Hyperspectral imaging for resection margin assessment during cancer surgery

Esther Kho[1], Lisanne L. de Boer[1], Koen. K. Van de Vijver[2,3], Frederieke van Duijnhoven[1], Marie-Jeanne T.F.D. Vrancken Peeters[1], Henricus J.C.M. Sterenborg[1,4], Theo J.M. Ruers[1,5]

[1]Department of Surgery, the Netherlands Cancer Institute, the Netherlands.

[2]Department of Pathology, the Netherlands Cancer Institute, the Netherlands.

[3]Department of Pathology, Ghent University Hospital, Belgium.

[4]Department of Biomedical Engineering and Physics, Academic Medical Center, the Netherlands.

[5]Technical Medical Centre, University of Twente, the Netherlands.

Running title: Hyperspectral imaging for resection margin assessment

Keywords: Hyperspectral imaging, resection margin assessment, breast-conserving surgery, diffuse reflectance, breast cancer

Financial support: This work was partly supported by the Dutch Cancer Society (grant KWF 10747)

Disclosure statement: The authors declare no potential conflicts of interest.

Corresponding author:

Name: Esther Kho

Address: Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands

E-mail: e.kho@nki.nl

Phone: +31 (0)20 512 7816

Word count: 5077 words

Number of figures and tables: 5 figures and 1 table. (4 supplementary figures)

Statement of translational relevance

Breast-conserving surgery in combination with adjuvant treatment is the preferred treatment for women with breast cancer. These surgeries, however, remain challenging because visualization of the tumor is difficult, and no direct feedback of tumor-positive margins, which indicate that tumor resection is incomplete, is possible. Therefore, additional surgery or a radiotherapy boost are often required to clear tumor tissue that is left behind during the initial operation. This study evaluates the use of hyperspectral imaging for tumor detection in fresh human breast tissue specimens. Thereby, a high diagnostic performance on sliced breast specimen was reported. In addition, hyperspectral imaging was much faster compared to currently available margin assessment techniques. Since data acquisition and analysis can be performed in the operation theater, it would be possible to provide near real-time feedback to the surgeon. This offers potential for re-excising suspected tumor-positive margins during the initial operation in the future.

Abstract

Purpose: Complete tumor removal during cancer surgery remains challenging due to the lack of accurate techniques for intraoperative margin assessment. This study evaluates the use of hyperspectral imaging for margin assessment by reporting its use in fresh human breast specimens.

Experimental design: Hyperspectral data was first acquired on tissue slices from 18 patients, after gross sectioning of the resected breast specimen. This dataset, which contained over 22.000 spectra, was well correlated with histopathology and used to develop a support vector machine classification algorithm and test the classification performance. In addition, we evaluated hyperspectral imaging in clinical practice by imaging the resection surface of six lumpectomy specimens. With the developed classification algorithm, we determined if hyperspectral imaging could detect malignancies in the resection surface.

Results: The diagnostic performance of hyperspectral imaging on the tissue slices was high; invasive carcinoma, ductal carcinoma *in situ*, connective and adipose tissue were correctly classified as tumor or healthy tissue with accuracies of 93%, 84%, 70% and 99%, respectively. These accuracies increased with the size of the area consisting of one tissue type. The entire resection surface was imaged within 10 minutes and data analysis was performed fast, without the need of an experienced operator. On the resection surface, hyperspectral imaging detected 19 out of 20 malignancies that, according to the available histopathology information, were located within 2 mm of the resection surface.

Conclusion: These findings show the potential of using hyperspectral imaging for margin assessment during breast-conserving surgery to improve surgical outcome.

Introduction

One of the major challenges of cancer surgery remains the complete removal of the tumor. Therefore, positive resection margins, which indicate that tumor resection is incomplete, are frequently found after various kinds of cancer surgery (1). Specifically for breast cancer, in up to 37% of the 500.000 women undergoing breast-conserving surgery in the United States and Europe each year, resection margins are tumor-positive (2-6). For these patients, additional surgery or a radiotherapy boost of the resection area is needed, which have a major impact on the cosmetic outcome, patients' quality of life, and health care costs (7,8).

Currently, evaluation of positive resection margins is performed by histopathological evaluation of the margins under the microscope and typically requires several days. The number of tumor-positive resection margins could be reduced by assessing resection margins during surgery. Methods currently available in the clinic for intraoperative resection margin assessment are frozen section analysis, touch preparation cytology, intraoperative ultrasound, and specimen radiography (9-12). However, since these methods are either time-consuming or exhibit a low accuracy, there is still a clear unmet need for a more effective intraoperative margin assessment technique.

Therefore, several other techniques have been proposed for margin assessment during breast-conserving surgery (9-25). Among these techniques are the use of various fluorescent imaging probes (13,14), optical coherence tomography (15), photoacoustic tomography (17), and optical spectroscopic methods like Raman spectroscopy (18,19), and diffuse reflectance spectroscopy (22-25). Despite the potential of these techniques, they have various drawbacks, like the need for exogenous contrast agents, excessive time to assess the entire resection surface, or a low diagnostic performance in detecting ductal carcinoma *in situ*, a potential precursor of invasive carcinoma.

A promising new approach for intraoperative margin assessment that eliminates these drawbacks is hyperspectral imaging. Hyperspectral imaging offers great potential for non-invasive tissue diagnosis because diffuse reflectance measurements can be performed fast, over a wide field of view, without tissue contact and without the need for exogenous contrast agents (26).

Here, we report for the first time on margin assessment with hyperspectral imaging, specifically for breast-conserving surgery. With a hyperspectral camera, diffuse reflectance images were obtained of fresh breast specimens. In these images, each pixel contained an entire spectrum from ~900-1700 nm, which reflects optical differences in tissue composition and morphology between tissue types. First, we acquired data from tissue slices that were obtained after gross sectioning of resected breast specimens. This dataset contained spectra of all different tissue types with accurate histopathological annotations and was used to develop and test a tissue classification algorithm. Next, we imaged the resection surface of lumpectomy specimens to evaluate hyperspectral imaging in a way similar to the intended application in the clinical workflow. In addition, we applied the developed classification algorithm to predict malignancies in these resection surfaces.

Materials and Methods

Hyperspectral imaging system

Hyperspectral images were obtained using a hyperspectral imaging system (VLNIR CL-350-N17E, Specim, Spectral Imaging Ltd., Finland) that captures light in the near-infrared (~900-1700 nm, 256 wavelength bands, 5 nm increments) with an InGaAs sensor (320 x 256 pixels). Previous research showed that this wavelength range contains distinctive information for a robust discrimination between tumor and healthy breast tissue (27,28). **Figure 1** shows the hyperspectral imaging setup. The tissue was placed upon a translation stage and illuminated with three halogen light sources that excite light over the full wavelength range of our camera, under an angle of 45 degrees. After penetration into the tissue, the light underwent multiple scattering and absorption events and was subsequently captured by the camera. The imaged scene was captured line-by-line by moving the translation stage. Each imaged line contained 320 pixels and on average 200 lines were obtained per tissue specimen. Thereby a 3D hypercube was created, in which each pixel in the 2D image contained one spectrum. The scanning speed was adjusted to 50 lines per second to match the spatial resolution of the imaged line (~0.5 mm/pixel). Due to the low sensitivity of the InGaAs sensor at the edges of the spectral range, only wavelengths between 953.25-1645.62 nm (210 wavelength bands) were used.

Study design

Measurements were performed on *ex vivo* fresh breast tissue from female patients that had primary breast surgery at the Antoni van Leeuwenhoek hospital and did not receive neoadjuvant treatment. This study was performed in compliance with the Declaration of Helsinki and approved by the Institutional Review Board of the Netherlands Cancer Institute/Antoni van Leeuwenhoek (Amsterdam, the Netherlands). According to Dutch law (WMO), no written informed consent from patients was required. The involved surgeons and pathologist were blinded to any hyperspectral information. In total, two datasets were obtained from different patients.

Data acquisition of tissue slices. In the first dataset, hyperspectral data was acquired of tissue slices to minimize spatial registration artifacts between hyperspectral images and histopathology information. Fresh lumpectomy specimens were transferred to the pathology department immediately after resection, where they were colored and sliced according to standard procedure. In consultation with the pathologist, one tissue slice per patient that contained both healthy and tumor tissue was selected. This slice was placed in a macrocassette, on top of black rubber, where it remained during the optical measurements. The black rubber highly absorbs light over the whole wavelength range and prevented that the macrocassette under the tissue was measured. In total, 18 tissue slices from different patients were measured. The age of these patients was 67 ± 11 years (mean ± standard deviation), and their ACR score, which reflects the breast density, was 2.18 ± 0.95 (mean ± standard deviation). The slices varied in size from 2 x 3.5 cm^2 to 4.5 x 6.3 cm^2 and were at least 2 mm thick. In addition to the hyperspectral image (HS), a white light (WL) image of the tissue was taken from the same point of view. All optical measurements took place within 10 minutes after collection of the tissue slice at the pathology department. After the measurements, the slices were placed in formalin and further processed according to standard protocol.

After a few days, H&E stained sections of the measured surface of the tissue slices were obtained and digitized (Aperio ScanScope, Leica Microsystems B.V., Amsterdam, The Netherlands). To correct for tissue deformations during histopathological processing, the H&E section were registered to the HS images using MATLAB 2015a (The Math Works Inc., Natick, Massachusetts, USA). **Supplementary**

Figure S1 shows this registration process. First, the whole H&E section was annotated with tissue classes after manually removing the background in the H&E section. These tissue classes were invasive carcinoma (IC), its potential precursor ductal carcinoma in situ (DCIS), adipose tissue, connective tissue, and healthy glandular ducts. IC, DCIS and glandular ducts were annotated on the H&E section using delineations drawn by the pathologist. The remaining tissue was annotated as connective or adipose tissue by thresholding all RGB channels of the H&E section in MATLAB at 90%: adipose tissue was washed away in the histopathological processing and was therefore white on the digitized H&E image, whereas connective tissue colored red/pink with the H&E stain. Second, a transformation matrix was obtained of the registration of the non-annotated H&E section to the WL image. The H&E section was resized to the size of the WL image and, based on distinctive features both visible in the H&E section and the WL image, control points were manually selected in both images (the minimum pair used was 61). Therefore, each point in the H&E section had a corresponding point in the WL image. Using a non-rigid local weighted means transformation with 12 neighboring control points, the H&E section was registered to the WL image. With this transformation matrix, also the annotated H&E section was registered to the WL image. Third, the annotated H&E section was registered to the HS image. For this, we used the transformation matrix that described the registration of the WL image to the HS image. Since the WL image was taken simultaneously to the HS image, the tissue did not deform and a simple rigid registration based on the edges of the tissue sufficed. Finally, with the H&E section, all pixels in the HS image were labeled as IC, DCIS, adipose or connective tissue. Since healthy glandular ducts were often smaller than the pixel size of our hyperspectral camera, they were included in the connective tissue class. Due to the highly inhomogeneous character of connective tissue, connective tissue was often observed in a hyperspectral pixel along with another tissue class. In the event that multiple tissue classes were observed in one hyperspectral pixel, this pixel was labeled as ‘unknown’.

Data acquisition of resection surface. The second dataset was used to evaluate hyperspectral imaging of lumpectomy specimens in the current clinical workflow. Our main goal was to evaluate the amount of time required to image the entire resection surface and to analyze the data. In addition, we evaluated whether hyperspectral imaging could detect malignancies within 2 mm of the resection surface.

In total, six fresh lumpectomy specimens were imaged immediately after resection. To ensure measuring tumor at the resection surface, lumpectomy specimens were selected that, according to the surgeon, were suspicious for tumor-positive resection margins. The age of these patients was 61 ± 8 years (mean ± standard deviation), their ACR score was 2.67 ± 0.52 (mean ± standard deviation), and the size of the specimens varied from 16 cm^3 to 118 cm^3. The resection surface of the specimen was imaged from six sides by repositioning the specimen. As guidance for repositioning, we used sutures that were placed on the specimen during surgery, which allow for reorientation by pathology. The resection specimen was placed on black polyoxymethylene material that, like the black rubber, highly absorbs light over the whole wavelength range. After the optical measurements, the specimens were brought to the pathology department. According to standard pathology protocol, they were painted and sliced from nipple to peripheral side. Of all specimens, 43 out of 66 slices (~3 mm thick) were further processed in H&E stained sections (~3 µm thick). Two sections per specimen, taken at the nipple- and peripheral side, were excluded from further analysis since they were taken parallel to the resection surface and thus did not provide histopathology information about the resection margin width. On the remaining 31 H&E sections, the pathologist delineated IC and DCIS so that we could calculate the shortest distance between the inked resection edge and the malignant tissue. Due to the limited number of H&E sections and the perpendicular orientation of the sections to the resection surface, it was not possible to verify the entire hyperspectral image of a resection side with histopathology. Therefore, we limited the analysis of hyperspectral margin assessment to verifying whether a malignant area, containing IC or DCIS, within 2 mm of the resection surface was detected with hyperspectral imaging. Two millimeter is the required tumor-free margin width for DCIS according to guidelines of the SSO-ASTRO-ASCO (29). We used the color of the histopathology paint at the edge near the malignant area to determine its location on the resection surface. A malignant area was considered correctly detected with hyperspectral analysis if at least one pixel in this location was classified as IC or DCIS.

Preprocessing

All data analysis and tissue classification were performed using MATLAB. Prior to spectral classification, hyperspectral data was pre-processed.

First, raw tissue images were converted into normalized diffuse reflectance to correct for spectral non-uniformity of the illumination setup and the influence of the dark current using:

$$R(x,\lambda) = \frac{1}{R_{ref}(\lambda)} \cdot \frac{\text{Image}_{tissue}(x,\lambda) - \text{Image}_{dark}(x,\lambda)}{\text{Image}_{white}(x,\lambda) - \text{Image}_{dark}(x,\lambda)} \cdot 100\% \quad (1)$$

Where $R(x,\lambda)$ is the normalized diffuse reflectance (in percentage), $R_{ref}(\lambda)$ the reference reflectance value of Spectralon (SRT-99-100, Lapsphere, Inc., Northern Sutton, New Hampshire), x the location of the pixel in the imaged line and λ the wavelength band. $\text{Image}_{white}(\lambda)$ and $\text{Image}_{dark}(\lambda)$ are the reference images acquired in addition to the tissue image. The white reference image was acquired on Spectralon 99% and for the dark reference image, we closed the shutter of the camera. Prior to normalization of raw tissue images, we applied a pixel-wise third-order correction model to the raw hyperspectral data to correct for the slight non-linearity of our InGaAs sensor. This model describes the sensors signal output with respect to the amount of light incident on the sensor.

Second, all spectra were pre-processed to eliminate spectral variability due to tissue morphology effects. The oblique illumination and non-flat surface of the tissue could cause uneven illumination of the tissue and nonspecific scatter. As shown in **Supplementary Figure S2**, this could result in variations between spectra that were not related to the physiological composition of the tissue. To correct for this, spectra were pre-processed using standard normal variate (SNV) (30,31). SNV normalized each individual spectrum to a mean of zero and a standard deviation of one and allowed for a robust and direct suppression of baseline drifts and multiplicative scaling (31).

Classification

Classification of tissue slices. The dataset acquired on the tissue slices contained ~51,000 spectra that were labeled with one of the tissue classes. Spectra were excluded if they were contaminated with pathology ink (~4,200 spectra), if the InGaAs sensor saturated (112 spectra), or if spectra were located within 0.5 mm of a tissue class border (~23,700 spectra). The latter causes a larger amount of spectra from smaller tissue class pockets, like DCIS and connective tissue, to be excluded. In addition, spectral outliers were detected and eliminated per tissue class (IC = 370 out of 5,193 spectra, DCIS = 7 out of 613

spectra, connective = 37 out of 652 spectra, adipose = 563 out of 16,948 spectra), using an outlier detection method (32). This method calculated the standard deviation of all Euclidean distances of a spectrum to the group mean and removed a spectrum when this distance was larger than three standard deviations.

A supervised classification model was developed using linear support vector machine (SVM). SVM is based on statistical learning theory and separates two classes by determining an optimal hyperplane that maximizes the margin between two classes in multidimensional space (33). With the one-against-one strategy, multiclass classification was realized so that each spectrum was classified into IC, DCIS, connective or adipose tissue (34). Of each spectrum, both the intensity of all 210 wavelength bands and the pathology labels were used as input for the SVM classification algorithm. SVM was implemented using the software tool PerClass (Academic version 5.0, PR Sys design, Delft, The Netherlands). To develop the algorithm with the largest amount of data, we first removed all spectra from one tissue slice from the dataset to use as test set. Second, with spectra from the remaining 17 slices, the algorithm was developed using inner-cross validation: the 17 slices were randomly split 5 times in a training set (2/3 of the slices) and a validation set (1/3 of the slices) while keeping spectra from one slice together. Finally, the trained SVM was applied on the test set for evaluation of the classification performance. To account for variability in the classification performance, this process was repeated until all tissue slices were used as test set. To correct for the unbalanced number of spectra per tissue class, each tissue class was assigned a different weight, which is the inverse ratio of the tissue class sizes (35).

Classification of resection surface. Hyperspectral images acquired on the resection surface were analyzed with a classification algorithm developed with all tissue slices. The same linear, weighted SVM algorithm as described above was used. However, instead of training the algorithm with 17 tissue slices, all 18 tissue slices were used. All the spectra in the hyperspectral image on the resection surface were SNV normalized and classified with the SVM algorithm as either IC, DCIS, connective or adipose tissue.

Classification performance

Classification performance of hyperspectral imaging on breast tissue was evaluated on the tissue slices that were used as test set. The number of spectra per tissue class varied between tissue slices. Therefore, we calculated classification accuracies over all spectra instead of averaging accuracies over the tissue slices. Both the overall accuracy to discriminate tissue classes and the accuracy per tissue class were calculated. ROC curves and the area under the ROC curve (AUC), however, were calculated per tissue slice, and subsequently averaged. Since classification with SVM was performed using the one-against-one strategy, one ROC curve for each pair of tissue classes was obtained. For the resection margin dataset, no accurate correlation with histopathology could be obtained with the whole hyperspectral image due to the limited number of H&E sections and their perpendicular orientation to the resection surface. As a result, the classification performance of resection margin assessment with hyperspectral imaging could not be evaluated. Nevertheless, we did assess whether hyperspectral analysis confirms tumor-positive margins found on the available H&E sections.

Results

Classification results from tissue slices

The 18 tissue slices comprised a hyperspectral dataset of more than 22,000 reliable spectra, which were classified as IC, DCIS, connective or adipose tissue using an SVM classification algorithm. **Supplementary Figure S3** shows an example of the tissue classes and the average diffuse reflectance spectra per tissue class before and after SNV normalization. Based on the average SNV normalized reflectance spectra, differences between tissue classes were most noticeable around absorption bands of the dominant chromophores in the near-infrared region. For adipose tissue, the large amount of fat in the tissue was discriminating, whereas spectral differences between connective tissue and the malignant classes were mainly related to the amount of water and collagen in the tissue. **Figure 2** shows the hyperspectral classification result of two representative examples of tissue slices that contain IC (top) and DCIS (bottom). **Table 1** shows the classification results of the hyperspectral analysis of all spectra in the dataset. Overall, we could discriminate the defined tissue classes with a diagnostic accuracy, averaged over all tissue slices, of 91%. Of these tissue classes, adipose tissue and IC were detected as,

respectively, healthy and tumor tissue in 99% and 93%. The more challenging tissue types were DCIS and connective tissue, which were classified as, respectively, tumor and healthy tissue in 84% and 70%. These numbers are in accordance with the AUCs shown in **Figure 3**. A high AUC is reported for both IC (0.90, 0.99 and 1.00) and adipose tissue (3 times 1.00), and a lower AUC (0.78 ± 0.12) is reported for the SVM classifier that distinguishes DCIS and connective tissue. The latter indicates that DCIS and connective tissue were more likely to be mistaken for each other. These classification results were based on all individual 22,000 spectra, without considering the spatial information of the hyperspectral images. However, the surroundings of a spectrum might have a strong influence on the classification accuracy. For example, in the classified HS image in **Figure 2D**, most of the incorrectly classified connective spectra are located near the border of the area with IC.

Smallest detectable size of tumor pockets

To evaluate the influence of the surroundings of a spectrum to its classification accuracy, we calculated for each pixel its distance to the closest tissue transition, based on the histopathological annotation of the slices. **Figure 4A** illustrates this by showing a tissue slice that contains IC and adipose tissue. As can be seen in **Figure 4B**, pixels near the IC-adipose tissue transition were classified with a lower accuracy. The inaccuracies that occur around these tissue transitions affect the smallest size of a pocket that can be detected. To assess this effect, the classification accuracy was calculated for each tissue class as a function of the smallest pocket size in which a pixel was present (**Figure 4C**). For each pixel, this smallest pocket size was estimated as twice the shortest distance to a tissue transition. The classification accuracy was determined for the ability of hyperspectral imaging to differentiate the malignant tissue types from the two healthy tissue types and vice versa. We found that for all tissue types, the classification accuracy per tissue class increased with an increase of the smallest pocket size in which the pixel is located. Specifically, pixels in all tissue type pockets larger than 2 x 2 mm were correctly classified in more than 93% of cases (**Figure 4C**). This also includes the two tissue classes, DCIS and connective tissue, which exhibited the lowest classification results in **Table 1** and **Figure 3**.

Hyperspectral imaging of lumpectomy specimens

To assess the entire resection surface, we imaged each specimen from six sides and applied the diagnostic classification algorithm that we developed using the tissue slices, to predict malignancies within 2 mm from the resection surface (**Figure 5**). Each resection side was imaged in 20 seconds and could be analyzed in 40 seconds. Imaging of the entire resection surface, including taking WL images and repositioning the specimen, took on average 9 minutes and 15 seconds.

All six specimens had at least one H&E section with a malignant area within 2 mm from the resection surface. In total, 31 H&E sections were available, of which 15 contained a malignant area within 2 mm from the resection surface. **Supplementary Figure S4** shows three of these H&E sections and the corresponding hyperspectral analysis of the resection side. Five sections contained two malignant areas that corresponded to hyperspectral images taken at two different resection sides. Of the 20 malignant areas, 10 were IC (distance to resection surface: 0 – 1 mm) and 10 were DCIS (distance to resection surface: 0-1.5 mm). Hyperspectral analysis confirmed all tumor-positive margins found on the available H&E sections, except for one IC spot that was smaller than 1 mm^2. Further studies are needed to evaluate the classification performance of resection margin assessment with hyperspectral imaging.

Discussion

Tumor positive resection margins are common after surgical procedures for many types of cancer (1). Residual tumor cells at the resection edge may increase the chance of local recurrence and may affect the ultimate patient outcome. Therefore, a technique for real-time intraoperative margin assessment is clearly needed. To comply with clinical needs during this type of surgery, a successful intraoperative margin assessment technique should be able to assess the entire resection surface almost real-time with high diagnostic performance, up to the required tumor-free margin width (36). In this study, we used hyperspectral imaging for tumor detection after breast-conserving surgery. We report on a high diagnostic performance of hyperspectral imaging on breast tissue slices and fast data acquisition and analysis of the entire resection surface of lumpectomy specimens. This approach could make near real-time assessment of tumor margins during surgery feasible in the future.

We trained and evaluated the diagnostic performance on fresh breast slices so that hyperspectral images were well correlated with histopathology. We defined four tissue classes by distinguishing tissue classes within the malignant and healthy classes. Specifically, within the malignant class, we made a distinction between IC and DCIS. Within the healthy class, we distinguished connective tissue (which included benign glandular tissue) and adipose tissue. Clinically, differentiating these two healthy classes is not relevant. However, due to their large spectral differences, this improved the classification algorithm. Both adipose tissue and IC were detected as either healthy or tumorous tissue types with high accuracies. More challenging was the differentiation between DCIS and connective tissue. Since IC is generally preceded by DCIS, which arises in benign glandular ducts embedded in connective tissue, we would expect DCIS to be optically the most similar to connective tissue. Nevertheless, DCIS and connective tissue were classified as tumor and healthy tissue in 84% and 70% respectively. In this study, individual spectra were classified without taking into account their surroundings. Thereby, 30% of the connective spectra were incorrectly classified as malignant. However, when these connective spectra are located near or in a tumor, misclassifying them as malignant tissue would be clinically less relevant. In **Figure 4**, we showed that classification inaccuracies are most likely to occur around tissue transitions. This can be explained by a difference in tissue information provided by the gold standard, the H&E sections, and the tissue measured with the hyperspectral camera. Based on diffusion theory, we image a volume up to a few millimeters in the tissue with our camera (37). However, the H&E sections reflect just a superficial portion of a few cell layers of the measured surface. As a result, hyperspectral data can be derived from a mixture of different tissue classes instead of the single tissue class given by the superficial layer of histopathology. Since our developed classification algorithm classifies each spectrum into one of the tissue classes, classification inaccuracies will be higher when hyperspectral imaging measures a mixture of tissue classes. This is more likely in small tissue pockets (like DCIS and connective tissue), or in tissue close to tissue transitions.

Nevertheless, pixels in tissue pockets larger than 2 x 2 mm were correctly classified in more than 93%. Whether this is clinically sufficient depends on the used definition of a positive resection margin. Over the last years, breast cancer guidelines have been progressing towards a more liberal attitude. For IC, the

tumor-free margin width changed from 2 mm to no tumor on the resection edge and for DCIS, a tumor-free margin width of 2 mm is now accepted (36). Even though we can detect tumors larger than 2 x 2 mm at the resection surface with our current hyperspectral setup, hyperspectral imaging has the potential to detect even smaller pockets by sampling a smaller volume. This can be achieved by, for example, using a method described by Jacques *et al.* (38) that uses polarization, or by developing classification algorithms using selected wavelengths at which the penetration depth of light is only superficial. Whether this is indeed necessary, should be investigated in a larger prospective study of hyperspectral imaging on resection specimens.

In this study, we classified hyperspectral data with a linear, weighted SVM algorithm. The usefulness of SVM for classifying hyperspectral data has been reported by previous studies in other research areas as well (39). On the lumpectomy specimens, we showed that hyperspectral data could be analyzed fast, without the need for an experienced operator. To evaluate the amount of time required to image the entire resection surface, we imaged the lumpectomy specimens by taking six images. However, a limiting factor in this approach was the correlation of hyperspectral analysis with histopathology due to the limited available histopathological information. Unlike hyperspectral imaging, resection margin assessment with H&E sections covered only a small fraction of the entire resection surface (**Figure 5c**). Therefore, we were not able to verify the whole measured surface, as we could do with the tissue slices dataset. In addition, the location of the H&E sections on the resection surface could not be precisely retrieved. We realize this is a limitation since this prevented us from obtaining a classification accuracy of hyperspectral imaging on analyzing resection margins. As a result, the classification performance of hyperspectral imaging on breast tissue was only calculated using the tissue slices dataset. Nevertheless, in the lumpectomy dataset, hyperspectral analysis confirmed 19 out of 20 malignancies found within 2 mm of the resection surface on the available H&E sections. One IC spot smaller than 1 mm^2 was missed, which is in accordance with the decrease in classification accuracy when tissue pockets get smaller (**Figure 4**). In this study, we focused on detecting all malignancies within 2 mm of the resection surface, without differentiating IC and DCIS. In addition, we were not able to confirm hyperspectral analysis of tumor-negative margins with histopathological information. For an evaluation of the clinical sensitivity of

hyperspectral imaging, a difference should be made between IC and DCIS in the resection margin. This, along with the specificity of hyperspectral imaging, needs to be further explored in a clinical study that evaluates the benefit of hyperspectral imaging in routine breast-conserving surgery. By analyzing hyperspectral images immediately after imaging, locations that are suspected to be tumor-positive as well as tumor-negative areas can be marked and retrieved after histopathological processing. Thereby, a direct correlation with histopathology can be obtained and the classification performance of hyperspectral imaging for resection margin assessment can be evaluated.

With the classification algorithm developed in this study, IC and DCIS spectra from the tissue slices were classified with an accuracy of 93% and 84%. This accuracy is much higher than can be achieved with currently available imaging techniques for margin assessment such as ultrasound and specimen radiography, on which DCIS often remains invisible (10). In contrast, our results resemble pathological margin detection techniques like frozen section analysis (accuracy of 84-98% (10)) and touch preparation cytology (accuracy of 78-99% (10)). Hyperspectral imaging, however, has the advantage over these techniques that it can scan an entire resection surface in a short amount of time. We imaged lumpectomy specimens in a similar set-up as the intended final application in the clinical workflow. In this specific setting, each resection side was imaged in less than 20 seconds and could be analyzed in 40 seconds. Repositioning of the specimen to image the next side required most of the time. Acquisition and analysis time, however, could still significantly be improved by further development of the hardware and optimization of the classification algorithm. Currently, we analyzed the whole resection margin after surgery without providing feedback. When resection sides would be analyzed immediately after imaging, during repositioning of the specimen, hyperspectral analysis would be much faster. In this way, it can easily outperform frozen section analysis and touch preparation cytology that require at least 15-30 minutes (11). In addition, hyperspectral imaging can be performed in the operating theatre, which could allow near real-time diagnostic feedback of the resected specimen to the surgeon in the future. This would make a re-excision of a tumor-positive margin during the initial operation possible.

In summary, we showed the potential of hyperspectral imaging for intraoperative margin assessment by reporting on its first use in human breast tissue. The diagnostic performance of hyperspectral imaging on

sliced breast specimen was high, and imaging of the entire resection surface was much faster in comparison with currently available margin assessment techniques. In addition, hyperspectral imaging in combination with SVM allowed for automated classification of the data, without the need for an experienced operator. These findings support that hyperspectral imaging can become a powerful clinical tool for real-time margin assessment during breast-conserving surgery.

Acknowledgments

The authors thank K. Jóźwiak for help with statistical analyses, B. Dashtbozorg and A.L. Post for attentive reading of the manuscript, G.C. Langhout for graphical assistance, and the NKI-AVL core Facility Molecular Pathology & Biobanking (CFMPB) for supplying NKI-AVL biobank material. The authors also thank all surgeons and nurses from the Department of Surgery and all pathologists and pathologist assistants from the Department of Pathology for their assistance in collecting specimens.

References

1. Orosco RK, Tapia VJ, Califano JA, Clary B, Cohen EE, Kane C*, et al.* Positive surgical margins in the 10 most common solid cancers. *Sci Rep* **2018**;8(1):5686.
2. Alrahbi S, Chan PM, Ho BC, Seah MD, Chen JJ, Tan EY. Extent of margin involvement, lymphovascular invasion, and extensive intraductal component predict for residual disease after wide local excision for breast cancer. *Clin Breast Cancer* **2015**;15(3):219-26.
3. Merrill AL, Coopey SB, Tang R, McEvoy MP, Specht MC, Hughes KS*, et al.* Implications of new lumpectomy margin guidelines for breast-conserving surgery: changes in reexcision rates and predicted rates of residual tumor. *Ann Surg Oncol* **2016**;23(3):729-34.
4. Merrill AL, Tang R, Plichta JK, Rai U, Coopey SB, McEvoy MP*, et al.* Should new “no ink on tumor” lumpectomy margin guidelines be applied to ductal carcinoma in situ (dcis)? A retrospective review using shaved cavity margins. *Ann Surg Oncol* **2016**;23(11):3453-8.
5. Langhans L, Jensen M-B, Talman M-LM, Vejborg I, Kroman N, Tvedskov TF. Reoperation rates in ductal carcinoma in situ vs invasive breast cancer after wire-guided breast-conserving surgery. *JAMA Surg* **2017**;152(4):378-84.
6. Stewart BW, Wild CP. World cancer report 2014. Lyon: IARC Press; 2014.
7. Munshi A, Kakkar S, Bhutani R, Jalali R, Budrukkar A, Dinshaw K. Factors influencing cosmetic outcome in breast conservation. *Clinical Oncology* **2009**;21(4):285-93.
8. Pleijhuis RG, Graafland M, de Vries J, Bart J, de Jong JS, van Dam GM. Obtaining Adequate Surgical Margins in Breast-Conserving Therapy for Patients with Early-Stage Breast Cancer: Current Modalities and Future Directions. *Ann Surg Oncol* **2009**;16(10):2717-30.
9. Keating JJ, Fisher C, Batiste R, Singhal S. Advances in intraoperative margin assessment for breast cancer. *Curr Surg Rep* **2016**;4(4):15.
10. St John ER, Al-Khudairi R, Ashrafian H, Athanasiou T, Takats Z, Hadjiminas DJ*, et al.* Diagnostic accuracy of intraoperative techniques for margin assessment in breast cancer surgery: a meta-analysis. *Ann Surg* **2017**;265(2):300-10.
11. Esbona K, Li Z, Wilke LG. Intraoperative imprint cytology and frozen section pathology for margin assessment in breast conservation surgery: a systematic review. *Ann Surg Oncol* **2012**;19(10):3236-45.
12. Versteegden D, Keizer L, Schlooz-Vries M, Duijm L, Wauters C, Strobbe L. Performance characteristics of specimen radiography for margin assessment for ductal carcinoma in situ: a systematic review. *Breast Cancer Res Treat* **2017**:1-11.
13. Koch M, de Jong JS, Glatz J, Symvoulidis P, Lamberts LE, Adams AL*, et al.* Threshold analysis and biodistribution of fluorescently labeled bevacizumab in human breast cancer. *Cancer Res* **2017**;77(3):623-31.

14. Tummers QRJG, Hoogstins CES, Gaarenstroom KN, de Kroon CD, van Poelgeest MIE, Vuyk J*, et al.* Intraoperative imaging of folate receptor alpha positive ovarian and breast cancer using the tumor specific agent EC17. *Oncotarget* **2016**;7(22):32144-55.
15. Nguyen FT, Zysk AM, Chaney EJ, Kotynek JG, Oliphant UJ, Bellafiore FJ*, et al.* Intraoperative evaluation of breast tumor margins with optical coherence tomography. *Cancer Res* **2009**;69(22):8790-6.
16. Allen WM, Chin L, Wijesinghe P, Kirk RW, Latham B, Sampson DD*, et al.* Wide-field optical coherence micro-elastography for intraoperative assessment of human breast cancer margins. *Biomed Opt Express* **2016**;7(10):4139-53.
17. Li R, Wang P, Lan L, Lloyd FP, Goergen CJ, Chen S*, et al.* Assessing breast tumor margin by multispectral photoacoustic tomography. *Biomed Opt Express* **2015**;6(4):1273-81.
18. Haka AS, Volynskaya Z, Gardecki JA, Nazemi J, Lyons J, Hicks D*, et al.* In vivo margin assessment during partial mastectomy breast surgery using Raman spectroscopy. *Cancer Res* **2006**;66(6):3317-22.
19. Keller MD, Vargis E, de Matos Granja N, Wilson RH, Mycek M-A, Kelley MC*, et al.* Development of a spatially offset Raman spectroscopy probe for breast tumor surgical margin evaluation. *J Biomed Opt* **2011**;16(7):077006.
20. Laughney AM, Krishnaswamy V, Rizzo EJ, Schwab MC, Barth RJ, Pogue BW*, et al.* Scatter spectroscopic imaging distinguishes between breast pathologies in tissues relevant to surgical margin assessment. *Clin Cancer Res* **2012**;18(22):6315-25.
21. Bigio IJ, Bown SG, Briggs G, Kelley C, Lakhani S, Pickard D*, et al.* Diagnosis of breast cancer using elastic-scattering spectroscopy: preliminary clinical results. *J Biomed Opt* **2000**;5(2):221-8.
22. de Boer LL, Hendriks BHW, van Duijnhoven F, Vrancken Peeters-Baas MJTFD, Van de Vijver KK, Loo CE*, et al.* Using DRS during breast conserving surgery: identifying robust optical parameters and dealing with inter-patient variation. *Biomed Opt Express* **2016**;7(12):5188-200.
23. Nachabé R, Evers DJ, Hendriks BH, Lucassen GW, van der Voort M, Rutgers EJ*, et al.* Diagnosis of breast cancer using diffuse optical spectroscopy from 500 to 1600 nm: comparison of classification methods. *J Biomed Opt* **2011**;16(8):087010.
24. Volynskaya Z, Haka AS, Bechtel KL, Fitzmaurice M, Shenk R, Wang N*, et al.* Diagnosing breast cancer using diffuse reflectance spectroscopy and intrinsic fluorescence spectroscopy. *J Biomed Opt* **2008**;13(2):024012.
25. Zhu C, Palmer GM, Breslin TM, Harter J, Ramanujam N. Diagnosis of breast cancer using fluorescence and diffuse reflectance spectroscopy: a Monte-Carlo-model-based approach. *J Biomed Opt* **2008**;13(3):034015.
26. Lu G, Fei B. Medical hyperspectral imaging: a review. *J Biomed Opt* **2014**;19(1):010901.

27. de Boer LL, Molenkamp B, Bydlon TM, Hendriks BHW, Wesseling J, Sterenborg HJ*, et al.* Fat/Water ratios measured with diffuse reflectance spectroscopy to detect breast tumor boundaries. *Breast Cancer Res Treat* **2015**;152(3):509-18.
28. Taroni P, Paganoni AM, Ieva F, Pifferi A, Quarto G, Abbate F*, et al.* Non-invasive optical estimate of tissue composition to differentiate malignant from benign breast lesions: A pilot study. *Sci Rep* **2017**;7:40683.
29. Morrow M, Zee KJV, 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. *J Clin Oncol* **2016**;34(33):4040-6.
30. Barnes R, Dhanoa MS, Lister SJ. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl Spectrosc* **1989**;43(5):772-7.
31. Vidal M, Amigo JM. Pre-processing of hyperspectral images. Essential steps before image analysis. *Chemometr Intell Lab* **2012**;117:138-48.
32. Burger J, Geladi P. Hyperspectral NIR image regression part II: dataset preprocessing diagnostics. *J Chemometr* **2006**;20(3-4):106-19.
33. Vapnik VN. The nature of statistical learning theory. New York: Springer Science & Business Media; 2013.
34. Hsu C-W, Lin C-J. A comparison of methods for multiclass support vector machines. *IEEE transactions on Neural Networks* **2002**;13(2):415-25.
35. Huang Y, Du S. Weighted support vector machine for classification with uneven training class sizes. 2005; Guangzhou, China, China. IEEE. p 4365-9.
36. Brouwer de Koning SG, Vrancken Peeters MTFD, Jóźwiak K, Bhairosing PA, Ruers TJM. Tumor resection margin guidelines in breast cancer – a systematic review and meta-analysis of current literature. *Clin Breast Cancer* **2018**;18(4):595-600.
37. Flock ST, Patterson MS, Wilson BC, Wyman DR. Monte Carlo modeling of light propagation in highly scattering tissues. I. Model predictions and comparison with diffusion theory. *IEEE Trans Biomed Eng* **1989**;36(12):1162-8.
38. Jacques SL, Ramella-Roman JC, Lee K. Imaging skin pathology with polarized light. *J Biomed Opt* **2002**;7(3):329-40.
39. Mountrakis G, Im J, Ogole C. Support vector machines in remote sensing: A review. *ISPRS J Photogramm* **2011**;66(3):247-59.

Author Manuscript Published OnlineFirst on March 18, 2019; DOI: 10.1158/1078-0432.CCR-18-2089
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Tables

Table 1. Classification results of hyperspectral analysis compared to pathology annotation

			Hyperspectral classification					
			TUMOR			HEALTHY		
			IC	DCIS	Total	Connective	Adipose	Total
Pathology annotation[a]	TUMOR	Invasive carcinoma (IC)	**87%**	6%	**93%**	7%	0%	7%
		Ductal carcinoma *in situ* (DCIS)	29%	**56%**	**84%**[b]	11%	4%	16%[b]
	HEALTHY	Connective	14%	16%	30%	**66%**	4%	**70%**
		Adipose	0%	1%	1%	1%	**98%**	**99%**

[a] IC (n = 4,823, p = 8), DCIS (n = 606, p = 9), connective (n = 615, p = 14) and adipose (n = 16,385, p = 18). With n = number of spectra and p = number of patients.

[b] due to rounding of percentages, the total percentage is not equal to the sum of IC and DCIS or connective and adipose tissue.

Figure legends

Figure 1. Hyperspectral data acquisition and hyperspectral data. A) The hyperspectral measurement setup. The breast specimen was placed upon the translational stage and imaged line-by-line by moving the translational stage. Thereby, an area of 20 x 10 cm^2 was scanned in 4 seconds, creating a 3D hyperspectral image (B), in which each pixel contains one spectrum (C).

Figure 2. Representative examples of two tissue slices. From left to right: (A, E) white light image, (B, F) histopathology H&E section, (C, G) annotated H&E section and (D, H) hyperspectral (HS) classification result using the support vector machine algorithm. The malignant tissue in the upper row (A-D) is mainly IC, whereas the malignant tissue in the bottom row (E-H) consists of malignant pockets of DCIS. White pixels in the classified HS image indicate pixels that could not be classified due to either high specular reflection or the presence of pathology ink. Grey pixels are pixels of which no histopathology was available. The arrow (D) points out the incorrectly classified connective spectra that are located near the border of the area with IC.

Figure 3. ROC curves of one-against-one support vector machine classification. ROC curves are calculated for (A) IC, (B) DCIS, (C) connective and (D) adipose tissue against the other classes. ROC curves and the area under the ROC curve (AUC) are calculated on the test set for each tissue slice and subsequently averaged. The AUC is given as the average ± standard deviation.

Figure 4. Relation between classification accuracy and pocket size. (A) Representative example of the histopathology annotation of one of the tissue slices that contains both invasive carcinoma (IC) and adipose tissue. For each pixel in the tissue slice, the shortest distance to the IC-adipose tissue transition is calculated. The scale bar indicates 4 mm. (B) The graph shows the hyperspectral classification accuracy of pixels in the white rectangle, which decreases towards the tissue transition. (C) The relation between the classification accuracy and the smallest pocket size in which a pixel was located, for all pixels in the dataset per tissue class.

Figure 5. Representative example of hyperspectral analysis on resection surface. Both white light (A) and hyperspectral images were taken from six sides so that the entire resection surface was imaged. (B) Classification result of the hyperspectral images using the support vector machine algorithm, developed with the tissue slices. The specimen was sliced according to standard pathology protocol and 6 H&E sections were processed for further analysis. (C) The orientation of these sections with respect to the 3-dimensional representation of the hyperspectral classified specimen. H&E sections 1 and 6 were taken parallel to the resection surface and hence do not provide information on the margin width. Therefore, these sections were excluded from analysis. (D) In H&E sections 2-5, an experienced pathologist annotated the tumor so that the resection margin width could be assessed.

Figure 1.

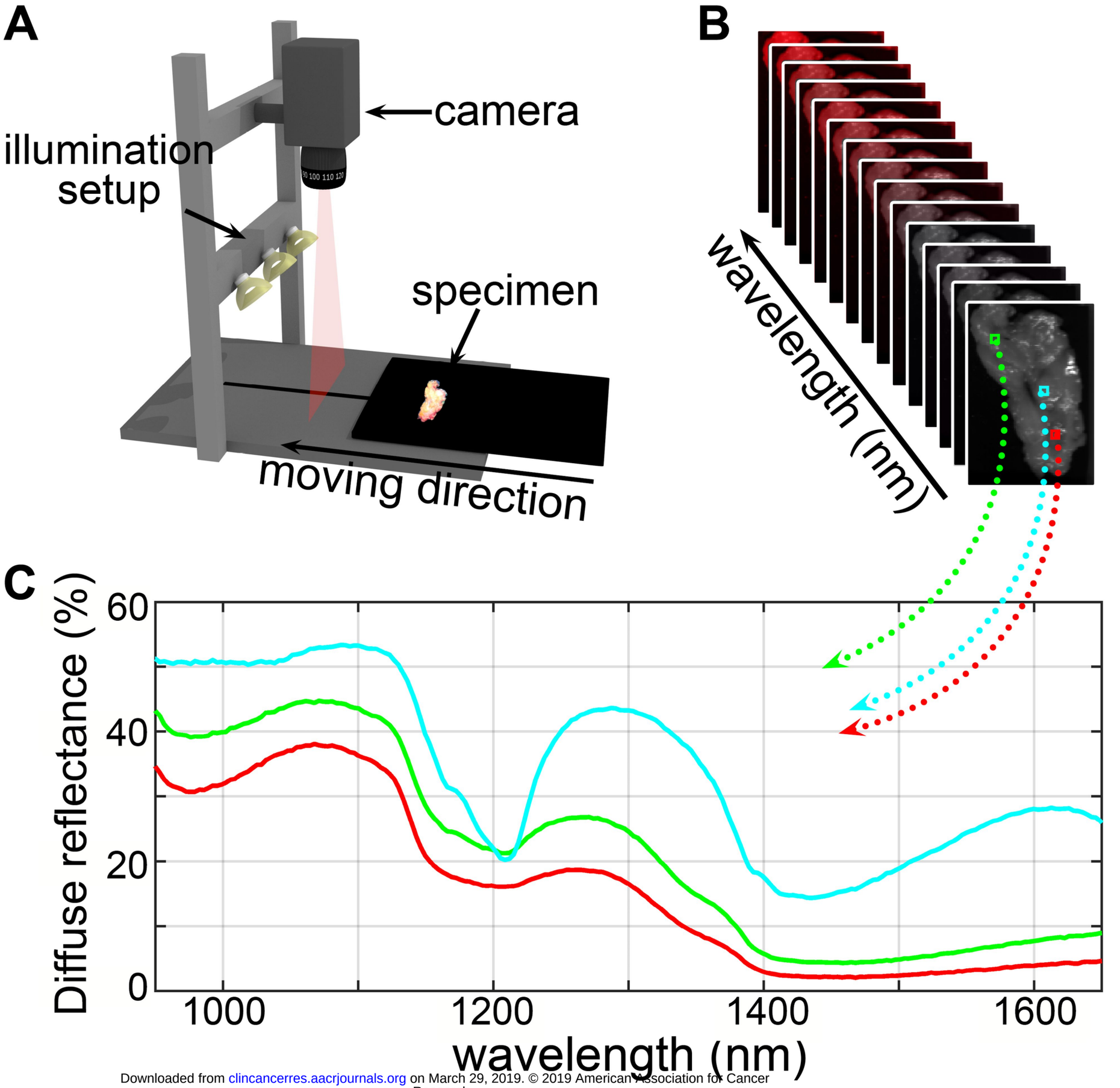


Figure 2.

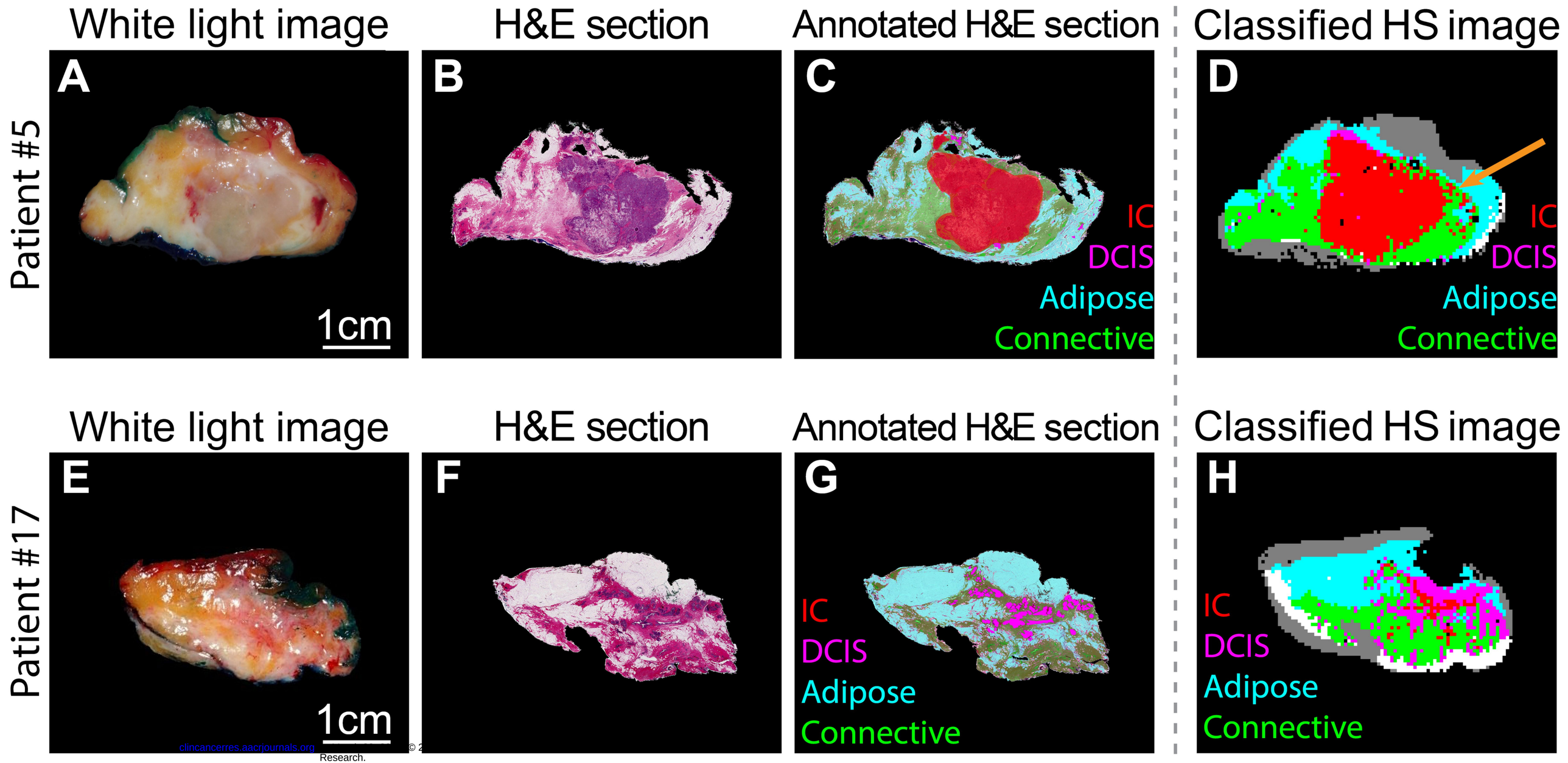
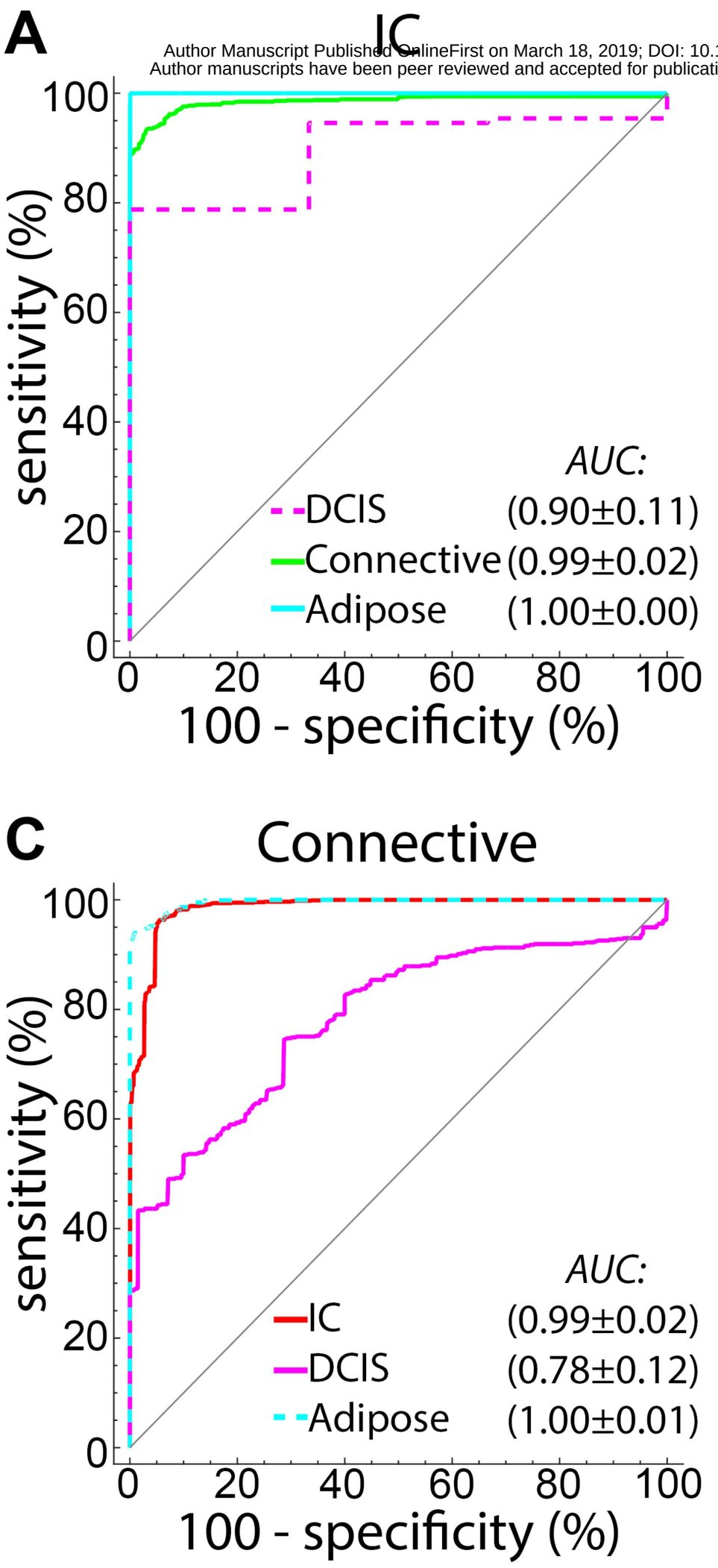


Figure 3.



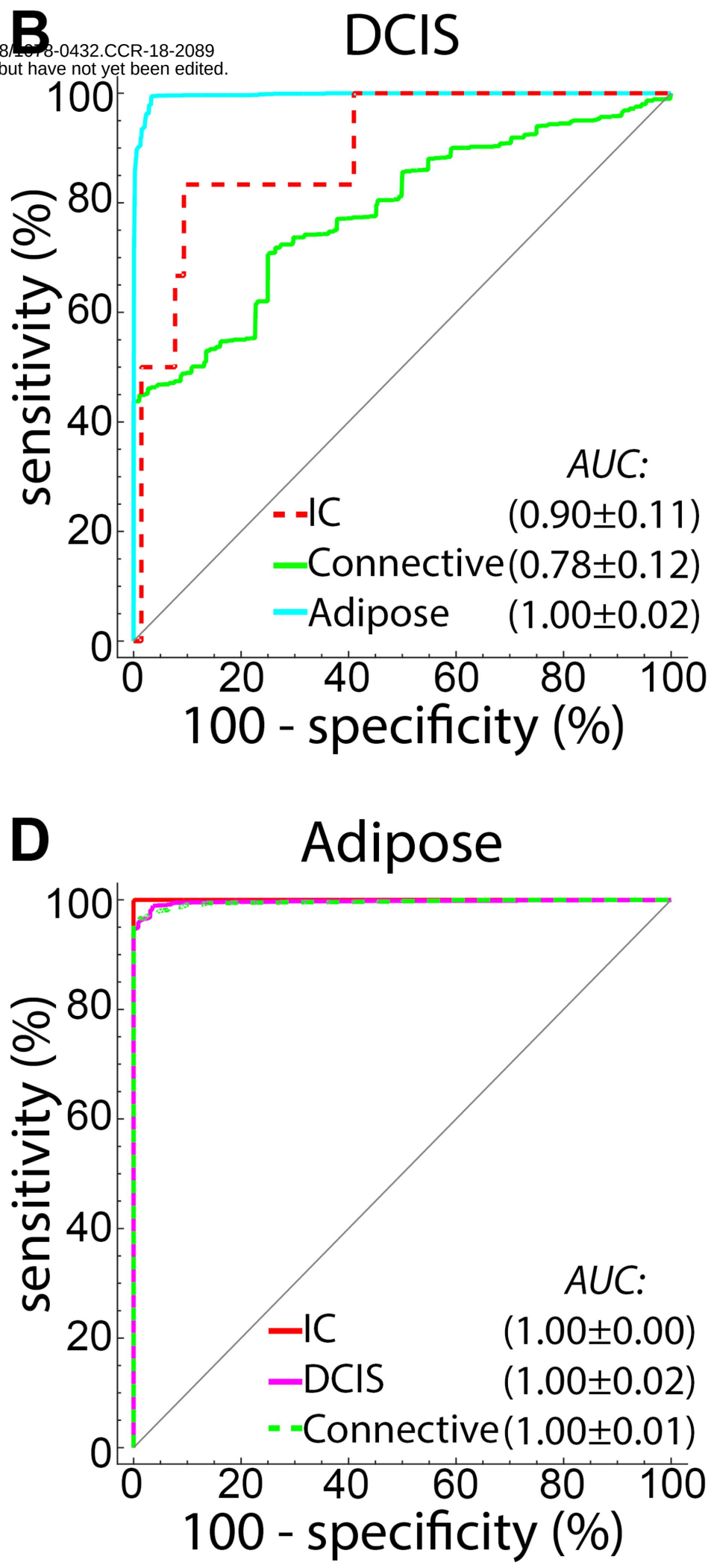


Figure 4.

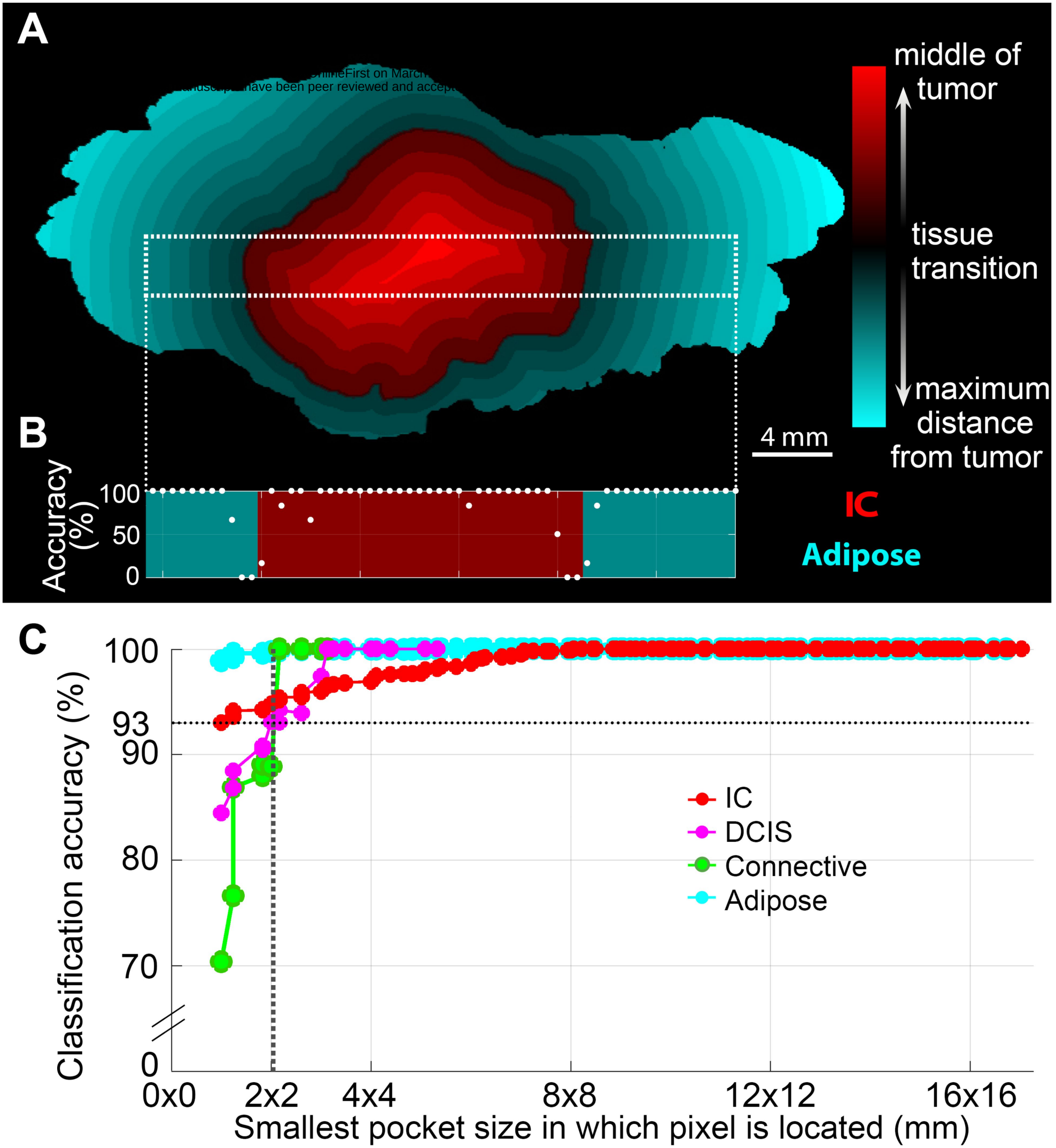
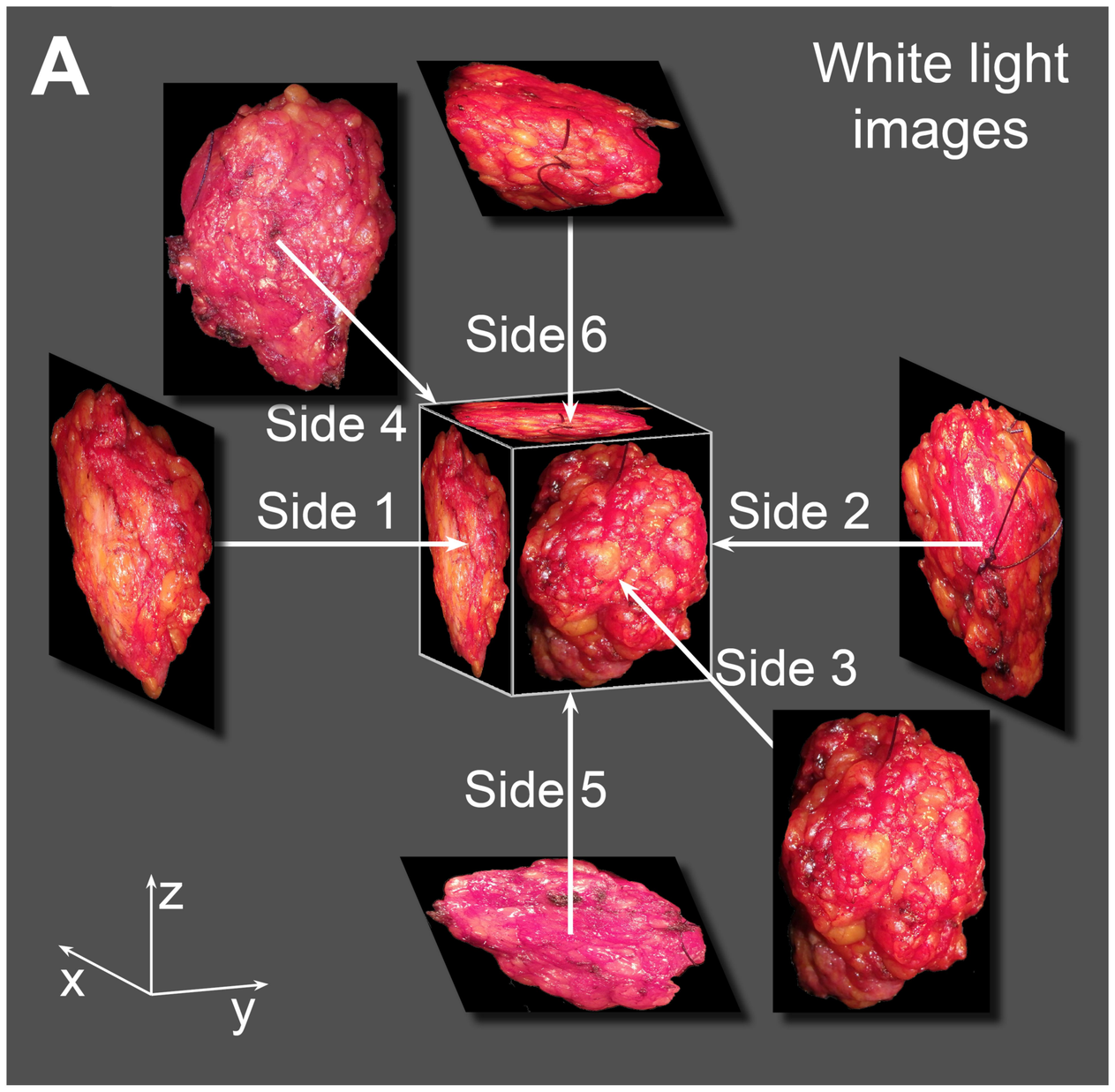
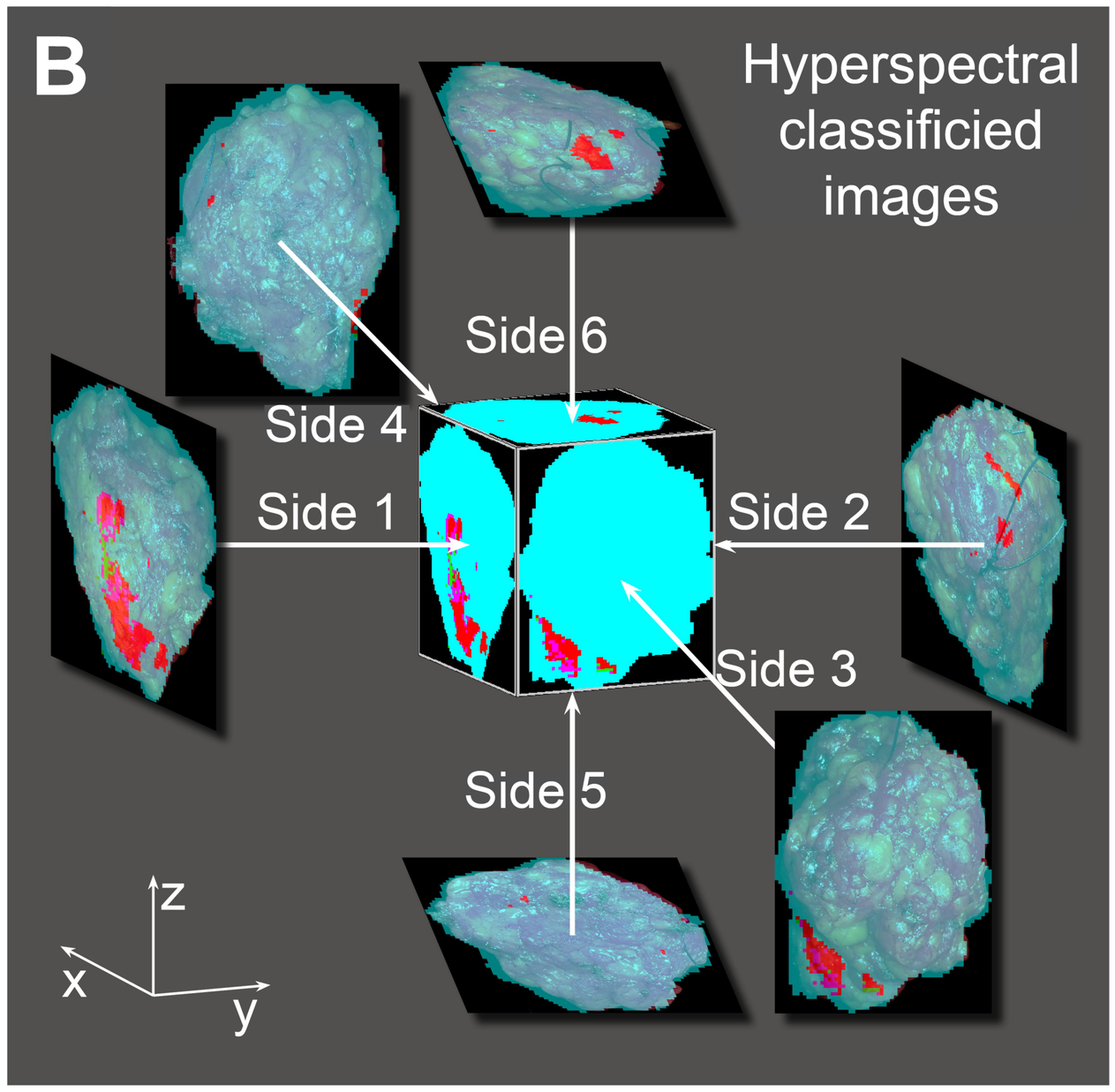
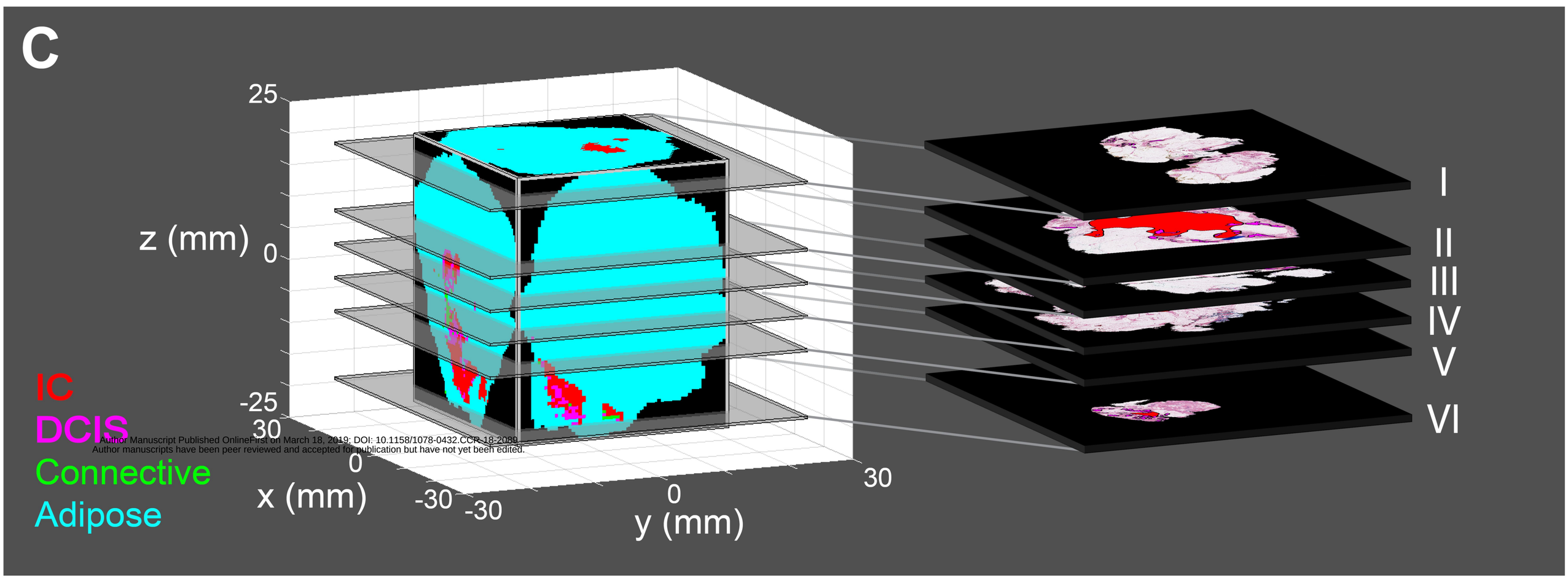
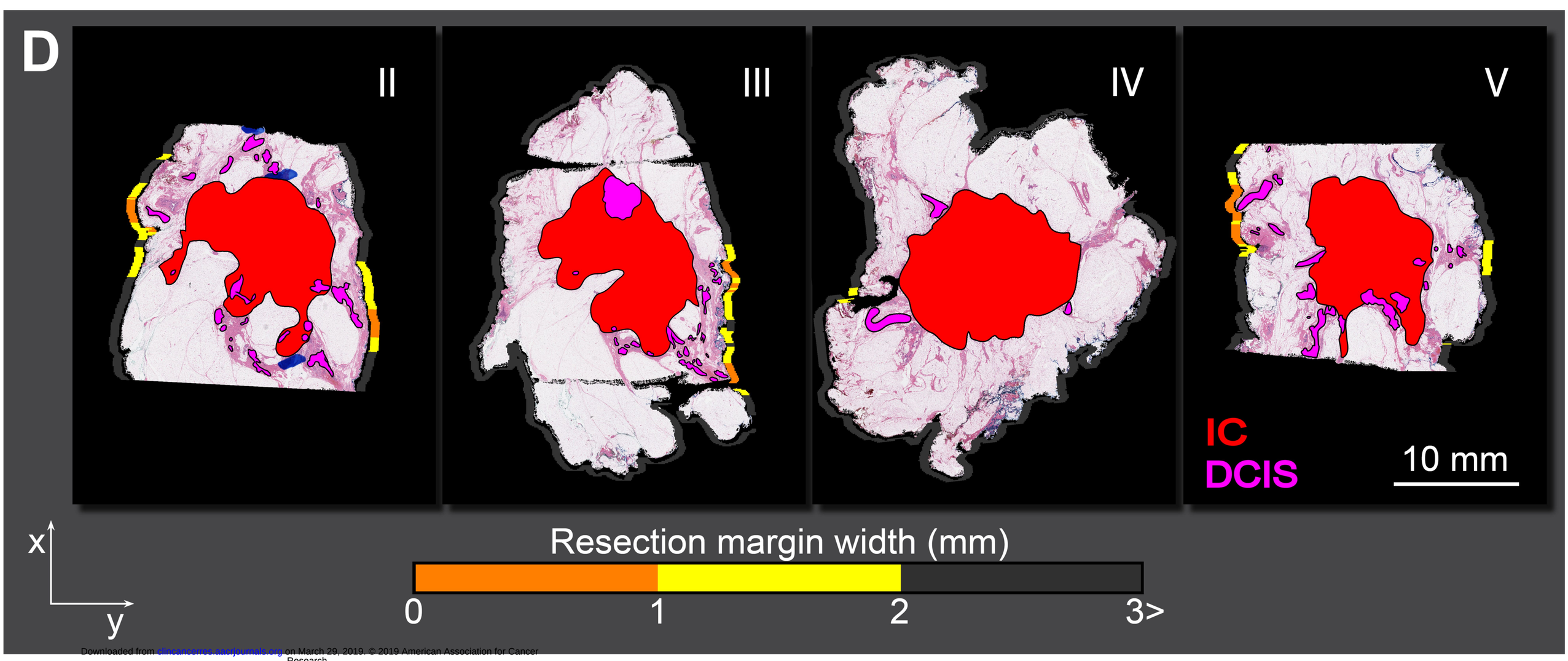


Figure 5.









AACR American Association for Cancer Research

Clinical Cancer Research

Hyperspectral imaging for resection margin assessment during cancer surgery

Esther Kho, Lisanne L. de Boer, Koen Van de Vijver, et al.

Clin Cancer Res Published OnlineFirst March 18, 2019.