Performance assessment of photon migration instruments: the MEDPHOT protocol

Pifferi, A; Torricelli, A; Bassi, A; Taroni, P; Cubeddu, R; Wabnitz, H; Grosenick, D; Moller, M; Macdonald, R; Swartling, Johannes; Svensson, Tomas; Andersson-Engels, Stefan; van Veen, RLP; Sterenborg, HJCM; Tualle, JM; Nghiem, HL; Avrillier, S; Whelan, M; Stamm, H

Published in:
Applied Optics

2005

Performance assessment of photon migration instruments: the MEDPHOT protocol

Antonio Pifferi, Alessandro Torricelli, Andrea Bassi, Paola Taroni, Rinaldo Cubeddu, Heidrun Wabnitz, Dirk Grosenick, Michael Möller, Rainer Macdonald, Johannes Swartling, Tomas Svensson, Stefan Andersson-Engels, Robert L. P. van Veen, Henricus J. C. M. Sterenborg, Jean-Michel Tualle, Ha Lien Nghiem, Sigrid Avrillier, Maurice Whelan, and Hermann Stamm

We propose a comprehensive protocol for the performance assessment of photon migration instruments. The protocol has been developed within the European Thematic Network MEDPHOT (optical methods for medical diagnosis and monitoring of diseases) and is based on five criteria: accuracy, linearity, noise, stability, and reproducibility. This protocol was applied to a total of 8 instruments with a set of 32 phantoms, covering a wide range of optical properties.

OCIS codes: 170.5280, 170.7050, 220.4840, 350.4800, 000.3110.

1. Introduction

In the past decade, the field of photon migration has grown rapidly, attracting the interest of researchers in a number of applications in the biomedical field, spanning from optical mammography to muscle and brain oximetry, from tissue spectroscopy to the study of bone and joint diseases, and from optical characterization of photosensitizers to molecular imaging.[1–3] In addition to *in vivo* applications, in which interest has been strong, other fields have been pioneered, such as nondestructive characterization of agricultural products[4] or quality assessment of pharmaceutical tablets.[5] All these applications have fostered the development of a wide collection of instruments based on the detection of light propagated through turbid media. Different techniques are exploited, most of which can be classified as time resolved, frequency domain, or space resolved, although mixed approaches are possible. These instruments are operated at a single wavelength, at a few discrete wavelengths or, in some cases, over a wide continuous spectral range. Active theoretical research has led to the development of various theoretical models and algorithms for data analysis that are generally—but not exclusively—based on the transport equation under the diffusion approximation. Measurements can produce average values with a single source–detector pair set at a given interfiber distance ρ, projection images with a scanning approach, as well as topographic or tomographic images that exploit multiple source–detector schemes. Also, the acquisition time is quite different, ranging from few milliseconds for instruments monitoring fast changes in the optical properties up to 1 h for fully tomographic systems.

From this brief overview it is clear that photon migration instruments are quite different from one another in terms of technical approach, performance, theoretical model used for the analysis, and finaliza-

A. Pifferi (antonio.pifferi@fisi.polimi.it), A. Torricelli, A. Bassi, P. Taroni, and R. Cubeddu are with the Laboratorio Nazionale per l'Ottica Ultra-rapida e Ultra-intensa, Dipartimento di Fisica e Istituto di Fotonicae Nanotecnologie, Consiglio Nazionale delle Ricerche, Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133 Milan, Italy. H. Wabnitz, D. Grosenick, M. Möller, and R. Macdonald are with Physikalisch-Technische Bundesanstalt, Abbestrasse 2-12, D-10587 Berlin, Germany. J. Swartling, T. Svensson, and S. Andersson-Engels are with the Department of Physics, Lund University Medical Laser Centre, Lund Institute of Technology, P.O. Box 118, 22100 Lund, Sweden. R. L. P. van Veen and H. J. C. M. Sterenborg are with the Department of Radiation Oncology, Erasmus Medical Center, University Medical Center, Rotterdam, Westzeedijk 118, 3000 CA, The Netherlands. J.-M. Tualle, H. L. Nghiem, and S. Avrillier are with Laboratoire de Physique des Lasers CNRS UMR 7538, Institut Galilée, Université Paris 13, 99 Avenue J.-B. Clément, 93430 Villetaneuse, France. M. Whelan and H. Stamm are with the Institute for Health and Consumer Protection, European Commission Joint Research Centre, Ispra I-21020, Italy.

Received 26 July 2004; revised manuscript received 5 November 2004; accepted 5 November 2004.

0003-6935/05/112104-11$15.00/0

tion to a specific application. Nonetheless, there is common ground unifying all these systems: the physics of photon migration, which does not depend on the way photons are detected, and possibly also the typical outcomes from the measurement of the absorption (μ_a) and the reduced scattering (μ_s') coefficients.

This common ground makes it possible to use the same phantom to test and characterize quite diverse systems. The most common phantoms for studying photon migration are water-based solutions and resin-based solid samples. The former are water solutions of a diffusive medium—typically Intralipid, a lipid suspension used for the nutrition of hospitalized patients[6,7]—together with inks or dyes as absorbers.[8] Using jellifying agents[9] or transparent films,[10] one can construct heterogeneous structures. The water phantoms are inexpensive, and are easy and fast to prepare; however, they are perishable, difficult to exchange among different laboratories, and may differ between batches. The latter phantoms are based on resin with a scatterer added—typically titanium dioxide or calibrated microspheres—and a resin-soluble absorber.[11,12] These phantoms are solid, durable, and easy to machine and exchange; however, they are more cumbersome to prepare and, in some cases, to characterize. There has been much research on the design, testing, and characterization of phantoms, some of which have been circulated among different institutions.[13]

Conversely, less attention has been paid to the definition of common protocols for the performance assessment of instruments, which compared with other more mature fields, has no consensus on the most relevant tests and figures for the performance evaluation of photon migration setups. Most often, the system specifications are expressed in terms of those parameters that are directly related to the hardware implementation, such as temporal resolution or phase sensitivity, but that are not easily related to the measured parameters and cannot be compared among instruments based on different techniques but used for the same application. Also, the effect of the theoretical model or the fitting algorithm on the recovered properties is not often taken into account.

Various needs could potentially be addressed by use a common protocol. Obviously, it could be used to assess the instrument performances in a measurable way, possibly with a direct relation to the application requirements and to the limitations of the instrument in use. Then it could provide quality assurance during routine operation of the instrument, particularly during clinical trials for which studywide data consistency is crucial. Furthermore a common protocol could be seen as an aid during the development of new instruments or during the upgrade of existing ones, permitting the quantification of the effect of the technical interventions on the final outcome of the measurements. Finally, it can serve as a common basis for the comparison of different instruments and consequently of the measurement results.

The issue of the formulation of a common protocol has been undertaken within the European Thematic Network MEDPHOT (optical methods for medical diagnosis and monitoring of diseases).[14] This project sets a common discussion floor among 21 European partners from 8 countries on the development, testing, and application of biomedical optics instruments. One task of the project was devoted to the quality assessment of photon migration instruments, and the protocol presented in this paper was discussed, designed, and tested during regular project meetings as well as during interlaboratory visits.

In this paper we present the general concepts on which the protocol was designed, identify the relevant measurables, and define the assays that constitute the protocol. Then we describe the phantom kit chosen and constructed for the implementation of the protocol. Finally, we provide some examples derived from the application of the protocol to a wide collection of photon migration instruments.

The methods that are presented in the following sections are not necessarily new to the photon migration community. Some of the proposed assays are routinely used by many research groups, others are straightforward implementations of metrology concepts. What is novel here is the cumulative use of these assays together in a well-defined way and—most of all—the consensus reached among many institutions on the protocol's adoption as a common platform.

2. Definition of the MEDPHOT Protocol

A. General Concepts

The general concepts that we agree as the basis for the design of the protocol are the following:

- to define general procedures applicable to the whole class of photon migration instruments;
- to characterize instruments in terms of measurement results and not of hardware specifications;
- to specify physical parameters instead of clinical ones; and
- to identify a few fundamental assays that probe the key features of photon migration instruments.

The first point was motivated by the need to cover a wide set of photon migration instruments independently of the type of application or measurement technique. The second point was meant to permit the performance assessment of different instruments on the basis of the final outcome of the measurement (e.g., μ_a and μ_s') and not of hardware-dependent specifications. The instrument is considered here as a "black box" with respect to both acquisition and the analysis tools. The hardware specifications of the instruments obviously determine the quality of the measurements, yet the characterization is given in terms of those measurables that are effectively used for the application. The third point simply states that the systems should be validated against reproducible and quantifiable assays, whereas the clinical parameters, such as sensitivity and sensibility, are specific to a particular clinical study (and may also be af-

fected by the specific clinical protocol). Finally, the protocol should be based on a limited set of fundamental assays, which should identify the distinct and fundamental features of photon migration instruments. Particular aspects pertinent only to specific applications should be left aside so as not to overly complicate the definition and the use of the protocol with a cumbersome set of special cases. Instead, dedicated assays could be devised as add-ons to the proposed protocol in order to meet the demands of particular functions.

B. Measurables

Different quantities can be considered as output of a photon migration measurement. The most obvious ones are the optical properties expressed in terms of absorption and reduced scattering coefficients (μ_a, μ_s', respectively). Other quantities often used either for imaging (e.g., mammography) or for oximetry are the detected intensities, expressed as total cw intensity (I_{cw}) or time-gated intensity ($I_{\Delta t}$) in the case of time-resolved instruments. Frequency-domain instruments often produce results directly in terms of amplitude or phase changes. Furthermore, other quantities can be derived, such as tissue-constituent concentrations or normalized intensities.

According to the first criteria expressed above, we designed the protocol by focusing on μ_a and μ_s' as the fundamental measurables. However, the proposed protocol could, in principle, be applied to any other measurable.

We indicate here with x any measurand of the photon migration instrument. In particular, we denote x_{meas} as the measured value of the measurand. The true value of the measurand x_{true} cannot be assumed to be known, thus the conventionally true value x_{conv} is chosen instead.[15] The conventionally true value represents an estimate of x_{true} as derived either by the independent measurements or by a comparison of different instruments operated under optimal experimental conditions. The assessment of a reasonable estimate for x_{conv} is particularly challenging for photon migration phantoms, since often the optical characterization of individual constituents can hardly be performed or cannot be performed at all, as is discussed in Section 5. For a synthetic description of the optical properties of a given phantom, the nominal value x_{nom} is introduced, which corresponds to the optical properties of the phantom as predicted at the design stage for a given optical property. It is only a rough estimate of the measurand to be used only for labeling purposes.

C. Assay Definitions

The MEDPHOT protocol is composed of five assays:

- accuracy
- linearity
- noise
- stability
- reproducibility

The accuracy of the measurement is defined as the capability of the instrument for obtaining a value for the measurable quantity x_{meas} as close as possible to the conventionally true value x_{conv} under optimal experimental conditions [e.g., high signal-to-noise ratio, optical properties of the sample well within the validity range of the theoretical model, and so on]. This figure can be quantified by use of the relative error of the measurement, defined as

$$\varepsilon = \frac{x_{meas} - x_{conv}}{x_{conv}}. \quad (1)$$

This parameter is important for absolute measurements, i.e., whenever the effective value of an optical property or of a constituent concentration is of interest (e.g., to classify the results or to discriminate pathological from healthy regions) or when a tissue must be characterized (e.g., to derive physiological information about the tissue).

A linearity assay is performed by measurement of a set of phantoms combining M values for the absorber concentration (A_i, $i = 1 \ldots M$), with N values for the scatterer concentration (S_j, $j = 1 \ldots N$). For each measurable x_{meas}, an $M \times N$ matrix of measured values is obtained:

$$x_{meas,\,i,j} = f(A_i,\ S_j). \quad (2)$$

If both μ_a and μ_s' are taken as measurands, a total of four linearity plots can be obtained, showing the dependence of the measured μ_a or μ_s' against the conventionally true μ_a or μ_s'. This assay is important for relative measurements to check whether the system can follow changes in a given parameter without distortions. Also, it is crucial for spectroscopy to assure that the shape of the spectrum is preserved, resulting in a correct estimate of the relative abundance of tissue constituents. On the other hand, it can reveal absorption-to-scattering coupling, which can produce artifacts and cause deformations of the scattering spectrum.

The noise assay concerns the variability due to random effects and can be performed by repeating a series of measurements on the same phantom. In particular, we study the noise as a function of the detected optical signal energy E_{out}. For each selected level of the light energy collected from the sample E_{out}, the coefficient of variation CV of a certain number of repeated measurements is derived as

$$\mathrm{CV}(E_{out}) = \frac{\sigma(x)}{\langle x \rangle}, \quad (3)$$

where σ denotes the standard deviation for x_{meas}, calculated for a series of repeated measurements, and $\langle x \rangle$ denotes the corresponding average value.

It is also useful to represent the CV against the injected energy E_{in}. The plot $\mathrm{CV}(E_{in})$ can be used to determine the minimum energy that must be injected in that particular phantom (with given optical properties) to obtain a fluctuation of the measurement

(CV) below a certain threshold. The main drawback in using E_{in} instead of E_{out} is that the instrument characterization is uniquely related to the chosen phantom. Yet E_{out} can be derived from E_{in}, calculating the light attenuation caused by photon migration in the sample by means of the diffusion equation. Finally, the plot of the noise against the number of counts per curve is useful when one is dealing with a time-correlated single-photon counting (TCSPC) system, although it is completely insensitive to the overall coupling and detector efficiencies.

The noise of the measurement directly determines the sensibility of the instrument. In fact, the lowest detectable change in x_{meas} is related to the noise of x_{meas}. Clearly, the noise level is reduced for higher signal intensities and detection efficiencies, and thus the sensibility of the instrument depends on the amount of signal (energy) collected per each measurement point.

The stability assay can be performed by repetition of the measurement on the same phantom many times at subsequent time instants t_i without changing the experimental conditions. The corresponding plot

$$x_{meas} = f(t_i) \tag{4}$$

can reveal short- or long-term drifts of the system as well as unwanted fluctuations. Clearly, the injected energy E_{in} for each measurement must be high enough to set a low CV for x_{meas}, as characterized in the noise assay.

The reproducibility assay is performed by repetition of the measurement on the same phantom under the same experimental conditions on different days. The instrumental reproducibility is expressed as the CV of these measurements. This figure quantifies how the system is self-consistent over different days and permits the correlation of results obtained in different measurement sessions. It is particularly important in the case of clinical studies and whenever measurements obtained over the course of years must be pulled together. In these cases the reproducibility assay should be performed regularly throughout the study.

3. Phantom Kit

The MEDPHOT protocol can be applied by use of any phantom with stable, homogeneous, and controllable optical properties. Yet a specific set of solid phantoms was made on purpose for the MEDPHOT project and was circulated among partners to test the protocol. The phantom is based on epoxy resin, with TiO_2 powder as the scatterer and black toner as the absorber. The phantom recipe was taken from the work of Firbank *et al.*,[11] with the improvements introduced by Swartling *et al.*[16] The scheme of phantom fabrication is depicted in Fig. 1. Briefly, the necessary amounts of toner (black 46/I, part 885 983 06; Infotec, France) and TiO_2 powder (T-8141; Sigma-Aldrich, St. Louis, Missouri) were dispersed in the hardener (H179B; Nils Malmgren AB, Ytterby, Sweden) by sonicating for 20 min. This suspension was added to the resin (NM500; Nils Malmgren AB, Ytterby, Sweden) and stirred manually for 30 min. The mix was poured in the mold and set aside for 1 day at room temperature. Then the phantom was cured in the oven at 50 °C for 12 h, after which it could be machined into the proper shape and polished.

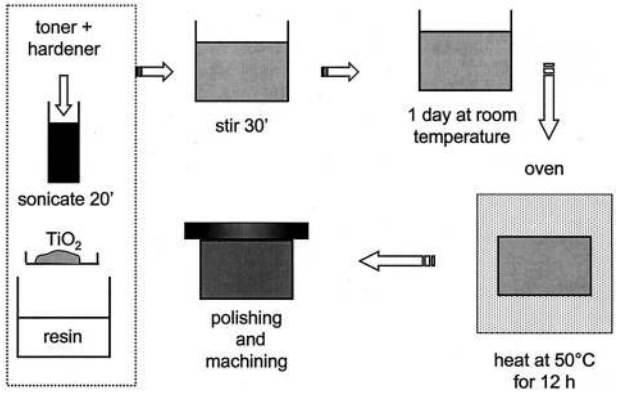

Fig. 1. Scheme of the fabrication of solid phantoms, following the recipe in Ref. 16.

A total of 32 homogeneous cylinders (4.5-cm height, 10.5-cm diameter) were constructed, combining 4 concentrations of TiO_2 powder with 8 concentrations of toner. The TiO_2 and toner concentrations were varied linearly in steps of $\sim 0.05\ cm^{-1}$ for μ_a and $5\ cm^{-1}$ for μ_s' at 800 nm. These phantoms were labeled with a letter and a number, in which the letter stands for the nominal scattering (*A*, *B*, *C*, *D* corresponding to $\mu_s' = 5, 10, 15, 20\ cm^{-1}$, respectively) and the number indicates the absorption (1, 2, 3, 4, 5, 6, 7, 8 correspond to $\mu_a = 0, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35\ cm^{-1}$, respectively).

In addition, three more phantoms were constructed with the same geometry and with identical nominal values of $\mu_a = 0.1\ cm^{-1}$ and $\mu_s' = 10\ cm^{-1}$ at 800 nm. These phantoms were labeled with T_x (T_a, T_b, T_c).

4. Systems Enrolled in the Study

A total of eight instruments, developed by research institutions in five countries, were enrolled in the first application of the MEDPHOT protocol. Table 1 summarizes the key aspects of the systems considered. The systems were grouped into three classes (spectroscopy, imaging, and monitoring) based on the finalization of the instrument-to-tissue optical characterization, on imaging through turbid media, or on monitoring of physiological changes of optical properties, respectively. Most of the instruments were implemented with time-resolved approaches. Two of them were based on cw techniques that exploit multidistance or interferometric measurements. Although no frequency-domain instrument was proposed for the test, the MEDPHOT protocol can be applied to that class of instrument as well. In the following paragraphs, we give a brief description of

Table 1. Classification, Measurement Technique, and Owner Partner of Instruments Characterized with the Proposed Protocol

Class	Technique	Partner	Instrument	Reference
Spectroscopy				
1	Time resolved	POLIMI[a]	Scanning tissue spectrometer	17
2	Space resolved	EMCR[b]	Multifiber tissue spectrometer	18
3	Time resolved	LLC[c]	Four-wavelength portable system	19
Imaging				
4	Time resolved	POLIMI[a]	Optical mammograph	20
5	Time resolved	PTB[d]	Optical mammograph	21
6	Time resolved	POLIMI[a]	Time-gated camera	22
Monitoring				
7	Interferometric	PARIS 13[e]	Oximeter	23
8	Time resolved	POLIMI[a]	Oximeter–functional imager	24

[a]Politecnico di Milano.
[b]Erasmus Medical Center Rotterdam.
[c]Lund University Medical Laser Center.
[d]Physikalisch-Technische Bundesanstalt.
[e]Université Paris 13.

each instrument, referring the reader to the appropriate references for more details.

Instrument 1 is a fully automated system for absorption and scattering spectroscopy of diffusive media.[17] It is based on mode-locked lasers (dye and Ti: sapphire lasers) that are continuously tunable in the 610–1050-nm range and on a detection stage for TCSPC. The maximum power is kept below 1 ÷ 10 mW, depending on the illuminating area, and the typical measurement time for a whole spectrum is ~15 min.

Instrument 2 is a portable cw-based system to derive absorption and scattering properties with a 0.6-nm resolution from 400 to 1100 nm.[18] Light from a 100-W halogen light source is coupled into the sample by means of an optical fiber. Diffuse remitted light from the sample is collected at nine different source–detection fiber distances and is coupled into a spectrograph and projected onto a cooled (−35 °C) CCD camera.

Instrument 3 is a time-domain system intended primarily for spectroscopy of biological tissue.[19] It is based on TCSPC technology and incorporates four pulsed picosecond diode lasers at 660, 786, 916, and 974 nm. The output power is 1–2 mW at each wavelength, and a single measurement is typically performed within 10–30 s.

Instrument 4 is a time-domain optical mammograph operated at four wavelengths (637, 785, 905, and 975 nm).[20] It produces images of the compressed breast both in craniocaudal and oblique projections by continuously scanning the breast at 1-mm intervals with the fibers in a tandem geometry. This instrument is based on picosecond diode lasers and two boards for TCSPC. The maximum laser power is a few milliwatts per each wavelength, and the acquisition time is 25 ms per measurement point, totaling ~5 min for a whole scan.

Instrument 5 is a time-domain optical mammograph for multiprojection imaging.[21] It is equipped with up to eight parallel detection channels, allowing one to measure transmittance through the compressed breast simultaneously on axis and for selected lateral offsets between sources and detectors. The breast is scanned sequentially in craniocaudal and mediolateral projection by use of a step size of 2.5 mm. The device employs four picosecond diode lasers (652, 684, 797, and 830 nm).

Instrument 6 is an imaging system based on a time-gated intensified CCD camera.[22] It permits the parallel acquisition of the time-dispersion curves either in reflectance or in transmittance geometry, over a wide area, and within a few seconds. It can be operated either with pulsed diode lasers or with mode-locked laboratory lasers.

Instrument 7 is an interferometric system that allows one to perform time-resolved measurements of the light scattered by the tissue.[23] The present version works at 780 nm, in reflectance geometry, with a source–detector separation of 1 cm.

Instrument 8 is a time-resolved tissue oximeter based on two diode lasers (690 and 820 nm) and PC boards for TCSPC.[24] It is implemented with 9 sources and 12 parallel detectors, permitting the acquisition of a whole combination of source–detector pairs in 1 s. The laser power is <1 mW per wavelength.

Data analysis is implemented by solving the transport equation under the diffusion approximation and by applying the extrapolated boundary conditions.[18,25,26]

5. Phantom Characterization

As a first step toward the optical characterization of the kit of solid phantoms, Fig. 2 shows the absorption coefficient [Fig. 2(a)] and the reduced scattering coefficient [Fig. 2(b)] obtained with all the instruments listed in Table 1 for the phantom B2, under the experimental conditions specified in the figure caption. There is a certain dispersion among the measurement points, although the spectral features of the absorption spectrum are similarly assessed by all the instruments. As expected, the absorption coefficient exhibits a rather flat plateau below 850 nm, owing mainly to the toner absorption, and higher peaks

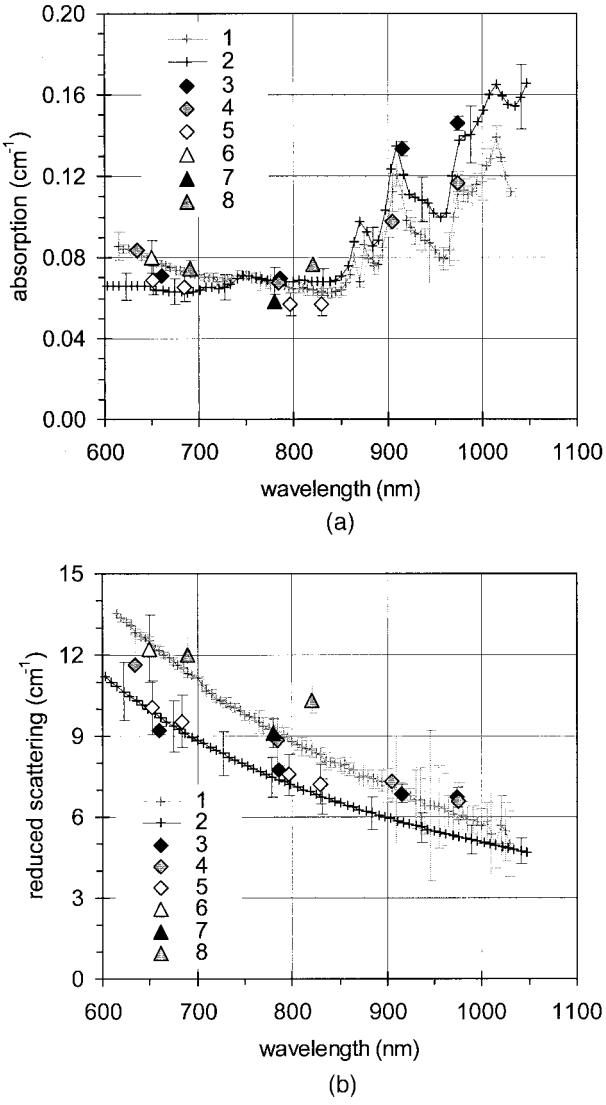

Fig. 2. Comparison of the estimate of (a) μ_a and (b) μ_s' obtained by the eight instruments on the phantom B2. The measurements are performed in transmittance geometry for instruments 4, 5, and 6; in reflectance geometry with $\rho = 2$ cm for instruments 1, 3, 7, and 8; and in reflectance geometry with ρ variable in the interval 0.2–1.8 cm in steps of 0.2 cm for instrument 2.

beyond 850 nm, ascribed to the resin contribution. The differences among instruments pertain mainly to the absolute values, whereas consistency in the spectral shape is maintained.

With respect to the scattering coefficient, all instruments detect, as expected, decreasing values of μ_s' with increasing wavelength. Again, there is a certain dispersion on the absolute estimate of μ_s', with minor variations in the slope of the spectrum.

Overall, comparing the results at wavelengths with at least three independent measurements within a 10-nm range, the average dispersion on μ_a is 10%, whereas the maximum discrepancy between two instruments is found at 970 nm, with a peak-to-peak difference of 32%. Correspondingly, the average dispersion on μ_s' is 13%, whereas the maximum discrepancy is 41% at 820 nm. The measurements were performed over a time period of ~6 months. The phantoms were not checked for long-term stability, and a small change in optical properties, particularly in the resin matrix, cannot be completely ruled out. Yet the intersystem differences in Fig. 2 are not chronologically correlated and thus cannot be ascribed to this problem.

The data in Fig. 2 are not sufficient for a robust characterization of the phantoms. The grand average of the measured optical properties does not necessarily converge to good conventionally true values, since most of the instruments are based on the same technical approach and might be affected by the same systematic errors. On the other hand, a direct assessment of the optical properties of the key phantom constituents is not straightforward. The reduced scattering coefficient produced by the TiO_2 particles is not easily derived. Also, the pure toner powder has nonnegligible scattering properties, which hamper a direct evaluation of its absorption coefficient. Yet we preferred it over other absorbers (e.g., organic dyes and inks) because it is not fluorescent and provides a rather flat absorption spectrum from 600 to 1100 nm.

The importance of obtaining robust and reliable estimates for the phantom optical properties forced us to explore other independent approaches as well as cross-validation tests. This task is nontrivial, owing to the need to derive the spectral properties continuously on a wide wavelength range and to the delicate procedure used to attain a reference standard. This research is still in progress, and the main goal of the present paper is to present the overall methodology of the five-assay protocol. So we decided to devote the characterization of the phantom kit to a specific study to be published in the near future.

To be able to show an example of the application of the accuracy assay, we derived an estimate of the phantom properties from the measurements obtained with system 1. This does not mean that we consider those results to be any better than the others. Neither do we assume that those numbers are conventional true optical properties. Rather, the use of the optical properties derived with the instrument 1 was performed with the sole purpose of producing an example of the accuracy test, as presented in the following paragraph. The absorption spectrum of resin was obtained by averaging the measurements of the phantoms with a null toner concentration (label 1). The toner contribution was assessed by subtracting and averaging the results of measurements performed on phantoms with close absorption values, e.g., average of $[\mu_a(B3) - \mu_a(B2)]$ and $[\mu_a(C3) - \mu_a(C2)]$. To this end, we considered only data obtained in the best experimental conditions in terms of signal and applicability of the diffusion approximation (phantoms B1–4, C1–4; interfiber distance 2–3 cm). The reduced scattering spectrum of the TiO_2 resin matrix was estimated by fitting the average spectra of a selection of phantoms with the power law spectral dependence of the scattering coefficient.[27,28] The optical properties of the phantom kit were calculated synthetically

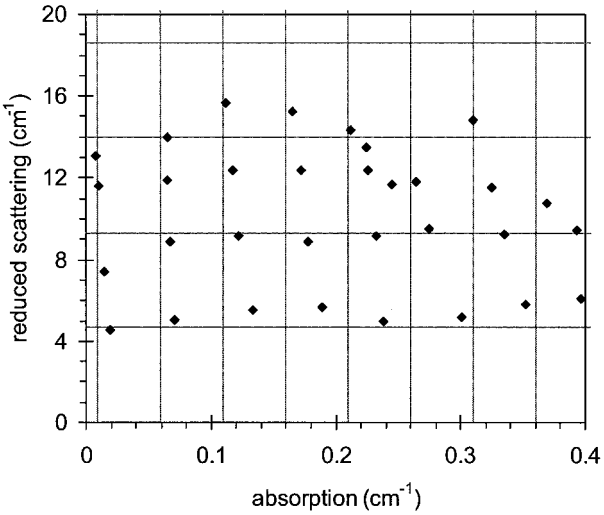

Fig. 3. Accuracy plot obtained with instrument 4 in a transmittance geometry at 785 nm on the whole phantom kit. Each diamond identifies the measured optical properties obtained for each of the 32 phantoms. The grid lines are set in correspondence to a first estimate of the conventionally true values for the phantom properties derived with the limitations described in Section 5.

every 5 nm by use of the optical properties of individual constituents estimated as described above. Thus the estimated μ_a and μ_s' are perfectly linear with the toner and TiO_2 concentrations, respectively.

6. Examples of Application of the MEDPHOT Protocol

The protocol was applied to all the instruments listed in Table 1. An example of a possible outcome of the accuracy assay is reported in Fig. 3, which shows the measured values of μ_a and μ_s' for the 32 solid phantoms, using instrument 4 at 785 nm in a transmittance geometry. The corresponding conventionally true values—obtained as described and with the limitations discussed in Section 5—are plotted as grid lines. Alternative representations are given in Tables 2 and Table 3, report the relative error ε for a measurement of μ_a and μ_s', respectively. In principle, the accuracy assay need not necessarily be applied to the whole kit of phantoms. It is sufficient to assess the

Table 2. Relative Error on the Estimate of μ_a [a]

	μ_s'			
μ_a	A **(4.7)**	B **(9.3)**	C **(14.0)**	D **(18.6)**
1 **(0.009)**	113%	66%	9%	−10%
2 **(0.059)**	19%	14%	9%	11%
3 **(0.109)**	22%	12%	7%	2%
4 **(0.159)**	19%	12%	8%	4%
5 **(0.209)**	14%	11%	8%	1%
6 **(0.260)**	16%	6%	2%	−14%
7 **(0.310)**	14%	8%	5%	−21%
8 **(0.360)**	10%	9%	3%	−14%

[a]The numbers in bold represent the conventionally true values for μ_a (phantoms 1–8) and μ_s' (phantoms A–D) in inverse centimeters.

Table 3. Relative Error on the Estimate of μ_s' [a]

	μ_s'			
μ_a	A **(4.7)**	B **(9.3)**	C **(14.0)**	D **(18.6)**
1 **(0.009)**	−2%	−20%	−17%	−30%
2 **(0.059)**	8%	−5%	−15%	−25%
3 **(0.109)**	18%	−2%	−11%	−16%
4 **(0.159)**	22%	−5%	−12%	−18%
5 **(0.209)**	7%	−2%	−11%	−23%
6 **(0.260)**	11%	2%	−16%	−28%
7 **(0.310)**	25%	−1%	−17%	−37%
8 **(0.360)**	30%	1%	−23%	−20%

[a]The numbers in bold represent the conventionally true values for μ_a (phantoms 1–8) and μ_s' (phantoms A–D) in inverse centimeters.

accuracy on a single phantom (e.g., any phantom T_x) that matches the optimal conditions for the instrument. This check provides the best system performance. Any deviation from this ideal behavior can be further explored with the linearity assay.

An example of a linearity assay is reported in Fig. 4. These data were obtained with instrument 3 at 786 nm in reflectance with $\rho = 2$ cm. If both μ_a and μ_s' are taken as measurands, a total of four linearity plots can be obtained, showing the dependence of the measured μ_a or μ_s' against the conventionally true μ_a or μ_s'. Each plot highlights a different aspect of the measurement. The plot of $\mu_{a,\,meas}$ versus $\mu_{a,\,conv}$ [Fig. 4(a)] displays the linearity characteristics of the system for absorption measurements. It is possible to derive information on the integral nonlinearity, the differential nonlinearity, and the linearity range of the system. Referring to the data presented in Fig. 4, one notes that the system is perfectly linear up to $\mu_a \leq 0.2$ cm^{-1}, and then it starts deviating from linearity, reaching a maximum displacement (integral nonlinearity) of ~20% for $\mu_{a,\,conv} = 0.36$ cm^{-1}. The plot of $\mu_{a,\,meas}$ versus $\mu_{s,\,conv}'$ [Fig. 4(b)] points out any coupling of the absorption coefficient to the scattering coefficient. With reference to the figure, we see that the trend lines are almost horizontal—at least for relative low-absorption properties, that is $\mu_a \leq 0.2$ cm^{-1}—permitting the exclusion of scattering-to-absorption coupling in this range. Conversely, the plot of $\mu_{s,\,meas}'$ versus $\mu_{a,\,conv}$ [Figure 4(c)] investigates the opposite coupling of the scattering to the absorption coefficient. In Fig. 4(c) the tendency of $\mu_{s,\,meas}'$ to increase with increasing values of $\mu_{a,\,conv}$ is a clear indication of scattering-to-absorption coupling. The variation here is not dramatic (an increase of ~30% in μ_s' on the A series for an increase of μ_a from 0.1 to 0.3 cm^{-1}), yet it can produce bumps in the scattering spectrum for large changes in μ_a (e.g., around the water absorption peak). Finally, the plot of $\mu_{s,\,meas}'$ versus $\mu_{s,\,conv}'$ [Fig. 4(d)] shows the scattering linearity. In this figure the trend lines show an almost negligible offset (the intercept of the vertical axis is <0.5 cm^{-1} for most series).

An example of the noise on μ_a plotted as a function

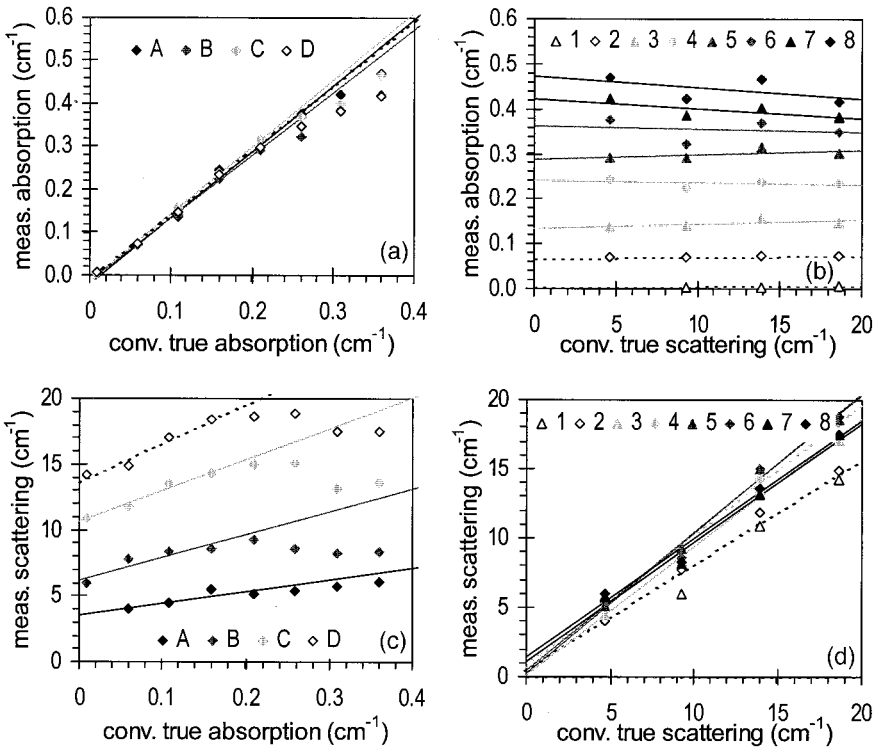

Fig. 4. Linearity plots obtained with instrument 3 in a reflectance geometry at 786 nm for $\rho = 2$ cm. Four different views of the data are presented, corresponding to the changes of $\mu_{a,\,\mathrm{meas}}$ against (a) $\mu_{a,\,\mathrm{conv}}$ and (b) $\mu_{s,\,\mathrm{conv}}'$, as well as to the changes of $\mu_{s,\,\mathrm{meas}}'$ against (c) $\mu_{a,\,\mathrm{conv}}$ and (d) $\mu_{s,\,\mathrm{conv}}'$. The letters and the numbers in the figure legends identify the scattering and absorption labels of the phantoms. The straight lines are linear interpolations on the first four points.

of the input energy E_{in} is shown in Fig. 5, again for instrument 4 at 635 nm, using the phantom T_a. In this case the assay is used to explore the consequence of using as-free parameters in the fitting procedure of both μ_a and μ_s' (dark gray); μ_a, μ_s' and a free time shift t_0 (black); or μ_a and t_0 while fixing μ_s' to a constant value (light gray). In the case of the (μ_a, μ_s') method, an energy of ~3.5 mJ is required to reach a noise level of 6%, which corresponds, for instance, to the typical absorption contrast foreseen in a given application. Conversely, using the free shift approach (μ_a, μ_s', t_0), the same noise level requires ~25 mJ (7 times more energy). The fixed approach (μ_a, t_0) is much more stable, requiring just 2 mJ, and can be of interest to follow small absorption changes under the assumption of a rather constant μ_s'.[29] Since the input power is often limited either by the safety regulations or by the available light power, the energy requirements can be easily related to the minimum acquisition time needed to achieve a given noise level. If the acquisition time is also fixed, the noise plot yields the noise level of the apparatus.

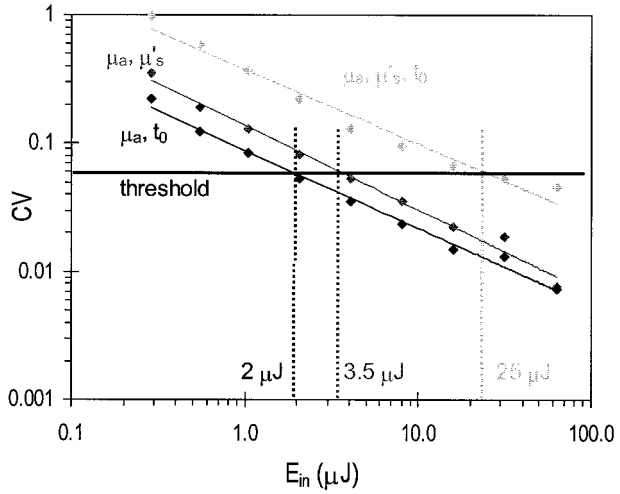

Fig. 5. Plot of the noise level for the measurement of μ_a expressed by the CV calculated for different values of the energy injected into the phantom. The data were obtained with instrument 4 in a transmittance geometry at 785 nm on the phantom T_a. The experimental measurements were fitted by use of as-free parameters μ_a and μ_s' (dark gray); μ_a, μ_s' and a free time shift t_0 (light gray); or μ_a and t_0 while fixing μ_s' (black).

Figure 6 shows an example of a stability assay obtained on the phantom T_a for instrument 8 operated at 690 nm in a reflectance geometry with an interfiber distance of 2 cm [Fig. 6(a)] and for instrument 4 at 635 nm in a transmittance geometry [Fig. 6(b)]. The time course is taken immediately after the instruments have been switched on, for a total of 2 h. The horizontal dashed lines represent a range of ±3% and ±10% with respect to the average value calculated in the last 30 min of the measurement period. For the first instrument [Fig. 6(a)], a reasonable warm-up time seems to be 30 min, after which the system is stable in the assessment of μ_a within ±3%, whereas the second one is still drifting after 1 h. Clearly, the stability requirements depend on the

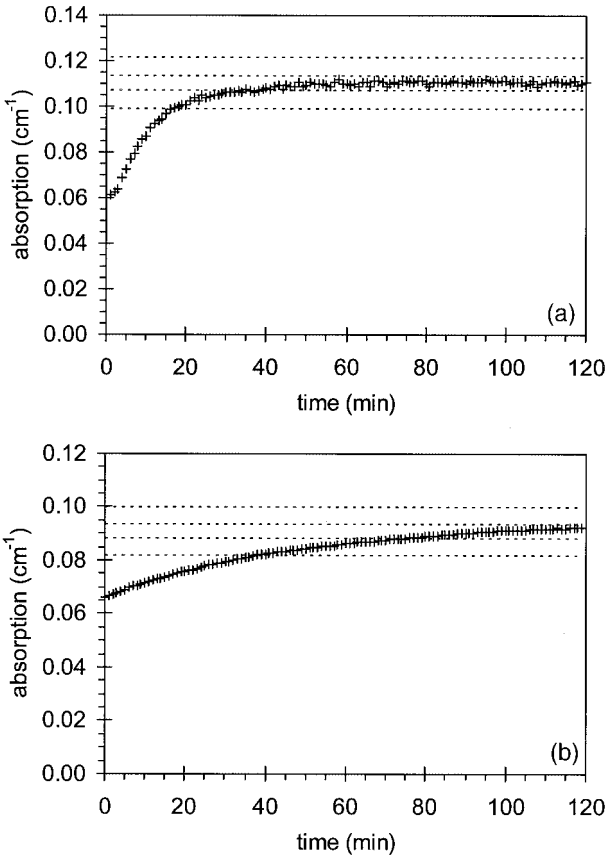

Fig. 6. Stability plot for μ_a obtained on the phantom T_a, using (a) instrument 8 at 690 nm, in a reflectance geometry, with $\rho = 2$ cm and (b) instrument 4 at 785 nm, in a transmittance geometry. The dashed lines correspond to $\pm 3\%$ and $\pm 10\%$ changes with respect to the average of μ_a over the last 30 min of measurement.

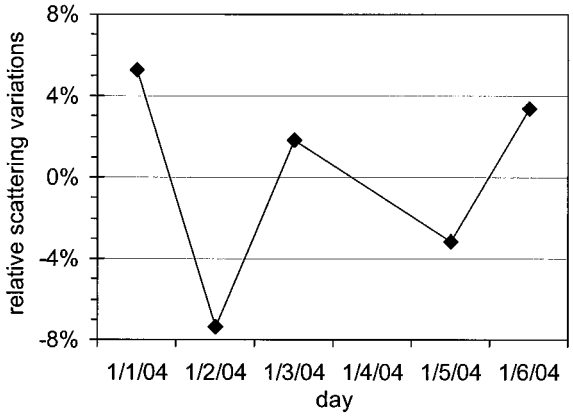

Fig. 7. Reproducibility plot for μ_a obtained on phantom T_a, using instrument 7 at 780 nm, in a reflectance geometry, with $\rho = 1$ cm. The plot represents the relative displacement of $\mu_{a,\,meas}$ obtained at each measurement day with respect to the average value calculated over 5 days.

overall duration of the measurement. Also, in the case of the second instrument, after 1 h the measurement is relatively stable within a few percent for a measurement session of 10 min. The ultimate cause of this deviation is the time-drift of the laser pulses. For constructing the plot of Fig. 6(b), we fitted the data by using the instrumental transfer function (ITF) recorded at the end of the trial. Thus this deviation can be compensated for by use of an ITF that is closer to the actual measurement period, or even better by continuous recording of the ITF as a reference pulse during the measurements.

Figure 7 displays an example of a reproducibility assay for a measurement of μ_s' over 5 different days. These data were obtained with instrument 7 on phantom T_a with $\rho = 1$ cm. All the experimental conditions were kept as constant as possible (e.g., allowing adequate warm-up time, keeping constant ambient light level, etc.). The data are presented as relative variations with respect to the average μ_s'. The average dispersion of the data is 5.2%, with a maximum displacement of 7.4%, which yields an indication of the day-by-day reproducibility and consistency of the measurements.

To permit an easy comparison of results among different instruments or even different releases of the same instrument, as well as to simplify the analysis and reporting tasks, we prepared a common reporting tool and implemented it as an Excel document. The fitted μ_a and μ_s' for the different assays are inserted together in a worksheet with some information related, e.g., to the phantom labels, assay type, measurement time, and so on. With minor actions from the user, a two-page printable summary report is produced, showing all relevant plots, as depicted in Fig. 8. The final section of the report contains some synthetic descriptors of the outcome of the assay, such as the median of the absolute error for the accuracy, the average deviation from linearity, the input energy required to yield a measurement with 1% noise, the slope and range of the stability plot after warm-up time, and the average value of the reproducibility. These numbers permit an immediate appraisal of the system performances as well as a fast and quantitative evaluation of an instrument upgrade.

It is clear that the proposed protocol does not cover all the features related to photon migration instruments. As specified in Section 1, some dedicated assays could be added to the protocol to properly assess issues specific to particular applications. This is the case, for instance, with imaging instruments, for which the aspect of spatial resolution is not encompassed by the MEDPHOT protocol and should be addressed with specific criteria and dedicated inhomogeneous phantoms. Nonetheless, it is also true that the problem of spatial resolution is somehow more linked to the physics of photon migration and to the algorithms used to produce the image rather than to the effective performances of the instruments, and it could possibly be derived from simulations or calculations. On the contrary, the visibility of a suspect lesion, quantified by the effective contrast, is directly related to the noise of the background, as defined the MEDPHOT protocol. Thus the assessment of the five

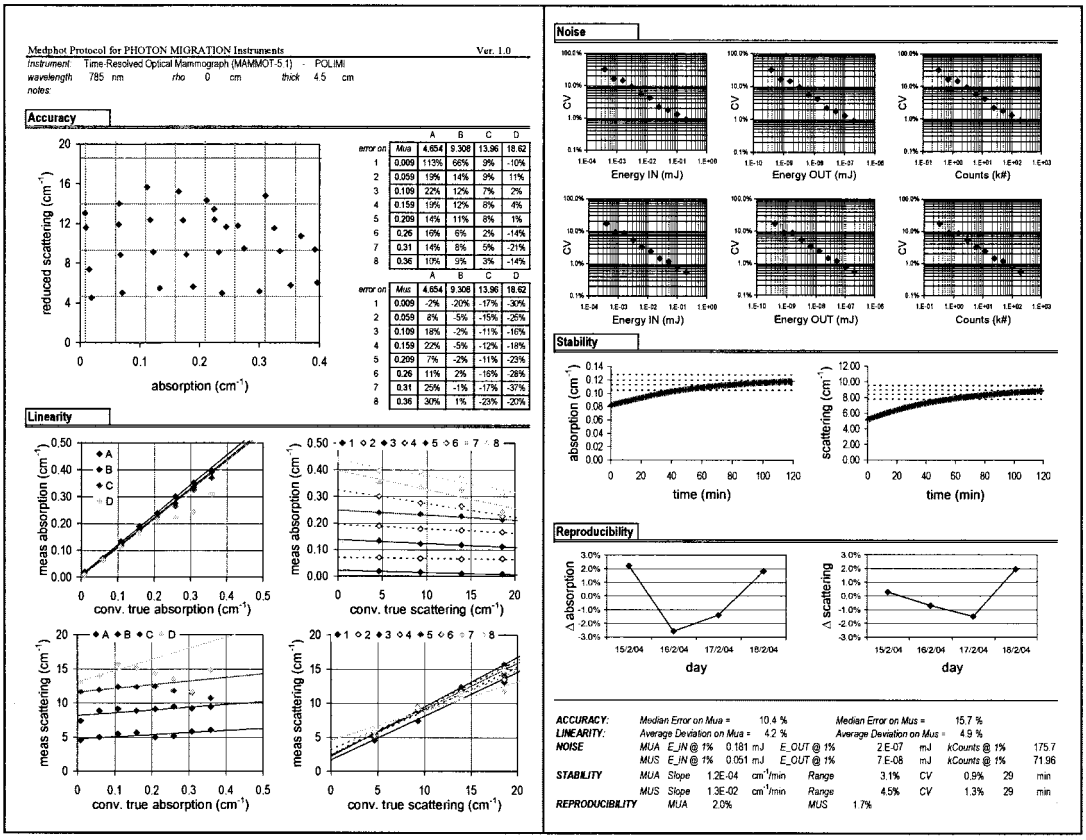

Fig. 8. Printout of the two-page reporting sheet, presenting general information on the instrument, the relevant figures of the MEDPHOT protocol, and some synthetic descriptors of the instrument performances derived from the five assays.

criteria described in this paper cover much of the key aspects of most instruments.

7. Conclusions

In conclusion, we have proposed a novel protocol for the performance assessment of photon migration instruments composed of five assays: accuracy, linearity, noise, stability, and reproducibility. The protocol was applied on a total of 8 instruments, from 5 different countries, using a kit of 32 solid phantoms covering a wide range of optical properties. We have shown examples of the applications of the protocol to encompass different aspects of a photon migration measurement that can be directly related to the needs of the specific application field of the instrument. A unified and quasi-automatic reporting tool permits objective, synthetic, and fast visualization of the protocol results. Research to obtain an accurate spectral characterization of the phantom kit is still in progress. As soon as this parallel study is finished, we will be able to circulate the phantom kit, the protocol specifications, and the reporting tool among the research groups interested in their use.

The study was partially supported by European Union grants QLG1-CT-2000-01464, QLG1-CT-2000-00690, and HPRI-CT-2001-00148.

References

1. *Biomedical Topical Meetings on CD-ROM* (Optical Society of America, Washington, D.C., 2004).
2. Special issue on Recent Developments in Biomedical Optics, Phys. Med. Biol. **49,** (2004).
3. Feature issue on Topics in Biomedical Optics, Appl. Opt. **10–11** (2005).
4. R. Cubeddu, C. D'Andrea, A. Pifferi, P. Taroni, A. Torricelli, G. Valentini, C. Dover, D. Johnson, M. Ruiz-Altisent, and C. Valero, "Nondestructive quantification of chemical and physical properties of fruits by time-resolved reflectance spectroscopy in the wavelength range 650–1000 nm," Appl. Opt. **40,** 538–543 (2001).
5. J. Johansson, S. Folestad, M. Josefson, A. Sparen, C. Abrhamsson, S. Andersson-Engels, and S. Svanberg, "Time-resolved NIR/Vis spectroscopy for analysis of solids: pharmaceutical tablets," Appl. Spectrosc. **56,** 725–731 (2002).
6. S. T. Flock, S. L. Jacques, B. C. Wilson, W. M. Star, and M. J. C. van Gemert, "Optical properties of Intralipid: a phantom medium for light propagation studies," Lasers Surg. Med. **12,** 510–519 (1992).
7. H. J. van Staveren, C. J. M. Moes, J. van Marle, S. A. Prahl, and M. J. C. van Gemert, "Light scattering in Intralipid-10% in the wavelength range of 400–1100 nm," Appl. Opt. **30,** 4507–4514 (1991).
8. S. J. Madsen, M. S. Patterson, and B. C. Wilson, "The use of India ink as an optical absorber in tissue-simulating phantoms," Phys. Med. Biol. **37,** 985–993 (1992).
9. R. Cubeddu, A. Pifferi, P. Taroni, A. Torricelli, and G. Valen-

tini, "A solid tissue phantom for photon migration studies," Phys. Med. Biol. **42,** 1971–1979 (1997).
10. S. Del Bianco, F. Martelli, F. Cignini, G. Zaccanti, A. Pifferi, A. Torricelli, A. Bassi, P. Taroni, and R. Cubeddu, "Liquid phantom for investigating light propagation through layered diffusive media," Opt. Express **12,** 2102–2111 (2004), http://www.opticsexpress.org.
11. M. Firbank, M. Oda, and D. T. Delpy, "An improved design for a stable and reproducible phantom material for use in near-infrared spectroscopy and imaging," Phys. Med. Biol. **40,** 955–961 (1995).
12. U. Sukowsky, R. Schubert, D. Grosenick, and H. Rinneberg, "Preparation of solid phantoms with defined scattering and absorption properties for optical tomography," Phys. Med. Biol. **41,** 1823–1844 (1996).
13. J. C. Hebden, D. J. Hall, M. Firbank, and D. T. Delpy, "Time-resolved optical imaging of a solid tissue-equivalent phantom," Appl. Opt. **34,** 8038–8047 (1995).
14. MEDPHOT Project, http://medphot.jrc.it.
15. ISO Draft Guide, "International Vocabulary of Basic and General Terms in Metrology (VIM)," ISO DGUIDE 99999 (2004).
16. 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,** 4612–4620 (2003).
17. R. Cubeddu, A. Pifferi, P. Taroni, A. Torricelli, and G. Valentini, "Non-invasive absorption and scattering spectroscopy of bulk diffusive media: an application to the optical characterization of human breast," Appl. Phys. Lett. **74,** 874–876 (1999).
18. R. L. P. van Veen, W. Verkruijsse, and H. J. C. M. Sterenborg, "Diffuse reflectance spectroscopy from 500 to 1060 nm by correction for inhomogenously distributed absorbers," Opt. Lett. **27,** 246–248 (2002).
19. J. Swartling, A. Pifferi, E. Giambattistelli, E. Chikoidze, A. Torricelli, P. Taroni, M. Andersson, A. Nilsson, and S. Andersson-Engels, "Rigorous characterization of time-resolved diffuse spectroscopy systems for measurements of absorption and scattering properties using solid phantoms," in *Photon Migration and Diffuse-Light Imaging*, D. A. Boas, ed., Proc. SPIE **5138,** 80–87 (2003).
20. A. Pifferi, P. Taroni, A. Torricelli, F. Messina, R. Cubeddu, and G. Danesini, "Four-wavelength time-resolved optical mammography in the 680–980-nm range," Opt. Lett. **28,** 1138–1140 (2003).
21. M. Moeller, H. Wabnitz, A. Kummrow, D. Grosenick, A. Liebert, B. Wassermann, R. Macdonald, and H. Rinneberg, "A four-wavelength multi-channel scanning time-resolved optical mammograph," in *Photon Migration and Diffuse-Light Imaging*, D. A. Boas, ed., Proc. SPIE **5138,** 290–297 (2003).
22. C. D'Andrea, D. Comelli, A. Pifferi, A. Torricelli, G. Valentini, and R. Cubeddu, "Time-resolved optical imaging through turbid media using a fast data acquisition system based on a gated CCD camera," J. Phys. D **36,** 1675–1681 (2003).
23. J.-M. Tualle, "E. Tinet, and S. Avrillier, "A new and easy way to perform time-resolved measurements of the light scattered by a turbid medium," Opt. Commun. **189,** 211–220 (2001).
24. A. Torricelli, V. Quaresima, A. Pifferi, G. Biscotti, L. Spinelli, P. Taroni, M. Ferrari, and R. Cubeddu, "Mapping of calf muscle oxygenation and haemoglobin content during dynamic plantar flexion exercise by multi-channel time-resolved near infrared spectroscopy," Phys. Med. Biol. **49,** 685–699 (2004).
25. M. S. Patterson, B. Chance, and B. C. Wilson, "Time-resolved reflectance and transmittance for the noninvasive measurement of tissue optical properties," Appl. Opt. **28,** 2331–2336 (1989).
26. R. C. Haskell, L. O. Svasaand, T. T. Tsay, T. C. Feng, M. S. McAdams, and B. J. Tromberg, "Boundary conditions for the diffusion equation in radiative transfer," J. Opt. Soc. Am. A **11,** 2727–2741 (1994).
27. A. M. K. Nilsson, C. Sturesson, D. L. Liu, and S. Andersson-Engels, "Changes in spectral shape of tissue optical properties in conjunction with laser-induced thermotherapy," Appl. Opt. **37,** 1256–1267 (1988).
28. J. R. Mourant, T. Fuselier, J. Boyer, T. M. Johnson, and I. J. Bigio, "Predictions and measurements of scattering and absorption over broad wavelength ranges in tissue phantoms," Appl. Opt. **36,** 949–957 (1997).
29. R. Cubeddu, A. Pifferi, P. Taroni, A. Torricelli, and G. Valentini, "Experimental test of theoretical models for time-resolved reflectance," Med. Phys. **23,** 1625–1633 (1996).