OXFORD

Breast carcinoma detection in *ex vivo* fresh human breast surgical specimens using a fast slide-free confocal microscopy scanner: HIBISCUSS project

Angelica Conversano[1,*] (ID), Muriel Abbaci[2,3] (ID), Paul van Diest[4], Aurélie Roulot[1], Giuseppe Falco[5], Malek Ferchiou[6], Saverio Coiro[5], Milan Richir[7], Pierre-Michel Genolet[8], Carine Clement[8], Odile Casiraghi[6], Aicha Ben Lahkdar[9], Nizard Labaied[6], Moira Ragazzi[10,11] and Marie-Christine Mathieu[6]

[1]Gustave Roussy, Département de Chirurgie, Université Paris-Saclay, Villejuif, France
[2]Gustave Roussy, Plate-forme Imagerie et Cytométrie, UMS 23/3655, Université Paris-Saclay, Villejuif, France
[3]Université Paris-Saclay, CEA, CNRS, Inserm, Laboratoire d'Imagerie Biomédicale Multimodale Paris Saclay, Orsay, France
[4]Department of Pathology, University Medical Center Utrecht, Utrecht, The Netherlands
[5]Department of Oncology and Advanced Technologies, Breast Surgery Unit, IRCSS Santa Maria Nuova Hospital, Reggio Emilia, Italy
[6]Gustave Roussy, Département de pathologie, Université Paris-Saclay, Villejuif, France
[7]Department of Surgery, University Medical Centre Utrecht, Utrecht, The Netherlands
[8]Espacegyneco, Pully, Switzerland
[9]SCM Bichat, Paris, France
[10]Pathology Unit, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy
[11]Department of Medical and Surgical Sciences for Children and Adults, University of Modena and Reggio Emilia, Modena, Italy

Correspondence to: Angelica Conversano, Gustave Roussy, Département de Chirurgie, Université Paris-Saclay, 114 rue Edouard Vaillant, 94805 Villejuif, France (e-mail: angelica.conversano@gustaveroussy.fr)

Abstract

Background: New generation ultra-fast fluorescence confocal microscopy allows the *ex vivo* intraoperative analysis of fresh tissue. The High resolution Imaging for Breast carcInoma detection in *ex vivo* Specimens after breast Conserving sUrgery by hiStolog Scanner (HIBISCUSS) project aimed to develop an online learning program to recognize the main breast tissue features on ultra-fast fluorescence confocal microscopy images and to evaluate the performance of surgeons and pathologists in diagnosing cancerous and non-cancerous breast tissue in ultra-fast fluorescence confocal microscopy images.

Methods: Patients who underwent conservative surgery or mastectomy for breast carcinoma (invasive or in situ lesions) were included. The fresh specimens were stained with a fluorescent dye and imaged using a large field-of-view (20 cm^2) ultra-fast fluorescence confocal microscope.

Results: One hundred and eighty-one patients were included. The images from 55 patients were annotated to generate learning sheets and images from 126 patients were blindly interpreted by seven surgeons and two pathologists. The time for tissue processing and ultra-fast fluorescence confocal microscopy imaging was between 8 and 10 min. The training program was composed of 110 images divided into nine learning sessions. The final database for blind performance assessment comprised 300 images. The mean duration for one training session and one performance round was 17 and 27 min respectively. The performance of pathologists was almost perfect with 99.6 per cent (standard deviation (s.d.) 5.4 per cent) accuracy. Surgeons' accuracy significantly increased ($P = 0.001$) from 83 per cent (s.d. 8.4 per cent) in round 1 to 98 per cent (s.d. 4.1 per cent) in round 7 as well as the sensitivity ($P = 0.004$). Specificity increased without significance from 84 per cent (s.d. 16.7 per cent) in round 1 to 87 per cent (s.d. 16.4 per cent) in round 7 ($P = 0.060$).

Conclusion: Pathologists and surgeons showed a short learning curve in differentiating breast cancer from non-cancerous tissue in ultra-fast fluorescence confocal microscopy images. Performance assessment for both specialties supports ultra-fast fluorescence confocal microscopy evaluation for intraoperative management.

Registration number: NCT04976556 (http://www.clinicaltrials.gov).

Introduction

Breast-conserving surgery associated with radiotherapy is the most frequently performed and safest long-term surgical treatment for women with early-stage breast cancer[1]. In such a surgical treatment, the surgeon needs to find the balance between removing the smallest possible breast lump containing the malignancy for optimal cosmesis and yet achieving cancer-free margins. This inevitably leads to a proportion of incomplete resections that may require re-excision, which has physical, psychological and economic sequelae and is associated with a higher incidence of postoperative complications. Furthermore, it may delay the administration of adjuvant therapy, and has been associated with an elevated risk of local

and distant disease relapse[2]. A meta-analysis suggested that the odds of local recurrence were 2.42 for positive *versus* negative margins in women undergoing breast-conserving surgery[3]. Considering the consensus of 2-mm-free margins for intraductal tumour[4] and no tumour at the inked margin in patients with invasive mammary carcinoma[3], at least 20 per cent of patients undergo more than one procedure to achieve acceptable margins as part of a breast-conserving strategy[5].

The development of reliable intraoperative assessment techniques would be an asset during the primary surgery to indicate promptly if the tumour bed requires additional re-excision. In the literature, the probe-based confocal laser endomicroscope (pCLE) for *in vivo* imaging was the only confocal system described for use in the operating room and is mainly applied in gastroenterology after fluorescent dye injection[6]. Proof of concept with pCLE on breast cancer imaging has been performed on *ex vivo* breast surgical specimens ($n = 13$) but the study was not followed by a clinical trial[7]. A confocal microscope (Vivascope, Munich, Germany) designed for *ex vivo* imaging has been used for various applications including breast cancer but its imaging time ($1\,cm^2$ in 4 min) is hardly compatible with a routine assessment of large specimens[8]. The Histolog Scanner (SamanTree Medical, Lausanne, Switzerland) is a new ultra-fast confocal microscope (UFCM) designed for operating room *ex vivo* imaging of large tissue specimens with high resolution and speed (<1 min for $20\,cm^2$) to guide intraoperative assessment and offer support in clinical decision-making. Recent studies in dermatology and breast cancer have provided the first supporting data for this new imaging device[9–13].

The motivation of the international project HIBISCUSS (High resolution Imaging for Breast carcInoma detection in *ex vivo* Specimens after breast Conserving sUrgery by hiStolog Scanner) was to identify a quicker and more effective intraoperative assessment of breast surgical specimens, regardless of centres' clinical workflow and intraoperative availability of a pathologist. The main objective was to demonstrate that both pathologists and surgeons can identify breast cancer tissue in UFCM images of breast surgical specimens. In this article, the authors present the design of an online training program available free of charge for clinicians, based on the data of 55 patients presenting the most frequent breast cancerous and non-cancerous features in UFCM images. They then describe the performance of the blind assessment by seven surgeons and two pathologists to differentiate cancerous from non-cancerous breast tissue in UFCM images from 126 patients.

Methods

Study design

The HIBISCUSS project was based on close international and multidisciplinary interactions with surgeons, pathologists and researchers specialized in photonics at Gustave Roussy (GR) Hospital (France), Reggio Emilia Hospital (Italy), University Medical Center of Utrecht (The Netherlands), Gynaecologic Centre of Pully (Switzerland) and the SamanTree Medical company (Switzerland).

The HIBISCUSS project was a prospective non-interventional preclinical single-arm study performed on 181 *ex vivo* human specimens from 181 patients following breast surgery for an untreated primary breast carcinoma planned at Gustave Roussy Hospital from June 2019 to January 2021.

The study was composed of three phases (*Fig. 1*): collection of specimen images and patient data ($n = 181$ patients); training of physicians (surgeons and pathologists) ($n = 55$ patients); evaluation of physicians' performance to blindly distinguish non-cancerous and cancerous tissue ($n = 126$ patients).

Patient selection

Patients from a single centre (GR) presenting preoperative histological diagnosis of invasive carcinoma of no special type (IC-NST), ductal carcinoma *in situ* (DCIS) or invasive lobular carcinoma (ILC) and undergoing mastectomy or lumpectomy without neoadjuvant chemotherapy were selected.

Surgical procedure and imaging protocol

Breast surgery was scheduled and performed according to the standard of care of GR. Surgical specimens were oriented with physical marks following standard surgical procedures and sent to the pathology department. Upon receipt of the specimen, an anonymized number was automatically assigned to each patient.
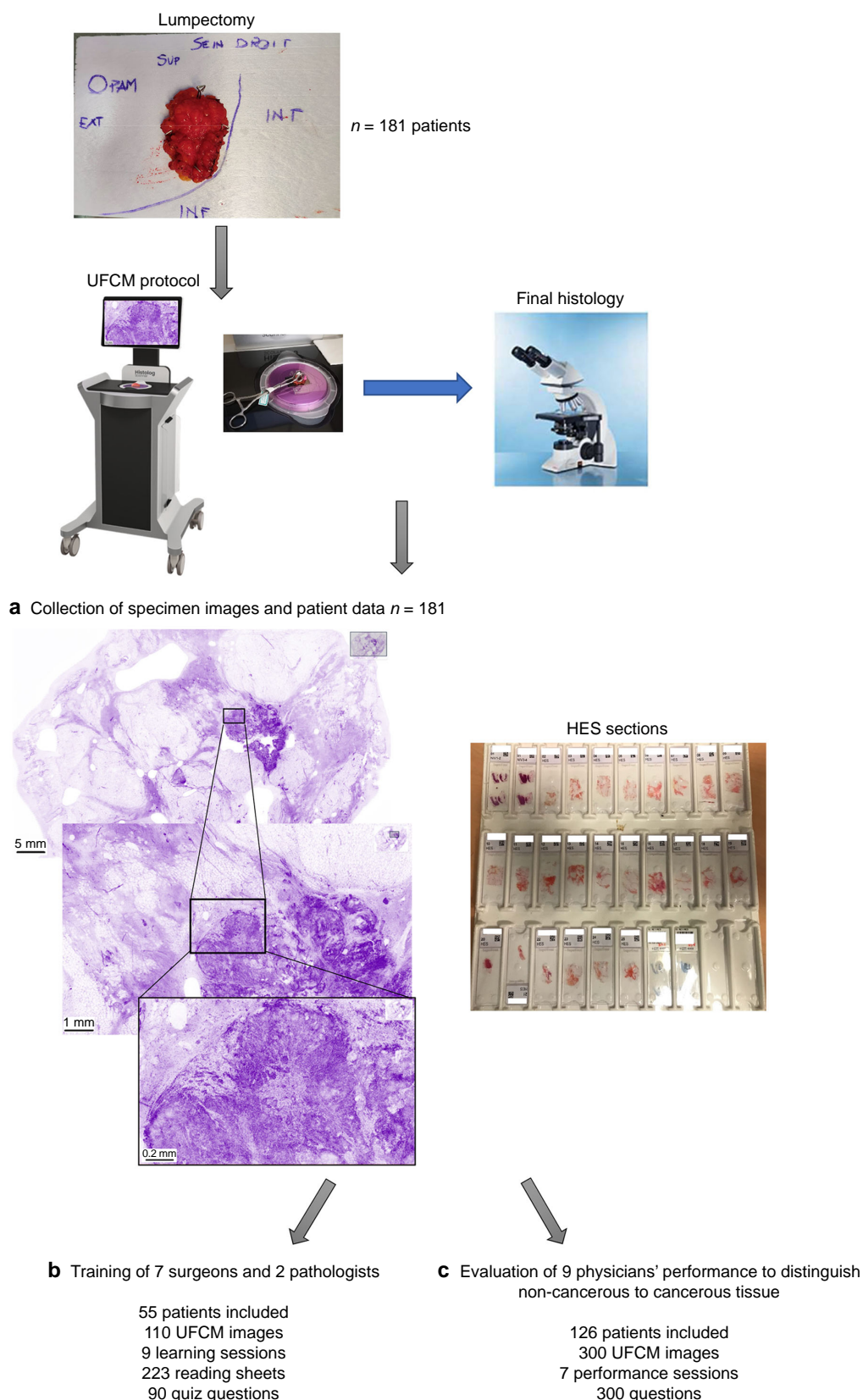
Specimen imaging

For lumpectomy specimens, a senior pathologist placed two single-use coloured plastic clips on the upper and outer locations to ensure the maintenance of anatomical orientation during tissue handling. The use of clips was mandatory as conventional inks could not be applied before UFCM imaging due to interferences with fluorescent dye light absorption. No freezing or fixing was a precondition for UFCM imaging. The fresh specimen was then sectioned into two parts according to a frontal plane: superficial half and deep half. Some fresh thick specimens were cut into three parts: superficial third, middle third and deep third. For fresh mastectomy specimens, the tissue was sectioned by a senior pathologist and a $4\,cm^2$ tumour tissue sample was collected in the tumour core for UFCM imaging. Intraoperative macroscopic and/or microscopic examination was performed before UFCM imaging if requested by the standard of care.

After drying with a surgical pad, the specimen was stained with the dedicated contrast agent Histolog Dip (SamanTree Medical, Lausanne, Switzerland), a diluted solution of the fluorescent dye acridine orange, which is a cell-permeant nucleic acid binding dye that emits green fluorescence when bound to DNA (λ excitation = 502 nm and λ emission = 525 nm). Previous studies have successfully described acridine orange dye for *ex vivo* microscopy for histologic diagnosis[9,10]. Consequently, Histolog Dip was optimized for imaging with the Histolog Scanner and certified for diagnostic use on human tissue specimens. The specimens were soaked for 10 s in Histolog Dip, then rinsed for 5 s in 0.9 per cent sodium chloride solution and dried on a surgical pad.

Next, each sample was placed over the imaging window. A surgical instrument could be gently applied onto the tissue to flatten out the imaged surface and optimize the imaging process. All sections of the lumpectomy specimen or sample from mastectomy were scanned with the UFCM (Histolog Scanner, SamanTree Medical, Lausanne, Switzerland).

The Histolog Scanner is an IVD (in-vitro diagnostic device) CE-marked fluorescence confocal microscope designed for use in a clinical setting on large tissue specimens. The Histolog Scanner is housed in a compact setup: (0.76 m × 0.76 m × 1.56 m). It is a stand-alone device that can be disinfected, and integrates a touch screen monitor to display the fluorescence confocal images and control device operation. The UFCM consists of a

Lumpectomy

n = 181 patients

UFCM protocol

Final histology

a Collection of specimen images and patient data *n* = 181

5 mm

1 mm

0.2 mm

HES sections

b Training of 7 surgeons and 2 pathologists

55 patients included
110 UFCM images
9 learning sessions
223 reading sheets
90 quiz questions

c Evaluation of 9 physicians' performance to distinguish non-cancerous to cancerous tissue

126 patients included
300 UFCM images
7 performance sessions
300 questions

Fig. 1 Synopsis of the study: from breast specimen collection to ultra-fast confocal microscopy images performance assessment
a A total of 181 patients were included in the HIBISCUSS project and the corresponding breast surgical specimens (lumpectomy or samples from mastectomy) were imaged by the ultra-fast confocal microscope (UFCM). Tissues were then processed for further conventional histology (haematoxylin, eosin and saffron (HES) sections) and final diagnosis was collected. **b** In total, 110 UFCM images from 55 patients were used to develop a training material presented to seven surgeons and two pathologists. **c** Finally, 300 UFCM images from 126 patients were interpreted by the nine physicians in a blind performance assessment study. HIBISCUSS, High resolution Imaging for Breast carcInoma detection in *ex vivo* Specimens after breast Conserving sUrgery by hiStolog Scanner.

laser diode operating at 488 nm for fluorophore excitation combined with fluorescence signal collection above 500 nm. The confocal device can eliminate the 'out-of-focus' signal at high magnifications, here from optical ×0.5 to digital ×40. The acquired images correspond to a field-of-view of 20 cm^2 with a lateral resolution of up to 2 µm and an axial resolution of 30 µm. The images are acquired at a single depth of 20 µm below the surface. The UFCM provides seamless images at once without additional image postprocessing. A fast-imaging mode (Preview) provides an image of lower resolution within 7 s and a second mode (Acquire) builds a high-resolution image within 50 s independently of the sample size. The Preview images thus enable the correct positioning of the tissue above the optical sensor, before launching the Acquire mode to provide high detail images for histological assessment.

The fluorescence images are displayed with an artificial purple colouring of the grey values, resembling the result of a mono reagent colouration such as haematoxylin observed by pathologists in conventional frozen sections, and were provided as *en face* view on a touchscreen monitor with magnification from ×0.5 to ×40 (*Fig. 1*).

For lumpectomy specimens larger than the UFCM imaging window, the tumour seen macroscopically was positioned at the centre of the window. Surgical specimens of non-palpable lesions were positioned on the imaging window based on the radiography of the specimen to locate the area containing the abnormality.

Histopathology

After completion of UFCM imaging, each sample underwent routine histopathology processing including margin identification with inks, drawing of the tissue specimen and 10 per cent buffered formalin fixation. Specimens were then transferred to standard tissue cassettes after macroscopic routine handling, embedded in paraffin, sectioned at 3 µm and stained with haematoxylin, eosin and saffron (HES). Histological sections were then scanned at 20× magnification with a NanoZoomer S210 Scanner (Hamamatsu Photonics, Massy, France) for data analysis and archiving. The NDPI (Hamamatsu NanoZoomer Digital Pathology Image file) format was converted into Tiff format with the ImageJ plugin 'NDPI tools' for subsequent analysis.

Final histopathological reports, location of tissue samples embedded in paraffin, and digital whole slide images of HES-stained sections were shared with pathologists for analysis and annotations of UFCM images.

Image annotations by a board of expert pathologists for the training program

A board of three senior breast pathologists with expertise in confocal/digital pathology (M.R., M.N. and P.J.vD.) independently assessed the UFCM images with the support of the corresponding digital HES sections and clinical data. Annotations were made in the UFCM images using dedicated Viewer software that allows the adjustment of the brightness and the switching from white and black to coloured purple mode. Each pathologist was asked to annotate both cancerous and non-cancerous regions-of-interest (ROIs) (at least two cancerous ROIs and one non-cancerous ROI for each image) that they considered most representative of malignant and normal breast tissue. Then, online meetings were periodically organized to allow pathologists to discuss their identification of ROIs. Only the ROIs reaching a consensus agreement of the three

pathologists were included in the reading sheets of the training and assessment program. Normal tissue areas were subclassified as fatty tissue, duct, lobule, vessel, connective tissue, nerve, inflammation or muscle, and cancers such as IC-NST, DCIS or ILC. Typical artefacts in UFCM images were also identified by the board of pathologists.

Reading and self-assessment sheets

Reading sheets were designed to present the ROIs of different UFCM images grouped into learning sessions. Each sheet focuses on a specific tissue pattern (for example fatty tissue, DCIS) identified with a high level of confidence by the board of pathologists. Each UFCM ROI was presented with short descriptions at different levels of magnification: low magnification corresponding to a complete UFCM field-of-view (5 per cent zoom level); intermediate magnification level to present tissue architecture (25 per cent zoom level); highest optical magnification to present cellular details (100 per cent zoom level). A final review of the reading sheets was performed by an independent senior pathologist expert in confocal imaging (A.B.L.) to ensure consistency of the content. Each learning session was focused on a specific topic such as normal ductal features or ILC and the sessions progressed from low complexity identification of normal parts (that is fatty tissue) to high complexity identification of cancer parts (that is ILC). The content of these reading sheets will be made available on the Zenodo free repository in accordance with European Union's Horizon 2020 research and innovation program guidelines (https://zenodo.org/).

Training session

The training program consisting of nine learning sessions was available on a web-based interface accessible throughout the training interval with a standard computer browser. Each training session was divided into two parts: 15 reading sheets containing annotated UFCM images and corresponding descriptive paragraphs; 9 or 10 questions per training session are presented to evaluate the understanding of the lesson in a self-assessment process. The time to complete the self-assessment sheets was monitored. Once completed, the accuracy of the answers was calculated and shared with the user with the possibility to review the correct answers (feedback) to promote self-training on the UFCM images. A lesson has to be completed to follow the next one. Two pathologists and seven surgeons undertook the training program to develop their expertise in interpreting UFCM images: an experienced pathologist in the histological assessment of breast cancer, an experienced pathologist in the histological assessment of breast cancer with prior experience in head and neck endomicroscopy[14,15] (both from GR, M.C.M. and M.F.) and seven surgeons without prior experience in microscopic optical imaging (two from France: A.C. and A.R., two from Italy: G.F. and S.C., one from The Netherlands: M.R., two from Switzerland: C.C. and P.M.G.).

Evaluation of physicians' performance on UFCM images

The performance assessment program was available on a web-based interface with a standard computer browser accessible during the interval of blind assessments. All nine trained physicians performed seven blind assessment rounds, almost one per month, in which 300 UFCM ROIs from 126 patients were randomly assigned. In each round, 45 ROIs presented at the three zoom levels (5 per cent, 25 per cent, 100 per cent of zoom level) were classified by each physician according to diagnosis interpretation. Clinical data including

macroscopic specimen imaging were not provided to the physicians during the evaluation interval so as not to influence UFCM image interpretation. Answers were interactively stored on the browser and the time to answer was monitored.

First, pathologists and surgeons had to classify the UFCM images into three groups: cancerous, non-cancerous or 'uncertain diagnosis'. Subsequently, an additional level of interpretation was requested only for pathologists to evaluate if breast cancer type was recognized that is IC-NST, ILC or DCIS. After evaluating each set of images, the overall accuracy and the correct diagnosis were revealed to the physicians to promote additional learning.

Statistical analysis

Results for seven review rounds (each containing up to 45 images) were extracted from original data files into a Microsoft® Excel workbook with one worksheet per round structured as a matrix (one row per image, one column per physician) together with a summary for each physician of the number and percentage of correct responses, incorrect responses, true positives, false positives, true negatives and false negatives. The study pathologists' reported cancer subtype was listed when it differed from the reference diagnostic.

The analysis covers the accuracy (agreement), sensitivity, specificity, and inability to answer for each surgeon and pathologist on each image summarized by study round and by individual physician. The evolution of individual surgeon's and pathologist's agreement, sensitivity and specificity by round was analysed using random-effects logistic regression models for binary data accounting for correlations between physicians who read the same set of images (Stata v 13.1, StataCorp, College Station, TX, USA).

Trends in performance were assessed by dividing the seven rounds into three predefined intervals covering two early (rounds 1 and 2), three middle (rounds 3, 4 and 5) and two late (rounds 6 and 7) intervals. Tests for differences between the three time intervals and for trends were assessed from the logistic regression model with *P* values computed using likelihood ratio tests. The performance criterion stated in the statistical analysis plan was 'After training, clinical centre surgeons and pathologists are expected to be able to correctly classify at least 90 per cent of the images from the Histolog Scanner confirmed by the senior reference pathologists' classification (accuracy).' The study was powered for a total of 300 images and a 10 per cent non-inferiority margin for the lower limit of the 95 per cent c.i. (that is lower two-sided 95 per cent confidence limit to be no less than 80 per cent).

Legal and regulatory aspects

The *ex vivo* HIBISCUSS project was approved by the institutional review board and registered as a non-interventional study to the French 'Institut National des données de santé' and on clinicaltrials.gov: NCT04976556. According to French Good Clinical Practices guidelines for non-interventional studies performed on *ex vivo* human tissues, written consent was not necessary and only an informative note had to be given to patients during preoperative visits and tracked into patient records. This study was conducted in compliance with the protocol, the current version of the Declaration of Helsinki, the ICH-GCP (International Council for Harmonisation–Good Clinical Practice) or ISO EN 14155 (as far as applicable), as well as all national legal and regulatory requirements.

Results

Collection of specimen images and patient data

Fifty-five patients with lumpectomy were imaged for the training phase but only 49 of 55 surgical specimens were diagnosed as having carcinoma (31 IC-NST, 11 ILC, 7 DCIS) at final histological examination. Then, 112 patients with lumpectomy and 14 patients with mastectomy were included in the performance assessment phase [41 IC-NST, 41 ILC, 40 DCIS, 1 lobular carcinoma *in situ* (LCIS), 3 without cancer]. All fresh tumour samples were correctly imaged with the UFCM. The duration for processing the two to three sections of the breast specimen using UFCM image acquisition was 8–10 min. Histopathological analysis was not compromised by the UFCM protocol since no remarkable changes were noted in the HES sections. Overall, the imaging process did not negatively impact the standard of care of specimen workflow.

Training of physicians (surgeons and pathologists)

A total of 110 UFCM images were acquired at high resolution (an average of two per specimen) and implemented in the training program. Cancerous tissue was not identified by the board of expert pathologists in 48 UFCM images (43 per cent) mainly due to carcinoma resection by preoperative biopsy or due to the location of the tumour that was not cut by the central sectioning of the specimen. For the other 62 UFCM images containing cancer patterns, 24 per cent were IC-NST, 14 per cent were DCIS, 35 per cent were both IC-NST and DCIS, and 27 per cent were ILC.

A total of nine learning sessions were planned: three for non-tumoural conditions, corresponding to fatty and connective tissue, lobules and ducts/vessels, inflammation and others; two for each cancer type (IC-NST, DCIS and ILC). Overall, 223 sheets were prepared for the training. Each session contained 15 reading sheets (*n* = 135) and an average of 9.7 self-assessment examination sheets (*n* = 88). Fifty-four per cent of the sheets were focused on non-cancerous features with the following distribution: 12 sheets on fatty tissue, 31 on connective tissue, 27 on ducts and vessels, 30 on lobules, 21 on inflammation and 2 on muscle. The remaining 46 per cent of the sheets were focused on cancerous patterns with the following distribution: 39 sheets on IC-NST, 31 on DCIS and 30 on ILC.

Table 1 Composition of the performance assessment program divided into seven rounds of image interpretation

|	Round 1	Round 2	Round 3	Round 4	Round 5	Round 6	Round 7	Total
Normal	15	15	15	15	13	13	14	100
IC-NST	11	10	10	10	9	9	5	64
DCIS	10	10	10	10	8	9	10	67
ILC	9	10	10	10	9	8	13	69
Total	45	45	45	45	39	39	42	300

IC-NST, invasive carcinoma of no special type; DCIS, ductal carcinoma *in situ*; ILC, invasive lobular carcinoma.
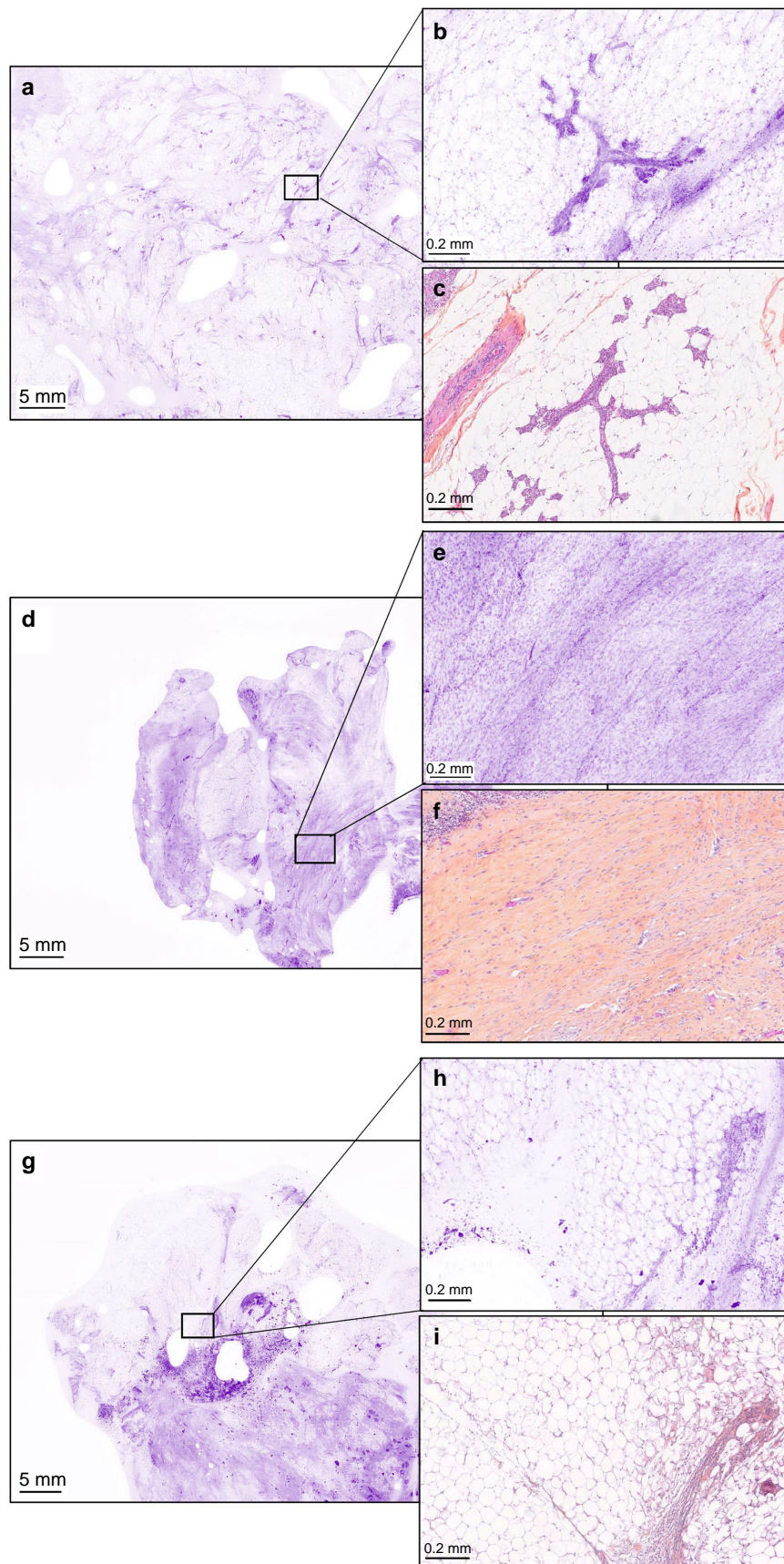
Fig. 2 Typical non-cancerous ultra-fast confocal microscopy images correctly identified by all seven surgeons and two pathologists at low magnification (lumpectomy section) and high magnification with corresponding haematoxylin, eosin and saffron sections (a, b, d, e, g, h UFCM images; c, f, i HES slides)
(a–c) Lobules (in the inset). (d–f) Fibrosis (in the inset). (g–i) Fatty tissue and inflammatory cells (in the inset). UFCM, ultra-fast confocal microscope; HES, haematoxylin, eosin and saffron.
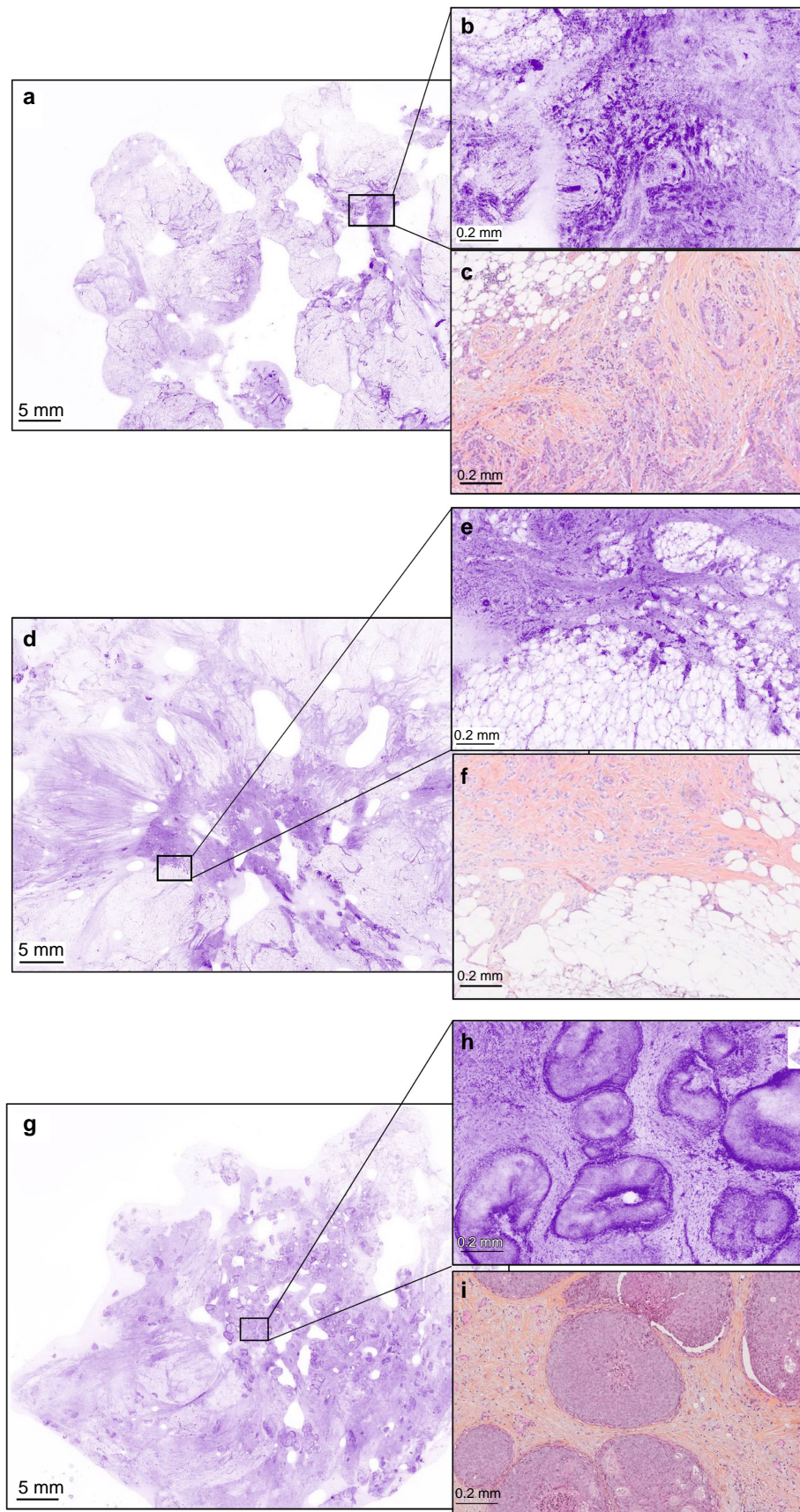
Fig. 3 Typical cancerous ultra-fast confocal microscopy images correctly identified by all the surgeons and pathologists at low magnification (lumpectomy section) and high magnification with corresponding haematoxylin, eosin and saffron sections (a, b, d, e, g, h UFCM images; c, f, i HES slides)
(**a–c**) Invasive carcinoma of no special type. (**d–f**) Invasive lobular carcinoma. (**g–i**) Ductal carcinoma *in situ*. UFCM, ultra-fast confocal microscope; HES, haematoxylin, eosin and saffron.
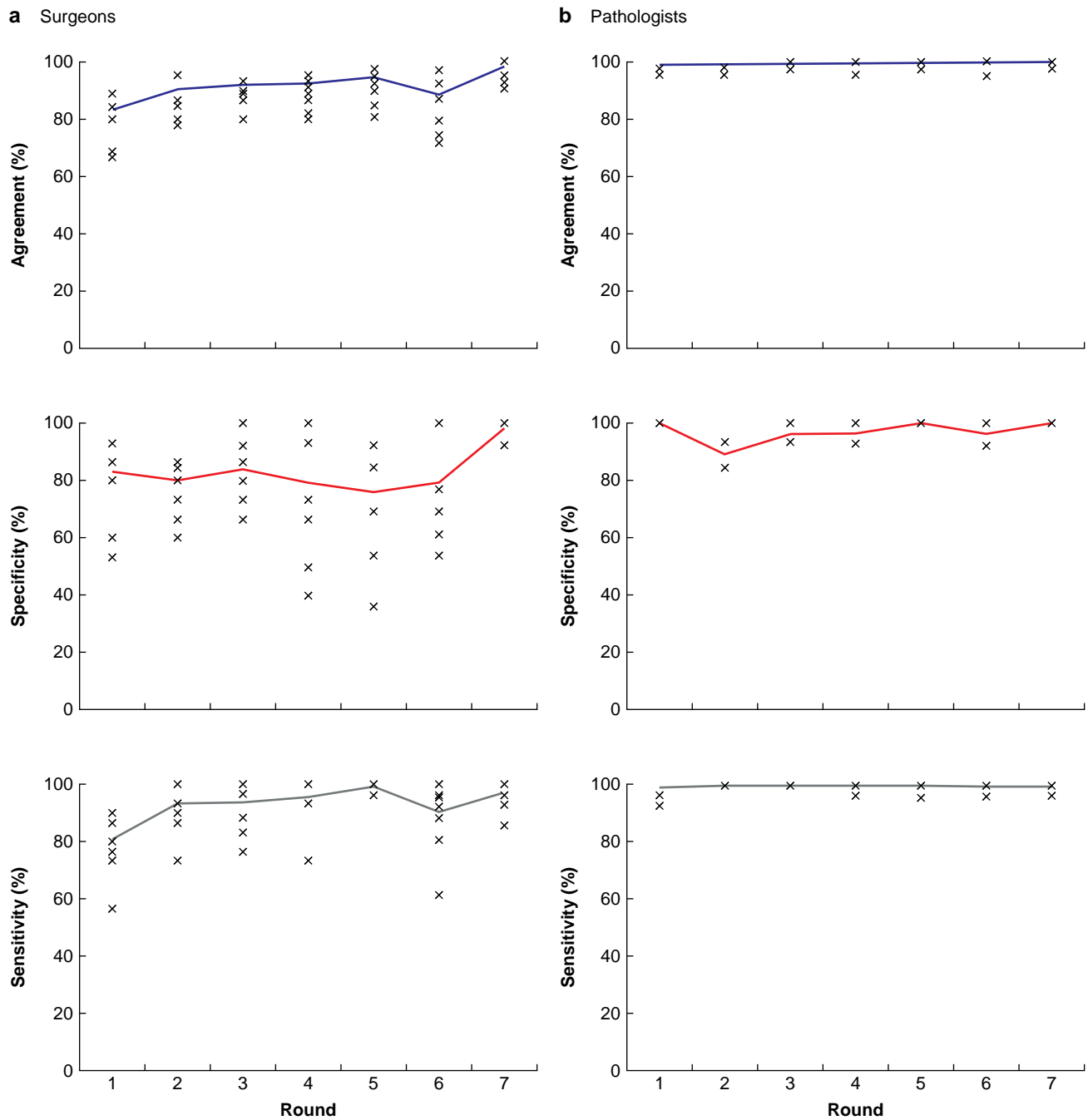
Fig. 4 Performance of physicians for breast cancer detection in ultra-fast confocal microscopy images
(a) Surgeons' and (b) pathologists' performance by round.

Table 2 Trend analysis of surgeon agreement, specificity and sensitivity (all surgeons)

	Early (R1, R2)	Mid (R3, R4, R5)	Late (R6, R7)	Logistic regression	Chi sq	d.f.	P
Agreement	87.5%	93.2%	94.6%	Homogeneity	11.59	2	0.003
c.i.	(82.7%, 91.1%)	(90.4%, 95.2%)	(91.5%, 96.6%)	Trend	10.70	1	0.001
Specificity	82.3%	80.9%	91.8%	Homogeneity	6.29	2	0.043
c.i.	(72.7%, 89.0%)	(72.5%, 87.1%)	(84.6%, 95.8%)	Trend	3.46	1	0.063
Sensitivity	88.4%	96.6%	94.8%	Homogeneity	17.93	2	<0.001
c.i.	(83.0%, 92.3%)	(94.5%, 98.0%)	(91.2%, 97.0%)	Trend	8.18	1	0.004

R1–7, round 1–7.

Table 3 Overall sensitivity of the seven blind assessments

	Sensitivity (%)			
	All cancer types	IC-NST	DCIS	ILC
All reviewers	93 (12)	96 (8)	91 (15)	91 (13)
Surgeons	91 (13)	95 (8)	89 (16)	90 (14)
Pathologists	98 (5)	100 (0)	96 (6)	97 (6)

Values in parentheses are standard deviations. IC-NST, invasive carcinoma of no special type; DCIS, ductal carcinoma *in situ*; ILC, invasive lobular carcinoma.

Each session was uploaded onto the website every 41 days on average (min. 29; max. 93). The mean duration of a training session was 17 min and physicians completed self-assessments in 5 min (min. 4; max. 7). The mean total time of training was 152 min per physician to complete the learning phase (the nine sessions). The authors noticed excellent results for both pathologists and surgeons in self-assessment examinations as 100 per cent of UFCM images were correctly classified by pathologists and 93 per cent by surgeons (min. 89 per cent; max. 99 per cent) leading to the start of the performance assessment phase.

Evaluation of physicians' performance

The final database for performance assessment was composed of 300 images from 126 large fields-of-view of surgical specimens including 100 images of normal breast tissue (fatty, collagenous) (33 per cent), 64 images of IC-NST (21 per cent), 67 images of DCIS (22 per cent) and 69 images of ILC (23 per cent). All images at three zoom levels were presented to the nine physicians in seven rounds of interpretation according to Table 1. The mean duration of rounds was 27 min (min. 6; max. 41) without differences between surgeons (mean 26 min) and pathologists (mean 27 min) although the two pathologists had to answer one more question describing which type of cancer they recognized (Fig. S1). Mean time per question was variable for each physician but almost constant during all rounds, from 1 to 6 min depending on the physician. Figures 2 and 3 present typical UFCM images from tumoural and non-tumoural tissues that were correctly identified by all participants.

The performance of surgeons by round is shown graphically in Fig. 4a. Surgeon performance was variable and a total of 34 'uncertain diagnosis' responses were given, 33 of which were by surgeon 1 in rounds 2 to 7 and one by surgeon 5 in round 1. The other surgeons did not provide this type of answer. A preliminary analysis of the percent agreement between surgeons working within the same centre provided no evidence that surgeons working in the same centre were more similar to each other than surgeons working in different centres. Percent agreement increased from 83 per cent (95 per cent c.i. 75–89 per cent) in round 1 to 98 per cent (96–99 per cent) in round 7 where specificity was 99 per cent (94–100 per cent) and sensitivity 97 per cent (94–99 per cent). There was a trend of increased agreement from the early to late time interval ($P = 0.001$) and a similarly significant trend in sensitivity ($P = 0.004$) (Table 2). Specificity increased with time, but the trend was just not significant ($P = 0.06$) (Table 2). There was also no apparent difference in sensitivity according to tumour type ($\chi^2 = 3.75$, d.f. = 2, $P = 0.153$) independently of the surgeon (Fig. S2).

The performance of pathologists by round is shown graphically in Fig. 4b. The combined pathologist agreement, specificity and sensitivity over all rounds were 99.6 per cent (94–100 per cent),
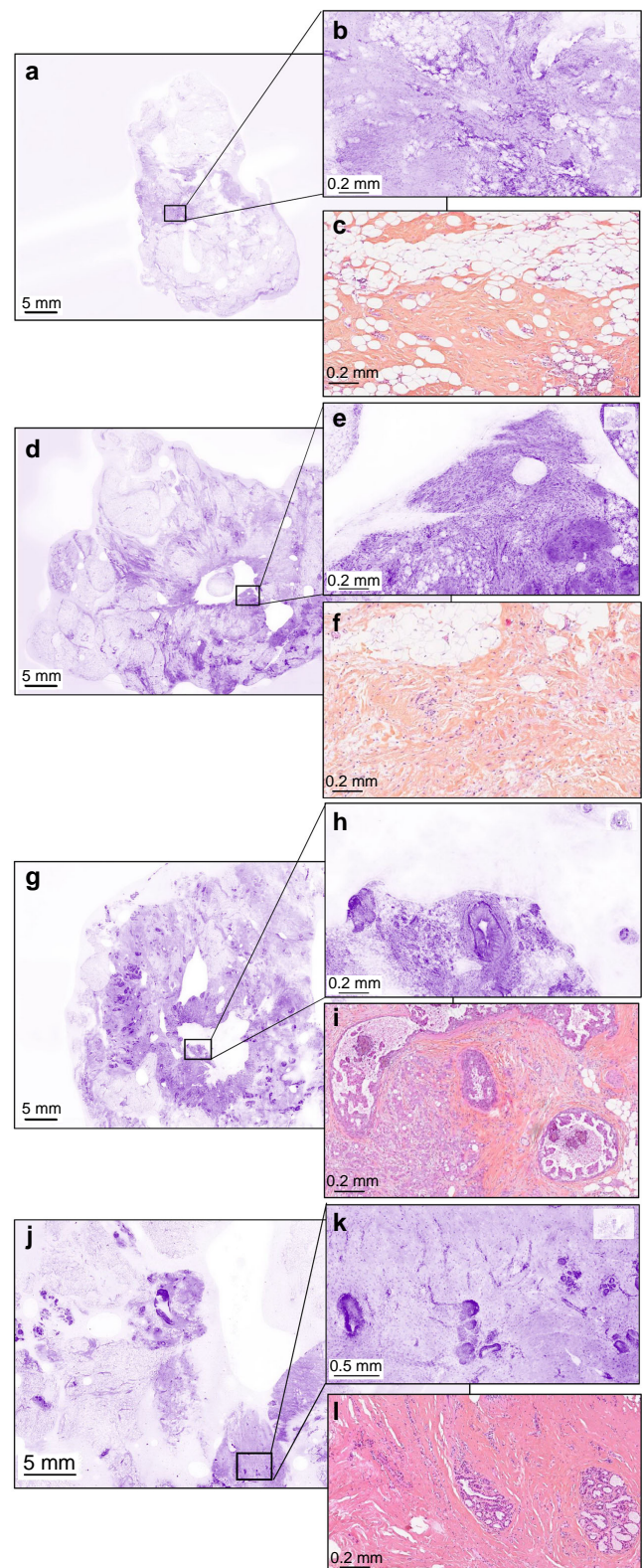
Fig. 5 Examples of ultra-fast confocal microscopy images not correctly classified by more than four surgeons at low magnification (lumpectomy section) and high magnification with corresponding haematoxylin, eosin and saffron section
(a–c) Non-tumoural tissue incorrectly classified as tumoural tissue by five surgeons and one pathologist. (d–f) Non-tumoural tissue incorrectly classified as tumoural tissue by six surgeons. (g–i) Invasive carcinoma and ductal carcinoma *in situ* incorrectly classified as non-tumoural tissue by four surgeons. (j–l) Non-tumoural tissue incorrectly classified as tumoural tissue by six surgeons.

97 per cent (93–99 per cent) and 99.9 per cent (98–99 per cent) respectively, with little variation between rounds (*Table S3*).

Sensitivity was higher for IC-NST diagnosis (96 per cent) than DCIS (91 per cent) and ILC diagnosis (91 per cent) for all physicians (*Table 3*). It was also noticed comparable sensitivity for DCIS and ILC according to the physician specialty: 89 per cent and 90 per cent for surgeons *versus* 96 per cent and 97 per cent for pathologists. In addition, standard deviation was always larger for surgeons than pathologists independently of the tumour type.

The authors noticed that some images were incorrectly classified (cancerous or not cancerous or 'uncertain diagnosis') by a large number of physicians, especially concerning ILC and DCIS (*Table S2*). Examples of frequent misinterpretations are presented in *Fig. 5*.

In addition to classifying each image as normal or cancerous tissue, pathologists were invited to classify the cancer subtype (ILC, IC-NST or DCIS) based on the UFCM image. The subtype was identified correctly in 81 per cent of the true positive diagnoses (155 of 191 for pathologist 1 or 156 of 193 cases for pathologist 1). The main error was the misclassification of ILC instead of IC-NST in 29 cases for the two pathologists (*Table S4*).

Discussion

To facilitate the adoption of the UFCM technology in breast surgery, the authors built a training program from *ex vivo* normal and most common types of cancerous breast tissue. The same approach has been described for pCLE in pulmonology, urology, gastroenterology and prostate surgery, as well as for confocal microscopy in oral cancer[16–18]. The HIBISCUSS project presented here aimed to evaluate the feasibility of interpreting UFCM images by pathologists and breast surgeons. To this end, a UFCM learning program was designed for the training and self-assessment of physicians on the morphological patterns of normal breast tissue, IC-NST, ILC and DCIS. The training phase was relatively short, with nine sessions of approximately 17 min, so the learning program could well be proposed to physicians without overloading their clinical duties planning. Physicians can monitor their progress autonomously and revise errors. Based on the analysis of the self-assessment tests, it was noted that both non-tumoural and tumoural histological features found in breast tissues were well recognized by all participants. Previous experience in optical imaging may therefore not be mandatory to successfully follow and complete the training program.

Several technologies have already been proposed for the intraoperative assessment of surgical margins, with very promising results such as radiofrequency spectroscopy to measure the dielectric properties of tissue[19], Raman spectroscopy[20] or quantitative microelastography[21]. A margin probe was designed to detect differences between dielectric properties of malignant and normal breast tissue adjacent to the probe's sensor. Used on *ex vivo* specimens, the approach previously showed limited sensitivity and specificity (both 70 per cent) when there are multiple tissue types in the explored region[22]. The effective measurement volume of the sensor is a 7-mm diameter disc, up to 1-mm thick. Each measurement takes approximately 1–5 s. Most techniques, including the margin probe, assess the margins of excised specimens, providing only an indirect indication of residual cancer in the cavity, and make it challenging to relate cancer identified in a specimen to its corresponding location in the cavity[23]. Contrary

to the small area evaluated by the margin probe at each acquisition, the UFCM image of 20 cm^2 can be achieved in 50 s. The complete surface of all lumpectomy margins might be processed and scanned in less than 12 min and images kept in patient medical files. Raman spectroscopy is an optical technique that can provide diagnosis based on quantitative molecular elements of the tissue[21]. The interpretation is based on quantitative properties, reducing the risk of subjective human reading from morphological data. The technology is highly sensitive, but with in vivo by point measurements, the spatial resolution is lower than in UFCM. Lastly, quantitative microelastography was described for intraoperative breast cancer management in a first *in vivo* clinical study by Gong *et al.*[24]. This optical imaging technique allows microscale imaging stiffness in three dimensions to depths of 1 mm in breast tissue without tissue processing. The tumour bed was analysed after lumpectomy and only selected, small local tissue regions of 129 mm^3 were imaged. Contrary to UFCM, zone selections have to be made by the surgeon. Currently, all the margins or the operative bed cannot be assessed in a short time with the above-mentioned technologies and UFCM might answer this problem. Another considerable advantage of UFCM is that by providing images at the histological level, imaging of the margins could be collected in the operating theatre, while diagnosis could be shared remotely with pathologists and it is directly comparable with the gold standard of margin assessment, that is, final pathological examination.

The authors' non-interventional study devoted to *ex vivo* tissue imaging demonstrated that the physician can accurately identify breast cancer tissues in UFCM images. Due to the spread of the COVID pandemic, both recruitment and training encountered certain difficulties, so the learning interval lasted more than 1 year with about one session every month. However, this technique appeared to be easy to learn with a sharply increased accuracy of surgeons' results during the performance assessment phase resulting in high final diagnostic values.

The right definition of positive margin in breast surgery has changed over the past few years, turning into no tumour at the inked margin for patients with invasive cancer, and no tumour left within 2 mm from the surface of lumpectomy for DCIS[25]. Nowadays, a frozen section is the standard microscopic technique for intraoperative margin assessment; however, it is tissue destructive, hardly applicable to fatty tissue, restricted to the analysis of only small pieces of tissue, time-consuming and it does not allow formal identification of some breast cancer types such as DCIS.

UFCM could represent a practical alternative for intraoperative margin assessment. This technique could allow the imaging at high magnification of the surface of the whole surgical specimen in a few minutes with conservation of tissue integrity allowing further histologic and immunologic investigations. In this study, the processing and imaging time was carried out in the pathology department and lasted approximately 4–5 min for a section of 20 cm^2.

The time to perform one session of performance assessment was relatively short (mean 27 min with almost 1 min per question), with a decreasing trend from round 1 to round 7. Variability in time per answer and physician performance reflected the individual skills and attitudes of each physician and their capacity to learn. Some physicians focused their attention only on the highest magnification of UFCM images while others spent more time reading and interpreting each magnification zoom. Nevertheless, time for image interpretation was not correlated to the accuracy

rate. Efficacy increased considering the decreasing time round after round, while performance remained steady. Generally, pathologists spent slightly less time than surgeons reading UFCM images and interpreting cancer subtype as expected.

The performance of pathologists was excellent with almost 100 per cent accuracy while surgeons achieved satisfactory values with an improved trend of agreement from 83 per cent (round 1) to 98 per cent (round 7) for cancer identification. No difference was found by the centre or country. Such a high rate suggests that UFCM images can be learned by both experienced pathologists familiar with HES sections and by physicians of other specialties without prior experience in histological interpretation.

Only two surgeons and one pathologist used the 'uncertain diagnosis' category. For the surgeons, it can be hypothesized that in real life they would have asked for the assistance of a pathologist in difficult cases, while for pathologists this category is known from frozen sections in which interpretation may be postponed to the definitive analysis.

Although surgeon sensitivity was not statistically different per tumour type and round, it was noted that half of the wrong answers were false diagnoses of DCIS and ILC, two types of breast cancer difficult to diagnose even in standard histological slides[26,27]. Nevertheless, difficult does not mean impossible and this study demonstrates that, after the training interval, sensitivity remains high for both physicians, especially for pathologists, proving that this system could help the diagnosis of intraoperative margins. ILC represented a challenging subtype in the first confocal breast cancer evaluation made with UFCM on core needle biopsies[11]. Another limit of UFCM images may be the interpretation of small foci of cells, above all in an inflammatory background. The other half of wrong answers provided by the surgeons were false positive answers and this may be related to the attitude the surgeon has to take during surgery. To achieve a satisfying outcome of the primary surgery, some surgeons may perform additional tumour bed excisions during the primary surgery to avoid reoperations. Similar results of good accuracy (more than 90 per cent) were also found for both pathologists and surgeons by Chang *et al.* using pCLE but with a higher trend of false positives[28]. In cases of doubt, surgeons preferred to diagnose cancer instead of potentially leaving neoplastic tissues in their study.

The HIBISCUSS project had some limitations. Image acquisition of the tumour core inside the lumpectomy specimen was completed after specimen section to confirm that physicians can diagnose cancer with UFCM. However, physician performance could be further evaluated for the analysis of tumour margins without cutting the specimen for intraoperative margin assessment purposes. The authors also recommend assessing whether the whole procedure is effective in terms of gain in time and money in the operative theatre workflow of lumpectomy procedures. Eight to ten minutes corresponded to sectioning the specimen in two parts, the staining and the imaging session of each part (50 s for one image at high resolution). For margins assessment, the specimen would not be cut (reducing the time of processing) and would be submerged in fluorescent dye and rinsed for a few seconds. To image all the margins of a lumpectomy specimen, the time required might be slightly increased (50 s per margin). Several variables could affect this timing, such as the experience of the operator and sample size exceeding 20 cm^2, and they should be evaluated in clinical practice.

UFCM images show morphological information at 20 μm below the surface with an axial resolution of 30 μm. They represent an optical section of limited cell layers from the tissue surface without any information deeper into the tissue in contrast to experimental confocal microscopes: the distance of 2 mm is required for DCIS and could not be seen at this depth in this technical configuration while it is not a limitation for invasive breast cancer defined as 'tumour at ink'.

To shorten the adoption of such technology in the clinical workflow, an automated approach such as a deep learning solution could be a major improvement to provide surgeons without previous histological education relevant decision-supportive information as proposed by Shavlokhova *et al.*[18] for fluorescence microscopic imaging in oral cancer and by Combalia *et al.* for automated detection of skin cancer[29]. The addition of immunofluorescence analysis directly on surgical specimens might also be developed in future studies to bring multi-information in UFCM images to support the image assessment. Finally, UFCM telepathology in hospitals could be developed to spread the UFCM adoption in breast cancer surgery and drive the diagnosis in case of less frequent histologies, both cancerous and non-cancerous, where the assessment of specialist breast pathologists would be valuable. This approach could also be proposed to medical centres with an external pathology laboratory for remote analysis in intraoperative time when extemporaneous analysis is not an option. Telepathology has already been described successfully during endomicroscopy for a digestive cancer surgical procedure[30]. The pathologist was not limited to the selection of samples provided by the surgeon and a new way of interactive real-time collaboration between pathologists and surgeons was highlighted.

The HIBISCUSS project shows that pathologists and surgeons from different countries, hospitals and backgrounds can quickly learn in 2.30 h to identify breast cancer at the cellular level in 300 UFCM images. This paves the way for the implementation UFCM for intraoperative assessment of specimen margins during breast lumpectomies to optimize surgeons' workflow and reduce the re-excision rate. A prospective study is under preparation for further validation.

Disclosure

The authors declare no conflict of interest. Histological assessment and data acquisition were carried out independently at Gustave Roussy Hospital.

Supplementary material

Supplementary material is available at *BJS Open* online.

Data availability

Data from the HIBISCUSS project are available from the authors and can be requested from Angelica Conversano.

Author contributions

Angelica Conversano (Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Resources, Validation, Visualization, Writing—original draft), Muriel Abbaci (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Validation, Visualization, Writing—original draft), Paul J. van Diest (Formal analysis, Methodology, Resources, Supervision, Validation, Visualization, Writing—review & editing), Aurélie Roulot (Formal analysis, Investigation, Methodology, Visualization, Writing—review & editing), Giuseppe Falco (Formal analysis, Investigation, Methodology, Visualization, Writing—review & editing), Malek Ferchiou (Formal analysis, Investigation, Methodology, Visualization, Writing—review & editing), Saverio Coiro (Formal analysis, Investigation, Methodology, Visualization, Writing—review & editing), Milan Richir (Formal analysis, Investigation, Methodology, Visualization, Writing—review & editing), Pierre-Michel Genolet (Formal analysis, Investigation, Methodology, Visualization, Writing—review & editing), Carine Clement (Formal analysis, Investigation, Methodology, Visualization, Writing—review & editing), Odile Casiraghi (Conceptualization, Investigation, Methodology, Writing—review & editing), Aicha Ben Lahkdar (Formal analysis, Methodology, Writing—review & editing), Nizard Labaied (Investigation, Methodology, Writing—review & editing), Moira Ragazzi (Conceptualization, Formal analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing—review & editing) and Marie-Christine Mathieu (Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing—original draft).

References

1. Veronesi U, Cascinelli N, Mariani L, Greco M, Saccozzi R, Luini A *et al.* Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer. *N Engl J Med* 2002;**347**:1227–1232
2. Kouzminova NB, Aggarwal S, Aggarwal A, Allo MD, Lin AY. Impact of initial surgical margins and residual cancer upon re-excision on outcome of patients with localized breast cancer. *Am J Surg* 2009;**198**:771–780
3. Houssami N, Macaskill P, Marinovich ML, Dixon JM, Irwig L, Brennan ME *et al.* Meta-analysis of the impact of surgical margins on local recurrence in women with early-stage invasive breast cancer treated with breast-conserving therapy. *Eur J Cancer* 2010;**46**:3219–3232
4. Morrow M, Van Zee KJ, Solin LJ, Houssami N, Chavez-MacGregor M, Harris JR *et al.* Society of Surgical Oncology-American Society for Radiation Oncology-American Society of Clinical Oncology Consensus Guideline on margins for breast-conserving surgery with whole-breast irradiation in ductal carcinoma in situ. *Ann Surg Oncol* 2016;**23**:3801–3810
5. McCahill LE, Single R, Ratliff J, Sheehey-Jones J, Gray A, James T. Local recurrence after partial mastectomy: relation to initial surgical margins. *Am J Surg* 2011;**201**:374–378
6. Pilonis ND, Januszewicz W, di Pietro M. Confocal laser endomicroscopy in gastro-intestinal endoscopy: technical aspects and clinical applications. *Transl Gastroenterol Hepatol* 2022;**7**:7
7. De Palma GD, Esposito D, Luglio G, Limite G, Accurso A, Sollazzo V *et al.* Confocal laser endomicroscopy in breast surgery: a pilot study. *BMC Cancer* 2015;**15**:252
8. Nackenhorst MC, Kasiri M, Gollackner B, Regele H. Ex vivo fluorescence confocal microscopy: chances and changes in the analysis of breast tissue. *Diagn Pathol* 2022;**17**:55
9. Krishnamurthy S, Brown JQ, Iftimia N, Levenson RM, Rajadhyaksha M. Ex vivo microscopy: a promising next-generation digital microscopy tool for surgical pathology practice. *Arch Pathol Lab Med* 2019;**143**:1058–1068
10. Karen JK, Gareau DS, Dusza SW, Tudisco M, Rajadhyaksha M, Nehal KS. Detection of basal cell carcinomas in Mohs excisions with fluorescence confocal mosaicing microscopy. *Br J Dermatol* 2009;**160**:1242–1250
11. Elfgen C, Papassotiropoulos B, Varga Z, Moskovszky L, Nap M, Güth U *et al.* Comparative analysis of confocal microscopy on fresh breast core needle biopsies and conventional histology. *Diagn Pathol* 2019;**14**:58
12. Grizzetti L, Kuonen F. Ex vivo confocal microscopy for surgical margin assessment: a histology-compared study on 109 specimens. *Skin Health Dis* 2022;**2**:e91
13. Kechrid N, Tonellotto L, Monnier S, Rossi SA, Ulrich F, Kuonen F. Ex vivo confocal microscopy for the intraoperative assessment of deep margins in giant basal cell carcinoma. *JAAD Case Rep* 2022;**27**:41–45
14. Abbaci M, Casiraghi O, Vergez S, Maillard A, Lakhdar AB, De Leeuw F *et al.* Diagnostic accuracy of in vivo early tumor imaging from probe-based confocal laser endomicroscopy versus histologic examination in head and neck squamous cell carcinoma. *Clin Oral Investig* 2022;**26**:1823–1833
15. Abbaci M, Breuskin I, Casiraghi O, De Leeuw F, Ferchiou M, Temam S *et al.* Confocal laser endomicroscopy for non-invasive head and neck cancer imaging: a comprehensive review. *Oral Oncol* 2014;**50**:711–716
16. Panarello D, Compérat E, Seyde O, Colau A, Terrone C, Guillonneau B. Atlas of ex vivo prostate tissue and cancer images using confocal laser endomicroscopy: a project for intraoperative positive surgical margin detection during radical prostatectomy. *Eur Urol Focus* 2020;**6**:941–958
17. Shavlokhova V, Flechtenmacher C, Sandhu S, Vollmer M, Vollmer A, Saravi B *et al.* Ex vivo fluorescent confocal microscopy images of oral mucosa: tissue atlas and evaluation of the learning curve. *J Biophotonics* 2022;**15**:e202100225
18. Shavlokhova V, Sandhu S, Flechtenmacher C, Koveshazi I, Neumeier F, Padrón-Laso V *et al.* Deep learning on oral squamous cell carcinoma ex vivo fluorescent confocal microscopy data: a feasibility study. *J Clin Med* 2021;**10**:5326
19. Hoffman A, Ashkenazi I. The efficiency of MarginProbe in detecting positive resection margins in epithelial breast cancer following breast conserving surgery. *Eur J Surg Oncol* 2022;**48**:1498–1502
20. Shipp DW, Rakha EA, Koloydenko AA, Macmillan RD, Ellis IO, Notingher I. Intra-operative spectroscopic assessment of surgical margins during breast conserving surgery. *Breast Cancer Res* 2018;**20**:69

21. Kennedy KM, Zilkens R, Allen WM, Foo KY, Fang Q, Chin L *et al.* Diagnostic accuracy of quantitative micro-elastography for margin assessment in breast-conserving surgery. *Cancer Res* 2020;**80**:1773–1783

22. Pappo I, Spector R, Schindel A, Morgenstern S, Sandbank J, Leider LT *et al.* Diagnostic performance of a novel device for real-time margin assessment in lumpectomy specimens. *J Surg Res* 2010;**160**:277–281

23. Pradipta AR, Tanei T, Morimoto K, Shimazu K, Noguchi S, Tanaka K. Emerging technologies for real-time intraoperative margin assessment in future breast-conserving surgery. *Adv Sci (Weinh)* 2020;**7**:1901519

24. Gong P, Chin SL, Allen WM, Ballal H, Anstie JD, Chin L *et al.* Quantitative micro-elastography enables in vivo detection of residual cancer in the surgical cavity during breast-conserving surgery. *Cancer Res* 2022;**82**:4093–4104

25. Pilewskie M, Morrow M. Margins in breast cancer: how much is enough? *Cancer* 2018;**124**:1335–1341

26. Reed AEM, Kutasovic JR, Lakhani SR, Simpson PT. Invasive lobular carcinoma of the breast: morphology, biomarkers and 'omics. *Breast Cancer Res* 2015;**17**:12

27. Pinder S. Ductal carcinoma in situ (DCIS): pathological features, differential diagnosis, prognostic factors and specimen evaluation. *Mod Pathol* 2010;**23**:S8–S13

28. Chang TP, Leff DR, Shousha S, Hadjiminas DJ, Ramakrishnan R, Hughes MR *et al.* Imaging breast cancer morphology using probe-based confocal laser endomicroscopy: towards a real-time intraoperative imaging tool for cavity scanning. *Breast Cancer Res Treat* 2015;**153**:299–310

29. Combalia M, Garcia S, Malvehy J, Puig S, Mülberger AG, Browning J *et al.* Deep learning automated pathology in ex vivo microscopy. *Biomed Opt Exp* 2021;**12**:3103

30. Fuks D, Pierangelo A, Validire P, Lefevre M, Benali A, Trebuchet G *et al.* Intraoperative confocal laser endomicroscopy for real-time in vivo tissue characterization during surgical procedures. *Surg Endosc* 2019;**33**:1544–1552