Rev.* **22**(1), 55–67 (2014).
29. 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).
30. H. Wabnitz et al., "A time-domain NIR brain imager applied in functional stimulation experiments," *Proc. SPIE* **5859**, 58590H (2005).
31. H. Wabnitz et al., "Time-resolved near-infrared spectroscopy and imaging of the adult human brain," *Adv. Exp. Med. Biol.* **662**, 143–148 (2010).
32. A. Hagen et al., "Development of a multi-channel time-domain fluorescence mammograph," *Proc. SPIE* **6434**, 64340Z (2007).
33. A. Sassaroli et al., "Monte Carlo procedure for investigating light propagation and imaging of highly scattering media," *Appl. Opt.* **37**(31), 7392–7400 (1998).
34. A. Liemert and A. Kienle, "Light diffusion in N-layered turbid media: frequency and time domains," *J. Biomed. Opt.* **15**(2), 025002 (2010).
35. A. Liebert et al., "Evaluation of optical properties of highly scattering media by moments of distributions of times of flight of photons," *Appl. Opt.* **42**(28), 5785–5792 (2003).

Biographies of the authors are not available.