Evaluation of an Integrated Spectroscopy and Classification Platform for Point-of-Care Core Needle Biopsy Assessment: Performance Characteristics from Ex Vivo Renal Mass Biopsies

Krishna Nand Keshavamurthy, PhD, Dmitry V. Dylov, PhD, Siavash Yazdanfar, PhD, Dharam Patel, BS, Tarik Silk, MD, Mikhail Silk, MD, Frederick Jacques, MD, Elena N. Petre, MD, Mithat Gonen, PhD, Natasha Rekhtman, MD, PhD, Victor Ostroverkhov, PhD, Howard I. Scher, MD, Stephen B. Solomon, MD, and Jeremy C. Durack, MD

ABSTRACT

Purpose: To evaluate a transmission optical spectroscopy instrument for rapid ex vivo assessment of core needle cancer biopsies (CNBs) at the point of care.

Materials and Methods: CNBs from surgically resected renal tumors and nontumor regions were scanned on their sampling trays with a custom spectroscopy instrument. After extracting principal spectral components, machine learning was used to train logistic regression, support vector machines, and random decision forest (RF) classifiers on 80% of randomized and stratified data. The algorithms were evaluated on the remaining 20% of the data set held out during training. Binary classification (tumor/nontumor) was performed based on a decision threshold. Multinomial classification was also performed to differentiate between the subtypes of renal cell carcinoma (RCC) and account for potential confounding effects from fat, blood, and necrotic tissue. Classifiers were compared based on sensitivity, specificity, and positive predictive value (PPV) relative to a histopathologic standard.

Results: A total of 545 CNBs from 102 patients were analyzed, yielding 5,583 spectra after outlier exclusion. At the individual spectra level, the best performing algorithm was RF with sensitivities of 96% and 92% and specificities of 90% and 89%, for the binary and multiclass analyses, respectively. At the full CNB level, RF algorithm also showed the highest sensitivity and specificity (93% and 91%, respectively). For RCC subtypes, the highest sensitivity and PPV were attained for clear cell (93.5%) and chromophobe (98.2%) subtypes, respectively.

Conclusions: Ex vivo spectroscopy imaging paired with machine learning can accurately characterize renal mass CNB at the time of tissue acquisition.

ABBREVIATIONS

AUC = area under the curve, CNB = core needle cancer biopsy, Hb = deoxyhemoglobin, HbO_2 = oxyhemoglobin, LR = logistic regression, MCC = Matthew correlation coefficient, PPV = positive predictive value, RCC = renal cell carcinoma, RF = random decision forest, SVM = support vector machine

The rate of molecular diagnostic discoveries has increased the number of solid tumor biopsies performed as well as magnified the importance of these specimens for classifying cancer and guiding treatment selection to optimize patient outcomes (1,2). Presently, the rate of failure to provide adequate material from needle biopsies for complete characterization by cell morphology, staining properties, and immunohistochemistry is substantial. In particular, genetic profiling is unsuccessful in over 50% of fine needle aspiration biopsies and over 30% of core needle biopsies (CNBs) (3). Rapid evaluation of specimen quality and tumor yield at the time of a biopsy could provide critical

Figures E1 and **E2** and **Appendices A** and **B** can be found by accessing the online version of this article on *www.jvir.org* and selecting the Supplemental Material tab.

J Vasc Interv Radiol 2022; 33:1408–1415
https://doi.org/10.1016/j.jvir.2022.07.027

RESEARCH HIGHLIGHTS

- Spectroscopy paired with machine learning had a high sensitivity and specificity, up to 96% and 92%, respectively, for renal mass core needle biopsy characterization.
- The algorithm differentiated between renal cell carcinoma subtypes, with the highest sensitivity and positive predictive value for clear cell (93.5%) and chromophobe (98.2%) subtypes, respectively.
- With direct feedback, biopsy procedures may be modified at the point of care to improve sampling, efficacy, and safety.
- A new performance standard for cancer biopsies may be possible through a combined hardware and software platform for rapid nondestructive tissue analysis.

feedback to the operator, leading to more effective and efficient tumor sampling.

Deployment of biopsy assessment technologies that facilitate rapid on-site evaluation in clinical settings has been hampered by factors such as lengthy analytic times, tissue degradation, expense, on-site tissue staining, requirement for interpretive expertise, and challenges to workflow integration. Additionally, many of these technologies analyze tissues in vivo, which may be helpful for determining intraoperative surgical margins but are less informative for ex vivo assessment of tissue acquired from a needle biopsy. To facilitate ex vivo CNB quality assessment, a transmission optical spectroscopy instrument was combined with machine learning methods to rapidly characterize CNB samples. This study aimed to evaluate the performance characteristics of the spectroscopy instrument paired with machine learning utilizing a renal mass biopsy training set.

MATERIALS AND METHODS

This study was approved by the institutional review board and human biospecimen utilization committee as a retrospective analysis.

Optical Spectroscopy and Machine Learning Platform Overview

A transmission optical spectroscopy instrument was designed to provide direct feedback to the operator including the analyzable specimen length, geometric proportion of the sample that contains malignant cells, and overall sample classification as predominantly malignant or benign. Based on this information, in clinical practice, an operator may choose to obtain additional biopsy samples, with an associated risk from additional needle passes, or conclude the procedure. After spectroscopic analysis and point-of-care feedback, unadulterated label or dye-free

STUDY DETAILS

Study type: Laboratory investigation

samples can then undergo standard of care diagnostic and molecular analysis. This workflow offers the opportunity to limit sources of preanalytical variation, such as the percentage of samples with little or no cancerous tissue, to review in the pathology laboratory. A comparison of the traditional on-site imprint cytology- and spectroscopy-based CNB assessment is depicted in **Figure E1** (available online on the article's Supplemental Material page at *www.jvir.org*).

Designed for use in procedure rooms or surgical suites or at the patient's bedside, the hardware and software platform acquires spectra along CNB samples that remain intact on the biopsy needle. Machine learning algorithms trained using ground truth histopathologic analysis of tumor samples classify individual transmission spectra along a biopsy sample. Output to the user includes the likelihood of the sample containing tumor tissue, geometric proportion of the malignant cells in the sample, and classification of the sample by histopathologic subtype. In this study, spectra obtained from every 0.75 mm of CNB tissue were used to train 3 different discriminative machine learning algorithms: (*a*) logistic regression (LR), (*b*) support vector machine (SVM), and (*c*) random decision forests (RFs). The CNB samples were obtained from human kidneys and kidney tumors. Classification sensitivity, specificity, and accuracy metrics were reported for individual spectra as well as for full CNBs to differentiate tumor biopsies from predominantly benign tissue biopsies.

Biopsy Protocol

Surgically excised human kidney specimens were biopsied ex vivo in a tissue procurement facility immediately after partial or complete nephrectomy for a renal tumor. The distinction between normal and tumor regions was made by gross examination of pathology specimens. Biopsies were obtained using 18-gauge side-notch, spring-loaded core needle devices with 20-mm-long sampling trays exposed (Temno Evolution or Adjustable Coaxial Temno; CareFusion Corporation, San Diego, California). Kidney tumor and normal renal parenchymal tissues distant from the tumor were sampled. In addition, renal sinus fat, visibly necrotic tissue, and blood-only samples were collected to evaluate classification in the presence of potential confounding tissue types. The final histopathologic tumor diagnosis was recorded for association with spectroscopy data. The optical spectroscopy, data acquisition, outlier rejection, principal component analysis, machine learning, and classification steps are succinctly described in the following and in more detail in **Appendices A** and **B** (available online at *www.jvir.org*).

Optical Spectroscopy

Biopsy samples were analyzed using a custom-built transmission optical spectrometer (**Appendix A** and **Fig E2**, available online at *www.jvir.org*). Up to 20 transmission spectra were collected at approximately 0.75-mm sampling increments. Data acquisition time was typically 45 seconds per biopsy, which includes multiexposure (high dynamic range) stitching to improve the signal-to-noise ratio. Spectra with a low signal-to-noise ratio due to light obstruction by the needle or 100% transmittance due to empty biopsy tray regions fell outside the interquartile range of the spectral distribution and were rejected as outliers.

Principal Component Analysis, Machine Learning, and Classification

The majority of spectral data were represented by 10 principal components. Machine learning classifiers, including LR, SVMs, and RFs (4,5), were trained on 80% of the data and tested on the remaining 20%. The training-test set split was randomized and stratified, and classifier hyperparameters were tuned using tenfold stratified cross validation on the training set.

Tissue classification was performed both at the individual spectrum and the full CNB sample levels. The ground truth pathology labels were assigned into renal cell carcinoma (RCC), normal renal parenchyma, renal sinus fat, blood, or necrotic tumor categories at the whole CNB level.

Performance Evaluation

Individual spectral and whole biopsy classifiers were quantitatively evaluated using ROC analysis and scalar performance metrics at their respective optimal decision thresholds (**Appendix B**, available online at *www.jvir.org*). The McNemar statistical test, Matthew correlation coefficient (MCC) and kappa statistics were used to compare the performance of RF, LR, and SVM algorithms for individual spectra and whole CNB classification. To facilitate data visualization, false color "heat maps" for biopsy samples were generated by color coding the likelihood of constituent spectra belonging to the tumor class.

RESULTS

Sample and Spectroscopy Characteristics

A total of 545 CNBs obtained from surgically resected kidneys from 102 patients undergoing partial or complete nephrectomy for renal tumors were analyzed with the transmission optical spectroscopy device (**Fig E2**, available online at *www.jvir.org*). **Table 1** summarizes the distribution of tissues sampled. The maximum scan duration using the spectroscopy platform was 60 seconds for a full-trough biopsy sample. Each CNB sample had up to 21 spectra (1 spectrum corresponds to 0.75 mm of tissue) depending upon the length of tissue obtained from a needle pass, and the number of spectra rejected from areas where no tissue was present on the needle. Analysis of the 545 CNB samples produced a library of 5,583 spectra. There was at least 1 spectrum obtained from each of the 545 biopsies (range, 1–20; mean, 10.24; median, 10 spectra per sample).

Table 1. Distribution of Biopsied Tissue and Total Number of Spectra Obtained from Each Sample Type

Tissue	Specimens	Spectra
Renal masses	322	3,176
Clear cell	194	1,847
Papillary	50	348
Chromophobe	37	440
Oncocytoma	15	220
TSC-associated RCC	3	51
Undifferentiated RCC	15	190
Adrenal cortical	4	52
Dedifferentiated sarcoma	4	48
Normal renal parenchyma	123	1,618
Renal sinus fat	44	370
Blood	37	242
Necrotic tumor	19	177
Total	545	5,583

RCC = renal cell carcinoma; TSC = tuberous sclerosis complex.

Classifier Performance for Individual Spectra

Classifier performance results for binary (tumor vs normal) and multiclass (confounders included) classifications based on individual spectra are presented in **Figure 1**. Of the 3 classification algorithms evaluated, the RF classifier achieved the highest area under the curve (AUC) of 0.98. The multiclass classification results were collapsed to produce a binary classification summary, with tumor and necrotic tissues considered positive and normal tissues (normal kidney, blood, and fat) considered negative classes. The RF again achieved the best performance with an AUC of 0.96. The comparative scalar performance metrics, including confidence intervals, at optimal receiver operating curve thresholds (P^{s}_{th} of 0.59 and 0.57) for binary and multiclass classification are shown in **Table 2**. Classifier discrimination was highest using the RFs algorithm, with sensitivities of 96% and 92% and specificities of 90% and 89% for the binary and multiclass analyses, respectively. The MCCs were 0.871, 0.689, and 0.718 for the RF, LR, and SVM algorithms, respectively. The kappa statistic was highly correlated with MCC with almost the same values (0.871, 0.688, and 0.717, respectively). **Figure 2** reveals the relative performance of the RF, LR, and SVM algorithms using the McNemar test. The difference between the performance of RF and the other algorithms was statistically significant across all compared metrics (error rate, sensitivity, and specificity). Taken together, the results of this study demonstrated statistically significant

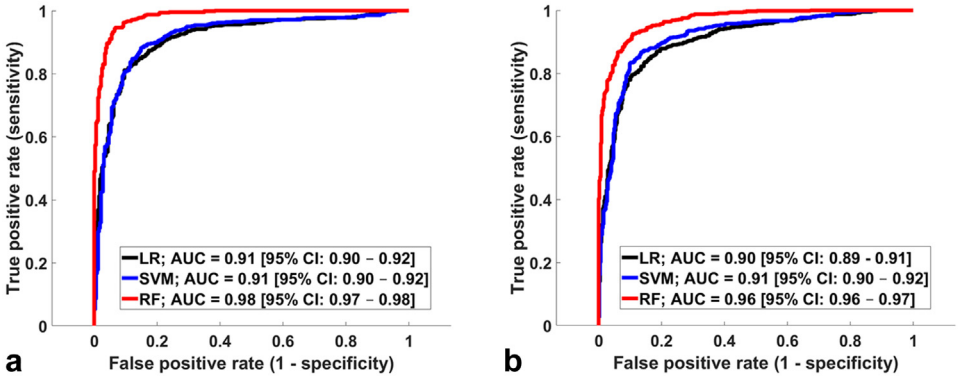

Figure 1. Receiver operating characteristic analysis for spectral classification: **(a)** tumor versus normal tissue. **(b)** Multiclass classification (collapsed to binary output) of tumor and normal tissues with 3 additional confounding tissue types: (*a*) blood, (*b*) fat, and (*c*) necrotic tissue (right). The machine learning algorithms included logistic regression (LR), support vector machine (SVM), and random decision forests (RFs). AUC = area under the curve; CI = confidence interval.

Table 2. Classifier Performance Metrics Comparing the Spectral Classifiers, Including LR, SVM, and RFs

Algorithm	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)
Tumor vs normal tissue spectrum classifier					
LR	90 (576/635) [95% CI, 87–92]	77 (251/323) [95% CI, 72–81]	86 (827/958) [95% CI, 84–88]	88 (576/648) [95% CI, 86–91]	80 (251/310) [95% CI, 76–85]
SVM	88 (564/635) [95% CI, 85–90]	84 (272/323) [95% CI, 80–87]	87 (836/958) [95% CI, 85, 89]	91 (564/615) [95% CI, 88–93]	79 (272/343) [95% CI, 74–83]
RF	96 (611/635) [95% CI, 95–97]	90 (293/323) [95% CI, 86–93]	94 (904/958) [95% CI, 92–95]	95 (611/641) [95% CI, 93–96]	92 (293/317) [95% CI, 89–94]
Tumor and necrotic tissues vs nontumor (fat, blood, and normal tissues) spectrum classifier					
LR	87 (589/670) [95% CI, 84–89]	80 (356/445) [95% CI, 76–83]	84 (945/1115) [95% CI, 82–86]	86 (589/678) [95% CI, 84–89]	81 (356/437) [95% CI, 77–85]
SVM	86 (581/670) [95% CI, 84–89]	86 (385/445) [95% CI, 83–89]	86 (966/1115) [95% CI, 84–88]	90 (581/641) [95% CI, 88–92]	81 (385/474) [95% CI, 77–84]
RF	92 (620/670) [95% CI, 90–94]	89 (397/445) [95% CI, 86–92]	91 (1017/1115) [95% CI, 89–92]	92 (620/668) [95% CI, 90–94]	88 (397/447) [95% CI, 85–91]

Note–The top half of the table reflects binary tumor versus normal tissue spectral classification; the bottom half shows multiclass classification with confounders collapsed to a binary output (positive class, tumor and necrotic tissue spectra; negative class, fat, blood, and normal tissue spectra). All metrics were computed at their respective optimal receiver operating characteristic thresholds.
CI = confidence interval; LR = logistic regression; NPV = negative predictive value; PPV = positive predictive value; RF = random decision forest; SVM = support vector machine.

superiority in terms of AUC, sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value for the RF algorithm over LR and SVM.

Table 3 shows the results of the multinomial classification for the predominant subtypes of RCC as a confusion matrix. The sensitivity and PPV values for each subtype were computed by considering each tumor subtype in the positive class and the remainder in the negative class. The highest sensitivity and PPV were attained for clear cell (93.5%) and chromophobe RCC (98.2%) subtypes, respectively. The majority of misclassification errors were associated with the chromophobe and papillary subtypes relative to the clear cell subtype.

Full CNB Classifier Performance

RFs, the best performing individual spectral classifier, were used to build the full CNB classifier. Figure 3 shows the results of classifying the full CNB specimen based on an optimal decision threshold ($p > P^{b}_{th}$). At the optimal ROC threshold of $P^{b}_{th} = 0.02$, the classification sensitivity and specificity were 93% and 91%, respectively. The MCC and kappa statistic were 0.8 and 0.807, respectively. All performance measures indicate highly accurate classification of full CNBs using the combined tissue spectroscopy and classification approach.

Heat Maps for Visualization of Classifier Results

Figure 4 shows representative false color heat maps that depict the probability of tumor at each site of spectral acquisition along CNB samples. True-positive samples are largely yellow to red because their constituent spectra were predicted to belong to the tumor category with high probability. Similarly, the true-negative biopsy samples are largely green to yellow because their constituent spectra had a low probability of belonging to the tumor category.

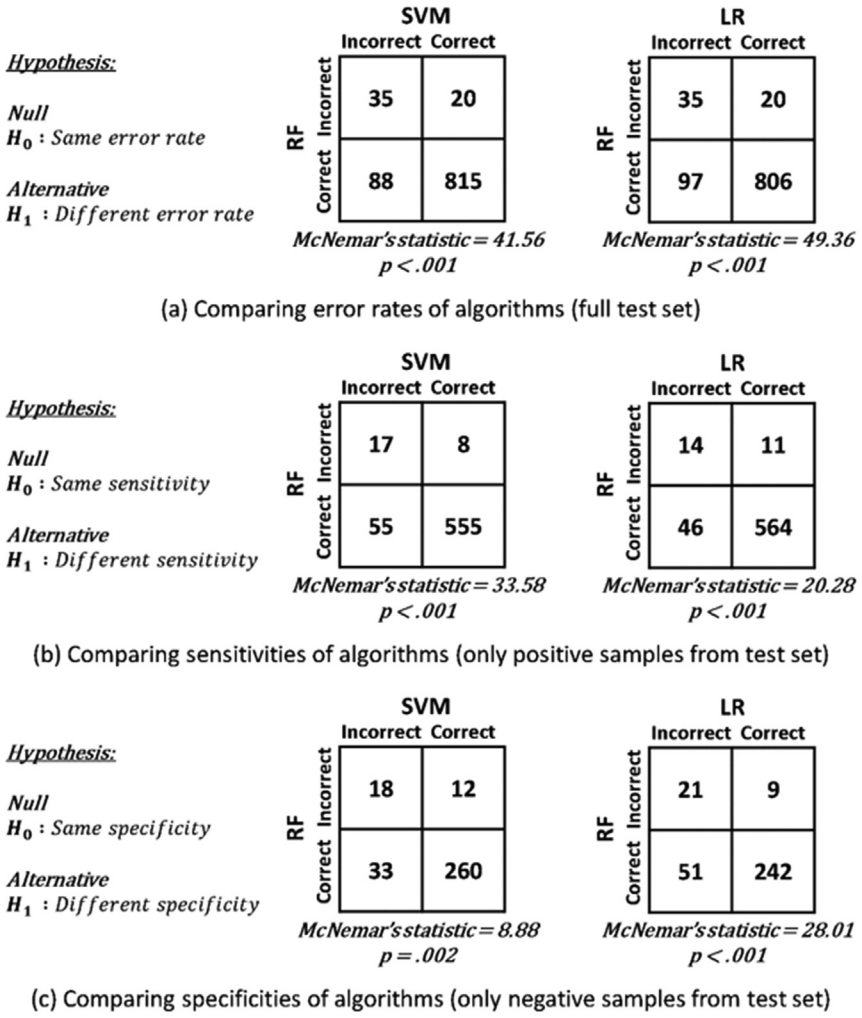

Figure 2. McNemar test for comparing random decision forests (RFs) with support vector machine (SVM) and logistic regression (LR) spectral classifiers for each of the compared metrics: error rate (top), sensitivity (middle), and specificity (bottom), all computed at their respective optimal receiver operating characteristic thresholds on the held-out test set.

Table 3. Spectroscopy-Based Tumor Subtype Classification Using the Random Forest Algorithm

	Test (spectroscopy)					Sensitivity
	Clear cell	Chromophobe	Oncocytoma	Papillary	Normal	
Gold Standard (pathology)						
Clear cell	345	0	0	9	15	93.5%
Chromophobe	22	55	0	1	10	62.5%
Oncocytoma	9	0	30	0	1	75.0%
Papillary	42	0	1	22	4	31.9%
Normal	22	1	0	1	299	92.6%
PPV	78.4%	98.2%	96.8%	66.7%	90.9%	

Note–The different classes include clear cell, chromophobe, oncocytoma, papillary, and normal spectra. The sensitivity and PPV values were computed by including each of the renal cell carcinoma subtypes in the positive class and the remainder in the negative class.
PPV = positive predictive value.

DISCUSSION

Automated biospecimen analytic platforms have transformed several aspects of modern clinical medicine. Pursuit of cancer biopsy quality improvement has led to the development of a variety of automated or semiautomated technologies and techniques for in vivo or ex vivo tissue characterization (6). These include optical spectroscopy (7), x-ray imaging (8), confocal microscopy (9), structured illumination microscopy (10), Fourier transform infrared

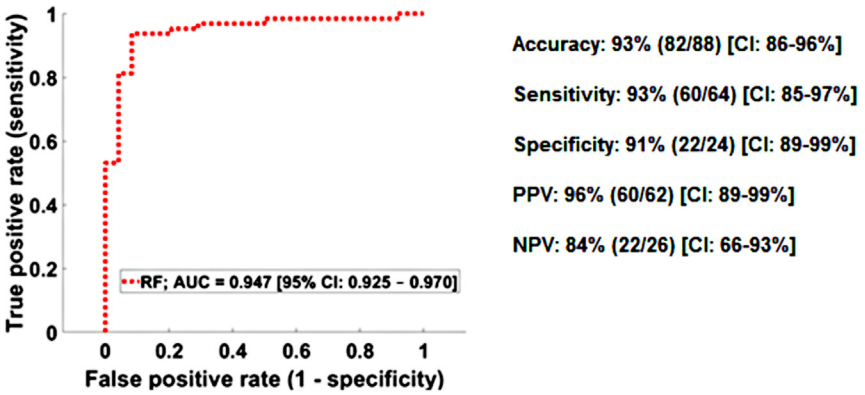

Figure 3. Whole biopsy classification using the random decision forests (RFs) spectral classifier on the test set. The receiver operating characteristic curve and the performance metrics at the optimal receiver operating characteristic threshold. AUC = area under the curve; CI = confidence interval; NPV = negative predictice value; PPV = positive predictive value.

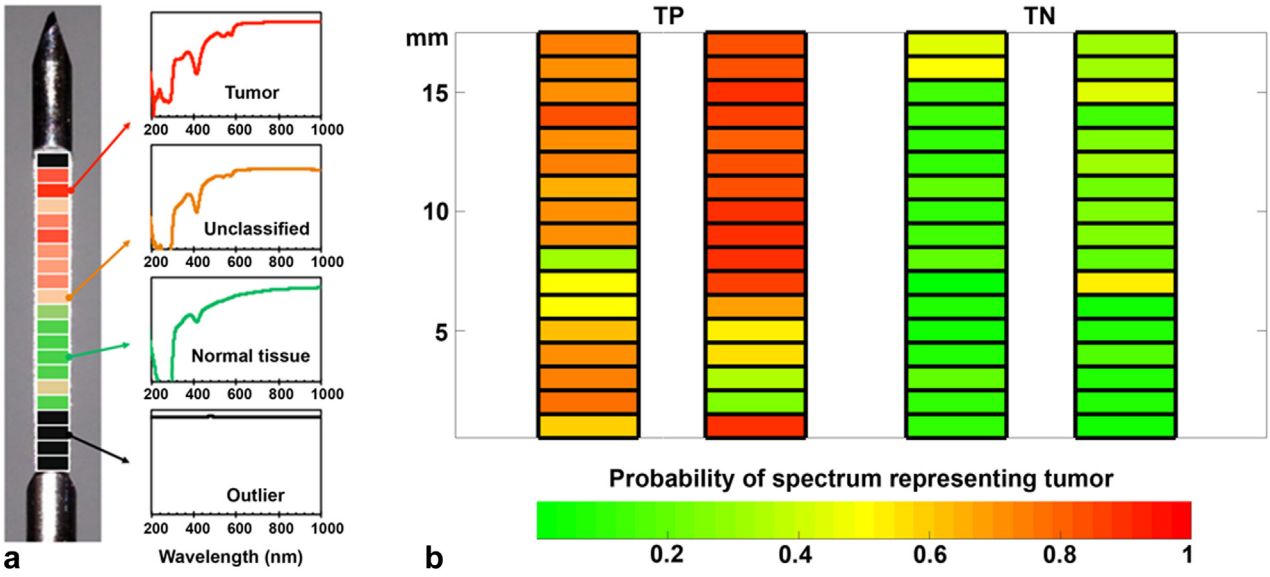

Figure 4. (a) False color heat map classification results displayed along the length of a biopsy sample obtained from the margin of a tumor and normal tissue. Normalized representative transmittance spectra are shown for tumor (red), unclassifiable (orange), normal tissue (green), and outlier (black) regions. In this example, the proximal region of biopsy trough did not contain any tissue, resulting in outlier spectra (black) that were not classified. **(b)** False color heat maps of biopsy samples selected to illustrate the variability of constituent spectra within standard core needle biopsy samples. Each horizontal bar represents approximately 0.75 mm of imaged tissue along a standard needle cancer biopsy sampling tray. The color of each bar reflects the probability of an individual spectrum representing tumor tissue. Automated classification of component spectra across full biopsy specimens enables grouping into true-positive, false-negative, true-negative, and false-positive categories. Two representative samples of true-positive (TP) and true-negative (TN) biopsies are displayed.

imaging (11), fluorescence microscopy (12–15), contrast-enhanced micrography (16), and diffuse reflectance, electrical impedance, and Raman spectroscopy (17–21). Differences between normal and cancerous tissue have been explored using these techniques and correlated with tissue morphology and histology (22–24). These technologies for ex vivo cancer specimen quality assessment possess advantages and disadvantages with respect to portability, infrastructure requirements, ease of use, and costs to develop or implement (5,6,25,26).

This study suggests that optical spectroscopy combined with machine learning–based tissue classification can quickly, accurately, and nondestructively characterize samples obtained for cancer diagnostics. The technology described here evaluates unstained tissues on the core biopsy needle itself, with no additional physical manipulation or risk of structural, cellular, or genetic damage prior to submission to the pathology laboratory.

For some time, immediate biopsy specimen assessment using traditional cytologic imprints (glass slide touch

preparations) under light microscopy has been shown to improve sampling accuracy (27,28). However, the resources and expertise necessary to provide rapid on-site sample assessment are often unavailable (29,30). Glass slide touch-preparation techniques can also be deleterious to downstream processing by substantially depleting the CNB of neoplastic cells or disrupting tissue architecture (31).

Several studies have examined the biological underpinnings of distinct optical imaging signatures from normal versus malignant tissues. Differences in tissue architecture, blood supply and oxygenation are commonly cited (32). For example, deoxyhemoglobin/oxyhemoglobin (Hb/HbO_2) absorption (Hb peak, approximately 350–400 nm; HbO_2 peak, approximately 550–600 nm) can impact the spectral slope within the near-infrared region, and the relationship between scattering and tissue composition has been outlined in other malignancies, such as breast cancer (22).

The combination of an imaging modality, a reference library of tissue characteristics, a computational platform, and data networking can effectively close the physical and temporal gap between operating rooms or biopsy suites (tissue acquisition) and pathology departments (tissue analysis), providing actionable results at the point of care. Relative to technologies that characterize tissues in vivo (prior to biopsy), the instrument studied here analyzes ex vivo material obtained from the targeted lesion. This offers the advantage of providing information regarding the actual sample destined for downstream processing, rather than a possible target within the vicinity of the device. For example, an operator can discontinue needle biopsies when readouts indicate high tumor content within acquired samples. Alternatively, an operator with a high degree of suspicion for disease may reposition the biopsy needle and obtain additional samples when the readout suggests only noncancerous tissues. Reducing undersampling could improve biopsy efficacy, whereas limiting oversampling may improve biopsy procedure safety.

Short scan durations and fast response times are important to minimize procedural delays, particularly when multiple biopsy samples are required for patients receiving sedation, or anesthesia. The present study data were acquired with scan times of <1 minute per sample based on focal spots every 0.75 mm. Combined with the computational time for spectral and whole biopsy classification of <1 second, this system is highly efficient relative to traditional imprint cytology slide preparation and review.

This study has a number of limitations. First, a common criticism of machine learning classification algorithms to automate diagnoses is that there may be limited insight into the specific determinants used for classification. Further efforts to improve classifier accountability may benefit from advances in the field of machine learning. At present, the precise etiology of false-positive and false-negative sample classification is uncertain. This may be due to the relatively small size of the spectroscopic data and lack of availability of truth at the level of individual spectra, requiring additional spectral data and further investigation of biological underpinnings to assess. Several studies have attempted to extract the basis of distinct optical imaging signatures from normal versus malignant tissues. Differences in tissue architecture, blood supply, and oxygenation are commonly cited (32). For example, Hb/HbO_2 absorption (Hb peak, approximately 350–400 nm; HbO_2 peak, approximately 550–600 nm) can impact the spectral slope within the near-infrared region, and the relationship between scattering and tissue composition has been outlined in other malignancies, such as breast cancer (28).

Necrotic tumor tissues, inflammation, fibrosis, and residual blood on tissue samples also present a challenge for classification and are potential confounders. In particular, necrotic tumors may be considered either tumor positive or tumor negative, even by highly trained histopathologists depending upon whether small clusters of viable tumor cells are detected. Further studies in real-world biopsy settings will also be necessary to determine whether ex vivo tissue sample spectroscopic profiles may have been impacted by the time delay between surgical removal of the tumors and imaging. The high sensitivity and specificity achieved in this study suggests that tissue desiccation or ischemia does not significantly alter the spectroscopic signature of tumors. Although studies to date demonstrate high instrument sensitivity and specificity in kidney, the testing domain has been limited and requires expansion to a range of tumor types to generalize performance claims. To this end, a clinical trial utilizing this technology in a wide array of tissue and tumor types would be beneficial. Clinical translation would also require attention to the "ASSURED" criteria (Affordable, Sensitive, Specific, User-friendly, Rapid and robust, Equipment-free, and Deliverable to end users), often cited to evaluate the potential of rapid point-of-care tests in resource limited environments (6,33). Following the ASSURED criteria, these technologies are promising based on performance characteristics, speed, and potential for integration into point-of-care environments. However, robust clinical validation, true cost-effectiveness assessment, and data reproducibility studies with histopathologic correlation at the individual spectral level will need to be performed.

Point-of-acquisition tissue assessment using instruments such as the optical spectroscopy and machine learning platform described in this study could help to establish a new performance standard for rapid cancer biopsy assessment and automated needle biopsy quality control. The potential of this technology is increasingly evident in the era of molecular oncology in which personalized cancer treatments require accurate and often repeated profiling of continually mutating cancer cell populations.

ACKNOWLEDGMENTS

This work was supported in part by the National Institutes of Health/National Cancer Institute P30 Cancer Center Support Grant (P30 CA008748) and a Prostate Cancer Foundation Young Investigator research grant to Jeremy C. Durack, MD.

AUTHOR INFORMATION

From the Department of Radiology (K.N.K., M.S., E.N.P, S.B.S., J.C.D.), Memorial Sloan Kettering Cancer Center, New York, New York; Skolkovo Institute of Science and Technology (D.V.D.), Moscow, Russia; Corning Incorporated (S.Y.), Corning, New York; Novartis Pharmaceutical Corporation (D.P.), East Hanover, New Jersey; New York University Langone Medical Center (T.S.), New York, New York; Montefiore Medical Center (F.J.), New York, New York; Department of Epidemiology and Biostatistics (M.G.), Memorial Sloan Kettering Cancer Center, New York, New York; Department of Pathology (N.R.), Memorial Sloan Kettering Cancer Center, New York, New York; GE Global Research (V.O.), Niskayuna, New York; and Department of Medicine (H.I.S.), Memorial Sloan Kettering Cancer Center, New York, New York. Received March 30, 2022; final revision received July 21, 2022; accepted July 29, 2022. Address correspondence to J.C.D., Department of Radiology, Memorial Sloan Kettering Cancer Center, 330 Madison Ave., Suite 200, New York, NY 10017; E-mail: jcdurack@gmail.com

K.N.K. and D.V.D. contributed equally to this work and are co-first authors.

J.C.D. is Scientific Advisor Board and Investor of Serpex Medical and Adient Medical and Chief Clinical Officer of Ajax Health and Cordis and reports equity interests in Cordis. S.B.S. reports consultancy of Varian, Johnson & Johnson, Adgero, XACT Robotics, Endoways, Aperture Medical, Microbot, and Olympus; grants from GE Healthcare, AngioDynamics, Elesta, and Johnson & Johnson; and Shareholder of Johnson & Johnson, Aperture Medical, Innoblative, Immunomedics, and Progenics. H.I.S. is Member of Board of Directors and reports stock options of Asterias Biotherapeutics, Consultancy/Advisory Board of Ambry Genetics Corporation, Konica Minolta, Inc., OncLive Insights, Physicians Education Resource, WCG Oncology (compensated); Amgen, ESSA Pharma Inc., Janssen Research & Development, LLC, Janssen Biotech, Inc. Menarini Silicon, and Sanofi Aventis (uncompensated); research funding to the institution from Epic Sciences, Illumina, Inc., Innocrin Pharma, Janssen, Menarini Silicon, and Thermo Fisher Scientific; and travel, accommodations, and expenses from Amgen, Asterias Biotherapeutics, Clovis Oncology, ESSA Pharma Inc., Genome Profiling, LLC, Menarini Silicon, OncLive Insights, Physicians Education Resource, Prostate Cancer Foundation, Sanofi Aventis, and WCG Oncology. V.O. is employed by GE Global Research. D.P. is currently employed by Novartis Pharmaceutical Corporation and was a Memorial Sloan Kettering Cancer Center employee at the time the study was conducted and did not have access to the data after leaving the institution. The other authors have not identified a conflict of interest.

From the 2014 SIR Annual Scientific Meeting, Abstract No. 250, “Rapid determination of core needle biopsy adequacy using optical spectroscopy.”

REFERENCES

1. Marshall D, Laberge JM, Firetag B, Miller T, Kerlan RK. The changing face of percutaneous image-guided biopsy: molecular profiling and genomic analysis in current practice. J Vasc Interv Radiol 2013; 24:1094–1103.
2. Ziv E, Durack JC, Solomon SB. The importance of biopsy in the era of molecular medicine. Cancer J 2016; 22:418–422.
3. Schneider F, Smith MA, Lane MC, Pantanowitz L, Dacic S, Ohori NP. Adequacy of core needle biopsy specimens and fine-needle aspirates for molecular testing of lung adenocarcinomas. Am J Clin Pathol 2015; 143: 193–200; quiz 306.
4. Huerta-Nuñez LFE, Gutierrez-Iglesias G, Martinez-Cuazitl A, et al. A biosensor capable of identifying low quantities of breast cancer cells by electrical impedance spectroscopy. Sci Rep 2019; 9:6419.
5. Patil PD, Hobbs B, Pennell NA. The promise and challenges of deep learning models for automated histopathologic classification and mutation prediction in lung cancer. J Thorac Dis 2019; 11:369–372.
6. Boppart SA, Richards-Kortum R. Point-of-care and point-of-procedure optical imaging technologies for primary care and global health. Sci Transl Med 2014; 6:253rv2.
7. Kennedy KM, McLaughlin RA, Kennedy BF, et al. Needle optical coherence elastography for the measurement of microscale mechanical contrast deep within human breast tissues. J Biomed Opt 2013; 18:121510.
8. Krupinski EA, Borders M, Fitzpatrick K. Processing stereotactic breast biopsy specimens: impact of specimen radiography system on workflow. Breast J 2013; 19:455–456.
9. Dobbs JL, Mueller JL, Krishnamurthy S, et al. Micro-anatomical quantitative optical imaging: toward automated assessment of breast tissues. Breast Cancer Res 2015; 17:105.
10. Wang M, Kimbrell HZ, Sholl AB, et al. High-resolution rapid diagnostic imaging of whole prostate biopsies using video-rate fluorescence structured illumination microscopy. Cancer Res 2015; 75: 4032–4041.
11. Tiwari S, Raman J, Reddy V, et al. Towards translation of discrete frequency infrared spectroscopic imaging for digital histopathology of clinical biopsy samples. Anal Chem 2016; 88:10183–10190.
12. Mueller JL, Harmany ZT, Mito JK, et al. Quantitative segmentation of fluorescence microscopy images of heterogeneous tissue: application to the detection of residual disease in tumor margins. PLoS One 2013; 8: e66198.
13. Rivenson Y, Wang H, Wei Z, et al. Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. Nat Biomed Eng 2019; 3:466–477.
14. Glaser AK, Reder NP, Chen Y, et al. Light-sheet microscopy for slide-free non-destructive pathology of large clinical specimens. Nat Biomed Eng 2017; 1:0084.
15. Fereidouni F, Harmany ZT, Tian M, et al. Microscopy with ultraviolet surface excitation for rapid slide-free histology. Nat Biomed Eng 2017; 1: 957–966.
16. Im H, Pathania D, McFarland PJ, et al. Design and clinical validation of a point-of-care device for the diagnosis of lymphoma via contrast-enhanced microholography and machine learning. Nat Biomed Eng 2018; 2:666–674.
17. Spliethoff JW, Prevoo W, Meier MA, et al. Real-time in vivo tissue characterization with diffuse reflectance spectroscopy during transthoracic lung biopsy: a clinical feasibility study. Clin Cancer Res 2016; 22: 357–365.
18. Shaikh R, Prabitha VG, Dora TK, et al. A comparative evaluation of diffuse reflectance and Raman spectroscopy in the detection of cervical cancer. J Biophotonics 2017; 10:242–252.
19. Orringer DA, Pandian B, Niknafs YS, et al. Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated Raman scattering microscopy. Nat Biomed Eng 2017; 1: 0027.
20. Hollon TC, Pandian B, Adapa AR, et al. Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. Nat Med 2020; 26:52–58.
21. van Manen L, Schmidt I, Inderson A, et al. Single fiber reflectance spectroscopy for pancreatic cancer detection during endoscopic ultrasound guided fine needle biopsy: a prospective cohort study. Int J Med Sci 2022; 19:205–212.
22. Kennedy S, Caldwell M, Bydlon T, et al. Correlation of breast tissue histology and optical signatures to improve margin assessment techniques. J Biomed Opt 2016; 21:66014.
23. Brown JQ, Wilke LG, Geradts J, Kennedy SA, Palmer GM, Ramanujam N. Quantitative optical spectroscopy: a robust tool for direct measurement of breast cancer vascular oxygenation and total hemoglobin content in vivo. Cancer Res 2009; 69:2919–2926.
24. Chang VT, Cartwright PS, Bean SM, Palmer GM, Bentley RC, Ramanujam N. Quantitative physiology of the precancerous cervix in vivo through optical spectroscopy. Neoplasia 2009; 11:325–332.
25. Taruttis A, Ntziachristos V. Translational optical imaging. AJR Am J Roentgenol 2012; 199:263–271.
26. Moriyama EH, Zheng G, Wilson BC. Optical molecular imaging: from single cell to patient. Clin Pharmacol Ther 2008; 84:267–271.
27. Tsou MH, Tsai SF, Chan KY, et al. CT-guided needle biopsy: value of on-site cytopathologic evaluation of core specimen touch preparations. J Vasc Interv Radiol 2009; 20:71–76.
28. Kubik MJ, Bovbel A, Goli H, Saremian J, Siddiqi A, Masood S. Diagnostic value and accuracy of imprint cytology evaluation during image-guided core needle biopsies: review of our experience at a large academic center. Diagn Cytopathol 2015; 43:773–779.
29. Balachandran I, Friedlander M. Cytology workforce study: a report of current practices and trends in New York State. Am J Clin Pathol 2011; 136:108–118.
30. Khurana KK. Telecytology and its evolving role in cytopathology. Diagn Cytopathol 2012; 40:498–502.
31. Rekhtman N, Kazi S, Yao J, et al. Depletion of core needle biopsy cellularity and DNA content as a result of vigorous touch preparations. Arch Pathol Lab Med 2015; 139:907–912.
32. Tromberg BJ, Shah N, Lanning R, et al. Non-invasive in vivo characterization of breast tumors using photon migration spectroscopy. Neoplasia 2000; 2:26–40.
33. Kettler H, White K, Hawkes S. Mapping the landscape of diagnostics for sexually transmitted infections: key findings and recommendations. Geneva: World Health Organization Special Programme for Research and Training in Tropical Diseases; 2004.

APPENDIX A

Optical Spectroscopy

Biopsy samples were analyzed using a custom-built automated transmission optical spectrometer (**Fig E1**, available online at *www.jvir.org*). A deuterium-tungsten light source (DH-2000-DUV; Ocean Optics, Dunedin, Florida) delivered broadband illumination in the ultraviolet (UV)-visible-near-infrared range (200–1,100 nm). Light passed through a correction filter (FSR-KG2; Newport, Irvine, California) for uniform output across the lamp's spectrum. To maximize transmission of UV signal, the illumination light was guided via a UV core multimode solarization-resistant fiber (QP600-1-SR; Ocean Optics) and focused on the sample using a 10-mm focal length fused silica lens (74-UV; Ocean Optics). Illumination intensity was controlled by neutral density filters (NDUV series; Thorlabs, Newton, New Jersey) selected from an addressable motorized filter wheel (FW102C; Thorlabs).

Imaging was performed directly across the core needle cancer biopsy sampling trough. A custom snap-in acrylonitrile butadiene styrene three-dimensional–printed needle holder was used to align the tip of the needle with the illumination beam prior to scanning. Light transmitted through the tissue sample was collected by fused silica lens (74-UV) and fiber-optically guided to a spectrometer coupled with a multichannel array detector (VS70; Horiba; spectral window, 190–1,000 nm; Kyoto, Japan).

Data Acquisition

Up to 20 transmission spectra were collected at approximately 0.75-mm sampling increments along the biopsy sample. The spot size of the beam incident on the sample was approximately 0.25 mm, verified by a knife-edge technique. This resulted in a sampling volume of approximately 0.17 mm^3 at each location. Similar to routine clinical core needle cancer biopsies, discontinuous samples (fragments) were commonly acquired, resulting in fewer than 21 spectra per sample. High dynamic range spectra were constructed by recording multiple exposures (10, 100, and 1,000 milliseconds of integration times) merged into a single spectrum optimized for signal-to-noise across the entire 190–1,000–nm detector sensitivity range. Data acquisition time was typically 45 seconds per biopsy, which includes multiexposure (high dynamic range) stitching to improve the signal-to-noise ratio. Instrument control and data acquisition were performed using a laptop computer with LabVIEW software (National Instruments, Austin, Texas).

Outlier Rejection

In this step, all data are concatenated, normalized, and distributed against the geometrical mean of the population. The acquired transmittance spectra were normalized to values between 0 and 1 by subtracting the dark spectrum and dividing by the reference spectrum. The spectra that fall outside the interquartile range of the distribution are rejected as outliers. In practice, these rejected spectra correspond to either low signal-to-noise ratio (due to light obstruction by the needle) or 100% transmittance (due to empty tray regions that are not covered with tissue). Samples falling out of the interquartile range were evaluated and confirmed to contain very little discriminative power. Specifically, these spectra were either predominantly flat across a wide range of wavelengths (saturated) or predominantly noise. Moreover, we performed separate classification experiments in which outliers were given a separate label and then classified against the main tissue types (normal and tumor). In these experiments, we confirmed that the outliers could be easily differentiated from the signals of interest, having a negligible impact on the sensitivity and specificity of tumor detection.

APPENDIX B

Classifier Performance For Individual Spectra

Machine learning classifiers were trained on the extracted principal spectral components to differentiate tumor spectra from normal tissue spectra using logistic regression, support vector machine, and random decision forest (RF) classification algorithms (1,2). The spectra were designated as either tumor or normal based on an optimal decision threshold on the probability of tumor ($p > P^{s}_{th}$). Each tissue spectrum was decomposed into statistically significant principal components by solving the eigenvalue problem for the concatenated data. The probability of a full core needle cancer biopsy (CNB) sample representing predominantly tumor or normal tissue was computed by combining the likelihood of each of its constituent spectrum via a signal fusion approach, where the spectra from the biopsy sample are treated as the output of different sensors. The naive Bayes assumption was used to compute the likelihood of each sensor output (spectra) belonging to either tumor or normal classes. To test the robustness of classification to confounders, the analysis was repeated with 3 additional types of tissue spectra commonly obtained during clinical kidney mass CNBs: (*a*) renal sinus fat, (*b*) blood, and (*c*) necrotic tumor regions. New multiclass classification models were trained to obtain probabilities of each spectrum belonging to one of the 5 classes.

Full Core Needle Biopsy Classifier Performance

The probability of a full CNB sample representing predominantly tumor or normal tissue was computed by combining the likelihood of each of its constituent spectra via a signal fusion approach and using the naive Bayes assumption. RF, the best performing individual spectral classifier, was used to build the full CNB classifier.

Performance Evaluation

The optimal decision thresholds, P^{s}_{th} and P^{b}_{th}, for the spectral and full CNB sample classifiers were computed using the Zweig and Campbell method (3), where the slope of the optimal receiver operating characteristic point is computed based on a combination of misclassification cost and prevalence analysis. The cost of misclassifying a tumor spectrum as normal tissue and vice versa was assumed to be equal. The McNemar statistical test, Matthew correlation coefficient, and kappa statistic were used to compare the performance of RF, logistic regression, and support vector machine algorithms on spectral classification and whole CNB performance. We tested the stability of our models with regard to imbalance using the synthetic minority oversampling technique for the underpopulated classes and by downsampling the overpopulated classes. The latter reduced the population size too much to train the classification model, whereas former confirmed stability of model training and performance.

REFERENCES

1. Huerta-Nuñez LFE, Gutierrez-Iglesias G, Martinez-Cuazitl A, et al. A biosensor capable of identifying low quantities of breast cancer cells by electrical impedance spectroscopy. Sci Rep 2019; 9:6419.
2. Patil PD, Hobbs B, Pennell NA. The promise and challenges of deep learning models for automated histopathologic classification and mutation prediction in lung cancer. J Thorac Dis 2019; 11:369–372.
3. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 1993; 39: 561–577.

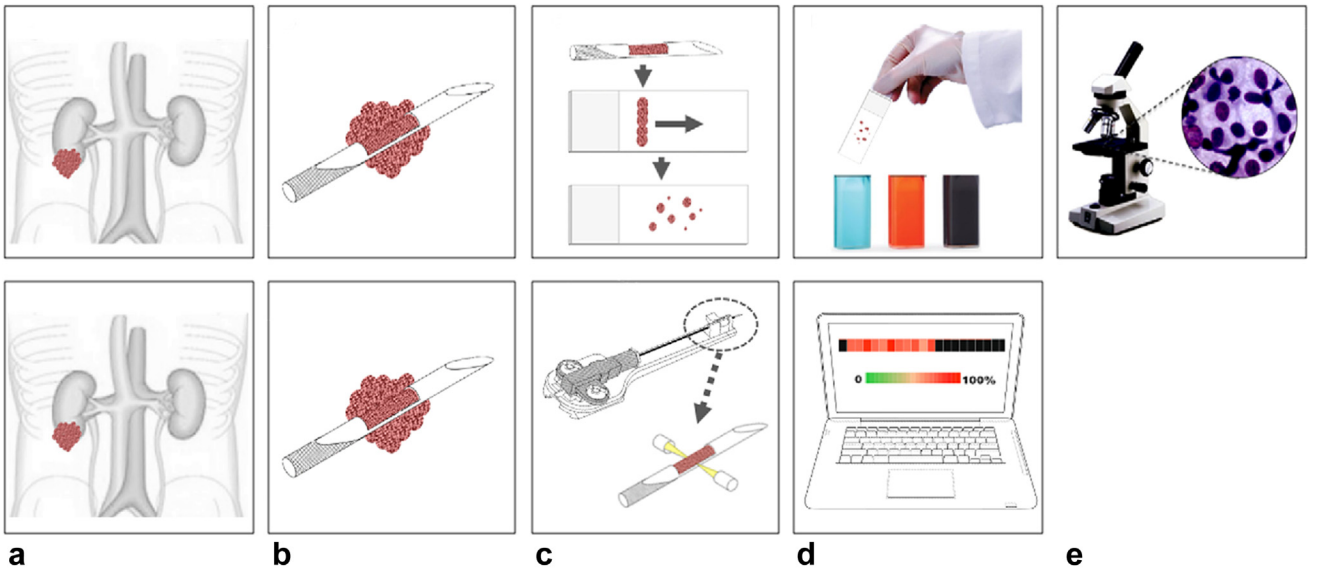

Figure E1. Comparison of glass slide imprint cytology versus transmission optical spectroscopy evaluation of core needle biopsy samples. Imprint cytology (top row, panels **a–e**) is based on the presence or absence of cytopathology on slides viewed by light microscopy. **(a)** Representative tumor in the lower pole of the right kidney, **(b)** core needle biopsy sample obtained from the tumor, **(c)** core sample physically advanced across glass slide to leave cytologic imprint, **(d)** imprinted slide dipped into a fixative agent followed by eosinophilic and basophilic solutions and finally rinsed in water (approximately 10–15 seconds per solution), and **(e)** evaluation of stained slide under a microscope by a cytopathologist. The typical time required for slide review per sample is 3–5 minutes (steps **c–e**). Based on a cytopathologist's interpretation of cell number, morphology, architecture, and staining characteristics, repeated needle passes may be recommended. The transmission optical spectroscopy (bottom row, panels **a–d**) workflow does not require a cytopathologist, sample transfer to slides, or staining methods that degrade the sample. **(a–b)** A needle biopsy sample was obtained; **(c)** the spectroscopy instrument scanned the sample on the needle and provided data to a computer for analysis; **(d)** an interpretative readout was immediately provided to the operator.

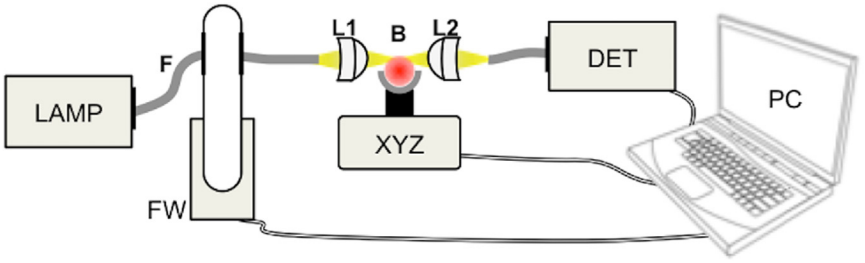

Figure E2. Schematic diagram of core biopsy spectroscopic imaging platform and subcomponents. Transmission spectroscopy systems. Deuterium/tungsten light source (LAMP), optical fiber (F), filter wheel (FW), lenses (L1 and L2), biopsy tissue core (B), motorized translation stage (XYZ), detector (DET), and laptop computer (PC).