nature photonics

PERSPECTIVE

https://doi.org/10.1038/s41566-018-0221-5

Tackling standardization in fluorescence molecular imaging

Maximillian Koch[1,2,3], Panagiotis Symvoulidis[1,2,3] and Vasilis Ntziachristos[1,2*]

The emerging clinical use of targeted fluorescent agents heralds a shift in intraoperative imaging practices that overcome the limitations of human vision. However, in contrast to established radiological methods, no appropriate performance specifications and standards have been established in fluorescence molecular imaging. Moreover, the dependence of fluorescence signals on many experimental parameters and the use of wavelengths ranging from the visible to short-wave infrared (400–1,700 nm) complicate quality control in fluorescence molecular imaging. Here, we discuss the experimental parameters that critically affect fluorescence molecular imaging accuracy, and introduce the concept of high-fidelity fluorescence imaging as a means for ensuring reliable reproduction of fluorescence biodistribution in tissue.

Human vision is the oldest medical visualization method, but is not an ideal approach for disease detection. Eyesight cannot penetrate beneath the tissue surface and shows low detection sensitivity and specificity, as cancer and other diseases do not always present strong or unique optical contrast over surrounding tissues. Despite these fundamental limitations, human vision remains a major detection strategy in important diagnostic and therapeutic procedures. Visual inspection of tissues by the naked eye guides critical decisions on tissue excision during open surgery. Colour video obtained by an endoscope during laparoscopy or gastrointestinal inspection guides detection and diagnosis. Likewise, dermatology and gynaecology widely employ optical inspection. However, the limitations of visual interrogation often dictate the need for random biopsies during an endoscopic process, which are later analysed with histopathology. Similarly, evaluation of complete surgical excision of cancer is available only through postoperative pathology confirmation. Radiological techniques such as X-ray or magnetic resonance imaging have been considered for intraoperative use but involve high cost and size, limited ability to detect microscopic disease and complex operation that challenges dissemination to every operating room. Ultrasound imaging can be made portable, but it offers limited contrast, requires physical contact with tissue and cannot intraoperatively survey large tissue areas. These features limit its use as an interventional modality, in contrast to the wide-field capability of optical methods. Therefore, the 3,000-year-old clinical use of visual inspection has not yet found a widely used surrogate technique for guiding interventions.

Fluorescence molecular imaging (FMI) — that is, imaging of fluorescent agents engineered to target specific disease biomarkers — is emerging as a technology with substantial potentials to assist visual inspection of tissue. Different fluorescent agents that target biochemical and molecular features of diseases are being approved for experimental clinical use[1–3]. Many FMI studies on patients have demonstrated greater detection sensitivity and abilities for subsurface visualization than can be achieved with human vision[1]. A folate receptor-targeting fluorescence agent has been used to improve detection of ovarian cancer in patients[3,4], and fluorescently labelled cetuximab has been employed to improve precision in oral cancer surgery[5]. A protease-activated fluorescent agent was also considered in humans for detecting soft tissue sarcoma[6]. Likewise, detection of colonic dysplasia[7] using a targeted heptapeptide and detection of Crohn's disease using a fluorescently labelled antibody targeting tumour necrosis factor[8] have been demonstrated in humans.

Overall, the use of fluorescent agents is expected to change the landscape of interventional guidance and of decision-making during interventions, improving disease detection, guiding biopsy and enabling theranostics. This encouraging progress with FMI translation brings a critical but unanswered question: are certain specifications and standards needed for clinical use and, if so, what should they be? Fluorescence imaging is not an established medical imaging modality at present: no guidelines or standards have been put in place that define what constitutes appropriate fluorescence detection performance. Different FMI systems are currently implemented using a variety of cameras, illumination sources, data processing methods and other experimental parameters. These variations in hardware and software may lead to systems with markedly different operational characteristics. Further complicating the definition of FMI specifications is the availability of fluorescence agents covering various spectral ranges, including the visible (450–650 nm), near-infrared (NIR; 650–900 nm), and NIR-II or short-wave infrared (SWIR; 900–1,700 nm).

A critical realization in fluorescence imaging is that clinical FMI findings do not depend only on the fluorescence agent employed but also on the detection methodology utilized. For example, the use of cameras of different sensitivity or dynamic range and the use of different data processing methods can lead to different clinical performance outcomes, even when imaging the same fluorescent agent and pathology. Consequently, different FMI systems and methods with dissimilar performance characteristics will report different sensitivity and specificity results for the same patient cohort and agent. Importantly, FMI performance depends not only on the system employed and data processing methodology but also on the tissue imaged and the overall experimental arrangement. For example, variation in the optical properties of the tissue, depth of the fluorescence activity or even ambient light affect the recorded image[1]. Such discrepancies may impede the acceptance and approval of FMI

[1]Chair for Biological Imaging, TranslaTUM, Technical University of Munich, Munich, Germany. [2]Institute of Biological and Medical Imaging, Helmholtz Zentrum München, Neuherberg, Germany. [3]These authors contributed equally to this work: Maximillian Koch, Panagiotis Symvoulidis. *e-mail: v.ntziachristos@tum.de

studies and illustrate the urgent need to develop processes that can ensure accurate and reliable FMI in clinical settings.

We review herein the parameters that affect the fluorescence image and advocate for improved FMI accuracy and standardization. In analogy to the term 'high fidelity' employed for audio systems, we introduce the concept of high-fidelity fluorescence imaging (HiFFI) to provide accurate clinical FMI. The term 'high fidelity' denotes the accurate reproduction of a recorded signal, with minimum distortions. This concept is proposed for clinical fluorescence imaging as a means for ensuring that different FMI implementations will reliably report on the actual biodistribution of a fluorescence agent in tissue, independent of the system, processing method, environmental conditions and optical properties or other parameters of the tissue imaged. Subsequently, we describe how standardization and reversion are necessary steps for the clinical implementation of HiFFI. We propose the use of advanced fluorescence standards, in the form of composite phantoms, to offer a holistic characterization of fluorescence imaging systems and system calibration, and we explain how calibration may be necessary for better understanding of system performance and processing data from different clinical studies. The overall aim of this analysis is to propose a basis for practising accurate FMI in the clinic.

Fluorescence imaging implementation and spectral range

FMI technology is based today on the principle of photography. A camera sensitive to fluorescence photons (Fig. 1a) is fitted with high-pass or band-pass filters to block photons at the excitation wavelength and allow the recording of photographs of emission (fluorescence) photons exiting the tissue surface. Different operational specifications and spectral regions spanning 400–1,700 nm (Fig. 1c) can be implemented, based on the particular light-sensitive technology employed, for example, a charge-coupled device (CCD), indium gallium arsenide (InGaAs) or a complementary metal-oxide–semiconductor (CMOS). Often, the fluorescence camera is combined with a colour camera, through a beam splitter (Fig. 1b). The colour image (Fig. 1d) provides a morphological reference that directly relates to the surgeon's or endoscopist's visual perception and allows registration of the fluorescence image with the colour image (Fig. 1e). A preferred method of superimposing the fluorescence image onto the colour image is to make pixels transparent if the fluorescence value is low, such that strong fluorescence signals appear in pseudocolour and weak fluorescence signals are invisible[9]. Typically, both the fluorescence camera and the colour camera visualize the same or similar field of view by using a common lens for image collection (Fig. 1a). Other optical systems, such as flexible endoscopes, can be employed instead of lenses[10].

While FMI has so far been performed primarily in the visible and near-infrared regions, the development of novel materials and the use of NIR fluorochromes that emit in the spectral ranges 950–1,100 nm (extended NIR) or 1,230–1,330 nm (NIR-II) or that emit at even longer wavelengths now allow imaging over broader spectral ranges[11–14]. Imaging with wavelengths beyond 950 nm offers advantages over imaging in the visible and NIR regions as scatter and autofluorescence decrease with increasing wavelength. Therefore, images generated at wavelengths beyond 950 nm can have higher resolution than images generated in the NIR region (Fig. 1f,g). At the same time, photon absorption in many spectral regions longer than 950 nm is greater than in the NIR region, which may reduce detection sensitivity and penetration depth. Nevertheless, advances in instrumentation and biocompatible fluorochromes may make the extended NIR or NIR-II a preferred FMI spectral region. NIR-III (Fig. 1c) offers even lower scatter than NIR-II; however, tissue absorption may be even greater in this spectral region than in the visible range, imposing significant sensitivity and penetration barriers.

Fluorescence imaging challenges

Photographic FMI (Fig. 1a,b) appears technologically simple and is relatively cost-effective. However, this technical simplicity comes with key performance limitations. The primary challenge of FMI is the dependence of the fluorescence signals collected on many different experimental conditions. The aim of a fluorescence image is to represent the underlying biodistribution and concentration of an administered fluorescent agent. This requirement implies that (1) there should be a linear relationship between signal strength and amount of fluorochrome, and (2) the fluorescence intensity recorded depends only on the concentration of the underlying fluorescent agent. In reality, many physical, operational and tissue-related parameters affect the fluorescence intensity and image appearance (Table 1)[15]. This means that for a given fluorochrome concentration in tissue, differences in the parameters in Table 1 result in different images and quantified readings (Fig. 2). Therefore, the apparent technical simplicity of FMI is rather misleading: accurate FMI operation requires technology that can consistently report the same fluorescence readings, independently of variation in the parameters in Table 1.

There are two general classes of parameters that affect a fluorescence image and are therefore implicated in FMI accuracy and fidelity.

The first class relates to invariable parameters, that is, the specifications of the hardware components employed and the overall design of the FMI system. This category primarily includes instrumentation parameters such as the photon detection sensitivity of the camera(s) employed, the resolution, the spectral region covered and other parameters as summarized in Table 1. Typically, these hardware-related features do not change during a measurement.

The second class of parameters relates to variable parameters, that is, parameters that can change during a measurement or from measurement to measurement. Variable parameters may relate to the environment (ambient light, distance of the camera from the tissue), hardware parameters (zoom and focus) or tissue features, such as optical properties of the lesion imaged and of surrounding tissue, autofluorescence or lesion depth. Changes in the tissue absorption and scattering occur due to different tissue constituents (for example, variation in vascular density, bleeding, tissue type) and the spectral region employed. Optical properties modify the signal strength and image appearance (Fig. 2). These parameters may change from patient to patient or even within the field of view imaged.

Sensitivity. FMI requires sufficient sensitivity to image the low concentrations typically attained by targeted fluorescence agents, in the range of sub-nM to 100 nM (ref. [1]). Video-rate FMI implementation further imposes tight sensitivity specifications; conventional 24 frames-per-second video allows an exposure time per frame of only ~42 ms. Camera sensitivity is one of the most critical parameters in achieving high signal-to-noise ratio fluorescence detection within tens of milliseconds. Highly sensitive cameras may reliably report fluorescence at agent amounts that are missed by low-sensitivity cameras (Fig. 2a–c). Other parameters that affect FMI sensitivity include the intensity and spectral response of the illumination employed for fluorochrome excitation and the ability to block excitation and other light from the fluorescence channel (cross-talk). Variations in the parameters that affect FMI sensitivity have consequences for the reliability of FMI data collected, as well as for the sensitivity and specificity values reported for a given agent. Moreover, FMI sensitivity affects the administered dose required. Low-sensitivity FMI implementations require higher doses than high-sensitivity implementations to compensate for the reduced detection performance. This practice raises safety and ethical concerns and also leads to higher costs because of the need to produce a larger amount of the agent administered.

The sensitivity that can be expected from an FMI system depends on the agent employed. Non-specific, vascular dyes such

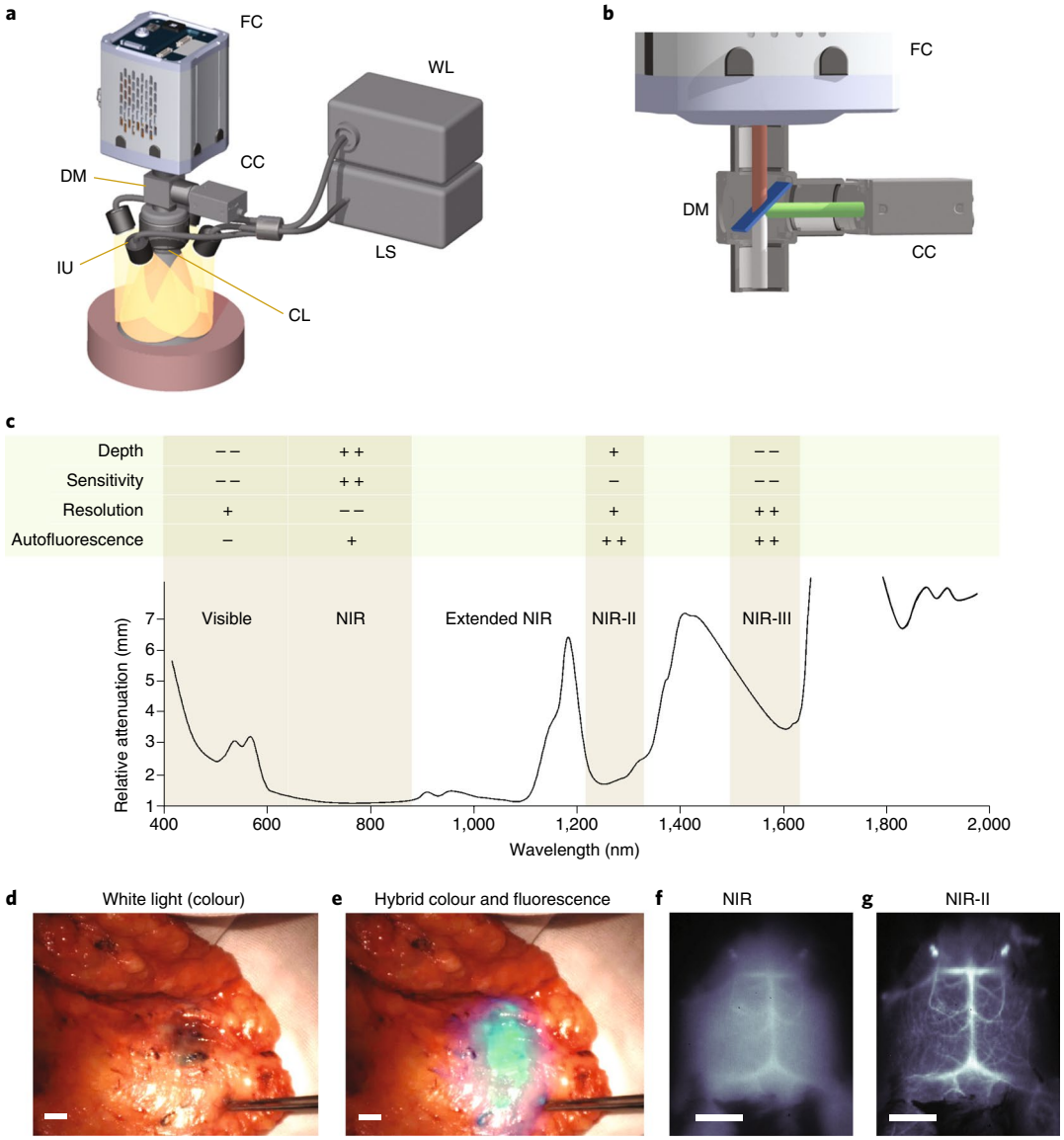


Fig. 1 | Fluorescence intraoperative imaging. a, Typical composite camera system using a highly sensitive fluorescence camera (FC) to collect fluorescence images and a colour camera (CC) to collect white-light images through a dichroic mirror (DM) and a common lens (CL). Different light sources may be used for white-light excitation (WL) and fluorescence excitation using a laser source (LS) and common illumination unit (IU). **b**, Detailed view of the optical paths collected through the dichroic mirror; visible light (green path) is directed to the colour camera and near-infrared light (brown path) is directed to the fluorescence camera. **c**, Absorption spectrum of tissue obtained by optoacoustic spectroscopy from a mouse skin in vivo. The spectrum shows the different spectral windows typically exploited by FMI. The '+' and '−' symbols indicate advantages and disadvantages of each window, respectively. **d**, Colour image collected intraoperatively from breast tissue of a breast cancer patient. **e**, Superposition of the colour image in **d** onto a fluorescence image in pseudo-cyan/green, identifying subsurface breast cancer. The fluorescence image was obtained after systemic administration of bevacizumab labelled with the NIR fluorescence dye CW800 (Licor). **f**,**g**, Image of a mouse brain in the NIR region (**f**) and the corresponding image obtained in the NIR-II region (**g**). Resolution is better in the NIR-II than NIR region because of less scatter. Scale bars, 5 mm. Figure reproduced from: **f**,**g**, ref. [11], Springer Nature Ltd.

as indocyanine green (ICG) are often injected in quantities of tens of milligrams, up to 25 mg when given systemically or ~1 mg when given intratumorally[1], reaching concentrations of several hundred nanomolars to several micromolars in tissues. In contrast, targeted agents are used in much lower concentrations[1,15,16], as most of the administered dye is cleared from tissue and only a small amount is present in the targeted tissue. Calculations based on clinical data have shown that due to targeting and clearance, the concentrations of targeted agents imaged may be five to six orders of magnitude lower than when imaging ICG[1,17].

Microdosing (phase 0) studies have also been suggested for accelerating clinical translation of novel agents[16]. Camera sensitivity plays a critical role in enabling microdosing (tracer-amount) protocols. This is evident in the recent clinical translation of fluorescently labelled bevacizumab[18], achieving a signal-to-noise ratio of ~12 dB for microdosing administration. This performance was achieved using cooled cameras with electronic noise detection of a few electrons, but not using less-sensitive cameras. Today, there is no definition or notion of the term 'sufficient sensitivity' in clinical fluorescence imaging.

Table 1 | Parameters affecting FMI performance

Parameter	Typical range	Effect on FMI performance	Calibration	Remedy
Invariable parameters				
Camera sensitivity	nM–pM	Dose of agent required Frame rate achieved Minimum fluorescence activity detected Phase 0/microdosing operation Sensitivity and specificity of clinical findings	Measure (calibrate) sensitivity with standard (FIS)	Use highly sensitive CCD/CMOS/InGaAs technology, current amplification methods, low-noise electronics and cooling technology to reduce noise
Electrical/read noise	$2e^-$–$20e^-$ per read operation	Same as in 'Camera sensitivity'	Same as in 'Camera sensitivity'	Same as in 'Camera sensitivity'
Resolution	10–500 μm	Minimum lesion size visible on white-light images	Register white-light and apparent diffusive resolution with standard	Match the number of pixels and field of view to the desired resolution
Dynamic range and dark current	10^4–10^6	Ability to differentiate between different amounts of distributed agent Saturation effects	Measure with standard	Select sensors with high full-well capacity
Frame capture speed	1–100 Hz	Quality of video	NA	Select camera with fast read-electronics and data transfer
Spectral coverage	400–1,700 nm	Resolution achieved Sensitivity achieved Depth achieved	Use fluorochromes (quantum dots) whose spectral responses are known	Select sensor technology with sufficient sensitivity in spectral range covered
Filters, cross-talk and ambient light	0.1–50% of excitation light	Reduction of sensitivity Increase background noise Increase image artefacts	Measure cross-talk and ambient light under control conditions	Select proper filters; condition light source; subtract reference light/time-share measurement
Illumination intensity and spectral profile	1–200 mW m^{-2} Upper value may be regulated by ANSI limits	Same as in 'Camera sensitivity'	Measure (calibrate) intensity with a standard (FIS), power meter and/or spectrometer	Employ illumination close to the ANSI limit and sources of optimal spectral response
Illumination homogeneity	Varies with system design	Shadowing effects on the images collected Accuracy (quantification) variations of different lesions Sensitivity and specificity of clinical findings	Measure the illumination pattern (see also Box 1)	Multi-angle illumination; normalize image with captured illumination pattern
Variable parameters				
Camera–tissue distance and field of view	15–100 cm	Variations in fluorescence intensity recorded Changes in focus Sensitivity	Record changes in field of view and distance	Real-time distance and field of view sensors or estimators
Depth of focus	1–10 cm	Reduced resolution with changes in tissue elevation and camera–tissue distance	Record iris and depth of focus settings	Use high depth of focus to avoid out-of-focus images; use an autofocus mechanism
Variation of optical properties	Scatter: 5–30 cm^{-1}; absorption: 0.05–0.5 cm^{-1}	Variations in fluorescence signal intensity Variations in apparent fluorescence distribution Variations in resolution and diffusion on the image	Record system performance as a function of optical property changes	Record variations in tissue absorption and scattering in real time
Autofluorescence	Varies with spectral region (see Fig. 1c)	Reduces detection sensitivity May lead to false positives	Record system performance as a function of background fluorescence	Use spectral differentiation of target fluorescence over background fluorescence
Lesion depth	0–2 cm	Attenuation of fluorescence intensity Variable diffusion and loss of resolution Spectral changes	Record system performance as a function of fluorescence depth	Tomography; depth reconstruction based on spectral changes

ANSI, American National Standards Institute; NA, not applicable.

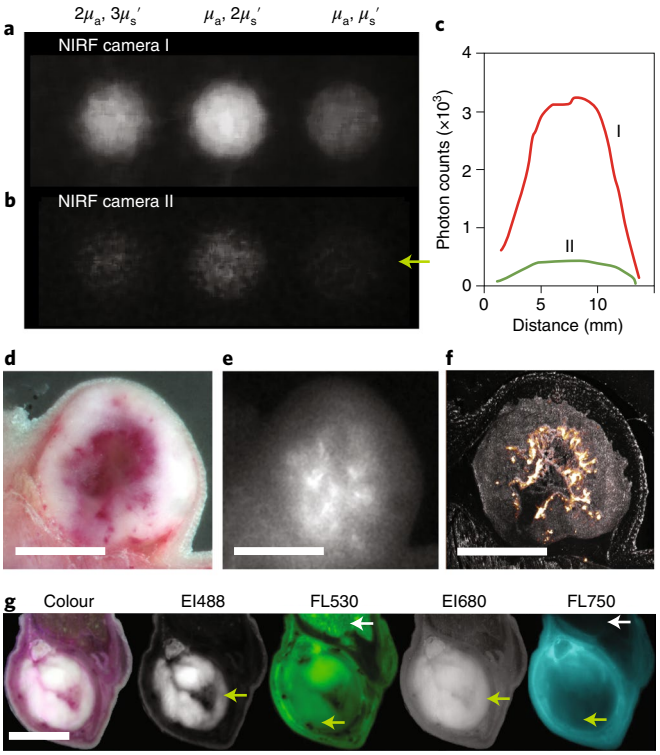


Fig. 2 | Fluorescence imaging challenges. Effects of system and tissue parameters on fluorescence images. **a**,**b**, Three fluorescence wells in the upper right quadrant of the composite phantom described in Box 1 were visualized using two different near-infrared fluorescence (NIRF) cameras with the same filters and illumination parameters. Camera I (iXon; Andor) offers tenfold higher sensitivity than camera II (Luca; Andor). Camera II nearly fails to detect a well (green arrow) that camera I resolves. All wells contained the same amount of fluorochrome; intensity variations are attributed to variation in optical properties (see also Fig. 3). **c**, Fluorescence profiles through the rightmost well imaged by the two cameras illustrate the marked detection difference between them. **d**, Colour image obtained by cryo-slicing through a 4T1 tumour subcutaneously grown on a nude mouse. **e**, Corresponding image at 750 nm after administration of liposomal ICG (24 h before euthanasia). **f**, A fluorescence image of a 100 μm slice obtained from **e** reveals a more detailed pattern of fluorescence distribution not prone to photon diffusion. Such images may contain cross-talk from the excitation channel (see Table 1), which can be eliminated using calibration. **g**, Effects of optical properties and spectral region on the fluorescence image. A mouse bearing an intramuscular 4T1 tumour was injected intravenously with equal amounts of AF488 and AF750 fluorescent dyes, then imaged in epi-illumination mode using white-light (colour) illumination. Excitation wavelengths were 488 and 680 nm (EI488, EI680); fluorescence was imaged around 530 and 750 nm (FL530; FL750). Arrows indicate regions of special interest: white arrows indicate an area of strong autofluorescence in the visible range, which disappears in the NIR range; green arrows indicate regions of high absorption that also disappear in the NIR region. Scale bars, 5 mm.

Other invariable system parameters. Several other camera parameters may affect image performance. (1) Dynamic range, that is, the difference between the largest and smallest signal that can be recorded by a camera, plays a key role in comprehensively recording the fluorescence response of a medium, that is, capturing both strong and weak fluorescence sources. (2) Camera spatial resolution, which is in a trade-off relationship with sensitivity for the same detector area, defines the smallest fluorescence lesion size that can be accurately recorded. (3) Homogeneity of the illumination pattern also plays a role in image accuracy. Spatial variations of the illuminating energy within the field of view will vary the fluorescence intensity independently of the underlying fluorochrome concentration. The particular illumination implementation may also accentuate or minimize shadowing effects when imaging curved surfaces. Ideally, a technology that offers homogeneous illumination that minimizes shadowing effects is preferred for HiFFI. (4) The maximum frame rate supported by a camera, together with the camera sensitivity, defines the ability to implement video-rate imaging, which is preferred in interventional applications. (5) The spectral range covered by the camera employed defines the range of fluorophores that can be imaged and the overall imaging performance (resolution, sensitivity, depth) that can be achieved.

Variable system parameters. In addition to parameters that are system-dependent and can be determined on a system basis, there is a range of parameters that may change during operation and will also affect image performance. In particular, variations in the camera focus, zoom and distance from tissue can modify the fluorescence intensity recorded, the minimum area resolved and the field of view visualized. Ambient light conditions may further affect camera performance, depending on its design. Changes in these parameters also modify the apparent resolution or sensitivity of the camera employed. Variable system parameters may change between patients and even within the field of view imaged.

Tissue properties. A particular set of variable parameters are tissue dependent as they relate to tissue optical properties. White-light (colour) imaging primarily visualizes photons back-scattered (reflected) from the tissue surface and sees only superficial structures. In contrast, fluorescence photons are generated within the tissue, not on the tissue surface. Fluorescent agents volumetrically concentrate within tissue structures and tumours, at depths ranging from micrometres to several centimetres under the tissue surface. Fluorescence photons are isotropically emitted within tissue and most of them typically undergo several scattering events before detection, even when produced by superficial lesions, often exhibiting diffusive behaviour (Fig. 2e–g). Therefore, the recorded fluorescence images depend not only on the fluorochrome concentration and FMI-system-specific parameters, but also on the optical properties of the tissue surrounding the fluorochrome imaged and on the depth at which fluorescence is generated. Typically, NIR fluorescence images exhibit stronger diffusion, such that visible fluorescence images show stronger dependence on photon absorption and appear to have higher resolution (Fig. 2g). NIR-II images also appear to have higher resolution, which is due to lower scatter than visible and NIR images.

The modification of fluorescence signals and images by tissue optical properties has been discussed in several studies[19–21] and presents a particular FMI challenge. First, fluorescence images do not report true fluorescent agent biodistribution (Fig. 2e–g) but rather a composite signal that depends on fluorescent agent concentration and underlying tissue optical properties. Second, diffusion reduces the resolution of fluorescence images and often makes the identification of disease borders and edges problematic (Fig. 2e).

Autofluorescence from tissue components such as collagen or nicotinamide adenine dinucleotide (NADH) may also reduce the contrast available from a targeted fluorochrome and alter the appearance of a fluorescence image (Fig. 2g). Fluorescence detected by native tissue fluorochromes can be mistaken for signal from the administered targeted agent. Further modification of fluorescence by tissue optical properties can exacerbate the effects of autofluorescence, such as when autofluorescence is present in a lesion of low attenuation or when a targeted agent is concentrated in an area of high attenuation (for example, a highly vascular lesion). Finally, the fluorescence signal recorded from a lesion shows a strongly nonlinear dependence on the lesion depth (Box 1). Therefore, variations in

lesion depth have a profound influence on the fluorescence intensity and the lesion appearance recorded on the image.

High-fidelity fluorescence imaging

Given the absence of established guidelines and standards, FMI data collected today uniquely depend on the particular system employed and the experimental conditions of the measurement, that is, on the particular values for the parameters in Table 1. Consequently, the statistical analysis and clinical performance of an agent reported may be study specific and not agent specific, with statistical metrics varying from system to system and study to study, especially when the parameters listed in Table 1 vary significantly between FMI implementations.

We introduce the principle of HiFFI, defined as the accurate representation of fluorochrome biodistribution in tissues, independent of the particular system or experimental and tissue conditions present during the measurement. HiFFI indicates that the fluorescence image recorded represents the true fluorochrome biodistribution in tissue and does not change when the parameters in Table 1 are modified.

HiFFI implementation requires an effective account of the parameters in Table 1. Addressing the effects of invariable parameters on image fidelity involves a different methodology than addressing the effects of variable parameters. This difference is because invariable parameters can be captured once or repeatedly over time to characterize system specifics, whereas variable parameters change during the imaging session and may necessitate real-time measurements throughout the fluorescence measurement. In the following, we discuss two principles, standardization and reversion, which are suggested for improving image fidelity in fluorescence imaging. Standardization relates to controlling a system in relation to the parameters in Table 1, whereas reversion relates to correcting for the effects of those parameters on the fluorescence image.

Standardization

Methodologies that introduce uniformity and quality control across FMI implementations are necessary for characterizing FMI performance specifications and ensuring that an FMI system performs according to its specification over time. Standardization can be achieved by utilizing a standard[22]. Several materials and objects have been proposed as fluorescence imaging standards. Box 1 outlines the ideal characteristics of a calibration standard and introduces the concept of a fluorescence imaging standard (FIS) in the form of a composite phantom that captures and calibrates all parameters in Table 1. The use of composite HiFFI phantoms can enable complete system characterization with a single image or a few images. We expect that the use of a commonly accepted phantom and its acceptance as a standard will promote good imaging practices in system development and clinical practice and will promote FMI as a high-quality, high-performance tool for improving interventional and endoscopic procedures. An FIS can be applied in the following processes.

(1) Guidelines: ensure a performance specification for all systems/cameras employed in clinical studies in regard to the areas summarized in Table 1.
(2) System calibration: measure and adjust system parameters to achieve desired performance. The measurements can be made in SI units to enable universal comparisons (see process 3). Calibration of FMI cameras in SI units of radiance has been reported[23].
(3) Referencing: register the operational characteristics of different cameras so that different systems can be compared with each other.
(4) Data consistency: register the parameters of Table 1 so that data produced by different systems can be referenced to one another or 'converted' to one standard.
(5) Quality control: ensure optimal operation of an imaging system before a study and over time.
(6) Algorithmic validation: imaging standards can also be employed for examining the performance of algorithms implemented in a system for improving an aspect of the system or data collected (see the next section on reversion).
(7) Absolute quantification: phantoms can assist studies requiring absolute quantification of the fluorescence concentration reported on an image, by providing reference signals of known fluorochromes or other optical property values.

Standardization may play a central role in the regulatory approval of clinical data and studies, and it may possibly provide a referencing scheme between clinical outcomes. For example, FMI characterization using an FIS could accompany the report of a clinical study submitted for approval and serve as a reference of detection performance for the particular technology employed in the study. In the absence of such a reference, the results of a study may be difficult to reproduce when using FMI implementations with different performance characteristics from the technology employed in the original study.

We note that standardization could be achieved for individual FMI components. For example, the camera of an FMI system can be accurately calibrated, as per process 2. However, HiFFI standardization refers to the final fluorescence image, which is affected by all parameters in Table 1, not only camera parameters. Moreover, we note that FIS could calibrate systems for absolute fluorochrome concentration (process 7) in concentration units and enhance disease detection based on absolute metrics, that is, on the absolute concentration of a fluorescent agent in a lesion. However, FMI calibration in absolute values is generally meaningful only when accurate reversion has been implemented (see the next section on reversion). Otherwise, variation of the parameters in Table 1 can modify the intensity and appearance of a fluorescence image, independently of the 'absolute' system calibration performed. For this reason, current practices detect disease based on relative metrics, for example, by assessing the intensity of a lesion over background signals (for example, Fig. 2g), over time or by comparing spectra from different tissue areas using ratiometric methods[21,24–26].

Reversion

FMI system parameters and their consistency over time can be measured using imaging standards. However, HiFFI requires accounting for the effect of the parameters in Table 1. Reversion is introduced herein as a set of HiFFI processes that can first capture and then correct fluorescence images for the effects of the parameters in Table 1, leading to superior representation of the underlying fluorochrome biodistribution than in raw images.

The basis of reversion is the assumption that there is a given fluorochrome distribution in tissue, which is a first 'true' state. The aim of FMI is to record this true fluorochrome state. However, this first state is converted through the effect of the parameters in Table 1 to a second state actually captured by FMI. Reversion refers to returning from this second state to the first state through various measures, some of which are summarized in Table 1 under the column 'Remedy'.

Reverting the effect of invariable parameters is relatively straightforward. For example, cross-talk between the intrinsic channel and fluorescence channel can be implemented by multiplying an excitation image by a coefficient, then subtracting the result from the fluorescence image. The coefficient can be determined through cross-talk measurements on a calibration phantom. Likewise, the effects of ambient light may be characterized using calibration measurements and subsequently controlled with hardware and data processing improvements[27,28]. Overall, FMI calibration can determine a number of invariable, system-specific parameters that can be

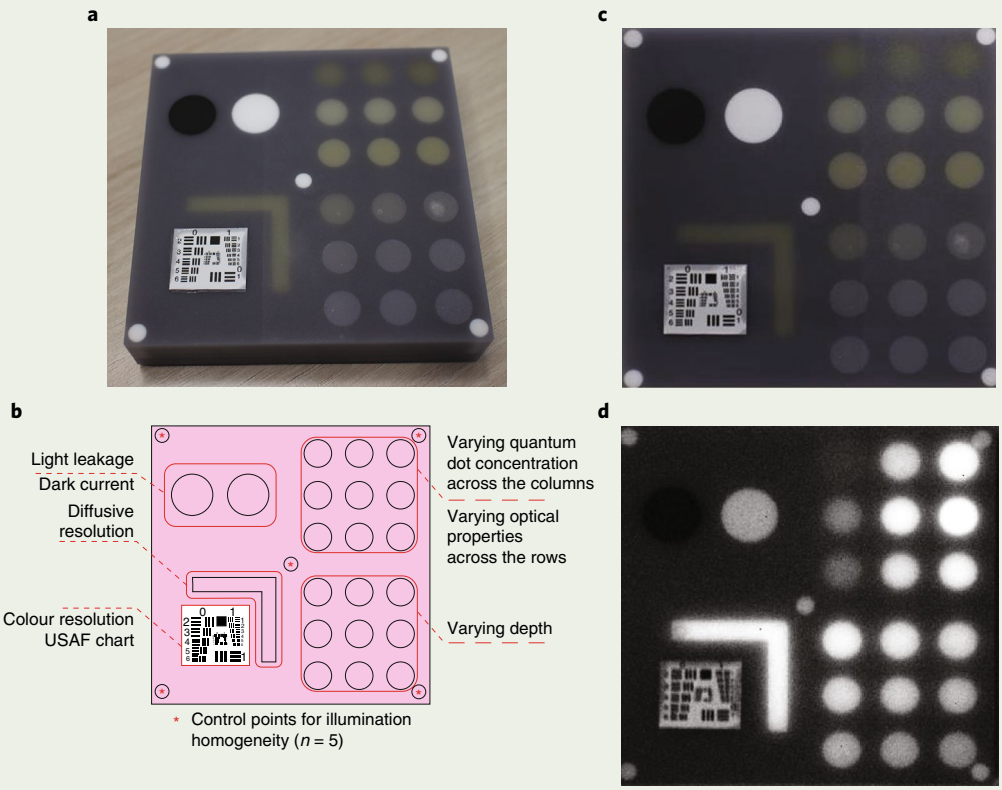
Box 1 | Phantoms

Imaging phantoms, that is, objects that contain features simulating tissue features with predetermined shapes and contrast, are common in radiological imaging for evaluating, confirming or measuring the performance of a device or method. Phantoms are generally employed in medical imaging for quality control and quality assurance. For optical imaging standardization, phantoms should meet three basic requirements: (1) allow the implementation of desired optical properties (absorption, scattering, fluorescence); (2) provide long-term photostability in diverse environmental conditions; and (3) assume a fixed shape (termed a 'solid phantom') that suffers no mechanical deformation over time.

To meet these requirements, optical phantoms typically use epoxy, polyester resin or polyurethane as base material. When cured, these materials can be machined into different shapes and volumes. Optical properties can be flexibly manipulated by adding absorbers, scatterers and fluorophores before curing. Typically, absorption properties are based on the addition of absorbing dyes (for example, India ink, nigrosin) and scattering properties are based on the addition of TiO_2 particles into the base material. Homogeneous distribution can be ensured with sonication. Quantum dots are typically preferred for fluorescence emission. Unlike organic fluorophores, which suffer from fast photobleaching, quantum dots provide better photostability, which is required in applications involving long-term imaging[40–43].

Several phantoms have been considered for fluorescence imaging studies. A typical parameter addressed is sensitivity, using different titrations of a known fluorochrome[44]. Other parameters have been considered, such as the effects of depth[45] or the cross-talk between excitation and emission channels[44]. However, a limitation of current phantoms is that they address only one or a few of the specifications in Table 1[44,46,47].

HiFFI standardization and reversion require the measurement and calibration of a large number of parameters (Table 1). Fortunately, modern cameras used for FMI contain more than 10^6 pixels, and the high pixel density means that each photograph may contain more than one million independent measurements. The actual number of independent features that an FMI camera captures is probably less given the presence of diffusion, the need for a minimum signal-to-noise ratio and other practical considerations; nevertheless, this number of features is sufficient to retrieve all invariable and variable parameters in Table 1. Consequently, we have recently introduced the concept of composite phantoms, that



Composite phantom for FMI standardization in the visible range. a, Colour photograph of the phantom, which has outer dimensions of 10 × 10 × 2.2 cm^3. **b**, Schematic explaining the function of the different parts of the phantom. 'USAF chart' refers to the 1951 US Air Force (USAF) resolution test chart. **c**, Reflectance image of the phantom obtained with the colour camera of a hybrid FMI system. **d**, Fluorescence image of the phantom. Background absorption was set to ~2.2 cm^{-1} by adding nigrosin dissolved in alcohol to the base material, while the reduced scattering coefficient was set to ~10 cm^{-1} by adding 1 mg g^{-1} TiO_2 particles (titanium(IV) oxide). The upper right quadrant of the phantom contains an array of nine fluorescent wells (10 mm diameter) that interrogates the sensitivity and fluorescence intensity variation as a function of optical properties. The wells contain a mix of cured polyurethane with organic quantum dots at varying concentrations (1, 5 and 10 nM) across the columns and varying hemin concentration (20, 20 and 40 μg g^{-1}) and TiO_2 amount (0.33, 0.66 and 1 mg g^{-1}) across the rows. The bottom right quadrant contains nine fluorescent wells embedded within the phantom at increasing depths below the phantom surface (0.2, 0.4, 0.6, 0.8, 1.0, 1.33, 1.66, 2.0 and 3.0 mm). The upper left quadrant examines camera dark-current offset and camera cross-talk, that is, excitation light leakage into the fluorescence channel. The lower left quadrant assesses the resolution of the fluorescence and visible images. Five identical reflective circular areas (5 mm diameter), made of 10 mg g^{-1} titanium oxide in polyurethane, sample the homogeneity of the light illumination employed by the camera system. Four circular areas lie at the corners of the phantom, and one area lies in the centre.

Box 1 | Phantoms (continued)

is, phantoms that enable the characterization of the parameters in Table 1 in a single snapshot (photograph). A first composite phantom (see Box figure) was constructed out of polyurethane and employed multiple targets, each exploring a different camera performance parameter. The construction of the phantom and its application have been described in detail elsewhere[48,49]. In brief, the background material and the different targets introduced are made of different mixtures of TiO_2 particles, photon-absorbing dyes and quantum dots. The phantom uses nigrosin as a generic absorber with a flat absorption spectrum, as well as hemin, an iron-containing porphyrin with a haemoglobin-like absorption spectrum that simulates the absorption of blood. This composite phantom enables the implementation of multiple features, including the measurement of sensitivity, cross-talk, illumination homogeneity, dark current, resolution, and the effects of depth and optical properties (see Box figure caption).

Composite phantoms could be employed as FIS for different standardization operations, as discussed in the main text, and for examining the performance of reversion methods. Preferably, phantoms should be relatively straightforward and inexpensive to manufacture. An ideal phantom should balance imaging features against manufacturing complexity. It is therefore possible that simpler or more complex features could be introduced, depending on the particular application. The preliminary phantom in the Box figure is not a comprehensive phantom for all parameters in Table 1; for example, it does not contain features for spectral measurements or calibrations. Next-generation phantoms can contain a larger number of wells that more accurately report on camera performance over a wider dynamic and spectral range. 3D printing techniques may significantly simplify manufacturing of complex phantoms for FMI standardization.

and/or to improve the hardware of the imaging system (for example, by improving ambient light or excitation light filters).

Reversion of variable parameters, in contrast, is more complex and typically requires dedicated hardware that collects additional information on the environment and tissue measured. In this case, the reversion should be adaptive, that is, responsive to the particulars of each measurement, not of each system. Implementation of variable parameter reversion also generally requires data-processing algorithms to process the measurements collected. The following section introduces original analysis, measurements and simulations that are necessary for explaining the complexity and the solutions that may be appropriate in the reversion of variable parameters.

Effects of tissue optical properties and uniqueness. It is possible that FMI records a fluorescence image of spatially varying intensity from a homogeneous fluorochrome distribution in tissue, if the tissue optical properties (or other parameters in Table 1) vary in a spatially related manner (Fig. 3a,b). Fluorescence images of different phantoms containing the same amount of ICG but varying amounts of intralipid (which scatters) and blood (which absorbs), as shown in Fig. 3a, exhibit a markedly different fluorescence intensity, which generally decreases with the addition of absorption and increases with the addition of scatter (Fig. 3b). This experiment quantitatively establishes that optical properties (absorption and scatter) modify the fluorescence intensity recorded. Even though the fluorescence measurements in Fig. 3a,b were performed in the NIR region, similar behaviour is expected in the extended NIR, NIR-II and NIR-III regions, where in this case the response is to variations in tissue fat and water concentration, which are the dominant tissue absorbers in the spectral region >950 nm.

Reversion of the effects of optical properties is therefore necessary for HiFFI but not possible when only fluorescence images are recorded. Similarly, the addition of colour images does not offer a straightforward way to correct for the effect of optical properties on fluorescence measurements. The intensity of the photon wave reflected from tissue illuminated with a plane wave, for example, signals collected in photographic mode (reflectance; Fig. 3e,f), depends on the ratio of the absorption coefficient to the reduced scattering coefficient (μ_a/μ_s')[29,30] and cannot independently separate absorption from scatter. We have further shown that the dependence of reflected intensity on μ_a/μ_s' also defines a non-uniqueness condition for spectral measurements: as long as the ratio of absorption to scatter remains constant, the spectrum collected from tissue shows the same intensity regardless of the levels of scatter and absorption[30] (Fig. 3g).

to reflectance measurements: in contrast to photographic (back-reflection) measurements (Fig. 3g), fluorescence intensity depends monotonically on the μ_a/μ_s' ratio (Fig. 3h), due to the differential effects of absorption and scatter shown in Fig. 3c,d. This leads to an apparent non-uniqueness paradox: it is possible that a fluorescence camera will record an image of spatially varying fluorescence intensity from a uniform fluorochrome distribution, even if the reflectance image shows constant intensity. This is possible when the region containing the uniform fluorochrome distribution has spatially variant optical properties but a spatially constant pixel-wise μ_a/μ_s' ratio.

Tissue autofluorescence may also alter the appearance of an FMI image (Fig. 2g). Autofluorescence is generated by intrinsic tissue molecules that fluoresce naturally. Images collected from tissue may have fluorescence contributions from the targeted agent administered as well as from intrinsic tissue fluorochromes, complicating specific detection. The effect of tissue optical properties on autofluorescence may further complicate the detection process, as the autofluorescence signals collected will also be modified by the background optical properties. Furthermore, the fluorescence intensity recorded depends nonlinearly on the depth of fluorochrome distribution in tissue. Therefore depth variations pose a challenge to high-fidelity imaging.

Fluorescence reversion for optical properties. The findings in the previous section suggest that fluorescence intensity cannot be quantified unless the effects of variable parameters are either accounted for or cancelled (minimized) through data-processing operations, that is, the process of reversion. We expect that, in the future, significant attention will shift from hardware calibration to reversion methods, that is, measurements and data-processing approaches that account for the effects of variable parameters on the fluorescence image. Phantoms such as the one described in Box 1 can play a key role in developing and comparing the accuracy and relative performance of different methods developed.

The importance of reversion is recognized in the literature, with a large range of methods suggested to improve fluorescence quantification[19]. Reversion generally requires the independent measurement of optical properties and other variable parameters and the consequent processing of fluorescence signals to move towards HiFFI. However, due to complications in accurately measuring the optical properties with planar-wave (wide-field) illumination[31,32], a more commonly applied approach considers ratiometric measurements for eliminating or minimizing the effects of variable parameters on fluorescence intensity. Effects of absorption variations can be compensated for by normalizing the fluorescence

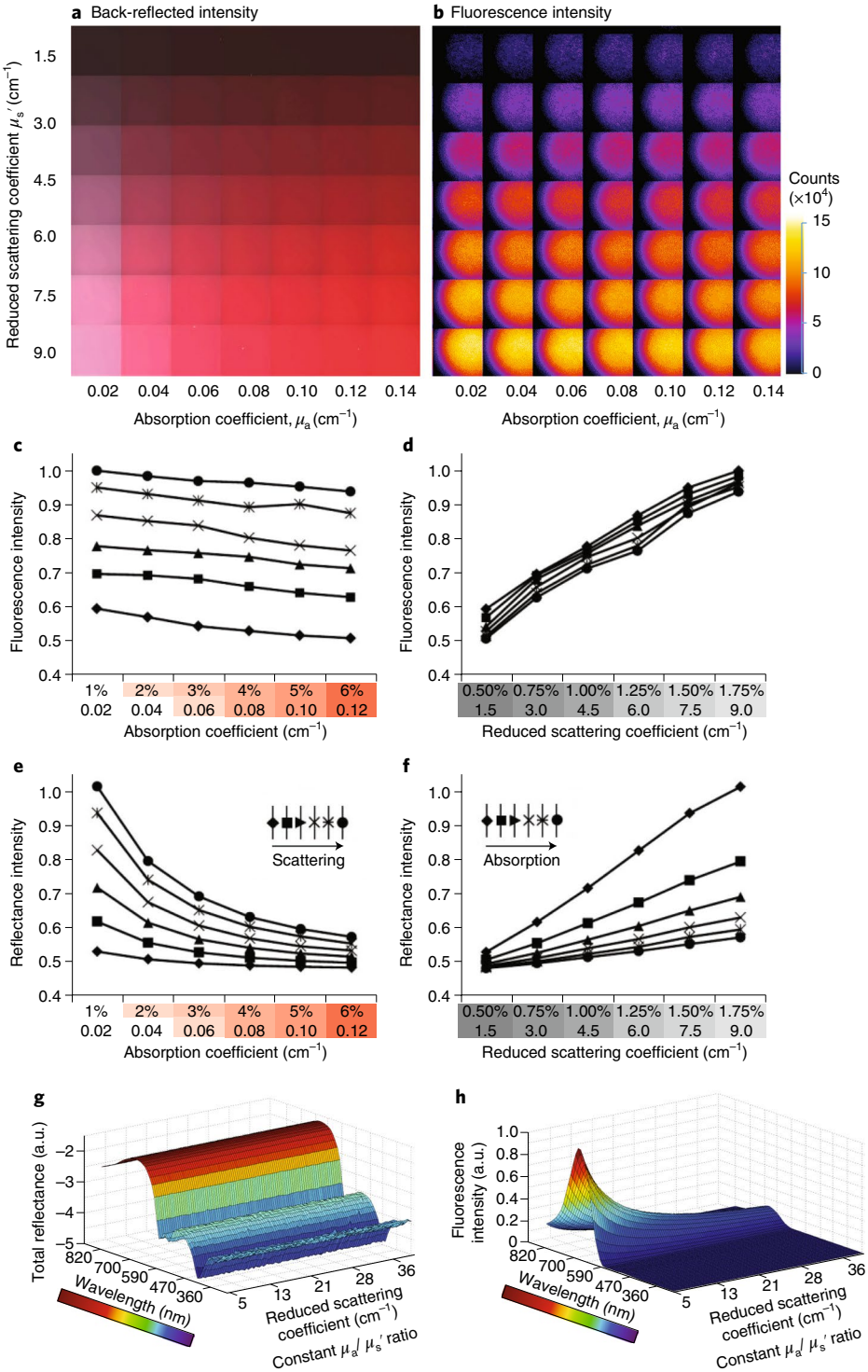


Fig. 3 | Effects of tissue optical properties on the fluorescence (and back-reflected) intensity. a, Phantom consisting of 7 × 7 wells showing different levels of absorption (*x* axis) and scatter (*y* axis). All wells contain 1 μM Alexa 750 fluorochrome. **b**, Fluorescence intensity collected from wells shown in **a**. Signal variation across the wells is due to inhomogeneous illumination. **c**, Fluorescence intensity recorded from the wells in **b** as a function of absorption coefficient. The symbols are defined in **e**. **d**, Fluorescence intensity recorded from the wells in **b** as a function of reduced scattering coefficient. The symbols are defined in **f**. **e**, Reflectance intensity recorded from the wells in **b** as a function of absorption coefficient. Symbols are organized with scatter increasing towards the right. **f**, Reflectance intensity recorded from the wells in **b** as a function of reduced scattering coefficient. Symbols are organized with absorption increasing towards the right. **g**, Reflectance intensity spectra generated by Monte Carlo simulations as scattering and absorption coefficients increase but their ratio remains constant. **h**, Fluorescence intensity spectra generated by Monte Carlo simulations as scattering and absorption coefficients increase but their ratio remains constant. Measurements in **g** and **h** have been experimentally confirmed (data not shown).

image with a back-scattered (reflection) image, provided that tissue scattering does not vary significantly. The back-scattered image can be obtained at the emission wavelength[24,26], or at the excitation wavelength when a fluorochrome with small Stokes shift is assumed[33]. Ratios that utilize at least two fluorescence images obtained in different spectral windows have also been proposed for

minimizing the effects of autofluorescence and scatter variation on the fluorescence image of interest[21,34]. In such cases, fluorochromes with a wide emission spectrum are preferred, such as protoporphyrin IX, so that fluorescence images collected in different spectral windows can be captured with sufficient signal-to-noise ratio. Spectroscopy methods such as single-point measurements[35] and spectral imaging have also been considered for improving quantification and reducing the effects of autofluorescence due to intrinsic tissue fluorochromes such as flavin adenine dinucleotide (FAD) and collagen[36].

We note that fluorescence ratios have also been employed in combination with fluorescent agents emitting in different spectral bands, further contributing to the concept of HiFFI and minimizing the effects of (some) parameters in Table 1. In a study based on a mouse model of pulmonary inflammation, animals were injected with a mixture of a protease-sensitive fluorescent agent that emits at 780 nm and a fluorescent agent with similar biodistribution and without tissue specificity that emits at 695 nm. The normalization of the 780 nm image with the 695 nm image minimized the effect of agent distribution and depth, and allowed for accurate quantification of biological activity (inflammation)[37]. The same principle was also applied for improving the quantification of epidermal growth factor receptor expression in a mouse model of metastatic breast cancer[25].

Conclusion

Several parameters in Table 1 affect FMI performance. Moreover, there are many different camera technologies, ranging from low-end consumer cameras to high-end scientific cameras, as well as different data-processing methods, all of which play a critical role in the clinical performance achieved. It, therefore, becomes necessary to create standards and specifications of FMI performance that take into account and possibly regulate the multitude of experimental settings. Composite phantoms can be used for different aspects of FMI standardization, including system calibration as well as quality control and quality assurance processes. Such phantoms can serve as the basis for a FIS. FIS measurements based on composite phantoms (Box 1) can simplify the quality-control process by recording multiple invariable parameters and provide a reference for system performance in terms of variable parameters. FIS measurements at different time points can be employed to ensure longitudinal system operation in accord with system specifications, allowing the detection of changes in system characteristics due to component failure or other reasons.

Accounting for variable parameters is a complex process as it requires real-time measurement of the experimental conditions and fast processing methods to revert (correct) raw fluorescence images to a high-fidelity state. Measurements of tissue optical properties, autofluorescence, lesion depth and camera–tissue distance are technically challenging and may require additional hardware and methodological solutions. By establishing gold standards in the form of FIS, one can evaluate FMI performance and benchmark different reversion methods. Data on FMI reproduction accuracy can lead to quantitative assessment of HiFFI performance. It may also be possible to devise a HiFFI score that can provide a quick reference of HiFFI performance, similarly to performance scores established for other complex systems such as computers. We note that FIS measurements can be used not only for fluorescence but also for white-light (colour) imaging standardization, spectral calibration or calibration of different illumination and detection techniques, for example time-domain systems using light pulses for time-resolved measurements, or light of modulated intensity[38–40].

New classes of targeted imaging agents, such as labelled antibodies, antibody fragments, peptides, mini/nanobodies or activatable agents are increasingly considered in phase 0 and I clinical trials. However, clinical translation of such agents through phase II and III clinical trials imposes a new set of considerations in regard to the regulatory process. In the absence of HiFFI and an accepted FIS, a common approach to conducting an imaging study for regulatory approval is to bundle a camera and an agent together, leading to the approval of the combination system and drug. Statistical outcomes of clinical trials in this case are not representative of the agent alone but of the combined performance of the agent and the system. Standardization with a FIS can play a pivotal role in disengaging the imaging system from the agent imaged. Even when the system and agent are to be evaluated together, a FIS can allow quality control and assurance, by measuring the long-term conformity and performance of systems and agents employed in clinical trials. By registering and calibrating operational parameters, it may become possible to compare data collected from different systems through appropriate post-processing. Overall, standardization can increase the quality of FMI data collection and ease the regulatory approval process, leading to accurate outcomes.

Received: 31 October 2016; Accepted: 3 July 2018;
Published online: 20 August 2018

References

1. Koch, M. & Ntziachristos, V. Advancing surgical vision with fluorescence imaging. *Annu. Rev. Med.* **67**, 153–164 (2016).
2. Zhang, R. R. et al. Beyond the margins: real-time detection of cancer using targeted fluorophores. *Nat. Rev. Clin. Oncol.* **14**, 347–364 (2017).
3. van Dam, G. M. et al. Intraoperative tumor-specific fluorescence imaging in ovarian cancer by folate receptor-alpha targeting: first in-human results. *Nat. Med.* **17**, 1315–1319 (2011).
4. Tummers, Q. R. J. G., Hoogstins, C. E. S. & Gaarenstroom, K. N. Intraoperative imaging of folate receptor alpha positive ovarian and breast cancer using the tumor specific agent EC17. *Oncotarget* **7**, 32144–32155 (2016).
5. Rosenthal, E. L. et al. Safety and tumor specificity of cetuximab-IRDye800 for surgical navigation in head and neck cancer. *Clin. Cancer Res.* **21**, 3658–3666 (2015).
6. Whitley, M. J. et al. A mouse-human phase 1 co-clinical trial of a protease-activated fluorescent probe for imaging cancer. *Sci. Transl. Med.* **8**, 320ra4 (2016).
7. Hsiung, P.-L. et al. Detection of colonic dysplasia in vivo using a targeted heptapeptide and confocal microendoscopy. *Nat. Med.* **14**, 454–458 (2008).
8. Atreya, R. et al. In vivo imaging using fluorescent antibodies to tumor necrosis factor predicts therapeutic response in Crohn's disease. *Nat. Med.* **20**, 313–318 (2014).
9. Glatz, J., Symvoulidis, P., Garcia-Allende, P. B. & Ntziachristos, V. Robust overlay schemes for the fusion of fluorescence and color channels in biological imaging. *J. Biomed. Opt.* **19**, 040501 (2014).
10. Glatz, J. et al. Concurrent video-rate color and near-infrared fluorescence laparoscopy. *J. Biomed. Opt.* **18**, 101302 (2013).
11. Hong, G. et al. Through-skull fluorescence imaging of the brain in a new near-infrared window. *Nat. Photon.* **8**, 723–730 (2014).
12. Hong, G. et al. Multifunctional in vivo vascular imaging using near-infrared II fluorescence. *Nat. Med.* **18**, 1841–1846 (2012).
13. Ghosh, D. et al. Deep, noninvasive imaging and surgical guidance of submillimeter tumors using targeted M13-stabilized single-walled carbon nanotubes. *Proc. Natl Acad. Sci. USA* **111**, 13948–13953 (2014).
14. Carr, J. A. et al. Shortwave infrared fluorescence imaging with the clinically approved near-infrared dye indocyanine green. *Proc. Natl Acad. Sci. USA* **115**, 4465–4470 (2018).
15. DSouza, A. V., Lin, H., Henderson, E. R., Samkoe, K. S. & Pogue, B. W. Review of fluorescence guided surgery systems: identification of key performance capabilities beyond indocyanine green imaging. *J. Biomed. Opt.* **21**, 080901 (2016).
16. Scheuer, W., van Dam, G. M., Dobosz, M., Schwaiger, M. & Ntziachristos, V. Drug-based optical agents: infiltrating clinics at lower risk. *Sci. Transl. Med.* **4**, 134ps11 (2012).
17. Marshall, M. V. et al. Near-infrared fluorescence imaging in humans with indocyanine green: a review and update. *Open Surg. Oncol. J.* **2**, 12–25 (2012).
18. US National Library of Medicine. VEGF-targeted fluorescent tracer imaging in breast cancer. ClinicalTrials.gov https://clinicaltrials.gov/ct2/show/NCT01508572 (2012).
19. Bradley, R. S. & Thorniley, M. S. A review of attenuation correction techniques for tissue fluorescence. *J. R. Soc. Interface* **3**, 1–13 (2006).
20. Themelis, G., Yoo, J. S., Soh, K.-S., Schulz, R. & Ntziachristos, V. Real-time intraoperative fluorescence imaging system using light-absorption correction. *J. Biomed. Opt.* **14**, 064012 (2014).

21. Moriyama, E. H., Kim, A., Bogaards, A., Lilge, L. & Wilson, B. C. A ratiometric fluorescence imaging system for surgical guidance. *Adv. Opt. Technol.* **2008**, 532368 (2008).
22. DeWerd, L. A. & Kissick, M. *The Phantoms of Medical and Health Physics* (Springer, New York, NY, 2014).
23. Zhu, B., Rasmussen, J. C., Litorja, M. & Sevick-Muraca, E. M. Determining the performance of fluorescence molecular imaging devices using traceable working standards with SI units of radiance. *IEEE Trans. Med. Imaging* **35**, 802–811 (2016).
24. Ntziachristos, V. et al. Planar fluorescence imaging using normalized data. *J. Biomed. Opt.* **10**, 064007 (2005).
25. Tichauer, K. M. et al. Microscopic lymph node tumor burden quantified by macroscopic dual-tracer molecular imaging. *Nat. Med.* **20**, 1348–1353 (2014).
26. Valdés, P. A. et al. Quantitative, spectrally-resolved intraoperative fluorescence imaging. *Sci. Rep.* **2**, 798 (2012).
27. Zhu, B., Rasmussen, J. C. & Sevick-Muraca, E. M. Non-invasive fluorescence imaging under ambient light conditions using a modulated ICCD and laser diode. *Biomed. Opt. Express* **5**, 562–572 (2014).
28. Sexton, K. et al. Pulsed-light imaging for fluorescence guided surgery under normal room lighting. *Opt. Lett.* **38**, 3249–3252 (2013).
29. Zonios, G. & Dimou, A. Modeling diffuse reflectance from semi-infinite turbid media: application to the study of skin optical properties. *Opt. Express* **14**, 8661–8674 (2006).
30. Garcia-Allende, P. B. et al. Uniqueness in multispectral constant-wave epi-illumination imaging. *Opt. Lett.* **41**, 3098–3101 (2016).
31. Saager, R. B., Cuccia, D. J., Saggese, S., Kelly, K. M. & Durkin, A. J. Quantitative fluorescence imaging of protoporphyrin IX through determination of tissue optical properties in the spatial frequency domain. *J. Biomed. Opt.* **16**, 126013 (2011).
32. Yang, B. & Tunnell, J. W. Real-time absorption reduced surface fluorescence imaging. *J. Biomed. Opt.* **19**, 090505 (2014).
33. Kanick, S. C. & Pogue, B. W. Why reflectance is an imperfect basis for the correction of fluorescence distortion due to optical properties. In *Biomedical Optics 2014* BS3A.35 (OSA, 2014).
34. Bogaards, A., Sterenborg, H. J. C. M. & Wilson, B. C. In vivo quantification of fluorescent molecular markers in real-time: a review to evaluate the performance of five existing methods. *Photodiagnosis Photodyn. Ther.* **4**, 170–178 (2007).
35. Yang, V. X. D., Muller, P. J., Herman, P. & Wilson, B. C. A multispectral fluorescence imaging system: design and initial clinical tests in intra-operative photofrin-photodynamic therapy of brain tumors. *Lasers Surg. Med.* **32**, 224–232 (2003).
36. Minamikawa, T. et al. Simplified and optimized multispectral imaging for 5-ALA-based fluorescence diagnosis of malignant lesions. *Sci. Rep.* **6**, 25530 (2016).
37. Haller, J. et al. Visualization of pulmonary inflammation using noninvasive fluorescence molecular imaging. *J. Appl. Physiol.* **104**, 795–802 (2008).
38. Godavarty, A. et al. Fluorescence-enhanced optical imaging in large tissue volumes using a gain-modulated ICCD camera. *Phys. Med. Biol.* **48**, 1701–1720 (2003).
39. Vervandier, J. & Gioux, S. Single snapshot imaging of optical properties. *Biomed. Opt. Express* **4**, 2938–2944 (2013).
40. Niedre, M. J. et al. Early photon tomography allows fluorescence detection of lung carcinomas and disease progression in mice in vivo. *Proc. Natl Acad. Sci. USA* **105**, 19126–19131 (2008).
41. Resch-Genger, U., Grabolle, M., Cavaliere-Jaricot, S., Nitschke, R. & Nann, T. Quantum dots versus organic dyes as fluorescent labels. *Nat. Methods* **5**, 763–775 (2008).
42. Gao, X. et al. In vivo molecular and cellular imaging with quantum dots. *Curr. Opin. Biotechnol.* **16**, 63–72 (2005).
43. Alivisatos, A. P., Gu, W. & Larabell, C. Quantum dots as cellular probes. *Annu. Rev. Biomed. Eng.* **7**, 55–76 (2005).
44. Zhu, B., Tan, I.-C., Rasmussen, J. C. & Sevick-Muraca, E. M. Validating the sensitivity and performance of near-infrared fluorescence imaging and tomography devices using a novel solid phantom and measurement approach. *Technol. Cancer Res. Treat.* **11**, 95–104 (2012).
45. Roy, M., Kim, A., Dadani, F. & Wilson, B. C. Homogenized tissue phantoms for quantitative evaluation of subsurface fluorescence contrast. *J. Biomed. Opt.* **16**, 016013 (2011).
46. Zhu, B., Rasmussen, J. C. & Sevick-Muraca, E. M. A matter of collection and detection for intraoperative and noninvasive near-infrared fluorescence molecular imaging: to see or not to see? *Med. Phys.* **41**, 022105 (2014).
47. Moffitt, T., Chen, Y.-C. & Prahl, S. A. Preparation and characterization of polyurethane optical phantoms. *J. Biomed. Opt.* **11**, 041103 (2006).
48. Anastasopoulou, M. et al. Comprehensive phantom for interventional fluorescence molecular imaging. *J. Biomed. Opt.* **21**, 091309 (2016).
49. Gorpas, D., Koch, M., Anastasopoulou, M., Klemm, U. & Ntziachristos, V. Benchmarking of fluorescence cameras through the use of a composite phantom. *J. Biomed. Opt.* **22**, 016009 (2017).

Acknowledgements

We thank A. Ghazaryan for help with optoacoustic measurements. V.N. acknowledges funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 687866 (INNODERM) and from the Deutsche Forschungsgemeinschaft (DFG), Germany (Gottfried Wilhelm Leibnitz Prize 2013; NT 3/10-1).

Competing interests

After the completion of the manuscript, but before the final submission, Bracco Imaging Deutschland, a company commercializing targeted fluorescence imaging, employed M.K.

Additional information

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence should be addressed to V.N.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.