FULL ARTICLE

Environment-specific spectral modeling: A new tool for the analysis of biological specimens

Giovanni Valbusa[1]* 🔟 | Martina Capozza[2] 🔟 | Chiara Brioschi[3] | Francesco Blasi[3] |

Simona Ghiani[3] | Alessandro Maiocchi[3]

[1]Ephoran Multi-imaging Solutions s.r.l, Turin, Italy

[2]Department of Molecular Biotechnologies and Health Sciences & Molecular Imaging Center, University of Torino, Turin, Italy

[3]Bracco Imaging Spa, Milan, Italy

*Correspondence
Giovanni Valbusa, Ephoran Multi-imaging Solutions s.r.l, Via Ribes 5, Colleretto Giacosa, Italy.
Email: giovanni.valbusa@bracco.com

The recent discovery of fluorescent dyes for improving pathologic tissues identification has highlighted the need of robust methods for performance validation especially in the field of fluorescence-guided surgery. Optical imaging of excised tissue samples is the reference tool to validate the association between dyes localization and

the underlying histology in a controlled environment. Spectral unmixing may improve the validation process discriminating dye from endogenous signal. Here, an innovative spectral modeling approach that weights the spectral shifts associated with changes in chemical environment is described. The method is robust against spectral shift variations and its application leads to unbiased spectral weights estimates as demonstrated by numerical simulations. Finally, spectral shifts values computed pixel-wise from spectral images are used to display additional information with potential diagnostic value.
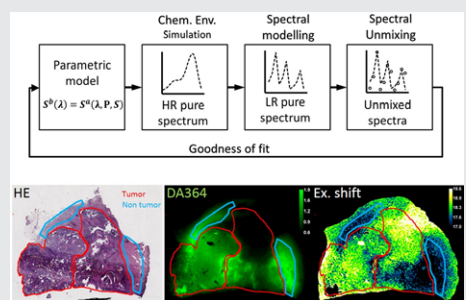
KEYWORDS
Cy5.5, fluorescence-guided surgery, formalin-fixed paraffin tissues, spectral unmixing, targeted contrast agent

1 | INTRODUCTION

Fluorescence-guided surgery (FGS) is an optical-based technique with the objective of identifying pathologic tissues during surgical procedures leading to a complete, "en-bloc," eradication of the disease, possibly reducing recurrences [1]. The objective of current FGS clinical trials is the evaluation of fluorescent contrast agents' performances in terms of tumor detection sensitivity and specificity. Dyes specificity for pathologic tissues is evaluated post-surgery by ex vivo optical imaging experiments on bread-loaf fresh tissue slices and on formalin-fixed paraffin embedded (FFPE) tissue

samples [2, 3]. Fluorescence intensity is compared with microscopy analyses on thin (5-10 μm) tissue sections, such as histopathology and tumor biomarkers immunostaining, for a final validation vs a gold standard [2, 3]. A second generation of agents [4, 5], engineered to overcome the low specificity for tumor tissues reported for clinically approved fluorescent agents [4, 6–9] (fluorescein, indocyanine green and methylene blue), requires robust methods to validate their performance. Fluorescence images of FFPE blocks, in particular, can provide precious information if associated to histological analyses, because tissue sections are sliced from the blocks with a geometry-conserving procedure that allows consistent images overlapping. However, the preparation

Giovanni Valbusa and Martina Capozza contributed equally to this work.

procedure of FFPE blocks and their final optical characteristics have effects on fluorophores spectral shape that can potentially bias the final results.

Spectral unmixing can discriminate fluorescence signals originating from exogenous (ie, contrast agents) and from endogenous (ie, autofluorescence, AF) sources. Spectral unmixing is based on the analysis of spectral fluorescence images in which each pixel correspond to a fluorescence spectrum. The pixel-wise contribution of each dye can be calculated by ordinary least squares under the assumption that individual dyes spectra sum linearly to the final spectrum and when individual dyes spectra are known [10]. However, measuring individual spectra (also referred as pure spectra) can be suboptimal, indeed, fluorescence is affected by samples chemical environment (eg, polarity, viscosity, pH) and by specific interactions of the dye with the biological matrix (ie, binding to albumin). Blind methods such as principal component analysis [11], nonnegative matrix factorization [12] and multivariate curve resolution [13, 14] have been proposed to overcome the need of measuring pure spectra. These methods provide as output both a list of calculated pure spectra and dyes contributions for each pixel of the spectral image. Recently, phasor analysis has been proposed for the real-time unmixing of spectral fluorescence images of biological specimens without requiring the knowledge of pure spectra [15]. This method uses the Fourier transform to provide images segmentation based on pixels with similar spectral properties [16].

Although the aforementioned approaches are effective in calculating dyes contributions to fluorescence intensity, they do not weight on spectral variations related to chemical environment usually affecting histopathology samples and other biological specimens. The analysis of chemical environment-related effects on fluorescence spectra is normally carried out using spectrofluorometers with high spectral resolutions (about 1 nm) and dedicated in vitro experiments on fluorophores solutions with variable physicochemical properties. Optical imaging devices typically used for the analysis of tissue samples, instead, have limited spectral resolution (about 20-30 nm). Additionally, the complex chemical environment of tissue samples can not be reproduced to provide reference pure spectra. Here, we propose an innovative method to correct biases in pure spectra shape by means of environment-specific spectral modeling. Moreover, the method is able to extract chemical environment information with potential diagnostic value. The method is applied to tissues obtained from a FGS study carried out on oncologic canine patients that received a tumor-specific fluorescent contrast agent (study results not published yet).

2 | MATERIALS AND METHODS

2.1 | Chemicals

Human serum albumin (HSA, lyophilized powder ≥97%) and high purity glycerol (GLY) were purchased from Sigma-Aldrich, Steinheim, Germany. The cyanine Cy5.5-labeled, integrin-targeted imaging probe DA364 (ex/em in water 675/694 nm) was synthesized as previously described [17]. Sterile batches were assessed for chemical identity (liquid chromatography mass spectroscopy), purity (high performance liquid chromatography, area of the peak >90%) and integrin binding efficacy (IC50 < 10 nM) before release.

2.2 | Sample preparation

DA364 solutions were freshly prepared by dilution in phosphate buffer saline (PBS, pH = 7.4) and GLY. DA364 was incubated (1 hour, room temperature) in HSA (final concentration: 500 μM). To minimize reabsorption, solutions with an absorbance lower than 0.1 were obtained [18].

Tissues were obtained from oncologic canine patients with spontaneous tumors administered intravenously 24 hours before surgery with the contrast agent DA364. Tumor and healthy biopsies were fixed in 10% buffered formalin for 24 hours, and then trimmed before being embedded in paraffin following standard protocol. Briefly, samples were immersed first in formol for 50 minutes, followed by serial dehydrations in isopropanol (55-60-75-90 minutes), then clarified in toluene for 3 steps (30-45-60 minutes), and finally embedded in paraffin at 65°C in 3 steps (45-80-90 minutes).

2.3 | Fluorescence measurements

Excitation and emission spectra of DA364 solutions in PBS, GLY and HSA were measured by a high-resolution spectrofluorometer FluoroMax-40 (Horiba Jobin Yvon, New York, NY). The sample was excited with a 150 W Xenon light source and fluorescent signal was detected by TBX-04 detector. To obtain the corresponding low-resolution spectra, the DA364 solutions were sampled 3 times in a black wellplate and imaged by the optical imager IVIS spectrum (Perkin Elmer, Waltham, Massachusetts).

Fluorescence images of the paraffin tissues (field of view 6 cm × 6 cm) were acquired in epi-illumination mode using combinations of excitation and emission filter pairs shown in Table 1. Additionally, reflectance images were acquired using the same excitation bandpass filters used for fluorescence acquisitions, and setting the emission filter to open position to implement the fluorescence correction method described below (see Table 1). Optimal exposure and f-stop values were automatically chosen by the system depending on the current filter choice and samples characteristics. Signal intensities were finally normalized considering optical parameters such as lamp power spectrum, spatial illumination inhomogeneity, zoom, f-stop, exposure, optical geometry and camera efficiency.

TABLE 1 Filter pair list for optical imaging spectral acquisitions on IVIS spectrum

Fluorescence images			Reflectance images		
Image ID	Excitation wavelength (nm)	Emission wavelength (nm)	Image ID	Excitation wavelength (nm)	Emission wavelength (nm)
1	605	740	35	605	Open
2	640	740	36	640	Open
3	675	740	37	675	Open
4	605	760	38	710	Open
5	640	760			
6	675	760			
7	710	760			
8	605	780			
9	640	780			
10	675	780			
11	710	780			

2.4 | Fluorescence correction

The flow of excitation and emitted light in turbid media such as biological tissues depends on its optical properties, namely absorption, μ_a and scattering, μ_s coefficients [19]. Local variations of μ_a and μ_s are hence responsible of fluorescence variations that are not dependent on fluorescence sources concentration. Since the final objective of spectral unmixing is to measure dyes affinity for specific tissue types through the spatial variation of fluorescence signal in the samples, a normalization procedure to correct the effect of μ_a and μ_s was required.

A simpler ratiometric normalization approaches have been previously proposed by Themelis et al. [20] In the cited work, "quantification errors of less than 8%, even at a 5-fold change of absorption variation" were reported. Briefly, the method is based on the use of reflectance images to characterize the optical response of the samples at the excitation wavelength. Despite the amount of reflected light at a specific wavelength has a complex relation with μ_a and μ_s, the normalization procedure is expressed by the following simple mathematical formula:

$$I_{x,y}(\lambda_p) = \frac{F_{x,y}(\lambda_p)}{R_{x,y}(\lambda_p)}, \tag{1}$$

where $F_{x,y}(\lambda_p)$ is the fluorescence signal measured at point (x, y) using the filter pair $\lambda_p = (WL_{ex}, WL_{em})$, WL_{ex} and WL_{em} are the central wavelengths of the excitation and emission filters, respectively, $R_{x,y}(\lambda_p)$ is the reflectance measured at point (x, y) acquired setting the emission filter to open position and $I_{x,y}(\lambda_p)$ is the normalized fluorescence intensity.

2.5 | Spectral unmixing method overview

The proposed method is a variation of linear unmixing, the most commonly used spectral unmixing method [21]. Linear unmixing assumes that the final fluorescence intensity, $I_{x,y}(\lambda)$, is the linear combination of pure spectra, $S_i(\lambda)$. C_i weights the contribution of ith spectrum to $I_{x,y}(\lambda)$ and $E(\lambda)$ is the residual error. λ represents the excitation/emission wavelengths pairs used to acquire the spectral data. Spectral weights C_i, are obtained pixel-wise by multivariate ordinary least squares (OLS, see Eq. (2)):

$$I_{x,y}(\lambda) = \sum_i C_i S_i(\lambda) + E(\lambda). \tag{2}$$

Differently from linear unmixing, the proposed method takes into account spectra dependency on chemical environment introducing additional parameters in the equation (Eq. (3)):

$$I_{x,y}(\lambda) = \sum_i C_i S_i(\lambda, P_{chem}) + E(\lambda), \tag{3}$$

where $S_i(\lambda, P_{chem})$ is a mathematical model of ith spectrum expressed as a function of the chemical environment parameters, P_{chem}. From Eq. (3), the optimal values of C_i and P_{chem} can be obtained by iterative numerical optimization methods by maximizing the quality of OLS fit (eg, by maximizing R^2 or other fitting parameters).

In this work, the dependency of DA364 spectra from chemical environment ($S_i^{high}(\lambda, P_{chem})$) was obtained from high-resolution spectral data obtained from dye solutions with variable chemical environments. To apply the proposed method (named CASUM, chemical aware spectral unmixing) on optical imaging dataset, a procedure to calculate low-resolution $S_i^{low}(\lambda, P_{chem})$ spectra was implemented. Standard linear unmixing is named LINUM in the following paragraphs. The CASUM method is outlined in Figure 1.

2.6 | Chemical environment modeling

Chemical environment is known to influence the shape of fluorescence spectra [22]. In the case of cyanine dyes like Cy5.5, contained in DA364, conformational restrictions in viscous solvents and upon binding to biological molecules are related to bathochromic shifts of fluorescence spectra [23]. Experimental results reported in this work showed that bathochromic shifts are associated with Stokes shift reduction. Minor spectral shape modifications were observed. This result suggested that DA364 fluorescence spectra in an
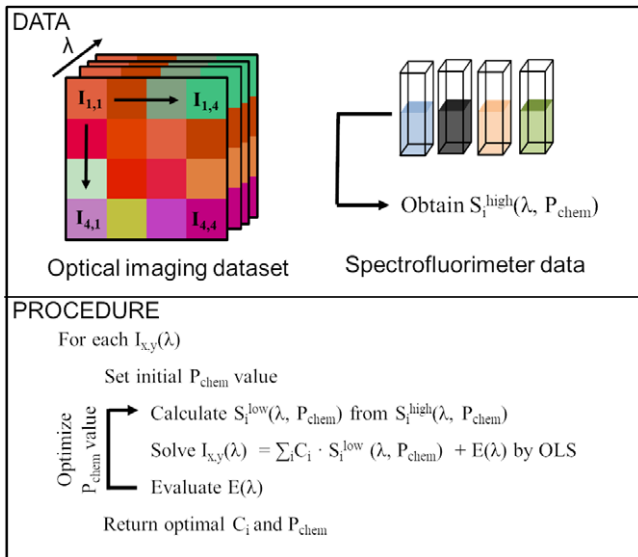
FIGURE 1 Pictorial description of the proposed CASUM (chemical aware spectral unmixing) method

unknown chemical environment can be modeled from the spectrum in a reference chemical environment applying simple spectral shift operations.

For convenience the expression of $S_{\mathrm{DA364}}^{\mathrm{high}}(\lambda, P_{\mathrm{chem}})$ is split for excitation and emission spectra:

$$
S_{\mathrm{DA364}}^{\mathrm{high}}(\lambda, P_{\mathrm{chem}}) = \begin{cases} S_{\mathrm{ex}}^{\mathrm{ref}}(\lambda + \mathrm{PS}) \text{ for excitation spectra}, \\\\ S_{\mathrm{em}}^{\mathrm{ref}}(\lambda + \mathrm{PS*SS}) \text{ for emission spectra}, \end{cases}
$$
(4)

where λ is the wavelength, $S_{\mathrm{ex}}^{\mathrm{ref}}$ and $S_{\mathrm{em}}^{\mathrm{ref}}$ are high-resolution excitation and emission spectra measured in a reference chemical environment, PS is a global peak shift parameter and SS is a multiplicative factor introduced to model Stokes shift reductions or increases. The product PS * SS represents the emission spectra shift.

2.7 | Low-resolution spectra modeling

Optical imaging devices are generally equipped with a set of excitation and emission bandpass optical filters characterized by increasing central wavelengths (WL) and large bandwidths (BW). Filters central wavelengths generally covers the visible spectra up to the near infrared in 20 to 30 nm steps. Spectral bandwidths are in the 20 to 30 nm range. In the description below, optical filters pairs λ_p are modeled as ideal excitation and emission bandpass filters with flat zero-attenuation passband from $\mathrm{WL} - \mathrm{BW}/2$ to $\mathrm{WL} + \mathrm{BW}/2$, and complete attenuation outside the passband. The efficiency of an excitation/emission filter in measuring fluorescence of a dye depends on the value of WL with respect to excitation/emission spectra maxima and on its BW. When high-resolution fluorescence spectra $S_i^{\mathrm{high}}(\lambda, P_{\mathrm{chem}})$ are known, low-resolution $S_i^{\mathrm{low}}(\lambda_p, P_{\mathrm{chem}})$ fluorescence spectra

can be easily calculated by integration. Excitation filter efficiency was calculated as follows:

$$
E_{\mathrm{ex}}(\lambda_p) = \int_{\mathrm{WL}_{\mathrm{ex}} - \mathrm{BW}_{\mathrm{ex}}/2}^{\mathrm{WL}_{\mathrm{ex}} + \mathrm{BW}_{\mathrm{ex}}/2} S_{\mathrm{ex}}^{\mathrm{ref}}(\lambda + \mathrm{PS}) d\lambda.
$$
(5)

Emission filter efficiency was calculated similarly substituting emission filter data as follows:

$$
E_{\mathrm{em}}(\lambda_p) = \int_{\mathrm{WL}_{\mathrm{em}} - \mathrm{BW}_{\mathrm{em}}/2}^{\mathrm{WL}_{\mathrm{em}} + \mathrm{BW}_{\mathrm{em}}/2} S_{\mathrm{em}}^{\mathrm{ref}}(\lambda + \mathrm{PS*SS}) d\lambda.
$$
(6)

The corresponding filter pair efficiency was finally calculated as follows:

$$
E(\lambda_p) = E_{\mathrm{ex}}(\lambda_p) * E_{\mathrm{em}}(\lambda_p).
$$
(7)

Here, the expression of filter pair efficiency was generalized to take into account chemical environment variations by introducing the PS and SS parameters and low-resolution normalized fluorescence spectra were calculated as follows:

$$
S_i^{\mathrm{low}}(\lambda_p, \mathrm{PS}, \mathrm{SS}) = \frac{E(\lambda_p, \mathrm{PS}, \mathrm{SS})}{\max(E(\lambda_p, \mathrm{PS}, \mathrm{SS}))}.
$$
(8)

Finally, for optical imaging datasets, Eq. (3) becomes:

$$
I_{x,y}(\lambda_p) = \sum_i C_i S_i^{\mathrm{low}}(\lambda_p, \mathrm{PS}, \mathrm{SS}) + E(\lambda_p).
$$
(9)

2.8 | Numerical simulations

In order to validate the performance of CASUM, the method was applied on simulated spectra obtained by linear combination of low-resolution AF and DA364 spectra. The DA364/AF ratio (S_{rat}) was varied between 0.1 and 2 to simulate increasing DA364 concentrations. DA364 simulated spectra were calculated from high-resolution DA364 spectrum in PBS using the integration method described above with PS = 0, 10, 20 and 30 nm and SS = 0.5 to resemble the typical values found in FFPE block. Instrumental error was simulated by adding Gaussian noise $N(\mu, \sigma^2)$ with $\mu = 0$ and $\sigma = 5\%$ of the fluorescence intensity. Instrumental noise was determined empirically with 10 consecutive multispectral optical imaging acquisitions of the same tissue block analyzing spectra variability pixel-wise. Simulated spectra were unmixed using CASUM and LINUM to obtain DA364 (C_{DA364}) and AF (C_{AF}) weights. Additionally, PS and SS estimates were obtained by CASUM. Estimation precision and accuracy of the 2 methods obtained under the different simulation conditions were finally compared. Precision was quantified as the percent ratio between SD and mean of the estimates distributions. This parameter is also referred as relative error. Precision represent the scattering of parameters estimates in the presence of noise. Accuracy was defined as the percent difference between the mean of estimates distributions and the true values. Accuracy measure the difference, positive or negative, of estimated parameters values

from the true (reference) values. For both the parameters, values closest to 0 represent better estimation performance.

2.9 | Spectral unmixing experiments

Unmixing of GLY and HSA solutions spectra was carried out using CASUM and LINUM to compare the 2 approaches. LINUM was applied using low-resolution DA364 spectrum measured in PBS as reference, adopting the standard approach that disregards environmental shifts. For CASUM, DA364 spectrum was calculated optimizing PS and SS. For both methods, equations were simplified removing reference to additional fluorescence sources. For CASUM, Eq. (9) was simplified as follows:

$$I(\lambda_p) = C_{\text{DA364}} * S^{\text{low}}(\lambda_p, \text{PS, SS}) + E(\lambda_p). \quad (10)$$

Since normalized spectra were used, the expected value of C_{DA364} was 1 for both the methods. C_{DA364} unmixing fitting quality (R^2) and unmixing errors obtained with the 2 methods were compared. Additionally, high-resolution spectra maxima calculated using CASUM (ie, from $S_{\text{DA364}}^{\text{high}}(\lambda, \text{PS,SS})$) were compared with the corresponding experimental spectra.

Spectral unmixing experiments on FFPE samples were carried out to compare LINUM and CASUM in the presence of secondary fluorescence sources (AF). Pure AF was obtained from noninjected canine FFPE blocks containing adenocarcinoma and muscle tissue. Briefly, FFPE blocks regions characterized by high and low DA364 concentrations (referred as high and low DA364 regions for brevity) have been identified in fluorescence images acquired using the 675/740 filter pair. The 675/740 filter pair is optimal for acquiring DA364 fluorescence and, neglecting AF, high fluorescence intensity is expected to be related to high DA364 content. High and low DA364 regions were identified blindly applying the Otsu thresholding method [24]. Average fluorescence intensity spectra measured in these regions were unmixed using the CASUM and LINUM methods and results were compared.

Finally, average PS and SS values calculated by CASUM from FFPE blocks were compared across different tissue types. To identify muscle, adenocarcinoma and mastocystoma tumors, regions of interest (ROI) were drawn on FFPE blocks by an expert pathologist using reference white light images of the samples. Finally, ROI were overimposed on FFPE blocks fluorescence images and PS and SS values were calculated pixel-wise in these regions by CASUM applying Eq. (11):

$$I_{x,y}(\lambda_p) = C_{x,y}^{\text{probe}} * S^{\text{low}}(\lambda_p, \text{PS}_{x,y}, \text{SS}_{x,y}) + C_{x,y}^{\text{AF}} * \text{AF}(\lambda_p) + E(\lambda_p).$$
$$(11)$$

2.10 | Chemical environment imaging

CASUM was used displayed spectral shifts inhomogeneity related to chemical environment within the samples.

Dedicated software was implemented to map $\text{PS}_{x,y}$ and $\text{SS}_{x,y}$ computed by CASUM to the corresponding x, y locations in parametric image maps. The software was written using the Python language. PS and SS maps were compared against reference hematoxylin and eosin (H&E) stained image where an expert pathologist identified pathologic and healthy tissue.

3 | RESULTS AND DISCUSSION

3.1 | Numerical simulations

The precision of C_{DA364} and C_{AF} estimation depends on the relative DA364/AF weight S_{rat}. C_{DA364} relative error is high for low S_{rat} values and the opposite was found for C_{AF}. On average, for low S_{rat} values (0.1) the relative error of C_{DA364} was $21.5 \pm 2.8\%$ and $21.35 \pm 1.2\%$ for the CASUM and LINUM methods, respectively. For high S_{rat} values ($S_{\text{rat}} = 2$), relative error drops to about $4.8 \pm 0.4\%$ and $3.9 \pm 0.16\%$ for the CASUM and LINUM methods, respectively. Conversely, C_{AF} relative error increase with S_{rat} from about 4% for $S_r = 0.1$, to 6% for $S_r = 2$ for both the methods (see Table S1). In general, CASUM and LINUM showed comparable precision in the estimation of C_{DA364} and C_{AF}. The 2 methods instead, showed differences in terms of accuracy with shifted DA364 spectra (PS = 10, 20, 30 nm and SS = 1, 0.5). On average CASUM provided unbiased C_{DA364} estimates across variations of the PS and SS parameters (accuracy = $0.7 \pm 1.5\%$) while LINUM estimates were underestimated on average of $-5.6 \pm 5.5\%$. Specifically, LINUM C_{DA364} values were underestimated on average by 1.2%, 5.8% and 10%, for PS values of 10, 20 and 30 nm, respectively and, in particular, with SS = 1, underestimation of 3.2%, 10.5% and 15.1% were observed. Similar results were obtained with C_{AF} where CASUM provided unbiased estimates with accuracy of -0.15 ± 0.4 while LINUM showed for SS = 1 underestimation of 5.6%, 3.9% and 1.3% for PS = 10, 20 and 30 nm, respectively. In conclusion, CASUM provided accurate estimates of C_{DA364} and C_{AF}, and comparable precision with the standard LINUM method. Moreover, CASUM is robust against spectral shift (PS) and Stokes shift (SS) variations, as demonstrated by numerical simulations results.

Precision and accuracy of PS and SS estimation by CASUM were analyzed for PS = 10, 20, and 30 nm (Table S2). As expected, also PS and SS estimation depends on S_r because it represented DA364 weight in the simulated spectra. In fact, PS average relative error dropped from 23% to 5% for S_r equal to 0.1 and 2, respectively. CASUM provided, on average, unbiased PS estimates for high S_r values. For low S_r values, PS can be overestimated or underestimated depending on the simulated PS value. With $S_r = 0.1$ estimated PS values were 15% lower and 15% higher than the reference value with true PS equal to 10 and 30, respectively. Estimation of SS demonstrated higher variability with respect to PS. SS average relative error was 331% and 28% for S_r

TABLE 2 Simulation results for CASUM (chemical aware spectral unmixing) in the estimation of PS and SS at different Exp_S_r levels

Exp_S_r	SS	Estimation performance with PS = 20	SS
0.1	0.5	$4.1 \pm 18.7\%$	$-14.6 \pm 117.2\%$
0.2	0.5	$0.8 \pm 8.0\%$	$-4.0 \pm 66.6\%$
0.4	0.5	$0.0 \pm 5.0\%$	$-3.7 \pm 40.7\%$
0.8	0.5	$0.1 \pm 4.3\%$	$-2.6 \pm 32.2\%$
1	0.5	$0.2 \pm 4.2\%$	$1.3 \pm 30.9\%$
1.5	0.5	$0.2 \pm 3.9\%$	$-2.2 \pm 28.4\%$
2	0.5	$-0.5 \pm 3.4\%$	$7.8 \pm 20.9\%$
0.1	1	$-0.2 \pm 25.1\%$	$-7.2 \pm 63.7\%$
0.2	1	$-0.9 \pm 7.8\%$	$2.7 \pm 28.8\%$
0.4	1	$0.1 \pm 5.8\%$	$1.6 \pm 17.3\%$
0.8	1	$-0.3 \pm 4.7\%$	$-1.0 \pm 13.6\%$
1	1	$-0.2 \pm 4.3\%$	$-0.3 \pm 13.1\%$
1.5	1	$-0.3 \pm 4.1\%$	$0.5 \pm 11.8\%$
2	1	$-1.1 \pm 4.1\%$	$5.1 \pm 9.8\%$

Estimation performances are expressed as accuracy $\pm$ precision.

TABLE 3 Unmixing results obtained using LINUM (linear unmixing) and CASUM (chemical aware spectral unmixing) on DA364 solutions containing human serum albumin (HSA) and glycerol (GLY)

|	LINUM		CASUM	
	HSA	GLY	HSA	GLY
C_{DA364}	0.96	0.95	0.95	0.95
R^2	0.86	0.95	0.99	0.99
Residuals analysis				
Mean	0.04	0.02	-0.02	-0.01
SD	0.18	0.10	0.05	0.04
Max	0.41	0.22	0.05	0.05
Normality test	Not normal	Not normal	Normal	Normal
Estimate of excitation peak (nm)	//	//	693 (693)	703 (705)
Estimate of emission peak (nm)	//	//	687 (685)	701 (698)

CASUM provided improved fitting quality demonstrated by R^2 increase, reduction of residuals SD and normal distribution of residuals ($P < .01$, Shapiro–Wilk test). CASUM provides estimates of excitation and emission maxima wavelength in good accordance with the experimental values (in parenthesis).

equal to 0.1 and 2, respectively. Similarly, to PS, SS accuracy depends on the true values of PS. For S_r equal to 0.1, estimated SS values for PS equal to 10, 20 and 30, were 454% higher and, −7 and −9% lower compared to the reference values. For $S_r = 2$, instead, SS estimates were slightly underestimated with percentage difference from the reference value of −4%, −2% and −1% for PS = 10, 20 and 30, respectively. In conclusion, from numerical simulations results that PS estimates are more accurate and precise than SS estimates. Moreover, CASUM precision and accuracy depends on S_r with better performances for increasing S_r values. For this reason the minimum required value of S_r to provide good PS and SS estimates was defined analyzing simulation data for PS equal to 20 nm (the condition resembling the value found in FFPE samples). Since the true value of S_r is generally not know, the ratio of C_{DA364} and C_{AF} (Ext_S_r) calculated by CASUM was used. For Ext_$S_r \geq 0.4$, CASUM provided unbiased PS estimation with precision in the 4% to 6% range that corresponds to a SD of 1 to 1.2 nm (Table 2). At the same Ext_S_r level, SS showed a relative error in the 10% to 40% range, that is, 1 to 8 nm error in the estimation of Stokes shift depending on the value of SS (note that Stokes shift is calculated as SS * PS). In conclusion, with peak shifts of 20 nm, CASUM provided reliable estimates of PS for Ext_$S_r \geq 0.4$ (SD: ~1 nm). Conversely, SS estimation performance is poor with SD in the derived Stokes shift approaching 2 nm only for Ext_$S_r \geq 2$.

3.2 | DA364 solutions experiments

Both low- and high-resolution spectra showed spectral red shift and Stokes shift reduction after HSA addition and dissolution in GLY (Figure 2).

CASUM provided improved fitting quality as can be seen by increase of R^2, reduction and normalization of fitting residuals (Table 3). Deviations of residuals from normal distribution found with LINUM demonstrated that fluorescence intensities are not modeled correctly using DA364 spectrum in PBS as pure spectrum. Namely, 2 fluorescence intensities were underestimated corresponding to 710/760 and 710/780 filter pairs. C_{DA364} values calculated by LINUM and CASUM did not display relevant differences. This is because the estimation of C_{DA364} was mostly influenced by 675/740 and 675/760 filter pairs as shown in the leverage plot (Figure 2). In conclusion, only 2 filter pairs demonstrated sensitivity against spectral shift in the tested range (excitation peak shift was 11 and 19 nm for GLY and HSA solutions, respectively). Despite that, CASUM provided estimates of excitation and emission maxima wavelengths with maximum error of 3 nm.

3.3 | Spectral unmixing experiments

AF spectra measured in muscle and tumor tissues belonging to noninjected dog showed high correlation (Pearson correlation >0.99, $P < .001$). Therefore, a common AF spectrum was defined averaging spectral data from both the pathological and nonpathological tissue.

Spectral unmixing experiments of fluorescence spectra measured from FFPE blocks regions with high and low DA364 fluorescence intensities are summarized in Figure 3 and Table 4. PS values calculated by CASUM were 19.1 and 18.07 nm in the low and high region, respectively, and SS was 0.4 in both the regions. S_r was 0.24 and 0.54 in regions with low and high DA364 content, respectively. C_{DA364} and C_{AF} values calculated by CASUM and LINUM were almost identical. Comparable results in the estimation of C_{DA364} and C_{AF} were expected from numerical simulation results. With PS ~20 nm, SS ~0.5 and S_r in the 0.2 to 0.4 range, the 2 methods provides mostly unbiased estimates with similar precision. Qualitatively, unmixing with LINUM resulted in systematic fitting errors for the 710/760 and
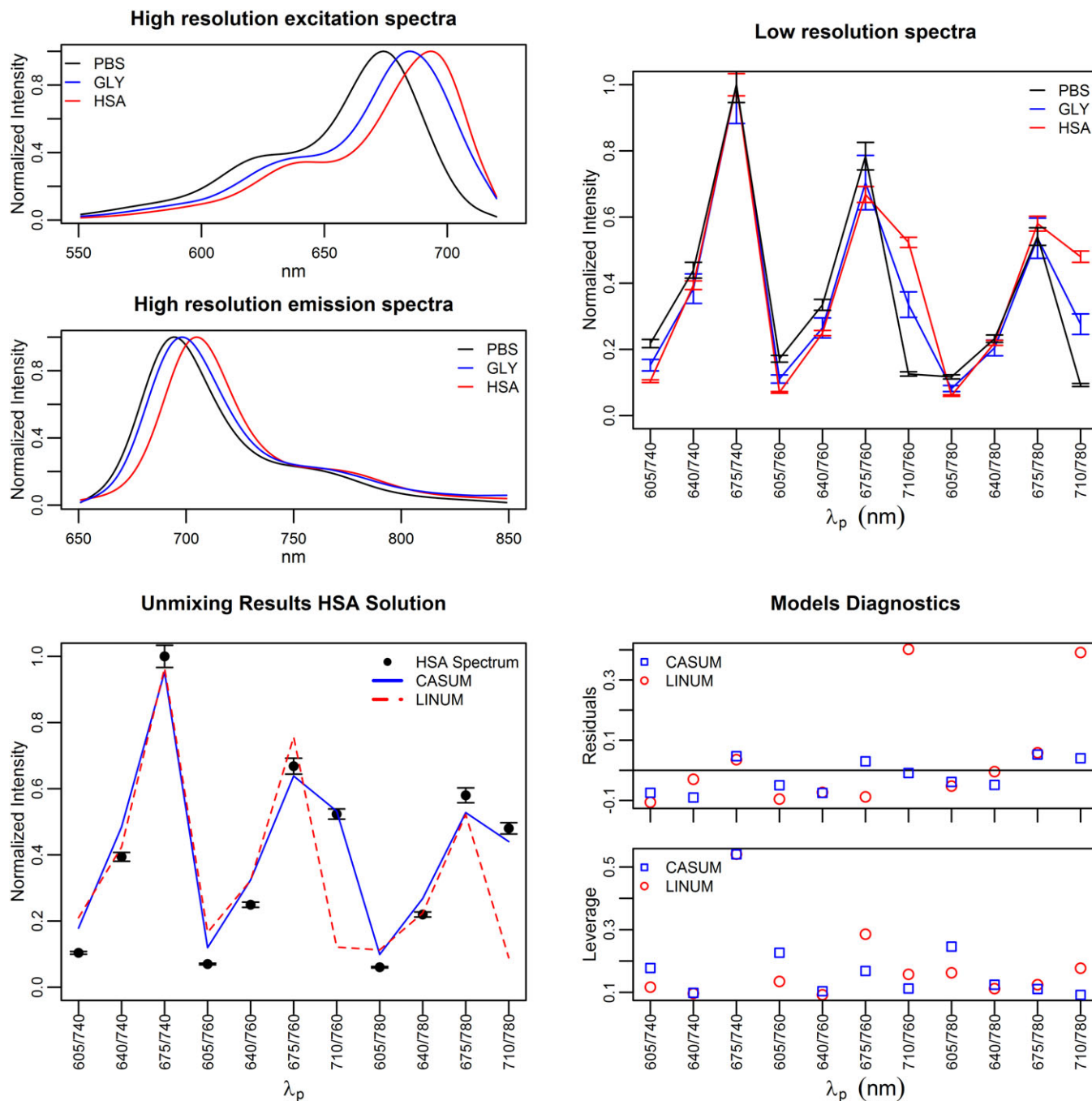
FIGURE 2   Upper panels: high- and low-resolution fluorescence spectra of DA364 measured in phosphate buffer saline (PBS), human serum albumin (HSA) and glycerol (GLY) are red-shifted in HSA and GLY solutions. Lower panels: spectral unmixing of low-resolution DA364 spectrum in HSA was carried out by LINUM (linear unmixing) and CASUM (chemical aware spectral unmixing). Unmixing results reported in the left panel show better fitting to experimental data when CASUM is applied. Lack-of-fit for the data points corresponding to 710/760 and 710/780 filter pairs is highlighted in the residuals plot for the LINUM method. The straight line in the residual plot represents unmixing models predicted value. Leverages plot shows that data points corresponding to 710/760 and 710/780 filter pairs have little effect on C_{DA364} estimation, explaining why, despite the fitting improvement, no differences were found between CASUM and LINUM estimates

710/780 filter pairs, especially in high DA364 region, consistently with the findings obtained with DA364 solutions. The fitting improvement obtained with CASUM suggested DA364 spectral shift related to chemical environment is the cause of the observed deviations. Absolute maximum relative error, indeed, dropped from 57% to 6% and from 34% to 6% in high and low DA364 regions.

The CASUM method was applied on a group tumor and muscle samples to investigate possibly tissue-related DA364 spectral shift. Average excitation and emission peak shifts (PS * SS) are reported in Table 5. PS values showed very narrow intra-sample distributions with SDs about 3% of the mean. PS differences were statistically significant ($P < .001$, ANOVA test) and post-hoc test indicate that every sample
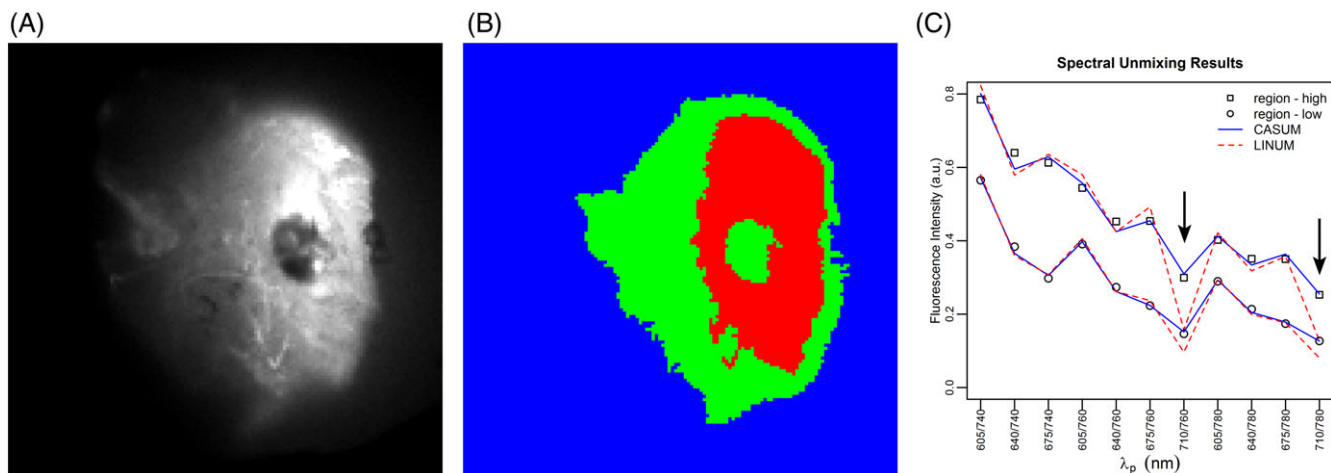
FIGURE 3 Panel A: example of raw fluorescence image (ex/em: 675/740). Panel B: images regions segmented by the Otsu thresholding method. Blue region corresponds to background fluorescence, green and red regions are characterized by low and high DA364 fluorescence. Panel C: graphical comparison of spectral unmixing with LINUM (linear unmixing) and CASUM (chemical aware spectral unmixing) on high and low intensity regions. Arrows indicate systematic fitting errors

was significantly different from all the others ($P < .001$, Tukey HSD). Differences were not related to pathology, for example, muscles did not show consistently higher or lower values with respect to tumor. Emission peak shifts (PS * SS) averages showed greater variance among samples and more scattered distributions with respect to excitation peak shifts consistently with the findings of numerical simulations. Also for emission peak shifts, values distributions showed statistically significant differences for all the possible comparison. Similarly to PS, differences were not related to tissue type.

In conclusion, PS and SS variability within the samples is smaller than differences found among the samples. This suggested the hypothesis that formalin blocks preparation variability could introduce DA364 spectra red-shift that hid tissue-related, potentially useful, variations.

3.4 | Chemical environment imaging

The extended spectral unmixing images dataset provided by CASUM comprising AF and DA364 weights map and excitation and emission spectra shifts maps was compared with
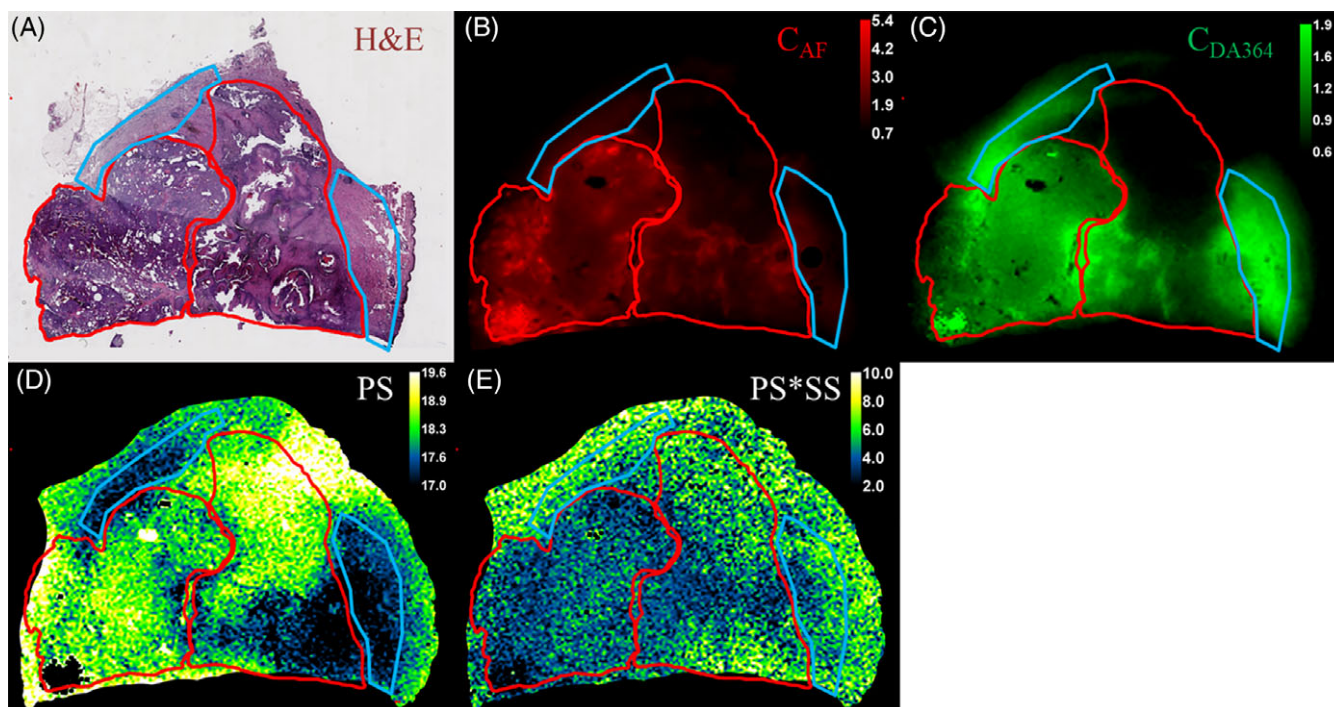
FIGURE 4 Panel A: reference histologic image obtained by H&E staining. Tumor regions are highlighted in red. Nontumor regions showing high DA364 signal are reported in blue. Panel B and C: C_{AF} and C_{DA364} (arbitrary units) weights obtained by spectral unmixing. Panel D and E: parametric images of excitation (PS) and emission peak shifts (PS * SS) expressed in nm obtained by CASUM (chemical aware spectral unmixing)

TABLE 4 Tabulated unmixing results obtained in image regions reported in Figure 3

	High		Low	
	LINUM	**CASUM**	**LINUM**	**CASUM**
C_{DA364}	0.4	0.4	0.13	0.13
C_{AF}	0.74	0.73	0.55	0.55
PS		18.07		19.06
SS		0.4		0.4
Sr		0.54		0.24

Abbreviations: CASUM, chemical aware spectral unmixing; LINUM, linear unmixing.

an H&E section sampled from the surface of an FFPE block (Figure 4). In this situation, the confounding effect of block preparation should be ruled out and spectral differences were only related to tissues differences within the block. H&E section showed that the sample contained 2 tumor masses, highlighted in red, separated by a thin layer of nontumor tissue. Healthy connective tissue was found adjacent to tumor masses (blue regions).

PS map reported spatial variations resembling, in some cases, tumor borders identified in H&E section (eg, the borders of the left tumor mass are visible in PS map). In other cases (right tumor mass), tumor borders were not associated with PS variations. SS map, instead, did not show any specific spatial relation with tumor borders. Direct comparison of spectral shifts parametric maps with spectral unmixing weights maps, showed that spectral shift images report information unrelated to fluorescence intensity. This was expected because spectral shift does not depend on the concentration of the dye. DA364 weight map, for example, showed comparable fluorescence intensity in tumor and nontumor tissue (Panel C, blue regions). Conversely, nontumor regions showed lower excitation peak shift (Panel D) compared to the average tumor regions. This is not generally true, the lower part of right tumor mass, in fact, showed PS values similar to nontumor regions. It is worth noting that H&E section represents only a thin layer (5-10 μm) of tissue while fluorescence light can penetrate into the sample hence PS maps could report information from deeper parts of tissue. In spite of the variance between the samples, that is probably introduced by the tissue preparation and processing, a discrimination of healthy vs pathological tissues in the same sample by PS and SS seems more promising. Nevertheless, the final assessment of the diagnostic potential of this information would require dedicated experiments carried out on samples more comparable to the in vivo conditions than FFPE blocks. From the computational point of view CASUM is more demanding than other methods. The calculation of PS and SS parametric maps, in fact, took minutes with the current Python implementation. Present execution times are not compliant for real-time in vivo applications, otherwise, when the method is applied ex vivo, this is not an issue.

In the proof-of-concept study reported herein the hypothesis that FFPE blocks viscosity and/or dye interactions with the biological matrix induced DA364 spectral shift was explored. With the development of CASUM, the systematic unmixing biases initially attributed to endogenous AF sources or spectral coloring due to wavelength-dependent light attenuation caused by formalin, were eventually explained by DA364 spectral shift. CASUM is robust against spectral shift variations; its application leads to unbiased spectral weights estimates as demonstrated by numerical simulations. Additionally, CASUM extracts from spectral images information related to chemical environment through the estimation of spectral shift parameters. The key points are that CASUM is a tool to study fluorescent dyes interaction with tissues, it exploits fluorophores spectral sensitivity to chemical environment and it can be used with the low-resolution imaging devices generally available in biology laboratories. The interest in chemical environment sensitivity has been widely reported in literature: fluorescent probes showing pH-dependent spectral variations have been proposed in the last years [25, 26] to detect the acidic tumor microenvironment recognized as a general hallmark of cancer [27–29]. Another work [30] described solvatochromic shifts of 10 to 40 nm in solvents with variable polarity for dyes of the same class of Cy5.5, the fluorescent moiety in DA364. In the cited work, dye interactions with chemical environment were detected in vitro by lifetime measurements and high-resolution spectral experiments. In this context, CASUM can measure spectral shift in the 10 to 40 nm range in biological samples, providing a tool for studying biological solvation systems, using a low-resolution optical imaging devices.

TABLE 5 PS and SS values calculated by CASUM (chemical aware spectral unmixing) in formalin-fixed paraffin embedded blocks containing different tissue types

	Tissue type					
	Mastocystoma	**Adenocarcinoma (Sample 1)**	**Adenocarcinoma (Sample 2)**	**Adenoma**	**Muscle (Sample 1)**	**Muscle (Sample 2)**
Excitation peak shift (PS)	18.1 ± 0.96 (N = 1408)	17.9 ± 0.35 (N = 1409)	17.4 ± 0.50 (N = 2314)	18.2 ± 0.34 (N = 645)	18.5 ± 0.69 (N = 2576)	17.7 ± 0.27 (N = 1149)
Emission peak shift (PS * SS)	7.9 ± 2.70 (N = 1430)	10.4 ± 1.33 (N = 1357)	5.2 ± 1.40 (N = 2408)	11.9 ± 1.28 (N = 630)	8.5 ± 1.86 (N = 2695)	10.1 ± 0.95 (N = 1132)

Average, SE and sample size (N = pixels number) for each region are summarized in the table. Muscle and adenocarcinoma Samples 1 and 2 were from 2 different dogs.

4 | CONCLUSION

In this work, a novel spectral unmixing procedure that weights fluorophores spectral variations related to chemical environment is proposed. The method is robust against spectral shifts and provides unbiased spectral unmixing parameters estimates. Additionally, the proposed method makes available to scientists a new tool to study fluorescent dyes interaction with tissues by exploiting fluorophores sensitivity to chemical environment. Finally, the method can be applied to images acquired with standard devices generally available in optical imaging laboratories.

Conflict of interest

Dr Valbusa and Dr Capozza declare that there is no conflict of interest, Dr Brioschi, Dr Blasi, Dr Ghiani and Dr Maiocchi are employees of Bracco Imaging Spa.

AUTHOR BIOGRAPHIES

Please see Supporting Information online.

ORCID

Giovanni Valbusa https://orcid.org/0000-0002-0046-6607
Martina Capozza https://orcid.org/0000-0001-9125-8004

REFERENCES

[1] B. E. Schaafsma, J. S. Mieog, M. Hutteman, J. R. van der Vorst, P. J. Kuppen, C. W. Löwik, J. V. Frangioni, C. J. van de Velde, A. L. Vahrmeijer, *J. Surg. Oncol.* **2011**, *104*, 323.

[2] L. E. Lamberts, M. Koch, J. S. de Jong, A. L. L. Adams, J. Glatz, M. E. G. Kranendonk, A. G. T. Terwisscha van Scheltinga, L. Jansen, J. de Vries, M. N. Lub-de Hooge, C. P. Schröder, A. Jorritsma-Smit, M. D. Linssen, E. de Boer, B. van der Vegt, W. B. Nagengast, S. G. Elias, S. Oliveira, A. J. Witkamp, W. P. T. M. Mali, E. Van der Wall, P. J. van Diest, E. G. E. de Vries, V. Ntziachristos, G. M. van Dam, *Clin. Cancer Res.* **2017**, *23*, 2730.

[3] N. J. Harlaar, M. Koller, S. J. de Jongh, B. L. van Leeuwen, P. H. Hemmer, S. Kruijff, R. J. van Ginkel, L. B. Been, J. S. de Jong, G. Kats-Ugurlu, M. D. Linssen, A. Jorritsma-Smit, M. van Oosten, W. B. Nagengast, V. Ntziachristos, G. M. van Dam, *Lancet Gastroenterol. Hepatol.* **2016**, *1*, 283.

[4] D. Holt, O. Okusanya, R. Judy, O. Venegas, J. Jiang, E. DeJesus, E. Eruslanov, J. Quatromoni, P. Bhojnagarwala, C. Deshpande, S. Albelda, S. Nie, S. Singhal, *PLoS One* **2014**, *9*, e103342.

[5] G. M. van Dam, G. Themelis, L. M. Crane, N. J. Harlaar, R. G. Pleijhuis, W. Kelder, A. Sarantopoulos, J. S. de Jong, H. J. Arts, A. G. van der Zee, J. Bart, P. S. Low, V. Ntziachristos, *Nat. Med.* **2011**, *17*, 1315.

[6] F. Acerbi, C. Cavallo, M. Broggi, R. Cordella, E. Anghileri, M. Eoli, M. Schiariti, G. Broggi, P. Ferroli, *Neurosurg. Rev.* **2014**, *37*, 547.

[7] T. Okuda, K. Kataoka, T. Yabuuchi, H. Yugami, A. Kato, *J. Clin. Neurosci.* **2010**, *17*, 118.

[8] J. T. Alander, I. Kaartinen, A. Laakso, T. Pätilä, T. Spillmann, V. V. Tuchin, M. Venermo, P. Välisuo, *Int. J. Biomed. Imaging* **2012**, *2012*, 940585.

[9] J. Yokoyama, M. Fujimaki, S. Ohba, T. Anzai, R. Yoshii, S. Ito, M. Kojima, K. Ikeda, *Onco. Targets Ther.* **2013**, *6*, 325.

[10] T. Zimmermann, J. Rietdorf, R. Pepperkok, *FEBS Lett.* **2003**, *546*, 87.

[11] I. T. Jolliffe, *Principal Component Analysis*, Springer Verlag, New York **1986**.

[12] R. A. Neher, M. Mitkovski, F. Kirchhoff, E. Neher, F. J. Theis, A. Zeug, *Biophys. J.* **2009**, *96*, 3791.

[13] H. Xu, B. W. Rice, *J. Biomed. Opt.* **2009**, *14*, 064011.

[14] H. Shirakawa, S. Miyazaki, *Biophys. J.* **2004**, *86*, 1739.

[15] F. Cutrale, V. Trivedi, L. A. Trinh, C. L. Chiu, J. M. Choi, M. S. Artiga, S. E. Fraser, *Nat. Methods* **2017**, *14*, 149.

[16] F. Fereidoun, Multi-spectral lifetime imaging: methods and applications. *Doctoral thesis*. Utrecht University **2013**.

[17] S. Lanzardo, L. Conti, C. Brioschi, M. P. Bartolomeo, D. Arosio, L. Belvisi, L. Manzoni, A. Maiocchi, F. Maisano, G. Forni, *Contrast Media Mol. Imaging* **2011**, *6*, 449.

[18] C. Würth, M. Grabolle, J. Pauli, M. Spieles, U. Resch-Genger, *Nat. Protoc.* **2013**, *8*, 1535.

[19] D. Fukutomi, K. Ishii, K. Awazu, *Opt. Rev.* **2016**, *23*, 291.

[20] G. Themelis, J. S. Yoo, K. S. Soh, R. Schulz, V. Ntziachristos, *J. Biomed. Opt.* **2009**, *14*, 064012.

[21] T. Zimmermann, J. Marrison, K. Hogg, P. O'Toole, *Methods Mol. Biol.* **2014**, *1075*, 129.

[22] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer, New York **2006**.

[23] J. Cao, T. Wu, C. Hu, T. Liu, W. Sun, J. Fan, X. Peng, *Phys. Chem. Chem. Phys.* **2012**, *14*, 13702.

[24] L. Ping-sung, C. Tse-sheng, C. Pau-choo, *J. Inf. Sci. Eng.* **2001**, *17*, 713.

[25] H. Xiong, P. Kos, Y. Yan, K. Zhou, J. B. Miller, S. Elkassih, D. J. Siegwart, *Bioconjug. Chem.* **2016**, *27*, 1737.

[26] J. T. Hou, W. X. Ren, K. Li, J. Seo, A. Sharma, X. Q. Yu, J. S. Kim, *Chem. Soc. Rev.* **2017**, *46*, 2076.

[27] S. A. Hilderbrand, K. A. Kelly, M. Niedre, R. Weissleder, *Bioconjug. Chem.* **2008**, *19*, 1635.

[28] Y. Wang, K. Zhou, G. Huang, C. Hensley, X. Huang, X. Ma, T. Zhao, B. D. Sumer, R. J. DeBerardinis, J. Gao, *Nat. Mater.* **2014**, *13*, 204.

[29] C. R. Drake, D. C. Miller, E. F. Jones, *Curr. Org. Synth.* **2011**, *8*, 498.

[30] M. Y. Berezin, H. Lee, W. Akers, S. Achilefu, *Biophys. J.* **2007**, *93*, 2892.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.
TABLE S1 Simulation results for LINUM (linear unmixing) and CASUM (chemical aware spectral unmixing) in the estimation of C_{DA364} and C_{AF}
TABLE S2 Simulation results for CASUM (chemical aware spectral unmixing) in the estimation of PS and SS