

Journal homepage: <https://modernpathology.org/>

Research Article

Rapid On-Site Histology of Lung and Pleural Biopsies Using Higher Harmonic Generation Microscopy and Artificial Intelligence Analysis

Laura M.G. van Huizen^{a,*}, Max Blokker^a, Johannes M.A. Daniels^b, Teodora Radonic^c, Jan H. von der Thüsen^d, Mitko Veta^e, Jouke T. Annema^b, Marie Louise Groot^a

^a Department of Physics, LaserLab Amsterdam, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; ^b Department of Pulmonary Diseases, Amsterdam University Medical Centers, Amsterdam, The Netherlands; ^c Department of Pathology, Amsterdam University Medical Centers, Amsterdam, The Netherlands; ^d Department of Pathology, Erasmus MC, Rotterdam, The Netherlands; ^e Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

ARTICLE INFO

Article history:

Received 27 February 2024

Revised 3 October 2024

Accepted 3 October 2024

Available online 16 October 2024

Keywords:

deep learning
higher-harmonic generation
microscopy
lung biopsy
lung cancer
multiinstance learning
multiphoton excited
autofluorescence

ABSTRACT

Lung cancer is one of the most prevalent and lethal cancers. To improve health outcomes while reducing health care burden, it becomes crucial to move toward early detection and cost-effective workflows. Currently, there is no method for the on-site rapid histologic feedback on biopsies taken in diagnostic, endoscopic, or surgical procedures. Higher harmonic generation (HHG) microscopy is a laser-based technique that provides images of unprocessed tissue. In this study, we report the feasibility of an HHG portable microscope in the clinical workflow in terms of acquisition time, image quality, and diagnostic accuracy in suspected pulmonary and pleural malignancy. One hundred nine biopsies of 47 patients were imaged and a biopsy overview image was provided within a median acquisition time of 6 minutes after excision. The assessment by pathologists and an artificial intelligence algorithm showed that image quality was sufficient for a malignancy or non-malignancy diagnosis in 97% of the biopsies, and 87% of the HHG images were correctly scored by the pathologists. HHG is therefore an excellent candidate to provide a rapid pathology outcome on biopsy samples, enabling immediate diagnosis and (local) treatment.

© 2024 THE AUTHORS. Published by Elsevier Inc. on behalf of the United States & Canadian Academy of Pathology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

Lung cancer is one of the most prevalent and lethal cancers, with 2.2 million cases being reported in 2020.¹ The incidence of new lung cancer cases is expected to increase to 3.6 million cases per year in 2040. To improve health outcomes while reducing the health care burden, it becomes crucial to move toward early detection, cost-effective workflows, and minimally invasive,

effective early therapy. An increasing number of countries have started or are preparing the implementation of national lung screening programs. A single hospital visit for diagnosis and (local) treatment would greatly reduce health care burden and improve patient outcomes. Providing an instant pathology outcome on biopsy samples will enable this.

Current on-site feedback techniques are frozen section analysis, which provides histologic information with a high diagnostic accuracy of 95% to 98%^{2,3} but takes at least 20 minutes to provide feedback to the operator and suffers from artifacts because of the freezing process,⁴⁻⁶ and rapid on-site evaluation (ROSE) is fast,⁷ but it allows only the assessment of cytologic and not histologic tissue samples.

* These authors contributed equally: Jouke T. Annema and Marie Louise Groot.

* Corresponding author.

E-mail address: l.m.g.van.huizen@vu.nl (L.M.G. van Huizen).



Higher harmonic generation (HHG) microscopy is a noninvasive laser-based imaging technique that provides high (subcellular)-resolution images within seconds, without a need for processing of the tissue, through third and second harmonic generation (THG/SHG) and 2-photon excited autofluorescence (2PEF) microscopy. Previous studies by us and other groups have shown the ability of HHG microscopy to successfully discriminate normal and malignant tissue based on morphologic information for several applications, such as for the brain, breast, skin, liver, pancreas,^{8–14} and lung.¹⁵ In lung tissue, HHG microscopy visualizes collagen (SHG), elastin (2PEF), and cellular structures, including tumor cells, pneumocytes, and leukocytes (THG).^{11,15–21} This allows the rapid identification of alveolar and bronchial structures and histopathologic hallmarks, including cell morphology and general tissue architecture (collagen and elastin organization), features that are necessary to diagnose lung tissue. Previous studies have explored the use of HHG microscopy on animal^{19,21} and human lobectomy samples,^{15,17,18,20} and canine core needle biopsies¹¹ in an off-site laboratory setting or during veterinary canine surgeries.

In this study, we have reported on the performance of a fast and portable HHG microscope, placed in the clinical workflow of the Amsterdam University Medical Center lung endoscopy department. We have focused on the following 2 key feasibility aspects: acquisition time and image quality. We have also assessed diagnostic accuracy in bronchoscopic and thoracoscopic biopsies, using both pathologist and artificial intelligence (AI) analysis, for suspected lung or pleural malignancy.

Materials and Methods

Patients

Patients undergoing bronchoscopy or thoracoscopy with histologic tissue samples for suspected lung cancer (recurrence) or pleural malignancy were eligible for study participation.

Study Flow and Higher Harmonic Generation Image Acquisition

Bronchoscopy and thoracoscopy were performed according to standard protocols, and the obtained biopsies were immediately imaged on-site, unprocessed, using a mobile HHG microscope. Following biopsy, a median of 2 (range, 1–5) fresh unprocessed biopsies were imaged consecutively. The biopsies were placed in a sample holder (35-mm-diameter μ-Dish with 0.17-mm-thick glass bottom; Ibidi GmbH) without a specific orientation. To prevent dehydration and shrinking of the biopsies until imaging, isotonic saline (in a tube or on a gauze) was used. After imaging, the biopsies were directly placed in 4% formaldehyde and processed according to routine pathology practice. This resulted in 3-μm-thick hematoxylin and eosin–stained slides, which were digitized with a slide scanner (UPS, Philips) and visualized with a pathology viewer (Philips, version 3.2). ([Supplementary Fig. S1](#)).

HHG images were acquired using the mobile FD1070 microscope (Flash Pathology B.V.) and consisted of the following 3 simultaneously detected nonlinear signals: THG, SHG, and 2PEF ([Fig. 1](#)). A 50-femtosecond 1070-nm laser source (Fidelity 2, Coherent) was used to generate the nonlinear signals. The microscope was equipped with an acousto-optic modulator to retain sufficient peak power to generate nonlinear optical signals but at the same time to avoid damaging of the tissue. Bunches of 5 to 10

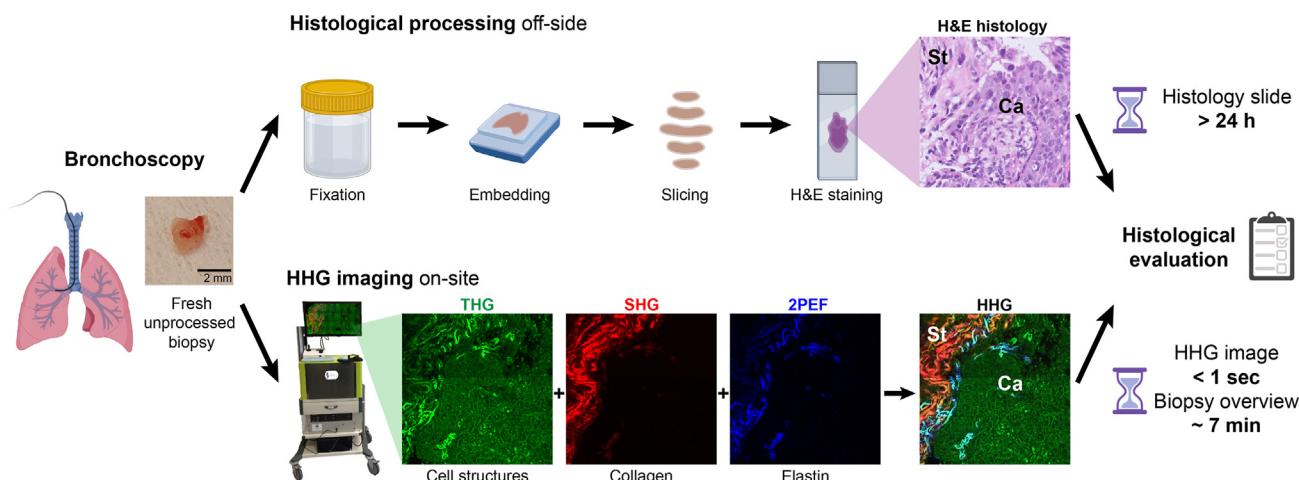
pulses at a repetition rate of 1 MHz out of the 70-MHz pulse train were selected to achieve a low average power of ~5 mW on the sample, with a pulse peak energy of ~5 nJ. No effects of HHG microscopy imaging were noticed in the frozen sections or processed histologic slides.

The THG, SHG, and 2PEF signals were collected in the back-scatter direction using analog photomultiplier tubes (H10720 and H7422; Hamamatsu). Dichroic mirrors and interference filters were used to separate the 3 signals, resulting in the following detection bandwidths: THG, 349 to 361 nm; SHG, 530 to 540 nm; and 2PEF, 562 to 665 nm.

A submicrometer focus was achieved with the use of an oil-immersion objective with a high numerical aperture (Super Fluor ×40/1.3-Numerical Aperture oil immersion; Nikon), resulting in an optical resolution of ~0.4 × 0.4 × 2.4 μm.³ The THG/SHG/2PEF images were acquired with a LabVIEW program (Flash Pathology B.V.), with an acquisition time of 1.0 second per image in a fast “inspection” mode (1 pixel/μm) and 3.6 seconds per image in a high-quality mode (2 pixels/μm). Mosaic images of larger areas were generated by moving a motorized stage and stitching the images together. For every biopsy, one or more “overview” images (with 1 pixel/μm) of maximum 5.0 × 5.0 mm² were generated, followed by one or more “higher quality” images (with 2 pixels/μm) of maximum 2.5 × 2.5 mm² at relevant depths and depth scans of maximum 0.5 × 0.5 mm² of relevant regions of interest ([Supplementary Fig. S1](#) and [Supplementary Video S1](#)). In all images, THG signals were displayed in green, SHG signals in red, and 2PEF signals in blue. All 3 signals, with a dynamic range of 2000 each, were intensity scaled using gamma correction ($\gamma = 0.7$) and were saved as 24-bit Red Green Blue (8-bit per color) bitmap using built-in software (Flash Pathology B.V.). Subsequently, the images were processed using local contrast enhancement with “Image J” software (version 1.53f51; NIH). For visualization purposes in this study (not applied to the HHG images assessed by the pathologists), radial correction was used in the THG images of the stitched “overview” images (with 1 pixel/μm) to correct for the decreased THG intensity toward the image periphery caused by optical aberration.

Higher Harmonic Generation and Corresponding Histology Assessment

Two experienced pulmonary pathologists from different hospitals (T.R. and J.H.v.d.T.) independently evaluated both HHG and histology images blinded for each other for study purposes. Both pathologists received a form of training on HHG microscopy. The first pathologist (T.R.) had collaborated in our earlier study.¹⁵ The second pathologist (J.H.v.d.T.) familiarized himself with the imaging technique by studying a data set of this previous study, consisting of corresponding HHG and histology images of lobectomy samples of both normal (alveolar and bronchial structures) cases and 4 different malignant cases (3 adenocarcinoma and 1 carcinoid). To establish the potential pitfalls in tissue imaging and interpretation, during this study, both pathologists discussed the correctly and incorrectly diagnosed HHG-imaged biopsies, including 32 nonmalignant biopsies (including 6 preinvasive), 13 malignant biopsies, and 5 biopsies for which an HHG-based diagnosis could not be provided. After this 3-hour discussion as additional training and a 10-week wash-out period, the second pathologist reassessed the following subset of biopsies: all indeterminately scored, all incorrectly scored, and an equal number of correctly scored biopsies. The results from this assessment are reported in this study.

**Figure 1.**

A flow chart of using higher harmonic generation (HHG) microscopy vs standard histopathology in the assessment of bronchoscopic tissue analysis. Standard histopathology requires processing of the tissue, which takes at least 24 hours. On-site HHG microscopy generates images of fresh unprocessed tissue that consist of the following 3 signals: THG showing all cellular interfaces revealing mainly cellular structures, SHG showing collagen, and 2PEF showing elastin. Single HHG images are acquired within 1 second, and an overview of the biopsy is provided within 7 minutes from the moment the biopsy is placed in the sample holder. HHG images can be immediately evaluated. This example shows a biopsy with squamous cell carcinoma with remaining stroma. 2PEF, 2-photon excited autofluorescence; H&E, hematoxylin and eosin; SHG, second harmonic generation; THG, third harmonic generation

The pathologists analyzed both HHG images and histology images blinded for each other using online pathology scoring software (Slide Score B.V.). First, all HHG images were assessed randomly per biopsy, followed by all histology images. For every assessment, the pathologists performed the following functions: (1) provided a diagnosis, with the classification based on the national standardized proforma reporting protocol, including the World Health Organization classification; (2) selected the lowest level of HHG image quality deemed sufficient for malignancy assessment, for example an “overview image” or a “higher quality image”; and (3) indicated other abnormalities, such as inflammation, fibrosis, and necrosis. For the case report form, see Supplementary List S1.

Higher Harmonic Generation Data Analysis: Feasibility and Diagnostic Accuracy

HHG microscopy feasibility was determined in terms of acquisition time and image quality. The time from placing the biopsy in the sample holder until providing the first “overview” image was recorded. The feasibility in terms of image quality was determined for both HHG microscopy and histology, by calculating the percentage of biopsies that could not be diagnosed owing to image quality.

To determine diagnostic accuracy, the diagnoses were grouped into “malignancy,” “non-malignancy” (including pre-invasive diseases), and “indeterminable.” These HHG-based diagnoses were compared with the consensus histology diagnoses. The diagnostic agreement, positive predictive value, and negative predictive value were calculated, together with the interobserver agreement. The diagnostic accuracy values were calculated for both pathologists, and for the second pathologist, the scores after reassessment were used, by replacing the incorrectly and indeterminately scored biopsies of the first round with those of the second round.

Nuclear Diameter

After pathologic assessment, to provide a potential tool to distinguish normal cell areas from malignant cell areas in the HHG images, the average cell nuclear diameter was determined. For 36 cell areas with different diagnoses (normal, metaplasia, dysplasia, carcinoma in situ, carcinoma, and carcinoid), 20 nuclear diameters per area of $250 \times 250 \mu\text{m}^2$ were measured using the ImageJ software (version 1.53f1). To compare the average cell nuclear diameter between normal and carcinoma cell areas, a 2-sample 2-tailed *t* test was performed.

Artificial Intelligence Analysis: Higher Harmonic Generation Data Preparation

For the proof-of-principle real-time diagnostic analysis, a deep learning pipeline was built and evaluated. An ensemble of AI models was engineered, mimicking the consensus of a team of pathologists and therefore, to improve performance, trained to perform binary (malignancy vs nonmalignancy) classification.

From the biopsies available, selection was made on which the consensus of our pathologists on the HHG imagery matched with the histology assessment. A total of 69 biopsies (61 bronchoscopic and 8 thoracoscopic) were included in the data set. These included 13 malignant biopsies, and 56 preinvasive and noninvasive biopsies. Opting for binary instead of multiclass classification kept the focus on the main task of distinct malignant tumor detection in the HHG data. The data set was split on the patient level into a training set (90%) and an independent testing set (10%), ensuring that each data set contained a balanced number of biopsies from both classes. Supplementary Table S1 lists the included biopsies and their respective data set split.

Per biopsy, all available higher quality images were included. To incur a high level of randomness while optimizing the deep learning models, we opted for 10-fold Monte-Carlo cross-validation. Instead of training 1 model, 10 models (forming an ensemble)

were trained with the exact same architecture and hyperparameters. Each model was trained on a different training-validation split obtained from the original training data set. Therefore, each of the ensemble's model had seen different training and validation data. The testing set was only involved after model optimization to perform independent evaluation of each model's performance. To combat data imbalance in the training and validation sets resulting from the Monte-Carlo splits, minority oversampling was applied.

Each included HHG image was split up into smaller tiles based on the original scanning settings. Each tile was resized to matching 0.5- μ m pixel size across the data set to keep cellular features in a fixed size during model optimization. To exclude noisy, almost black, noninformative tiles resulting from scanning biopsy surroundings, an exclusion criterion was defined based on the mean tile intensity. To induce further model regularization, tile augmentation was employed during batch generation on both the training and the validation data, including random contrast, brightness, and rotation augmentations.

Artificial Intelligence Analysis: Model Engineering and Optimization

Owing to the dynamic image input sizes, a fully convolutional model was engineered. Initially, the model was built in varying sizes of stacked blocks of convolution, batch normalization, rectified linear unit activation, and 2×2 max pooling to find the optimal model size. For our purpose, 5 stacked blocks proved most optimal based on validation performance. Each convolution layer performs a 5×5 convolution operation, except for the last layer, which performs 1×1 convolution to mimic a fully connected layer. The convolution layers of these stacked blocks comprise 32, 64, 128, 256, and 64 filters, respectively. All filter weights were initialized with He uniform variance scaling and were L2 regularized with factor 0.01. Multiple-instance learning on the image level was enabled using Noisy-AND²² after the fifth convolution block, closing the model with a dense layer with 1 output neuron with sigmoid activation for the final binary classification output. All models were optimized by minimization of the binary cross-entropy loss with the Adam optimizer with 0.0001 learning rate. Early stopping was employed after 20 epochs if the validation loss did not improve. Given the Monte-Carlo cross-validation folds, 10 models were uniquely trained given these hyperparameters. The pipeline was built with TensorFlow²³ (version 2.10; Google) and executed on a 4-Graphics Processing Unit workstation (Lambda Quad; Lambda Labs Inc).

Artificial Intelligence Analysis: Evaluation and Anomaly Detection

After training, the model ensemble was evaluated on the independent testing set. Each model was tested with the same, identical HHG imagery, which was not available during model optimization. The testing set was prepared similarly to the training set, with exception for the noninformative tile exclusion, to mimic clinical practice and avoid introducing a bias by removal of data.

The final classification (malignancy vs rest) was obtained by applying a decision threshold on each model's sigmoid output, which ranged between 0 and 1. The optimal model decision threshold was found by computing a receiver operating characteristic area under the curve graph for the entire validation set per model and using Youden J statistic. Plurality voting was applied to

obtain the final prediction of the test image. Based on these, the testing accuracy, area under the curve, sensitivity, and specificity were calculated of the AI ensemble.

Aside from these metrics, we constructed an anomaly detection scheme by matching individual tile plurality votes with the overview vote, and annotating mismatches as anomalies.

Results

Forty seven patients were included, resulting in 109 HHG images and corresponding histology images of bronchoscopic ($n = 94$) and thoracoscopic ($n = 15$) biopsies, of which 2 bronchoscopic biopsies were excluded because the corresponding histology slides were not available. The histopathologic consensus diagnoses based on the standard hematoxylin and eosin staining analysis were pulmonary malignancy ($n = 26$; adenocarcinoma, squamous cell carcinoma, carcinoid, and epithelioid mesothelioma), preinvasive disease ($n = 10$; carcinoma in situ and squamous dysplasia), nonmalignancy ($n = 58$; normal bronchus/pleura and squamous metaplasia), and indeterminable ($n = 13$) (specified diagnoses are shown in [Supplementary Table S2](#)). Reasons to score as "indeterminable" were "not representative" ($n = 7$), "insufficient tissue or poor scan quality" ($n = 5$), and "indeterminable for dysplasia" ($n = 1$).

Higher Harmonic Generation Microscopy Imaging and Feasibility

A transportable HHG microscope codeveloped with Flash Pathology B.V., with dimensions of $0.6 \times 0.8 \times 1.5$ m³ and a weight of 40 kg, was placed in the endoscopy suite ([Fig. 1](#)). Biopsies were handed to the operator (L.M.G.v.H.) immediately after extraction and placed in a small container with an optical window at the bottom. The container was placed in the HHG microscope and nonlinear images were generated by scanning a 50-femtosecond laser over the sample from below, employing 5-mW power. SHG, THG, and 2PEF signals were collected in back scatter geometry and spectrally separated on 3 detectors. Single HHG images were acquired within 1 second, enabling almost real-time scanning. Larger view images were made by moving the sample with a translation stage. The first overview image of the biopsies was provided in a median acquisition time of 6 minutes (IQR = 3 minutes), with the overview images itself generated within 1 to 2 minutes. A higher quality image was acquired within 2 to 3 minutes. Immediate on-site evaluation of the acquired images was possible. [Figure 1](#) shows the difference in clinical workflow between the standard histopathology and using HHG microscopy. Samples were handed back after imaging and entered the normal pathology workflow.

Images were assessed by 2 experienced lung pathologists (T.R. as pathologist 1 and J.H.v.d.T. as pathologist 2). The feasibility score in terms of image quality for HHG microscopy was 96.8%, according to the pathologists, with an interobserver agreement of 95.3% (Cohen kappa = 0.26; 95% CI [-0.18, 0.71]) ([Table 1](#) and [Supplementary Table S3](#)). Reasons to rate the HHG image quality as too low were either insufficient cellular detail or an incomplete image. This image quality score was comparable with the average image quality score of the histology slides of 95.8%. However, in histology, there was an almost perfect interobserver agreement of 99.1% (Cohen kappa = 0.88; 95% CI [0.66, 1.00]) ([Table 1](#) and [Supplementary Table S4](#)). Reasons to rate the histology slide quality as too low were either insufficient or denaturized tissue.

Table 1

Feasibility of HHG image quality to assess lung biopsies compared with standard histology

Total = 107	Biopsies with sufficient image quality	
	HHG	Histology
Pathologist 1	104 (97.2%)	103 (96.3%)
Pathologist 2	103 (96.3%)	102 (95.3%)
Interobserver agreement	95.3% (kappa = 0.26; 95% CI [-0.18, 0.71])	99.1% (kappa = 0.88; 95% CI [0.66, 1.00])

Feasibility, of both HHG and histology images, defined as suitability to make a classifying diagnosis, was determined per biopsy and per pathologist, including their interobserver agreement (with 95% CI). Reasons to rate an image as of insufficient quality were insufficient cellular detail or an incomplete image (HHG) or insufficient or denaturized tissue (histology). Lung biopsies include 92 bronchoscopic and 15 thoracoscopic biopsies. HHG, higher harmonic generation.

Higher Harmonic Generation Microscopy: Diagnostic Accuracy for Malignancy

The HHG-based diagnoses, malignant or nonmalignant (including noninvasive and preinvasive), compared with the consensus histology diagnoses (Table 2 and *Supplementary Table S5*), were in a substantial agreement of 85.6% (Cohen kappa = 0.63; 95% CI [0.44, 0.81]) for the first pathologist and of 89.1% (Cohen kappa = 0.73; 95% CI [0.58, 0.89]) for the second pathologist. The interobserver agreement for HHG microscopy was 88.5% (Cohen kappa = 0.72; 95% CI [0.56, 0.87]) and that for histology was 96.6% (Cohen kappa = 0.92; 95% CI [0.83, 1.00]). Specific individual histology diagnoses of both pathologists are shown in *Supplementary Table S6*.

Important to note is that of the 13 “indeterminable” biopsies based on histology, 46% (pathologist 1) and 69% (pathologist 2) had an HHG-based diagnosis that was in agreement with the final multidisciplinary diagnosis. Only one of these biopsies received a false-negative score from the first pathologist and no false positives (*Supplementary Table S5*).

Correctly diagnosing the presence of malignancy was possible because the HHG images reveal similar tissue components as standard histology. This includes cellular structures, such as squamous and respiratory epithelium and immune cells (macrophages, granulocytes, and lymphocytes), and stroma components, such as collagen and elastin fibers, glands, fat cells, blood vessels, and cartilage. This allows the identification of normal mucosa and pleura, squamous cell metaplasia, fibrosis, and elastosis (Fig. 2). Histopathologic hallmarks to distinguish nonmalignant from malignant tissue are increased cellularity, cellular pleomorphism, and growth pattern, which are visualized using THG. In addition, the identification of the stroma components, collagen (SHG) and elastin (2PEF) fibers, enabled assessment of the stroma for cellular invasion. Figure 3 shows examples of increased cellularity (panels

I-III), cellular pleomorphism (panels I and II), abnormal cellular growth patterns (panel III), and invasion in the stroma (panels I and II) (more examples are shown in *Supplementary Figs. S2* and *S3*). Besides cellular morphology, THG signals reveal cell nuclear morphology (Fig. 4). Abnormal cell areas showed enlarged cell nuclei and more prominently present cell nucleoli.

In future application, measuring the average cell nuclear diameter could serve as a helpful tool to indicate abnormal cells. The HHG images reveal a significant difference between carcinoma and normal cell diameters (10.6 μm , SD = 1.2 μm vs 7.4 μm , SD = 0.7 μm ; $P < .001$) (Fig. 4B) (the cell nuclear diameters were measured after pathologic reassessment, and therefore did not influence the diagnostic accuracy). Note that carcinoids were excluded from the carcinomas, as carcinoids are well-differentiated neuroendocrine tumors with cells quite regular in size and shape, and therefore have smaller measured average cell nuclear sizes (8.7 μm , SD = 1.7 μm). In case of carcinoids, a good discriminative feature is the distinct cellular growth pattern (Fig. 3III and Fig. 4D).

To account for potential pitfalls, for both operators and pathologists, when HHG microscopy is introduced in the clinic as a novel diagnostic tool, we looked in more detail at the misdiagnosed cases (*Supplementary Fig. S4*). First, in histologic processing, care is taken to have both stroma and epithelium in the same plane during the embedding of the tissue in paraffin. In this study, owing to the random biopsy orientation for HHG imaging, some biopsies were imaged from the epithelial side, resulting in homogeneous cellular areas and no or little connective tissue, analogous to tangential cutting artifacts in histopathology. The example in *Supplementary Figure S4I* shows that taking HHG images deeper into the tissue revealed the presence of collagen and clarified that the homogenous cell area was a superficial cell layer only. In addition, measuring the average cell nuclear diameter provided support that the homogeneous cell area consisted of normal cells.

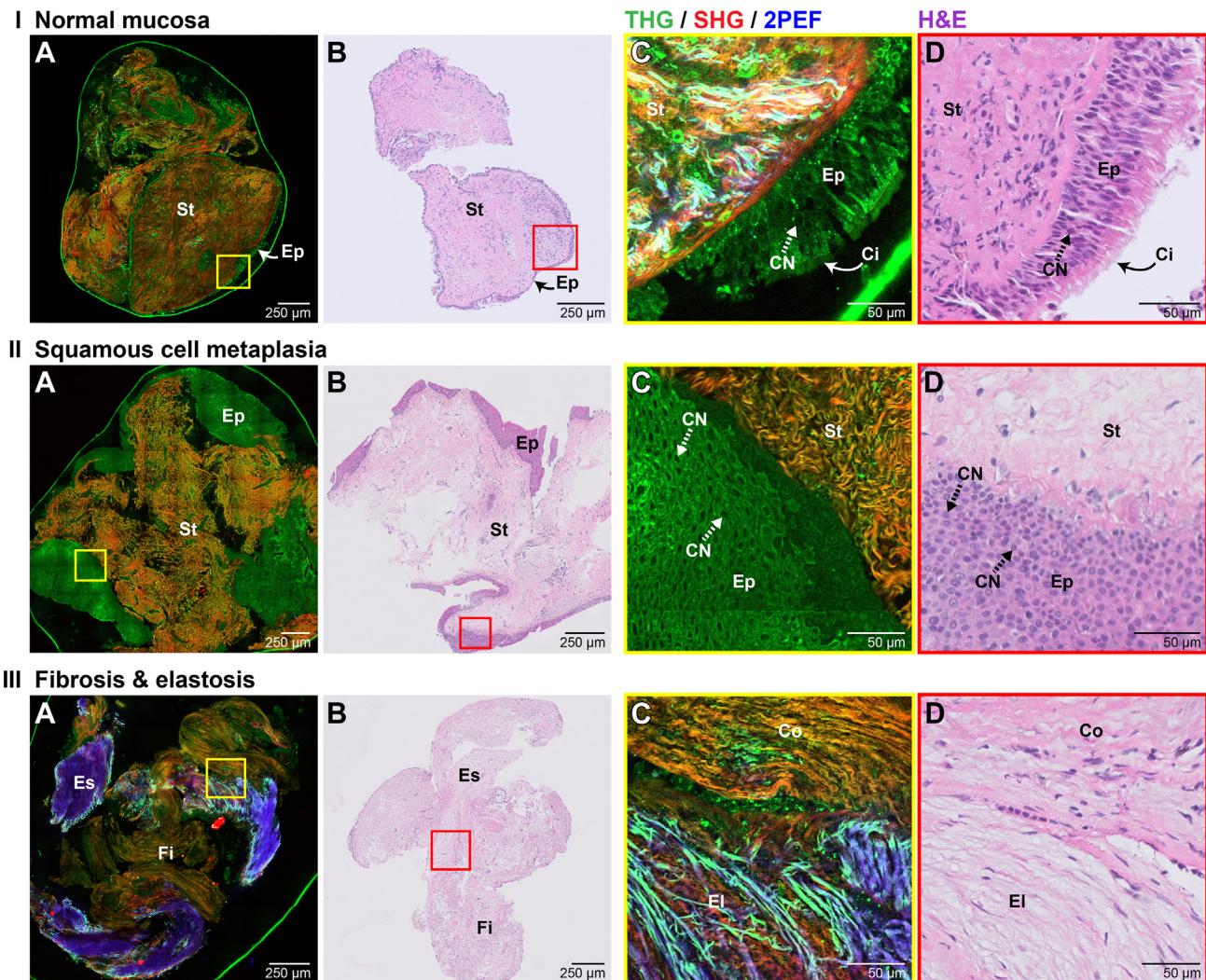
Table 2

HHG microscopy diagnostic accuracy for the presence of malignancy

HHG diagnostic accuracy			
Pathologist 1	NPV	60/67	89.6%
	PPV	17/23	73.9%
	Overall agreement	77/90	85.6% (kappa = 0.63; 95% CI [0.44, 0.81])
Pathologist 2	NPV	61/64	95.3%
	PPV	21/28	75.0%
	Overall agreement	82/92	89.1% (kappa = 0.73; 95% CI [0.58, 0.89])

The diagnosis of malignancy based on the HHG images by pathologist 1 and pathologist 2 with the corresponding histology consensus diagnoses as reference standard. Of the biopsies provided with a diagnosis (excluding 13 out of 107 biopsies scored as “indeterminable” based on histology, and 4 and 2 extra biopsies scored as “indeterminable” based on HHG by pathologist 1 and 2, respectively; see *Supplementary Table S5*), the diagnostic accuracy is determined per pathologist in terms of NPVs, PPVs, and overall agreement of HHG vs histology diagnosis (with 95% CI). NPV = true negatives/(true negatives + false negatives); PPV = true positives/(true positives + false positives).

HHG, higher harmonic generation; NPV, negative predictive value; PPV, positive predictive value.

**Figure 2.**

Higher harmonic generation images of nonmalignant biopsies with corresponding histology. (A, B) Higher harmonic generation microscopy images show stroma and epithelial cells, in which the epithelium layer is (panel I) thin in case of normal mucosa and (panel II) thickened in case of squamous cell metaplasia. In addition, (panel III) elastosis and fibrosis can be identified by an increased amount of elastin and collagen, respectively. (C, D) Magnified images show the cellular morphology of epithelial cells and their nuclei and cilia, and show the stroma components collagen and elastin. Image acquisition times of the “higher quality” overview images were (A) 1.8 minutes (panel I), 2.5 minutes (panel II), and 1.6 minutes (panel III). 2PEF, 2-photon excited autofluorescence; Ci, cilia; Co, collagen; El, elastin; Ep, epithelial cells; Es, elastosis; Fi, fibrosis; H&E, hematoxylin and eosin; St, stroma; SHG, second harmonic generation; THG, third harmonic generation.

Second, a number of the interpretive errors on the HHG biopsies were due to correct recognition of abnormal cells but incorrect assessment of invasion (Supplementary Fig. S4II). Preinvasive and malignant cells both show an increased nuclear diameter (Fig. 4B), and distinction was difficult in cases in which little or no connective tissue was present in the HHG images. Therefore, samples should be oriented by an experienced operator, as in classic histology, or HHG images should be taken deeper into the tissue to clarify if an underlying superficial cellular layer of connective tissue is present, to make an adequate assessment of invasiveness.

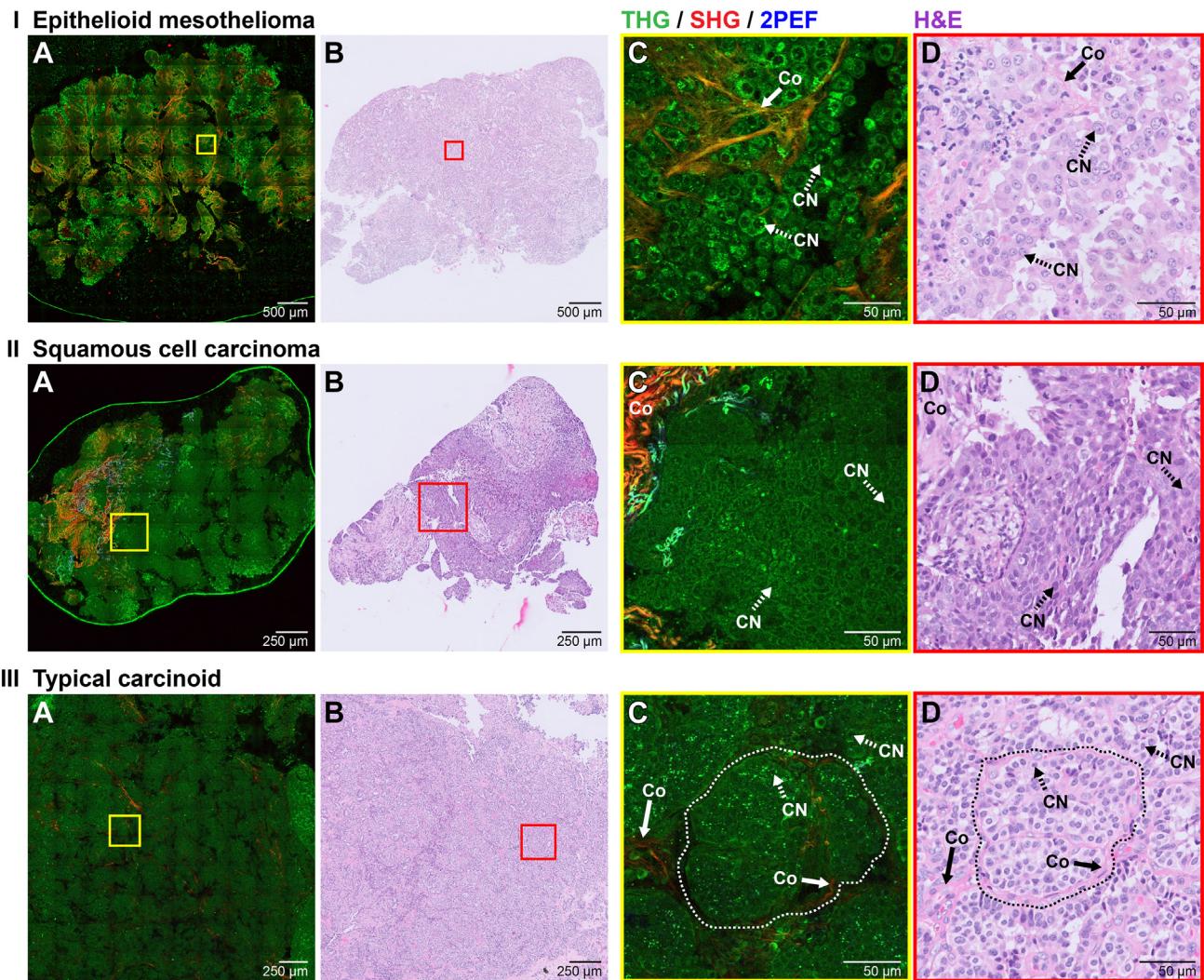
Third, control by pathologists over the separate contrast settings of the 3 different imaging modalities could have helped in interpreting the images (Supplementary Figs. S4III and S5). The HHG images have a large dynamic range of 2000 per image modality and were saved by the operator with specific contrast settings. However, in some cases, an alternative SHG contrast setting would have revealed the presence of collagen to identify blood vessels or to aid in diagnosing malignancy. In addition, adjusting the THG contrast

could have helped in the identification of necrosis and granulocytes, which generate strong THG signals, and identification of lymphocytes, which generate weak THG signals.²⁴

Finally, THG signals are generated by all tissue interfaces and therefore do reveal not only cellular structures but also interfaces of other tissue components, such as interfaces of cellular debris in necrosis and interfaces of fibrin fibers (Supplementary Fig. S4IV). These tissue structures were assessed as indeterminable or mistakenly scored as malignancy, because no cellular details were visible and they were not included in the training set, and therefore were unknown for the pathologists.

Artificial Intelligence Assistance to Interpret Higher Harmonic Generation Microscopy Images

The efficiency of the diagnostic process could be further aided by having a deep learning–based algorithm available. Such an

**Figure 3.**

Higher harmonic generation images of malignant biopsies with corresponding histology. (A, B) Higher harmonic generation microscopy images show malignant cells invading the connective tissue, diagnosed as epithelioid mesothelioma (panel I), squamous cell carcinoma (panel II), and typical carcinoid (panel III). (C, D) Magnified high-resolution images show histologic hallmarks, including cell and nuclear morphology (cell nuclei, CN), and cell growth patterns (panel III, dashed marked area). In addition, the remaining collagen (Co) is revealed. For the (panel IA) "overview" image and (panels IIA and IIIA) higher quality images, the acquisition times were 2.4 minutes, 2.0 minutes, and 2.5 minutes, respectively. For visualization purposes in this study, radial correction was used on the THG image of panel IA. H&E, hematoxylin and eosin; 2PEF, 2-photon excited autofluorescence; SHG, second harmonic generation; THG, third harmonic generation.

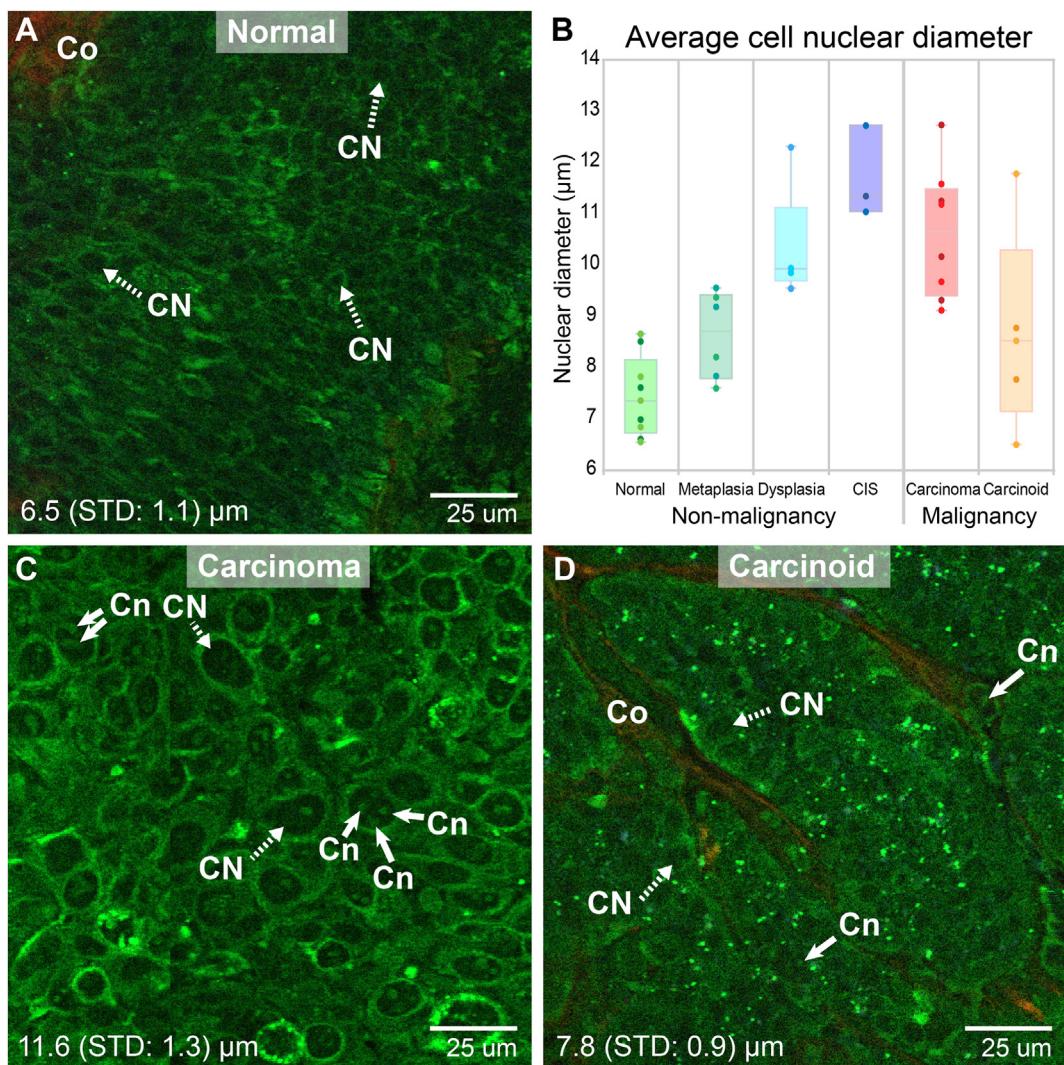
algorithm should be able to flag with high certainty that a biopsy contains tumor, to help the endoscopist decide to stop taking more biopsies. In addition, by identification of anomalies, it could point toward regions of interest in the biopsy. Deep learning allows for real-time tumor and anomaly detection by means of pattern detection from a representative training data set. Therefore, we used 69 biopsies to train an algorithm based on the pathologists' consensus on the HHG data matching the histology, from which 6 biopsies were set apart for independent testing. By employing minority oversampling (Fig. 5I), ensemble training, multiinstance learning (Fig. 5II), data augmentation, and plurality voting (Fig. 5III), this resulted in a binary classification ensemble that scored the individual tiles of the overview images into "0" for nonmalignant (including preinvasive and noninvasive) and "1" for malignant tumors. We obtained 71.1% accuracy, 0.71 area under the receiver operating characteristic curve, 68.3% sensitivity, and 72.7% specificity on the independent test set (Fig. 5IV). An

additional benefit was the ability to detect "anomalies," in which a single tile diagnosis did not match the overall image diagnosis ([Supplementary Fig. S6](#)).

Clinical Case

As an example of how rapid on-site histologic feedback can be of benefit in selected clinical cases by assessing the biopsy representativeness, we describe in this study an additional case, not included in the analysis above, as an indication of the added value of instant pathology with HHG.

A 70-year-old woman presented with symptoms of fatigue and dyspnea on exertion. Chest computed tomography showed left-sided unilateral pleural effusion with discrete pleural thickening and slight shrinkage of left hemithorax (Fig. 6A). The patient underwent thoracoscopy for diagnostic purposes. During

**Figure 4.**

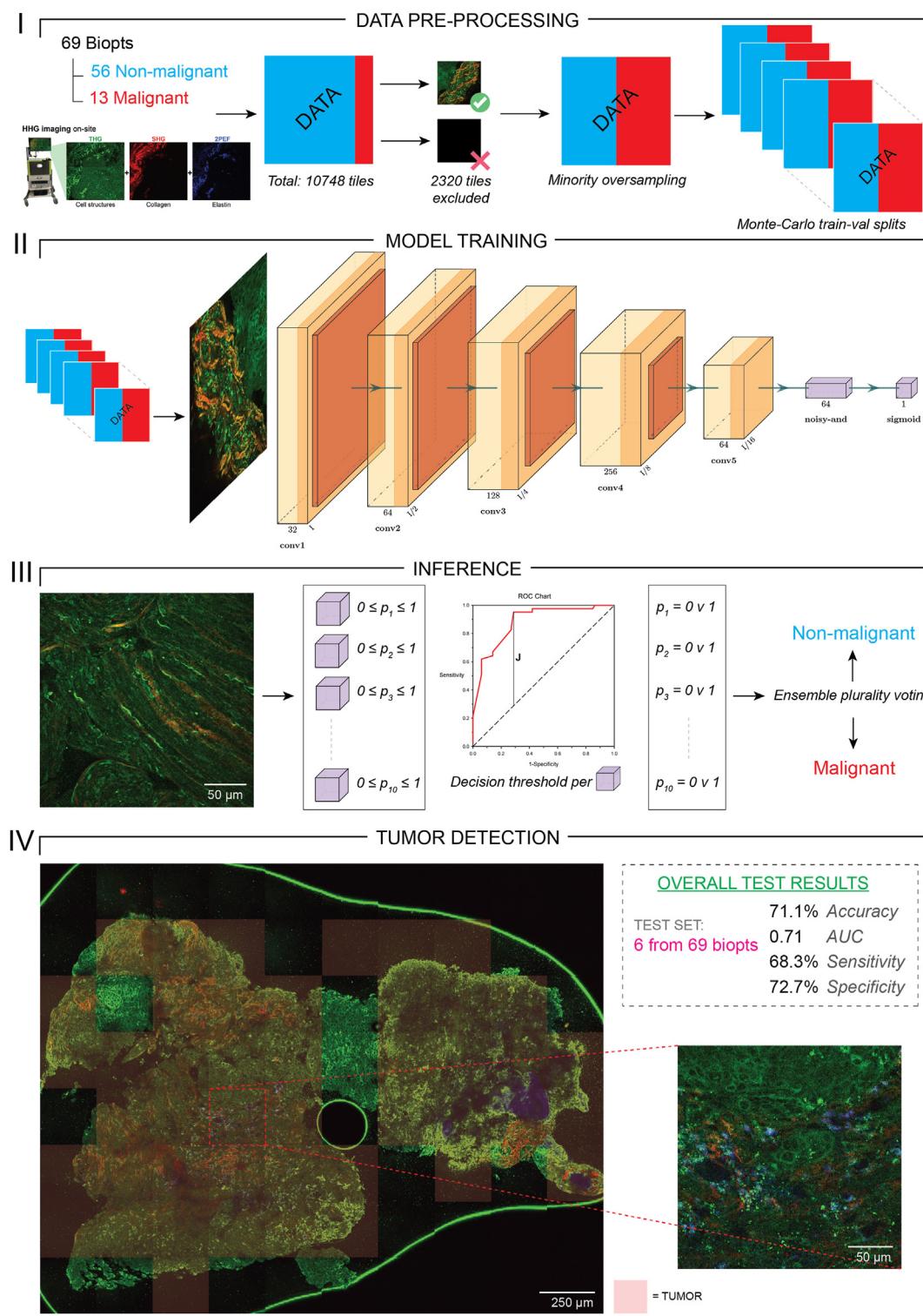
Higher harmonic generation microscopy reveals morphologic and quantifiable subcellular features potentially useful to indicate abnormal cells. (A, C, and D) Higher harmonic generation microscopy images of cell areas with a consensus histology diagnosis of no abnormalities, non-small cell carcinoma, and typical carcinoids, respectively. Cell nuclei (CN) in the carcinoma area are enlarged compared with those in the normal and carcinoid area. In the CN, cell nucleoli (Cn) are more prominently present in the carcinoma and carcinoid cells, in which some carcinoma cells show multiple cell nucleoli. In addition, collagen (Co) is revealed, highlighting the particular growth pattern of carcinoid cells. (B) The average cell nuclear diameter is determined by measuring 20 cells per area of $250 \times 250 \mu\text{m}^2$ from different cases with different diagnoses, showing an increased cell nuclear diameter for abnormal cells. CIS, carcinoma in situ.

thoracoscopy, a diffuse whitening of the parietal pleura was visualized without clear signs of tumor deposits. The normal architecture of ribs and intercostal spaces was absent (Fig. 6B). Based on white light inspection of the parietal pleura, there was a differential diagnosis of diffuse benign pleuritis, but pleural malignancy could not be excluded. Multiple biopsies were taken, and it was decided not to perform talc pleurodesis, as malignant pleural disease was not evident.

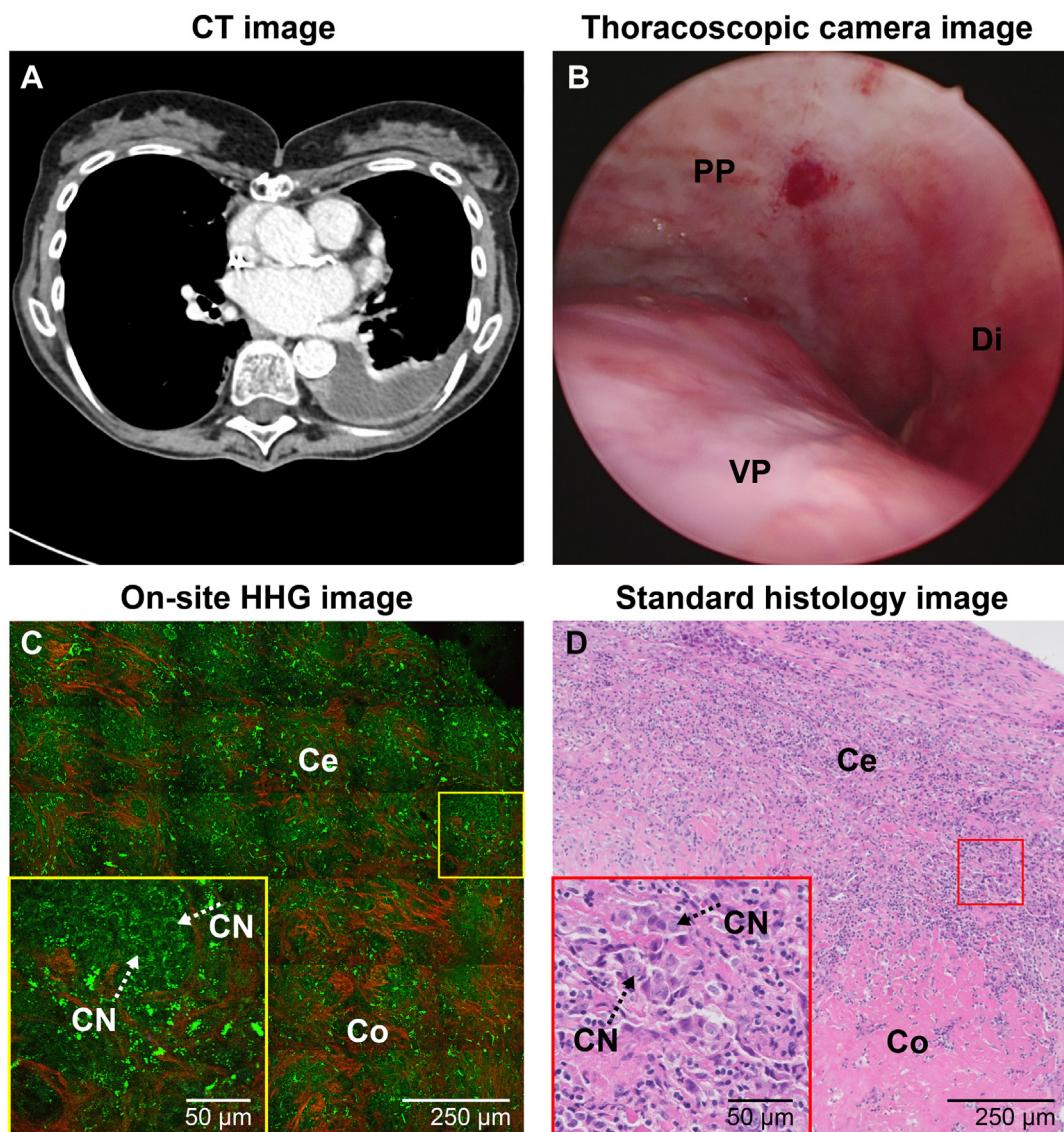
Unprocessed thoracoscopic biopsies were imaged with HHG microscopy, showing abnormal cells typical for pleural malignancy within 4 minutes (Fig. 6C). Histology, 10 days later, confirmed the diagnosis of epithelioid mesothelioma (Fig. 6D). Using the result of HHG microscopy for patient management could have been of benefit in this case, by providing immediate feedback on the presence or absence of malignancy leading to a decision in favor of immediate talc pleurodesis, thus preventing a repeat procedure for the patient.

Discussion

Lung cancer is a burden for the patient but also on the health care system, with rising costs and a growing shortage in staff. Innovations in diagnosis and treatment are needed to increase efficiency in the face of rising lung cancer cases. In this study, we demonstrated that with HHG microscopy, rapid on-site assessment of the presence of pulmonary or pleural malignancy in unprocessed bronchoscopic and thoracoscopic biopsies is feasible with a high diagnostic accuracy. The time between the receipt of the biopsy and feedback based on the first biopsy image was 6 minutes. This is significantly faster than the conventional histologic diagnostic work-up (days). The final average diagnostic agreement between HHG microscopy and histology of 87% was high, knowing that in histology, an acceptable concordance between individual pathologists is 80% to 90%. Compared with the current on-site feedback techniques, HHG microscopy

**Figure 5.**

A schematic overview of the deep learning classification pipeline and results. (Panel I) Higher harmonic generation microscopy data ($n = 69$) are preprocessed for model training ($n = 63$ for training and $n = 6$ for testing). Random training-validation splits ($n = 10$) are generated for Monte-Carlo cross-validation. Noninformative tiles are excluded based on the mean gray value. Minority oversampling is applied to obtain balanced training and validation sets for each Monte-Carlo fold. (Panel II) A fully convolutional classification network is trained per Monte-Carlo fold with identical hyperparameters. Multiinstance learning (Noisy-AND) is applied on the tile level. Sigmoid activation results in the binary classification output. (Panel III) A decision threshold per model is computed using Youden J statistic on the testing set receiver operating characteristic curve. (Panel IV) The decision thresholds translate average model predictions into model votes. Plurality voting is applied to produce an ensemble vote. (Panel IV) During inference, the ensemble vote is obtained for any testing set image, before being calculated on the image's individual tiles. This enables fast and accurate tumor detection within the overview. Red opaque tiles indicate tumor detections, which are counted toward the ensemble's vote for this test overview case. Overall evaluation set results are listed for the independent testing set. AUC, area under the curve; 2PEF, 2-photon excited autofluorescence; SHG, second harmonic generation; THG, third harmonic generation.

**Figure 6.**

A clinical case in which rapid on-site feedback using higher harmonic generation (HHG) microscopy could have been beneficial. (A) Chest CT of a patient showing left-sided unilateral pleural effusion with discrete pleural thickening and slight shrinkage of the left hemithorax. (B) White light inspection visualized the visceral pleura, diaphragm, and diffuse whitening of the parietal pleura without clear signs of tumor deposits. From this area, multiple biopsies were procured. (C) On-site HHG images of the first biopsy showed collagen-rich areas and cell-rich areas, with abnormal cells with increased cell nuclei within the connective tissue of the pleura, suggesting pleural malignancy. Acquisition time of this HHG image was 54 seconds. (D) Corresponding conventional histology images showed similar histopathologic features and confirmed the diagnosis of epithelioid mesothelioma, which was provided 10 days after the procedure. Ce, cell-rich areas; Co, collagen; CN, cell nuclei; CT, computed tomography; Di, diaphragm; PP, parietal pleura; VP, visceral pleura.

has a lower diagnostic accuracy than reported for frozen section analysis (95%–98%)^{2,3} but is significantly faster (6 minutes vs at least 20 minutes). HHG microscopy is equally fast as ROSE, but literature is contradictory on the ROSE diagnostic accuracy, ranging from 73% to 97%^{25–27} and does only provide cytologic information, in which HHG microscopy provides histologic information.

It is important to note that the number of preinvasive lesions in this cohort is high, as Amsterdam University Medical Center is a referral center for patients with such lesions, and it is known that pathologic assessment of these lesions is difficult.^{28,29} When excluding the biopsies with histology consensus diagnosis pre-invasive malignancy, the diagnostic accuracy performance

improves to 90%, and the positive predictive value increases from 74% to 86%.

The results from our AI classification algorithm indicate the further clinical feasibility of implementing computer-assisted diagnosis (CAD) to reduce clinical workload and improve patient outcomes. The portable microscope setup is perfectly suited for CAD on-site. We showed on a small data set how an ensemble of deep learning models is capable of detecting tumor features from our information-rich HHG imagery. This provides a foundation for further data acquisition and model engineering. The current results were obtained by incorporating image-level pathologist annotations, instead of tile-level annotations. A more detailed local comparison and extensive training

process on a larger data set will enable a higher accuracy and, we expect, also a higher level diagnosis, in scoring malignancy subtypes. Noninformative tiles were present in both classes, which introduced some noise in the data set but allowed for improved generalization during training. With tile-level annotations, the models would benefit from a better distinction between malignant and nonmalignant tiles (both informative and noninformative). By switching to 3-dimensional overview scanning in the future, our deep learning models can take the advantage of recognizing histologic features in an extra spatial dimension. Compared with classical histologic assessment, in addition to the speed, this could prove to be a game changer in on-site biopsy assessment.

As this is a first-of-its-kind study, better diagnostic performance could likely be achieved by improving tissue-handling processes and image processing, or by putting in the hands of pathology assistants, because HHG scores in this study were limited by biopsy orientation, 2-dimensional imaging, and fixed-image contrast. Further developments in the optical and electronic design of the HHG microscope are also envisioned, which could lead to a 4-fold reduction in imaging time. In follow-up studies, pathologists can be more systematically trained by the extensive image atlas provided by this study. This allows investigation of the dependency of the diagnostic performances on the training of operators and assessors. In addition, the diagnostic performances of assessors without thoracic expertise, such as general pathologists, can be included, which will show the ease of adaptation of HHG microscopy in the clinic.

The digital HHG images, even without CAD, provide the opportunity for pathologists to assess tissue images remotely from their desk while they are being recorded, and physical presence of pathologists is not required, which is relevant for regions with a shortage of pathologists. We envision that CAD and remote assessment by pathologists may be integrated by the “flagging” of suspect areas by the CAD model, which are then assessed by the pathologist. Alternatively, in most cases, a simple assessment on whether the biopsy is representative, that is, whether it contains tumor cells or not, can be done by the CAD algorithm or the operator. The HHG microscopy stand alone, or combined with CAD, for assessment of surgical and endoscopic biopsies could therefore reduce the number of biopsies and thereby lower the risk of complications and mortality. In addition, the fast feedback on biopsies provided by HHG microscopy enables a single hospital visit for diagnosis and (local) treatment, which would greatly reduce health care burden and improve patient outcomes. Furthermore, it could prevent repeat procedures as illustrated in the discussed case.

We conclude that HHG is an excellent candidate technique for the rapid assessment of biopsies during bronchoscopy and thoracoscopy procedures. Images are generated within minutes after extraction and can be assessed for malignancy with a high diagnostic accuracy. Quantification of the potential clinical impact of HHG microscopy, that is, reduction of procedure time, number of biopsies, the rate of nondiagnostic procedures as well as the need for repeat or additional diagnostic procedures, will be the focus of future studies.

Acknowledgments

The authors thank Frank van Mourik (Flash Pathology B.V.) for the technical support and the nursing team of the Behandelkamer at Amsterdam University Medical Center for their help with logistics and acquisition of biopsies. This publication is part of the

project InstantPathology (project number: 15825) of the research program Applied and Engineering Sciences, which is (partly) financed by the Dutch Research Council. Furthermore, this project has received funding from the European Union's Horizon 2020 research and innovation program under grants 654148 and 871124 (Laserlab-Europe). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

L.M.G.v.H., M.B., J.M.A.D., J.T.A., and M.L.G. conceived and designed the study. L.M.G.v.H. performed project administration. L.M.G.v.H., M.B., T.R., and J.H.v.d.T. performed data acquisition. L.M.G.v.H. and M.B. performed development of methodology, data analysis, interpretation of data, and writing of the original draft. J.M.A.D., T.R., J.H.v.d.T., M.V., J.T.A., and M.L.G. performed review and revision of the paper. J.M.A.D., J.T.A., and M.L.G. provided material support. J.T.A. and M.L.G. provided supervision. M.L.G. provided funding acquisition. All authors read and approved the final paper.

Data Availability

The data, model, and code files that support the findings of this study are available from the DataverseNL database (accession number: DOI:10.34894/ZZYU9M).

Funding

This publication is part of the project InstantPathology (project number: 15825) of the research program Open Technology by the domain which is (partly) financed by the Dutch Research Council (Applied and Engineering Sciences) and Horizon 2020 Framework Program grants 654148 and 871124.

Declaration of Competing Interest

M.L.G. declares to have indirect interest in Flash Pathology B.V. L.M.G.v.H. received technical support from Flash Pathology. Flash Pathology was not involved in the design of the study and data acquisition and analysis. Other authors have no conflict of interest.

Ethics Approval and Consent to Participate

The Medical Ethics Review Committee of the VU University Medical Centre confirmed that this study did not fall within the scope of the Medical Research Involving Human Subjects Act, as the diagnostic procedures were performed for clinical purposes (2020.223). The study is in line with Dutch license procedures and the declaration of Helsinki. Written consent was provided by all included patients.

Supplementary Material

The online version contains supplementary material available at <https://doi.org/10.1016/j.modpat.2024.100633>.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209–249.
2. Winther C, Graem N. Accuracy of frozen section diagnosis: a retrospective analysis of 4785 cases. *APMIS.* 2011;119(4-5):259–262.
3. Ferreiro JA, Myers JL, Bostwick DG. Accuracy of frozen section diagnosis in surgical pathology: review of a 1-year experience with 24,880 cases at Mayo Clinic Rochester. *Mayo Clin Proc.* 1995;70(12):1137–1141.
4. French JMR, Betney DT, Abah U, et al. Digital pathology is a practical alternative to on-site intraoperative frozen section diagnosis in thoracic surgery. *Histopathology.* 2019;74(6):902–907.
5. Laakman JM, Chen SJ, Lake KS, et al. Frozen section quality assurance: using separate frozen section slide preparation times and interpretative time measurements to improve process. *Am J Clin Pathol.* 2021;156(3): 461–470.
6. Spencer LT, Bancroft JD. Microtomy: paraffin and frozen. In: Bancroft JD, Gamble M, eds. *Theory and Practice of Histological Techniques.* 6th ed. Churchill Livingstone; 2008:93–104.
7. Cai G. Facility, equipment, specimen preparation, and stains. In: Cai G, Adeniran AJ, eds. *Rapid On-Site Evaluation (ROSE)-A Practical Guide.* Nature Switzerland AG: Springer; 2019.
8. Zhang Z, de Munck JC, Verburg N, et al. Quantitative third harmonic generation microscopy for assessment of glioma in human brain tissue. *Adv Sci (Weinh).* 2019;6(11):1900163.
9. Kuzmin NV, Wesseling P, Hamer PCW, et al. Third harmonic generation imaging for fast, label-free pathology of human brain tumors. *Biomed Opt Express.* 2016;7(5):1889–1904.
10. van Huizen LMG, Kuzmin NV, Barbé E, van der Velde S, Te Velde EA, Groot ML. Second and third harmonic generation microscopy visualizes key structural components in fresh unprocessed healthy human breast tissue. *J Biophotonics.* 2019;12(6):e201800297.
11. Yang L, Park J, Marjanovic M, et al. Intraoperative label-free multimodal nonlinear optical imaging for point-of-procedure cancer diagnostics. *IEEE J Sel Top Quantum Electron.* 2021;27(4): 6801412.
12. Gavgiotaki E, Filippidis G, Tsafas V, et al. Third harmonic generation microscopy distinguishes malignant cell grade in human breast tissue biopsies. *Sci Rep.* 2020;10(1):11055.
13. Wu P-C, Hsieh T-Y, Tsai Z-U, Liu T-M. In vivo quantification of the structural changes of collagens in a melanoma microenvironment with second and third harmonic generation microscopy. *Sci Rep.* 2015;5:8879.
14. Pham T, Banerjee B, Cromey B, et al. Feasibility of multimodal multiphoton microscopy to facilitate surgical margin assessment in pancreatic cancer. *Appl Opt.* 2020;59(22):G1–G7.
15. van Huizen LMG, Radonic T, van Mourik F, et al. Compact portable multi-photon microscopy reveals histopathological hallmarks of unprocessed lung tumor tissue in real time. *Transl Biophotonics.* 2020;2(4): e202000009.
16. van Huizen LMG, Kalverda KA, Bugiani M, et al. Rapid on-site pathology visualization of COVID-19 characteristics using higher harmonic generation microscopy. *Am J Respir Crit Care Med.* 2023;208(2):199–200.
17. Jain M, Narula N, Aggarwal A, et al. Multiphoton microscopy: a potential “optical biopsy” tool for real-time evaluation of lung tumors without the need for exogenous contrast agents. *Arch Pathol Lab Med.* 2014;138(8): 1037–1047.
18. Golaraei A, Mostaço-Guidolin LB, Raja V, et al. Polarimetric second-harmonic generation microscopy of the hierarchical structure of collagen in stage I-III non-small cell lung carcinoma. *Biomed Opt Express.* 2020;11(4):1851–1863.
19. Sun C-K, Yu C-H, Tai S-P, et al. In vivo and ex vivo imaging of intra-tissue elastic fibers using third-harmonic-generation microscopy. *Opt Express.* 2007;15(18):11167–11177.
20. Wang C-C, Li F-C, Wu R-J, et al. Differentiation of normal and cancerous lung tissues by multiphoton imaging. *J Biomed Opt.* 2009;14(4):044034.
21. Pavlova I, Hume KR, Yazinski SA, et al. Multiphoton microscopy and micro-spectroscopy for diagnostics of inflammatory and neoplastic lung. *J Biomed Opt.* 2012;17(3):036014.
22. Kraus OZ, Ba JL, Frey BJ. Classifying and segmenting microscopy images with deep multiple instance learning. *Bioinformatics.* 2016;32(12):i52–i59.
23. Abadi M, Agarwal A, Barham P, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. *arXiv.* Published online March 14, 2016. <https://doi.org/10.5281/zenodo.4724125>
24. van Huizen LMG, Blokker M, Rip Y, et al. Differentiation of leukocytes in bronchoalveolar lavage fluid samples using higher harmonic generation microscopy and deep learning. *PLoS One.* 2023;18(6). <https://doi.org/10.1371/journal.pone.0279525>
25. Chen X, Wan B, Xu Y, et al. Efficacy of rapid on-site evaluation for diagnosing pulmonary lesions and mediastinal lymph nodes: a systematic review and meta-analysis. *Transl Lung Cancer Res.* 2019;8(6):1029–1044.
26. Mondoni M, Carlucci P, Di Marco F, et al. Rapid on-site evaluation improves needle aspiration sensitivity in the diagnosis of central lung cancers: a randomized trial. *Respiration.* 2013;86(1):52–58.
27. Verhoeven RJL, Vos S, van der Heijden EHF. Multi-modal tissue sampling in cone beam CT guided navigation bronchoscopy: comparative accuracy of different sampling tools and rapid on-site evaluation of cytopathology. *J Thorac Dis.* 2021;13(7):4396–4406.
28. Keith RL, Miller YE, Gemmill RM, et al. Angiogenic squamous dysplasia in bronchi of individuals at high risk for lung cancer. *Clin Cancer Res.* 2000;6(5): 1616–1625.
29. Venmans B, Linden J, Elbers H, et al. Observer variability in histopathologic reporting of bronchial biopsy specimens. *J Bronchology Interv Pulmonol.* 2000;7(3):210–214.