Towards automated spectroscopic tissue classification in thyroid and parathyroid surgery

ORIGINAL ARTICLE

Towards automated spectroscopic tissue classification in thyroid and parathyroid surgery

Rutger M. Schols[1,2]*
Lejla Alic[2]
Fokko P. Wieringa[2,3]
Nicole D. Bouvy[1]
Laurents P. S. Stassen[1]

[1]*Department of Surgery, Maastricht University Medical Center & NUTRIM School for Nutrition, Toxicology and Metabolism, Maastricht University, Maastricht, The Netherlands*

[2]*van't Hoff Program on Medical Photonics, Netherlands Organization for Applied Scientific Research TNO, Delft, The Netherlands*

[3]*Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands*

*Correspondence to: Rutger M. Schols, Department of Surgery, Maastricht University Medical Center & NUTRIM School for Nutrition, Toxicology and Metabolism, Maastricht University, P. Debyelaan 25, 6229 HX Maastricht, The Netherlands.
E-mail: rutgermschols@hotmail.com

Abstract

Background   In (para-)thyroid surgery iatrogenic parathyroid injury should be prevented. To aid the surgeons' eye, a camera system enabling parathyroid-specific image enhancement would be useful. Hyperspectral camera technology might work, provided that the spectral signature of parathyroid tissue offers enough specific features to be reliably and automatically distinguished from surrounding tissues. As a first step to investigate this, we examined the feasibility of wide band diffuse reflectance spectroscopy (DRS) for automated spectroscopic tissue classification, using silicon (Si) and indium-gallium-arsenide (InGaAs) sensors.

Methods   DRS (350–1830 nm) was performed during (para-)thyroid resections. From the acquired spectra 36 features at predefined wavelengths were extracted. The best features for classification of parathyroid from adipose or thyroid were assessed by binary logistic regression for Si- and InGaAs-sensor ranges. Classification performance was evaluated by leave-one-out cross-validation.

Results   In 19 patients 299 spectra were recorded (62 tissue sites: thyroid = 23, parathyroid = 21, adipose = 18). Classification accuracy of parathyroid–adipose was, respectively, 79% (Si), 82% (InGaAs) and 97% (Si/InGaAs combined). Parathyroid–thyroid classification accuracies were 80% (Si), 75% (InGaAs), 82% (Si/InGaAs combined).

Conclusions   Si and InGaAs sensors are fairly accurate for automated spectroscopic classification of parathyroid, adipose and thyroid tissues. Combination of both sensor technologies improves accuracy. Follow-up research, aimed towards hyperspectral imaging seems justified. Copyright © 2016 John Wiley & Sons, Ltd.

Keywords   diffuse reflectance spectroscopy; thyroid and parathyroid surgery; automated tissue classification; parathyroid; thyroid; adipose tissue

Introduction

To perform safe and efficient thyroid and parathyroid surgery, a high level of surgical skills is required together with understanding of natural variations in head and neck anatomy (e.g. thyroid, parathyroid, lymph nodes, and recurrent laryngeal nerves). Detecting the small sized parathyroid glands and

performing re-operative thyroid surgery can be especially challenging and time consuming. This visual task can be even more challenging due to anatomical variations in location, especially for the lower parathyroid glands (1). During these complicated surgical procedures iatrogenic injury to the parathyroid glands may occur (2). In a retrospective analysis of 5104 primary and 685 secondary thyroidectomies, temporary hypocalcaemia associated with hypoparathyroidism occurred in respectively 7.1% and 5% of cases (permanent in 1.8% and 2.5%). The rate of permanent complications was found to be significantly higher in re-operative surgery (3). Intraoperative identification of parathyroid glands before removal of the thyroid gland is of great importance to prevent these complications. Therefore, a tool for improved intraoperative parathyroid gland detection is desirable. Exploring the capabilities of spectroscopy beyond the limitations of the human eye offers a possible roadmap towards such a tool.

Several innovative optical techniques have been under investigation for their potential in differentiating benign from malignant cells in thyroid and parathyroid specimens: multispectral image analysis (4,5), Raman spectroscopy (6) and elastic scattering spectroscopy (7). Fluorescence imaging after peripheral infusion of aminolevulinic acid (8,9) or methylene blue (10) is reported for intraoperative detection of parathyroid adenomas. In addition, near-infrared auto-fluorescence incorporates potential for real-time parathyroid tissue localization (11). Furthermore, optical coherence tomography is reported as a tool for parathyroid gland identification (12,13).

For color vision, the human eye contains only blue, green and red cones, which also partly overlap in sensitivity (poor channel separation). Yet, the trained human eye can discern quite subtle color differences within the visible range (400–780 nm). Normal surgical video cameras mimic human vision and also just capture red, green and blue light. Hyperspectral imaging, however, discerns a multitude of well-separated spectral bands for each pixel, including the near-infrared range, which is (by definition) invisible to the human eye. Since aerospace science already routinely combines hyperspectral camera technology with pre-acquired reference-library spectra recorded on the earth surface, to generate satellite images for discovering places of interest, for example, agricultural purposes and military and homeland security applications, we reasoned this technique might also hold potential for image-guided surgery (14). Although hyperspectral imaging is still far from being a routine surgical tool, it has already shown potential for noninvasive intraoperative assessment of tissue oxygen saturation (15,16), for intraoperative enhancement of anatomical structures (17,18) and for intraoperative tumor detection (19). Medical hyperspectral imaging-systems typically use silicon (Si)

and occasionally indium gallium arsenide (InGaAs) camera chips. The wavelength range 400–1000 nm is covered by Si, whereas InGaAs is typically sensitive in the 900–1700 nm wavelength region (and depending on chip composition even up to 2500 nm) (14).

Diffuse reflectance spectroscopy (DRS) was studied for tissue differentiation as basis for a feedback system to enhance nerve preservation in oral and maxillofacial surgery (20), in head and neck surgery and in carpal tunnel release surgery (21). DRS has, however, not yet been studied for thyroid or parathyroid differentiation.

Nachabé *et al.* (22,23) showed the benefit of extending visualization 'beyond silicon' up to a wavelength of 1600 nm, providing additional information regarding tissue concentrations of the biological chromophores water and lipid.

The present translational study is a first fibre-optic spectrometric exploration with the ultimate long-term goal of obtaining parathyroid gland-specific spectral signatures as a basis for further research on image enhancement. *In vivo*, wide band (350–1830 nm) diffuse reflectance spectra of human thyroid, parathyroid and surrounding adipose tissue were collected and assessed for the presence of endogenous contrasts that might enable future tissue-specific contrast enhancement. Therefore we applied tissue classification using spectral features that are largely independent of signal intensity using a previously reported approach (21).

Materials and methods

All *in vivo* data acquisition was performed at the Surgery Department of Maastricht University Medical Center (MUMC, Maastricht, The Netherlands) during thyroid and parathyroid surgery (surgeon N.D.B.). Prior to measurements, the local institutional review board of Maastricht University Medical Center (registration number METC 10-4-035) granted approval and preoperative written informed consent was obtained from all patients.

Material

Diffuse reflectance spectra were acquired using custom developed ETO-sterilized disposable optical fibre probes (TNO, Eindhoven the Netherlands and Light Guide Optics, Rheinbach Germany), a modified Xenon light source (D-light C, Karl Storz, Tuttlingen Germany), and a spectrometer (Analytical Spectral Devices, Inc., Colorado USA) covering the range 350–1830 nm. The spectrometer is equipped with two sensor technologies: a silicon (Si) based sensor and an indium gallium arsenide (InGaAs) based sensor. The cross-over point between

these sensors is at 1000 nm. The fibre probe is the medium to transport light from source to tissue, and to transport reflected light from tissue to spectrometer. The system was installed on a compact trolley to facilitate *in vivo* tissue data acquisition during routine surgery. The setup was tested and approved according to the essential requirements of IEC 60601-1 to assure patient safety. The equipment has previously been described in more detail (24).

Figure 1 shows the fibre probe tip composed of eight optical fibres: one central receiving fibre (400 um diameter, NA 0.22 ± 0.02) and seven illuminating fibres (300 µm diameter, NA 0.22 ± 0.02). The rigid stainless steel probe tip has a 2 mm diameter and a length of 10 mm. The remaining length of the sterile fibre probe is highly flexible.

In vivo data acquisition

During parathyroid and thyroid resections, *in vivo* wide-band diffuse reflectance spectra (350–1830 nm, 1 nm spectral resolution) were collected. For each tissue type, five spectra per site (taking 30 s per site) were acquired, covering at least one site per tissue

Figure 1. Intraoperative fibre probe tissue measurement. Intraoperative spectral fibre probe measurement performed during a left hemithyroidectomy procedure. The sterile optical fibre probe is gently brought into contact with one of the parathyroid glands. Right lower corner: close-up of the fibre probe tip. A ring light of seven fibres (illuminated white) injects light into the tissue, one central fibre (here for illustration purposes illuminated yellow) transports the diffusely reflected light to the spectrometer

type; this resulted in 1.5 to 3 additional min per surgical procedure (see also Table 1). The sterile fibre probe was handled by the surgeon and gently brought into direct contact with one of the designated tissues (see Figure 1). If blood was visibly present on the tissue surface, it was dapped away using a sterile gauze. Between the measurements on different locations, the probe tip was swiped with a clean sterile gauze wetted with saline. Acquired data were labeled according to the tissue type description of the attending surgeon. Since extirpation of the measured 'healthy' tissue types could logically not be conducted in this explorative *in vivo* human study, histopathological confirmation of the measurements was not possible.

For inter-patient comparability, all raw *in vivo* spectra were calibrated, after the completion of *in vivo* spectroscopy, by using the dark current (14) and white reference phantom (Optical-grade spectralon reference; Labsphere, Inc., North Sutton, New Hampshire USA). The integration times of the silicon and InGaAs sensor were individually optimized during the Spectralon calibration. No correction for ambient light was performed.

Data processing

The calibration normalizes the radiance spectrum to yield the reflectance, which manages the problem of spectral non-uniformity of the illumination device and influence of the dark current (26). Since both reflectance intensity and spectral shape are related to the composition of the tissue, no further normalization steps were performed. To identify possible distinctive features for tissue-specific enhancement, 36 features (i.e. 18 gradients and 18 amplitude differences at predefined points in the tissue spectra) were extracted based on known wavelengths related to characteristic absorption features for blood, water and fat (27–29). Gradients are 'slopes' between two predefined points in the tissue spectra, i.e. $(DR_2–DR_1)/(\lambda_2–\lambda_1)$. Amplitude differences are 'intensity differences' between two predefined spectral points, i.e. $(DR_2–DR_1)$. DR = diffuse reflectance; λ = wavelength. These features have also been described in our previous publication (21). Figure 2 illustrates the characteristic wavelengths and features in a mean spectrum for human adipose tissue (anatomical region: neck). All data processing was performed by in-house developed software (using MATLAB environment Version 7.7.0, MathWorks Inc., Natick, Massachusetts, USA).

To evaluate the results for application in imaging technology for identification of parathyroid within its
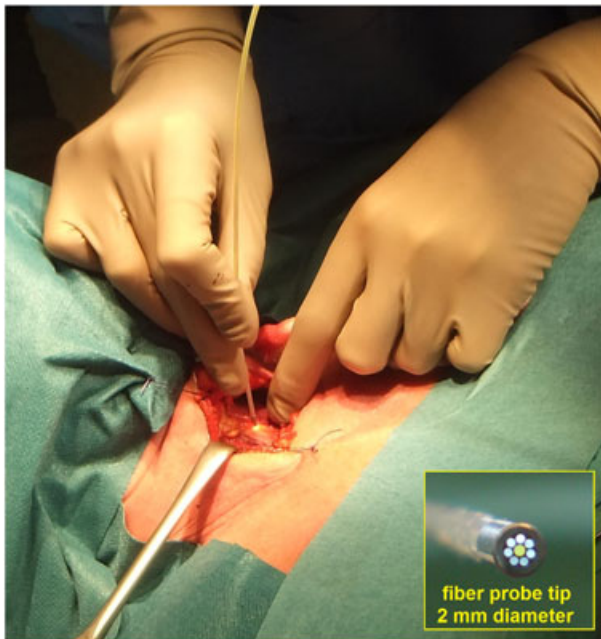
Table 1. Study subject characteristics and measured sites per tissue type

Subject#	Gender	Age	BMI	Indication for surgery	Surgical procedure	Thyroid	Parathyroid	Adipose
						Sites per tissue type		
1	M	67	24	Bethesda 3 * node thyroid left	Left HT	2	2	0
2	F	55	32	PH	Right PAR	2	1	2
3	M	54	31	MEN-2a syndrome, medullar thyroid carcinoma left	TT + radical LND neck, right axillary LND	0	0	1
4	F	14	21	PH	Left PAR	2	2	2
5	F	67	37	Status after hemithyroidectomy right, in which a follicular carcinoma was found	Resection of remaining thyroid (left HT)	0	2	1
6	F	53	34	Multinodular goiter	TT	2	2	1
7	F	56	44	Bethesda 2 * node thyroid right	Right HT	0	0	1
8	F	67	30	Multinodular goiter	Left HT	1	0	0
9	F	60	43	Tertiary hyperparathyroidism after kidney transplantation	Subtotal parathyroid resection	1	1	0
10	F	55	-	Multinodular goiter	Left HT + subtotal right thyroidectomy	1	0	0
11	F	50	18	Multinodular goiter	Left HT	1	1	1
12	M	80	26	PH	Left PAR	1	0	1
13	F	54	35	PH	Right PAR	1	1	1
14	M	74	29	PH	Right PAR	1	0	1
15	F	60	31	Multinodular goiter	Left HT	1	1	1
16	F	46	30	Multinodular goiter	Left HT	1	2	1
17	F	53	34	Bethesda 3 * node thyroid right	Right HT	2	2	2
18	F	23	23	Graves' disease	TT	2	2	1
19	F	49	19	Multinodular goiter	Left HT	2	2	1
					Sites per tissue type (grand total 62)	23	21	18
					Spectra per tissue type (grand total 299)	112	97	90

*The Bethesda system was used for reporting thyroid cytopathology (25).

M = male; F = female; BMI = body mass index; PH = primary hyperparathyroidism; PAR = parathyroid adenoma resection; HT = hemithyroidectomy; TT = total thyroidectomy; LND = lymph node dissection.

natural surroundings, we performed three separate classification steps: i.e. Si-sensor (≤1000 nm) features, InGaAs-sensor (≥900 nm) features, and features covering the whole range (350–1830 nm). Considering the number of subjects, the number of extracted features (n = 36) is too large to perform a statistically meaningful classification, as the extracted features could be redundant in the information they retain. Therefore, using combinations of all 36 features to build a classifier would result in a dimensionality problem and over-fitting (30). To omit this problem, we identified the most distinctive features, for classification of parathyroid in relation to surrounding adipose tissue or adjacent thyroid tissue, by using binary logistic regression (SPSS Inc., Chicago, IL, USA) for both wavelength regions separately. This statistical technique allows for the prediction of categorical dependent variables (here the *tissue type*: parathyroid or adipose /thyroid) using a set of independent variables (here the 36 *spectral features*) (31).

In accordance with our previous report on classifying nerve tissue within adipose surroundings (21), we trained a support vector machine (SVM) classifier (32) which attempts to find an optimum line to separate the training data groups with minimum risk (30). To estimate classification performance and to prevent overly optimistic results (20,33,34), we implemented a cross-validation (CV) approach. With a goal to obtain the classification accuracy as a performance measure, the data set is divided into a training set (to train the classifier) and a test set (to validate the classifier). The CV approach uses leave-one-out (LOO; in this case: leave-one-'tissue-site'-out) validation of the collected parathyroid and adipose/thyroid hyperspectral data. This approach utilizes the same data set for both training and testing purposes and is very useful in cases of a relatively small data sample.

In-house developed classifiers (using MATLAB environment Version 7.7.0, MathWorks Inc., Natick, Massachusetts, USA) were used to estimate the classification performance by calculating sensitivity, specificity, positive predictive value, negative predictive value and accuracy for both wavelength regions (i.e. Si-sensor and InGaAs-sensor detection range) separately and in combination.

Bₙ = Blood related point   Wₙ = Water related point   Fₙ = Fat related point

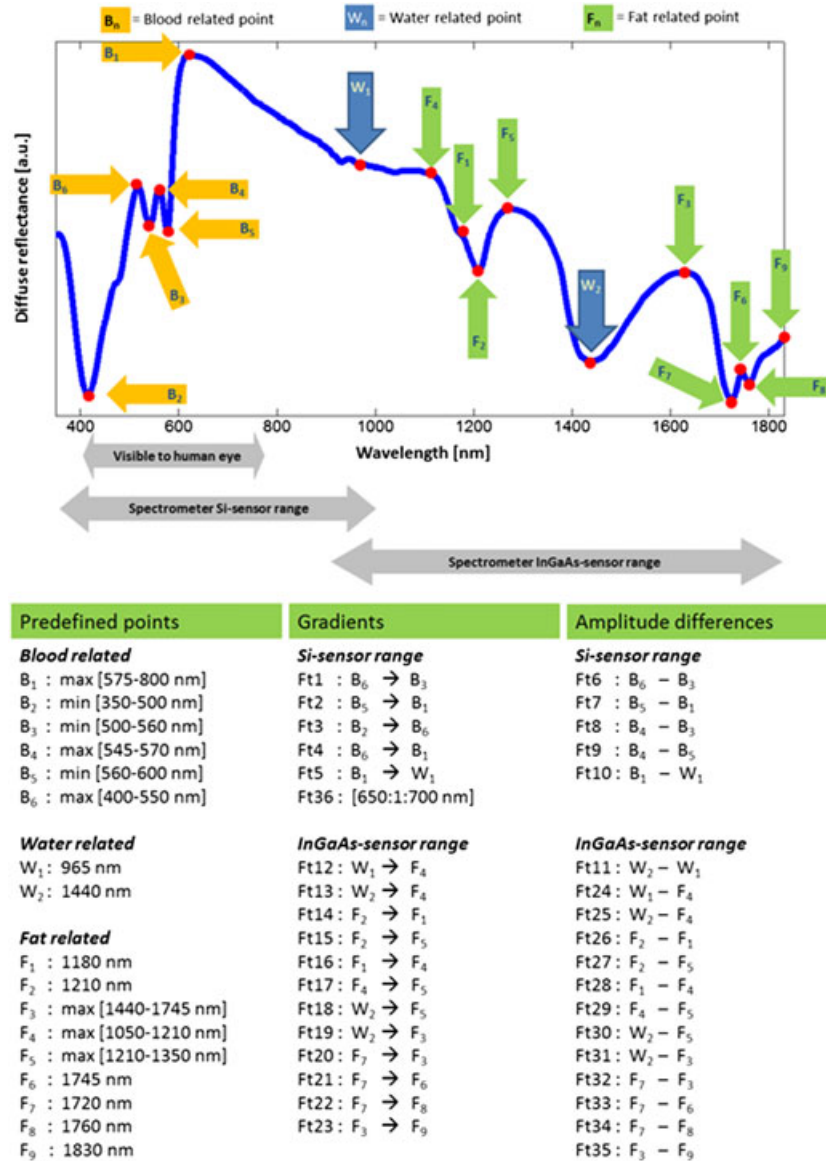| Predefined points	Gradients	Amplitude differences
Blood related	**Si-sensor range**	**Si-sensor range**
B_1 : max [575–800 nm]	Ft1 : $B_6 \rightarrow B_3$	Ft6 : $B_6 - B_3$
B_2 : min [350–500 nm]	Ft2 : $B_5 \rightarrow B_1$	Ft7 : $B_5 - B_1$
B_3 : min [500–560 nm]	Ft3 : $B_2 \rightarrow B_6$	Ft8 : $B_4 - B_3$
B_4 : max [545–570 nm]	Ft4 : $B_6 \rightarrow B_1$	Ft9 : $B_4 - B_5$
B_5 : min [560–600 nm]	Ft5 : $B_1 \rightarrow W_1$	Ft10 : $B_1 - W_1$
B_6 : max [400–550 nm]	Ft36 : [650:1:700 nm]	
Water related	**InGaAs-sensor range**	**InGaAs-sensor range**
W_1 : 965 nm	Ft12 : $W_1 \rightarrow F_4$	Ft11 : $W_2 - W_1$
W_2 : 1440 nm	Ft13 : $W_2 \rightarrow F_4$	Ft24 : $W_1 - F_4$
	Ft14 : $F_2 \rightarrow F_1$	Ft25 : $W_2 - F_4$
Fat related	Ft15 : $F_2 \rightarrow F_5$	Ft26 : $F_2 - F_1$
F_1 : 1180 nm	Ft16 : $F_1 \rightarrow F_4$	Ft27 : $F_2 - F_5$
F_2 : 1210 nm	Ft17 : $F_4 \rightarrow F_5$	Ft28 : $F_1 - F_4$
F_3 : max [1440–1745 nm]	Ft18 : $W_2 \rightarrow F_5$	Ft29 : $F_4 - F_5$
F_4 : max [1050–1210 nm]	Ft19 : $W_2 \rightarrow F_3$	Ft30 : $W_2 - F_5$
F_5 : max [1210–1350 nm]	Ft20 : $F_7 \rightarrow F_3$	Ft31 : $W_2 - F_3$
F_6 : 1745 nm	Ft21 : $F_7 \rightarrow F_6$	Ft32 : $F_7 - F_3$
F_7 : 1720 nm	Ft22 : $F_7 \rightarrow F_8$	Ft33 : $F_7 - F_6$
F_8 : 1760 nm	Ft23 : $F_3 \rightarrow F_9$	Ft34 : $F_7 - F_8$
F_9 : 1830 nm		Ft35 : $F_3 - F_9$

Figure 2. Example of spectrum (adipose tissue) with characteristic wavelengths and investigated features. Qualitative landmarks for oxygenated/reduced hemoglobin (characteristic 'W-shape' between 500 and 600 nm), water (absorption peaks at 965 and 1440 nm) and human fat (absorption peaks at 1210 and 1720 nm) (27–29) are indicated by the arrows. The spectral detection ranges of respectively the human eye and of Si-sensor and InGaAs-sensor based camera chips are indicated. (This Figure was previously published as Figure 2 in: 'Differentiation between nerve and adipose tissue using wide-band (350–1830 nm) *in vivo* diffuse reflectance spectroscopy', by Schols RM, ter Laan M, Stassen LP, Bouvy ND, Amelink A, Wieringa FP, Alic L; Lasers in Surgery and Medicine, 2014 Jun 4. Epub ahead of print, doi: 10.1002/lsm.22264, © 2014 Wiley Periodicals, Inc. Wiley-Blackwell. Used with permission.)

Results

In 19 patients (4 male, 15 female) undergoing parathyroid or thyroid surgery, 299 *in vivo* tissue spectra were recorded on 62 tissue sites (thyroid n = 23, parathyroid gland n = 21, adipose tissue n = 18). Table 1 summarizes patient characteristics and the number of measured sites per tissue type. Figure 3 shows mean diffuse reflectance spectra and corresponding standard deviations for thyroid, parathyroid and surrounding adipose tissue.

Classification of spectral data in Si-sensor range

From the 36 extracted features (see Figure 2) 11 were located within the silicon detection range. Given the study sample of an average of 20 spots per tissue type, inclusion
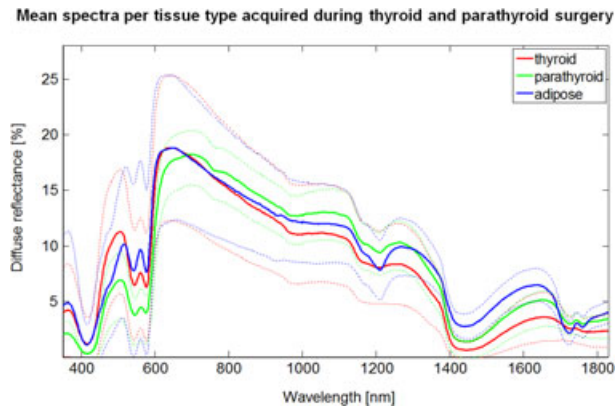
Figure 3. **Mean spectra per tissue type.** Average tissue spectra for thyroid (red), parathyroid (green) and adipose tissue (blue). Dashed lines in the corresponding colors indicate the respective standard deviations

of maximum 2 features is statistically responsible. Binary logistic regression identified gradients Ft3 (B_2–B_6) and Ft36 (650–700 nm) as most promising combination for differentiation of parathyroid from surrounding adipose tissue. Gradient Ft36 (650–700 nm) solely, also appeared to be the best distinctive feature for differentiation of parathyroid from thyroid tissue. Figure 4(A) shows a scatter plot (Ft36 and Ft3) for classification of parathyroid in adipose surrounding. Figure 4(B) shows a boxplot (Ft 36) for classification of parathyroid in thyroid surrounding.

The quantitative results of classification performance are listed in Table 2. Leave-one-out cross-validation (LOO CV) is based on Si-sensor range data from thyroid and parathyroid surgery for train and test purposes.

Classification of spectral data in InGaAs-sensor range

From the 36 extracted features 25 were located within the spectral detection range of InGaAs (see Figure 2). After binary logistic regression, gradients Ft12 (W_1–F_4) and Ft17 (F_4–F_5) were selected as the most promising combination for differentiation of parathyroid from surrounding adipose tissue. Gradients Ft12 (W_1–F_4) and Ft14 (F_2–F_1) were identified as best distinctive feature combination for differentiation of parathyroid from thyroid tissue. Figure 5(A) and 5(B) show scatter plots for these InGaAs-sensor based features, extracted for respectively parathyroid versus adipose tissue and parathyroid versus thyroid tissue.

The quantitative results of LOO CV classification performance are listed in Table 3. LOO CV is solely based on InGaAs-sensor range data from thyroid and parathyroid surgery for train and test purposes.

Combining Si-sensor and InGaAs-sensor ranges for spectral data classification

Classification performance was also evaluated when combining the Si-sensor and InGaAs-sensor detector ranges, including up to 3 spectral features. After binary logistic regression, gradients Ft12 (W_1–F_4), Ft17 (F_4–F_5) and Ft36 (650–700 nm) were selected as the most promising combination for differentiation of parathyroid from surrounding adipose tissue.

Gradients Ft27 (W_1–F_4) and Ft32 (F_2–F_1) and Ft36 (650–700 nm) were identified as best distinctive feature combination for differentiation of parathyroid from thyroid tissue.

Figure 6(A) and 6(B) show scatter plots for these combined Si-sensor and InGaAs-sensor based features, extracted for respectively parathyroid versus adipose tissue and parathyroid versus thyroid tissue.

The quantitative results of LOO CV classification performance are listed in Table 4. LOO CV is based on the whole wide-band spectral range covering both Si-sensor and InGaAs-sensor range data from thyroid and parathyroid surgery for train and test purposes.

Discussion

The wide band (350–1830 nm) diffuse optical reflectance fingerprints of human thyroid, parathyroid and surrounding adipose tissue were identified in this explorative study. These spectra covered silicon (Si) and indium gallium arsenide (InGaAs) detector ranges, thereby exceeding wavelength boundaries (1600 nm or 1700 nm) reported by preceding work (22,23,25,26). Spectroscopic measurements on human skin samples in the wavelength range 1000–2200 nm have previously been reported (35), but not regarding any of the tissues included in our study.

Although the presented reflectance spectra (Figure 3) show great similarity between the three different tissue types, performance of automated tissue-specific classification was fairly accurate in both sensor detection ranges (i.e. Si and InGaAs) to discriminate parathyroid tissue from either surrounding adipose tissue or adjacent thyroid tissue. Based on the classification accuracies (Tables 2 and 3) we can conclude that Si and InGaAs sensors are equally suited for automated discrimination between parathyroid glands and surrounding adipose tissue or adjacent thyroid tissue. As tissue-classification accuracies of around 80% are far from clinically relevant, we also investigated combining the two spectral sensor ranges, which resulted in improved classification performance regarding differentiation of parathyroid tissue from surrounding adipose tissue.
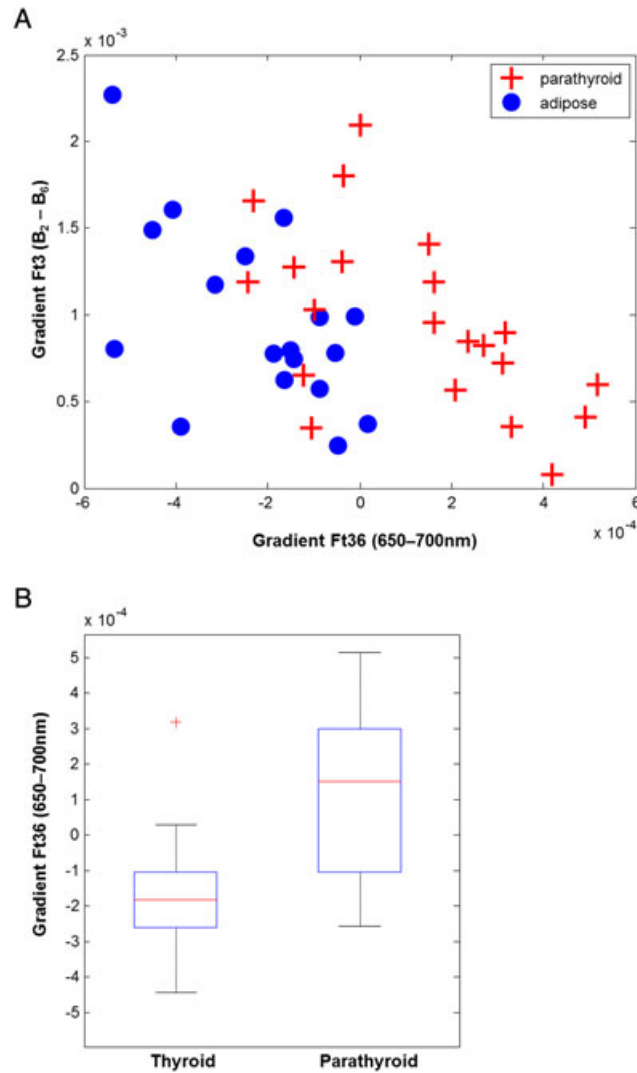
Figure 4. (A) Parathyroid versus adipose tissue: scatter plot of two selected features within Si-range. Each data point in this scatterplot represents the selected features (gradients Ft3 and Ft36) estimated using DRS data from 21 parathyroid and 18 adipose tissue sites. (B) Parathyroid versus thyroid: box plot of selected feature within Si-range. Box plot showing selected feature (gradient Ft36) estimated using DRS data from 23 thyroid and 21 parathyroid tissue sites

Table 2. Classification performance of selected Si-sensor features

LOO CV	TP	TN	Sensitivity	Specificity	PPV	NPV	Accuracy
Parathyroid – Adipose	15/21	16/18	71 (48–88)	89 (64–98)	88 (62–98)	73 (50–88)	79
Parathyroid – Thyroid	15/21	20/23	71 (48–88)	87 (65–97)	83 (58–96)	77 (56–90)	80

TP = true positive; TN = true negative ➔ numbers indicate identified tissue spots. A positive test is defined as the tissue observed being parathyroid gland; a negative test is defined as the tissue observed being adipose tissue / thyroid.
Sensitivity; specificity; PPV = positive predictive value; NPV = negative predictive value; accuracy ➔ numbers are percentages; numbers in parentheses indicate 95% confidence interval.
LOO CV = leave-one-out cross–validation.

Regarding the spectral signatures of the three tissue types (Figure 3) we observed a clear opposite course (i.e. ascending slope between 650 and 700 nm) for parathyroid compared with the spectra of adipose tissue and thyroid (i.e. descending slope). From this, feature Ft36 was established. Another blood-related feature
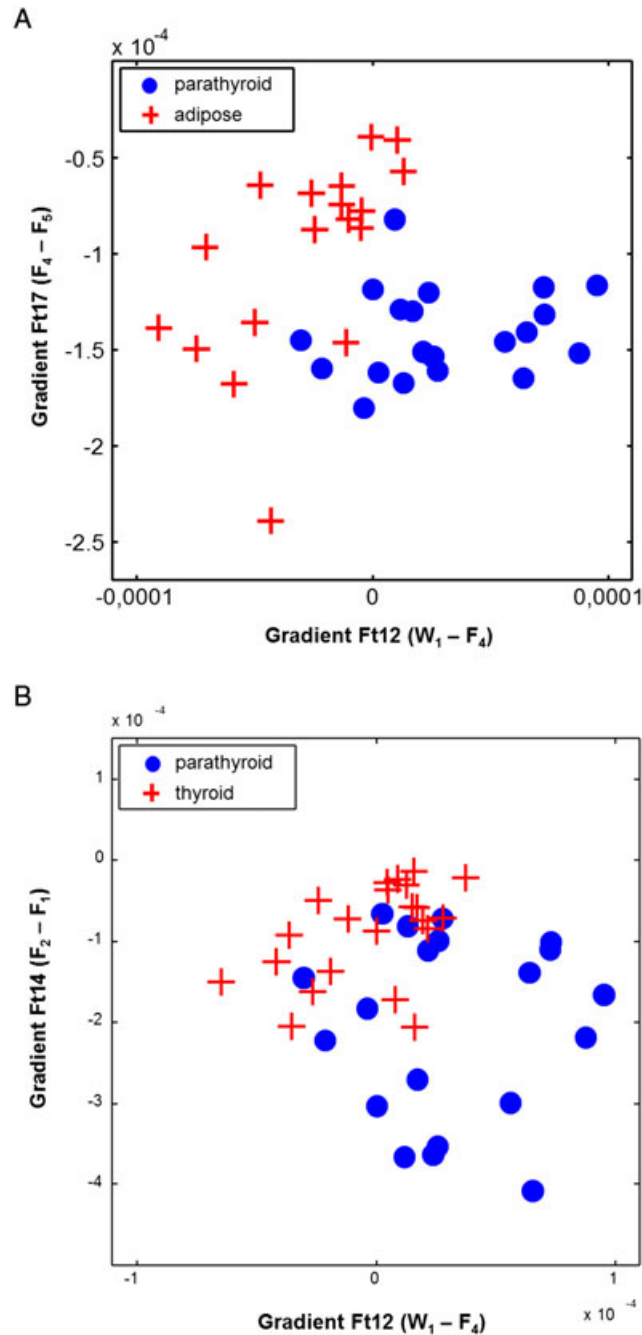
Figure 5. (A) Parathyroid versus adipose tissue: scatter plot of two selected features within InGaAs-range. Scatter plot showing two selected features (gradients Ft12 and Ft17) estimated using DRS data from 21 parathyroid and 18 adipose tissue sites. (B) Parathyroid versus thyroid tissue: scatter plot of two selected features within InGaAs-range. Scatter plot showing selected features (gradients Ft12 and Ft14) estimated using DRS data from 21 parathyroid and 23 thyroid tissue sites

(Ft3) in the Si sensor range was complementary for distinguishing parathyroid from surrounding adipose tissue (Figure 4). Within the InGaAs sensor range fat-related features (Figure 5, respectively, Ft12 and Ft14, and Ft12 and Ft17) performed best for enabling automated tissue classification. Combining Si and InGaAs sensor features resulted in improved tissue classification; this was especially profound when discriminating parathyroid from adipose tissue. The statistically different spectral features used for classification of parathyroid versus adipose tissue or thyroid are probably due to differences in chemical composition (e.g. water/lipid content) of the investigated tissues and structures.

Table 3. Classification performance of selected InGaAs-sensor features

LOO CV	TP	TN	Sensitivity	Specificity	PPV	NPV	Accuracy
Parathyroid – Adipose	19/21	13/18	90 (68–98)	72 (46–89)	79 (57–92)	87 (58–98)	82
Parathyroid – Thyroid	16/21	17/23	76 (52–91)	74 (51–89)	73 (50–88)	77 (54–91)	75

TP = true positive; TN = true negative ➜ numbers indicate identified tissue spots. A positive test is defined as the tissue observed being parathyroid gland; a negative test is defined as the tissue observed being adipose tissue / thyroid.
Sensitivity; specificity; PPV = positive predictive value; NPV = negative predictive value; accuracy ➜ numbers are percentages; numbers in parentheses indicate 95% confidence interval.
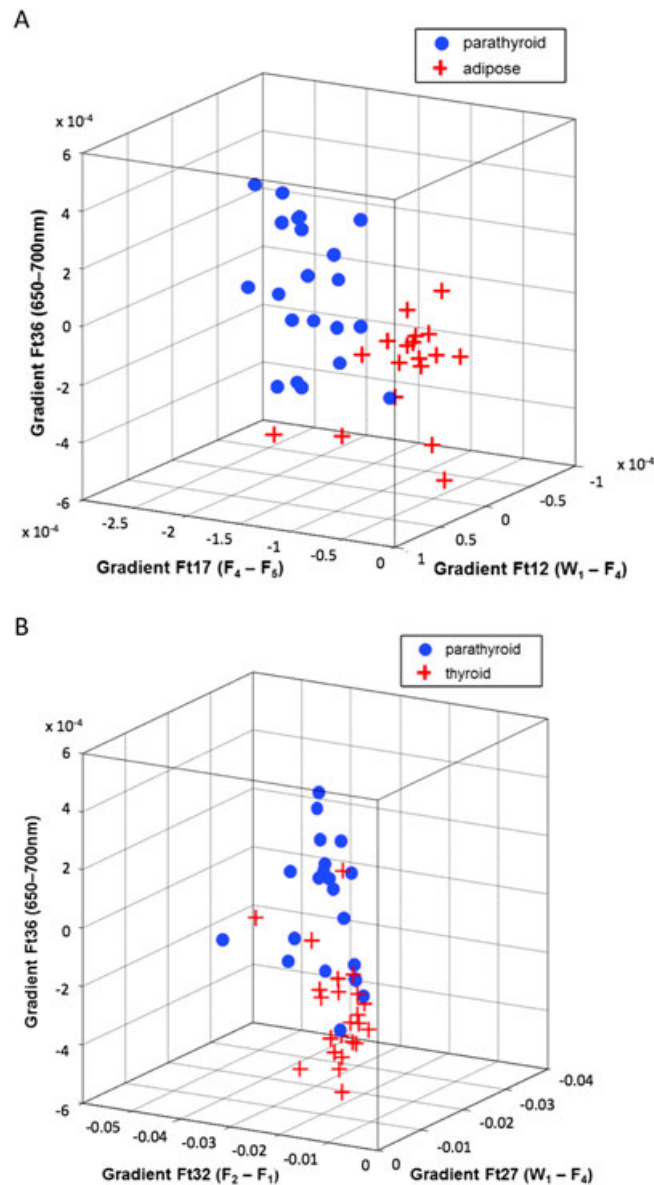LOO CV = leave-one-out cross–validation.

Figure 6. (A) Parathyroid versus adipose tissue: scatter plot of three selected features. Three-dimensional scatter plot showing selected features (gradients Ft12, Ft17 and Ft36) estimated using DRS data from 21 parathyroid and 18 adipose tissue sites. (B) Parathyroid versus thyroid tissue: scatter plot of three selected features. Three-dimensional scatter plot showing selected features (gradients Ft27, Ft32 and Ft36) estimated using DRS data from 21 parathyroid and 23 thyroid tissue sites

Table 4. Classification performance of combined Si-sensor and InGaAs-sensor features

LOO CV	TP	TN	Sensitivity	Specificity	PPV	NPV	Accuracy
Parathyroid – Adipose	21/21	17/18	100 (81–100)	94 (71–100)	95 (75–100)	100 (77–100)	97
Parathyroid – Thyroid	18/21	18/23	86 (63–96)	78 (56–92)	78 (56–92)	86 (63–96)	82

TP = true positive; TN = true negative ➔ numbers indicate identified tissue spots. A positive test is defined as the tissue observed being parathyroid gland; a negative test is defined as the tissue observed being adipose tissue / thyroid.
Sensitivity; specificity; PPV = positive predictive value; NPV = negative predictive value; accuracy ➔ numbers are percentages; numbers in parentheses indicate 95% confidence interval.
LOO CV = leave-one-out cross–validation.

Experimental methods such as near-infrared fluorescence imaging after peripheral infusion of aminolevulinic acid (8,9) or methylene blue (10) (i.e. exogenous contrast agents) could also be used for parathyroid gland or parathyroid adenoma localization. The reflectance spectra in our study, which were the basis to extract spectral features, originate from intrinsic tissue properties (i.e. endogenous chromophores) that do not require dye administration. Consequently there is no risk of potential toxicity or allergy to a contrast agent.

As the LOO cross-validation method inherently produces relatively optimistic classification results, external validation remains essential before classification models can be implemented in clinical practice (36). Such validation would need to be performed on newly acquired data in a multi-center study.

In this study, we use physiological knowledge on composition of different tissue types, and thereby use the pre-defined spectral regions of interest covering hemoglobin, water and fat. Tissue differentiation on the basis of completely automatically extracted features from a larger data set might achieve better results, and should be explored in future work. In addition, as the acquired reflectance spectra are specific to the probe geometry, more research is needed on the relation between these reflectance spectra and *in vivo* intrinsic tissue biological properties, which potentially could be captured by optical tissue properties. This is needed to obtain a better understanding of the nature of discrimination performance: i.e. whether parathyroid glands can be optically distinguished from surrounding adipose tissues or adjacent thyroid tissue based on differences in light scattering behaviour (related to structural differences) or due to differences in absorption behaviour (related to chromophore concentrations such as blood, water and fat). Such an approach would also be tissue-specific and robust to inter-patient and multi-centre variability.

A restriction of this study is that histopathological confirmation of the *in vivo* measured tissues was not possible, since these tissues were not removed from the body. Therefore, the gold standard used in this study was the visual judgment of an experienced surgeon with a track record of over 1000 thyroid surgeries. This judgment is based on color (spectral) information, on the recognition of anatomical position as well as on spatial structure of a specific tissue, and on palpation. For our technique to add most to this judgment, it should be extended from spot-wise probe-measurements to imaging of the complete surgical field of view.

In contrast to previous *ex vivo* experiments using wide-band diffuse reflectance spectroscopy (24), these measurements were performed during surgery. *In vivo* circumstances such as vascular filling and oxygenation were thus not disrupted.

Lymph node tissue was not included in this initial explorative study in the field of head and neck surgery, but certainly should be a subject of future research, as lymph node differentiation from parathyroid tissue can also be rather challenging during neck surgery.

In conclusion, in this study we have identified that Si and InGaAs sensors are quite accurate for automated spectroscopic classification of parathyroid glands and surrounding adipose tissue or adjacent thyroid tissue; but when applied separately not yet with clinically relevant levels of accuracy. Combination of these two sensor technologies, however, improves accuracy especially regarding parathyroid–adipose tissue classification. Note that an imaging system embodiment would provide spectral classification for each pixel, thus offering additional information about anatomical structures of the whole displayed image (which the present fibre probe spot measurements lack). Without such structural clues, also the human eye would be at loss for distinction. Hence, automated spectroscopic tissue classification in (para-)thyroid surgery seems feasible. Our next research will be to enable moving from single pixel contact mode to non-contact mode imaging. Having provided the proof that the hyperspectral imaging technique contains clinically relevant information opens a new path towards its integration in existing surgical tools like camera systems for open surgery or even endoscopes or robot-assisted systems. In addition, a clinically relevant camera system needs further development in terms of its physical size, energy consumption, total cost and operational speed. Recognizing this potential, the European Union has funded two research projects on surgical hyperspectral imaging: HELICoiD (HypErspectraL Imaging Cancer Detection) (37), and EXIST (Extended Image Sensing Technologies) (38).

Acknowledgements

Conflicts of interest

None declared.

Ethical approval

Prior to measurements, the local institutional review board of Maastricht University Medical Center (registration number METC 10-4-035) granted approval and preoperative written informed consent was obtained from all patients.

References

1. Mohebati A, Shaha AR. Anatomy of thyroid and parathyroid glands and neurovascular relations. *Clin Anat* 2012; **25**: 19–31.
2. Christou N, Mathonnet M. Complications after total thyroidectomy. *J Visc Surg* 2013; **150**: 249–256.
3. Lefevre JH, Tresallet C, Leenhardt L, *et al.* Reoperative surgery for thyroid disease. *Langenbecks Arch Surg* 2007; **392**: 685–691.
4. Hahn LD, Hoyt C, Rimm DL, Theoharis C. Spatial spectral imaging as an adjunct to the Bethesda classification of thyroid fine-needle aspiration specimens. *Cancer Cytopathol* 2013; **121**: 162–167.
5. Mansoor I, Zalles C, Zahid F, *et al.* Fine-needle aspiration of follicular adenoma versus parathyroid adenoma: the utility of multispectral imaging in differentiating lesions with subtle cytomorphologic differences. *Cancer* 2008; **114**: 22–26.
6. Das K, Stone N, Kendall C, *et al.* Raman spectroscopy of parathyroid tissue pathology. *Lasers Med Sci* 2006; **21**: 192–197.
7. Suh H, A'Amar O, Rodriguez-Diaz E, *et al.* Elastic light-scattering spectroscopy for discrimination of benign from malignant disease in thyroid nodules. *Ann Surg Oncol* 2011; **18**: 1300–1305.
8. Prosst RL, Willeke F, Schroeter L, *et al.* Fluorescence-guided minimally invasive parathyroidectomy: a novel detection technique for parathyroid glands. *Surg Endosc* 2006; **20**: 1488–1492.
9. Prosst RL, Weiss J, Hupp L, *et al.* Fluorescence-guided minimally invasive parathyroidectomy: clinical experience with a novel intraoperative detection technique for parathyroid glands. *World J Surg* 2010; **34**: 2217–2222.
10. van der Vorst JR, Schaafsma BE, Verbeek FP, *et al.* Intraoperative near-infrared fluorescence imaging of parathyroid adenomas using low-dose methylene blue. *Head Neck* 2014; **36**: 853–858.
11. Paras C, Keller M, White L, *et al.* Near-infrared autofluorescence for the detection of parathyroid glands. *J Biomed Opt* 2011; **16**: 067012.
12. Ladurner R, Hallfeldt KK, Al Arabi N, *et al.* Optical coherence tomography as a method to identify parathyroid glands. *Lasers Surg Med* 2013; **45**: 654–659.
13. Conti de Freitas LC, Phelan E, Liu L, *et al.* Optical coherence tomography imaging during thyroid and parathyroid surgery: a novel system of tissue identification and differentiation to obviate tissue resection and frozen section. *Head Neck* 2014; **36**: 1329–1334.
14. Lu G, Fei B. Medical hyperspectral imaging: a review. *J Biomed Opt* 2014; **19**: 10901.
15. Holzer MS, Best SL, Jackson N, *et al.* Assessment of renal oxygenation during partial nephrectomy using hyperspectral imaging. *J Urol* 2011; **186**: 400–404.
16. Olweny EO, Faddegon S, Best SL, *et al.* Renal oxygenation during robot-assisted laparoscopic partial nephrectomy: characterization using laparoscopic digital light processing hyperspectral imaging. *J Endourol* 2013; **27**: 265–269.
17. Zuzak KJ, Naik SC, Alexandrakis G, *et al.* Intraoperative bile duct visualization using near-infrared hyperspectral video imaging. *Am J Surg* 2008; **195**: 491–497.
18. Wieringa FP, Mastik F, Duncker DJ, Bogers AJ. Contrast enhancement of coronary arteries in cardiac surgery: a new multispectral stereoscopic camera technique. *Euro Intervention* 2006; **2**: 389–394.
19. Panasyuk SV, Yang S, Faller DV, *et al.* Medical hyperspectral imaging to facilitate residual tumor identification during surgery. *Cancer Biol Ther* 2007; **6**: 439–446.
20. Stelzle F, Zam A, Adler W, *et al.* Optical nerve detection by diffuse reflectance spectroscopy for feedback controlled oral and maxillofacial laser surgery. *J Transl Med* 2011; **9**: 20.
21. Schols RM, ter Laan M, Stassen LP, *et al.* Differentiation between nerve and adipose tissue using wide-band (350–1830 nm) *in vivo* diffuse reflectance spectroscopy. *Lasers Surg Med* 2014; **46**: 538–545.
22. Nachabe R, Hendriks BH, Desjardins AE, *et al.* Estimation of lipid and water concentrations in scattering media with diffuse optical spectroscopy from 900 to 1600 nm. *J Biomed Opt* 2010; **15**: 037015.
23. Nachabe R, Hendriks BH, van der Voort M, *et al.* Estimation of biological chromophores using diffuse optical spectroscopy: benefit of extending the UV-VIS wavelength range to include 1000 to 1600 nm. *Biomed Opt Express* 2010; **1**: 1432–1442.
24. Schols RM, Dunias P, Wieringa FP, Stassen LP. Multispectral characterization of tissues encountered during laparoscopic colorectal surgery. *Med Eng Phys* 2013; **35**: 1044–1050.
25. Cao Q, Zhegalova NG, Wang ST, *et al.* Multispectral imaging in the extended near-infrared window based on endogenous chromophores. *J Biomed Opt* 2013; **18**: 101318.
26. Akbari H, Kosugi Y, Kojima K, Tanaka N. Blood vessel detection and artery-vein differentiation using hyperspectral imaging. *Conf Proc IEEE Eng Med Biol Soc* 2009; **2009**: 1461–1464.
27. Zijlstra WG, Buursma A, Meeuwsen-van der Roest WP. Absorption spectra of human fetal and adult oxyhemoglobin, deoxyhemoglobin, carboxyhemoglobin, and methemoglobin. *Clin Chem* 1991; **37**: 1633–1638.
28. Anderson RR, Farinelli W, Laubach H, *et al.* Selective photothermolysis of lipid-rich tissues: a free electron laser study. *Lasers Surg Med* 2006; **38**: 913–919.
29. Roggan A, Friebel M, Do Rschel K, *et al.* Optical properties of circulating human blood in the wavelength range 400–2500 nm. *J Biomed Opt* 1999; **4**: 36–46.
30. Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory, 1992: 144–1452.
31. Tabachnick BG, Fidell LS. *Using Multivariate Statistics* (5th edn). Allyn-Bacon: Boston, MA, 2007.
32. Mercier G, Lennon M. Support vector machines for hyperspectral image classification with spectral-based kernels. Geoscience and

Remote Sensing Symposium, 2003 IGARSS '03 Proceedings 2003 IEEE International. 2003; **1**: 288–290.

33. Stelzle F, Tangermann-Gerk K, Adler W, *et al.* Diffuse reflectance spectroscopy for optical soft tissue differentiation as remote feedback control for tissue-specific laser surgery. *Lasers Surg Med* 2010; **42**: 319–325.

34. Stelzle F, Adler W, Zam A, *et al. In vivo* optical tissue differentiation by diffuse reflectance spectroscopy: preliminary results for tissue-specific laser surgery. *Surg Innov* 2012; **19**: 385–393.

35. Troy TL, Thennadil SN. Optical properties of human skin in the near infrared wavelength range of 1000 to 2200 nm. *J Biomed Opt* 2001; **6**: 167–176.

36. Bleeker SE, Moll HA, Steyerberg EW, *et al.* External validation is necessary in prediction research: a clinical example. *J Clin Epidemiol* 2003; **56**: 826–832.

37. http://cordis.europa.eu/project/rcn/111274_en.html (accessed 14 February 2016).

38. http://cordis.europa.eu/project/rcn/198017_en.html (accessed 14 February 2016).