



ANNUAL REVIEWS **Further**

Click [here](#) to view this article's online features:

- Download figures as PPT slides
- Navigate linked references
- Download citations
- Explore related articles
- Search keywords

Emerging Themes in Image Informatics and Molecular Analysis for Digital Pathology

Rohit Bhargava¹ and Anant Madabhushi²

¹Departments of Bioengineering, Chemical and Biomolecular Engineering, Electrical and Computer Engineering, Mechanical Science and Engineering, and Chemistry, and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana–Champaign, Urbana, Illinois 61801; email: rxb@illinois.edu

²Center for Computational Imaging and Personalized Diagnostics; Departments of Biomedical Engineering, Urology, Pathology, Radiology, Radiation Oncology, General Medical Sciences, Electrical Engineering, and Computer Science; and Case Comprehensive Cancer Center, Case Western Reserve University, Cleveland, Ohio 44106; email: axm788@case.edu

Annu. Rev. Biomed. Eng. 2016. 18:387–412

The *Annual Review of Biomedical Engineering* is online at bioeng.annualreviews.org

This article's doi:
10.1146/annurev-bioeng-112415-114722

Copyright © 2016 by Annual Reviews.
All rights reserved

Keywords

digital pathology, algorithms, computational, precision medicine, chemical imaging, infrared spectroscopic imaging, FT-IR spectroscopy, microenvironment, stainless staining, diagnosis, prognosis, outcome

Abstract

Pathology is essential for research in disease and development, as well as for clinical decision making. For more than 100 years, pathology practice has involved analyzing images of stained, thin tissue sections by a trained human using an optical microscope. Technological advances are now driving major changes in this paradigm toward digital pathology (DP). The digital transformation of pathology goes beyond recording, archiving, and retrieving images, providing new computational tools to inform better decision making for precision medicine. First, we discuss some emerging innovations in both computational image analytics and imaging instrumentation in DP. Second, we discuss molecular contrast in pathology. Molecular DP has traditionally been an extension of pathology with molecularly specific dyes. Label-free, spectroscopic images are rapidly emerging as another important information source, and we describe the benefits and potential of this evolution. Third, we describe multimodal DP, which is enabled by computational algorithms and combines the best characteristics of structural and molecular pathology. Finally, we provide examples of application areas in telepathology, education, and precision medicine. We conclude by discussing challenges and emerging opportunities in this area.

Contents

1. INTRODUCTION	388
2. DIGITAL PATHOLOGY METHODS	389
2.1. Instrumentation for Digital Pathology	389
2.2. Computational and Big Data Analytic Methods in Digital Pathology	389
2.3. Machine Learning–Based Classification for Disease Presence, Grading, and Outcome Prediction	391
2.4. Applications in Digital Pathology	393
3. MOLECULAR DIGITAL PATHOLOGY	394
3.1. Instrumentation and Computing Resources for Label-Free Molecular Digital Pathology	394
3.2. Recent Trends in Infrared Chemical Imaging for Molecular Digital Pathology	395
3.3. Informatics Methods for Chemical Imaging	398
3.4. Computational and Big Data Analytic Methods for Molecular Digital Pathology	398
4. MULTIMODAL DIGITAL PATHOLOGY	400
4.1. Informatics Methods for Combining Functional and Morphologic Parameters ..	400
4.2. Applications in Precision Medicine	400
5. CONCLUDING REMARKS	404

1. INTRODUCTION

Digital pathology (DP) can be broadly defined to include all aspects of acquisition, process management, and data interpretation to yield pathology information from a digitized pathology sample's image. First, the field is largely based on utilizing the advantages of modern computer and electronic (digital) systems (1–4) to improve and enhance pathology practice. Second, the development of whole-slide imaging (WSI) technology has enabled digitization of tissue slides, giving rise to new areas such as telepathology (5) and remote diagnosis (6, 7). Third, informatics and big data analytic methods (8) are providing unprecedented detail about data from the subcellular to the tissue level [in, e.g., nuclei (9–13), mitoses (14, 15), and lymphocytes (12, 16)]. Moreover, recently proposed machine learning–based approaches, in conjunction with “subvisual” image biomarkers of disease architecture, could provide information about the state of aggressiveness of the disease and enable prognostic prediction of therapeutic outcome (17).

A major shift is also taking place by including markers derived from molecular biology or chemical analysis of tissues. Whereas long-used stains such as hematoxylin and eosin (H&E) are fairly nonspecific, chemically or molecularly specific stains are increasingly being employed to understand disease progression (18). Immunohistochemical (IHC) or molecular imaging is becoming increasingly quantitative (19) and is evolving for multiplexed analysis (20). More recently, label-free methods for pathology have been developed using spectral imaging (19, 21) due to advances in instrumentation (22–25). Direct recording of chemical composition eliminates the need for dyes or stains and could be much more informative because it is not restricted to the imaging of known molecular species. The complex spectral data, however, need informatics approaches in order to be interpreted and used robustly in routine analysis (26).

The focus of this review is on the two major themes of structural and molecular trends in DP enabled by informatics. Complementing related reviews (27–31), we first discuss advances in

computer algorithms. Next, we describe the state of the art in instrumentation and informatics that enable molecular DP (mDP). Because excellent reviews and monographs are available for IHC imaging and spectral analysis in molecular imaging, we focus on recent, fast-moving developments in label-free spectral imaging and the role of informatics therein. Finally, using illustrative examples, we discuss opportunities for both DP and mDP in the context of applications areas such as telepathology, education, and precision medicine. We also briefly discuss the regulatory and infrastructural challenges involved in incorporating these technologies into routine practice, both in the United States and globally.

2. DIGITAL PATHOLOGY METHODS

2.1. Instrumentation for Digital Pathology

Instrumentation for DP consists of a microscope and computing, storage, and visualization hardware. The microscope and computing hardware have been extensively described. Enhancements for fluorescence or multispectral optical imaging are commercially available. Instrumentation has been transformed from static snapshots of specific fields of view taken with camera-equipped microscopes to scanning of whole glass slides with integrated, often quantitative histologic evaluation. The primary challenge involves ensuring high quality and fidelity of digital images so that pathologists' diagnostic performance is not compromised when they rely on digital data. Whereas pathologists have traditionally had to pore over stained tissues on glass slides, technology will enable them to use the same basic equipment but will advance information content, visualization, and analytical tools to greatly improve decision making.

2.2. Computational and Big Data Analytic Methods in Digital Pathology

Computerized image analysis is at least 50 years old. Although face recognition algorithms (32) were proposed in the early 1970s, image processing and pattern recognition algorithms began to be applied to medical data with the digitization of radiology 30 years ago (33). Similarly, although analysis algorithms for microscopic cellular images (34) were put forward nearly 50 years ago, only recently, with the introduction of WSI, did tools to aid the pathologist (for, e.g., mitosis and nuclear counting) become available. These approaches aim to extend human cognition by predicting disease outcome and aggressiveness (35, 36), thereby providing true decision support.

In this review, we focus first on the emerging theme of image informatics as histologic companion diagnostics. Although visually reading routine pathologic slides can allow pathologists to make diagnoses, sophisticated computer-aided image analysis can uncover more-revealing subvisual attributes from morphology (e.g., texture, shape, architecture). Researchers have shown that these histologic biomarkers are correlated with disease progression independently of existing clinical and pathologic features. Careful processing steps are needed in order to capture subvisual cues from DP that may be associated with tumor heterogeneity and disease outcome, as described in the following subsections.

2.2.1. Color standardization. The visual appearance of digital histopathology slides has many sources of variance, two of which are the staining protocol and the choice of digital slide scanner. Although pathologists are unaffected by these factors when rendering a diagnosis from H&E images, these variances often limit their ability to generalize computer-based approaches. Consequently, development of color standardization and normalization algorithms to help improve the performance of subsequent image analysis algorithms (13, 37) has been a primary goal of recent research. Often, a single image with optimal tissue staining and visual appearance is designated

the template. Intensity distributions of subsequent images are mapped to match the distribution of the template image. Several studies (38–40) have suggested that matching distributions of tissue subtypes (e.g., epithelium, nuclei, stroma) is more optimal than aligning global image distributions. In the context of histopathology, this process might involve identifying stromal tissue, nuclei, lymphocytes, fatty adipose tissue, and cancer epithelium, within both the target and template images, and then specifically establishing correspondences between the tissue partitions in the two images. This two-level matching process can be performed either by images with one contrast mechanism or by multimodal imaging and alignment (41).

2.2.2. Detection and segmentation of histologic primitives. Automated counting and enumeration of histologic primitives such as nuclei (42) in H&E images have attracted significant attention. Recently, there has been interest in grand challenge competitions in digital pathology image analysis, in which groups compete with their algorithms on specific problems by using a set of common data sets. These grand challenge competitions have involved detection algorithms to identify nuclei (29), lymphocytes (43), and mitoses (14). In addition to aiding pathologists in manual disease grading (9, 44–47), these identifications are also critical for subsequent automated feature analysis algorithms. Several of these detection algorithms have employed well-established computer vision and image processing approaches with heuristic rules such as active contours (48), level sets (49), and active shape models (50). Deep learning approaches, which involve the use of multiple layers of convolutional neural networks for unsupervised learning of features in order to identify the object of interest from a set of training images, led their creators to win recent grand challenges in DP image analysis (51). **Figure 1** shows results of a deep learning model used to distinguish areas of invasive cancer from benign stromal areas in breast cancer H&E images (52).

2.2.3. Quantitative histomorphometry and feature analysis. Quantitative histomorphometry (QH) involves computerized image analysis tools for quantitatively assessing cancer tissue and

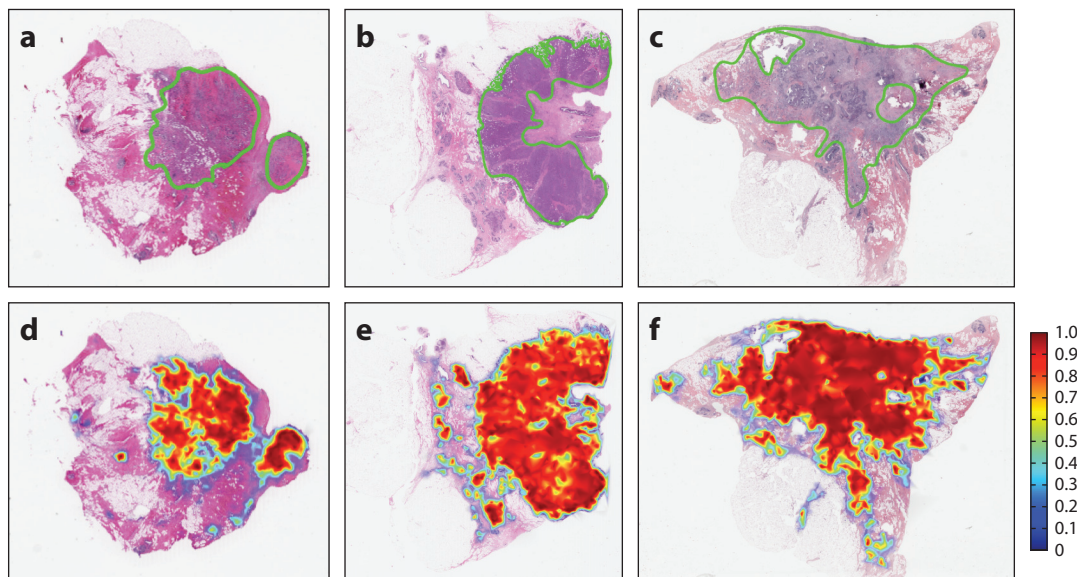


Figure 1

(Top) Probabilistic output of a deep learning classifier for regions of invasion. (Bottom) Corresponding hematoxylin and eosin images with a pathologist's mark of the extent of cancer extent. Note the concordance between the two rows.

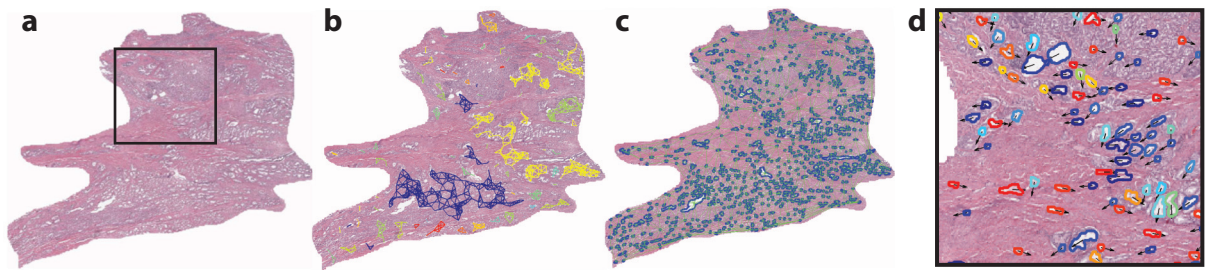


Figure 2

(*a*) Prostate cancer tumor region. The region of interest (ROI) is outlined in blue. (*b*) Cluster graphs establish localized gland networks. (*c*) Delaunay triangulation reveals a global graph which traverses stromal and epithelial boundaries, whereas co-occurring gland tensors compute localized features from the gland networks. (*d*) The ROI from panel *a*. The color map of the gland orientations (0° , 180°) demonstrates the variation in local gland orientation. Gland orientations are architecturally differently arranged in tissue from patients with and without disease recurrence (58).

non-cancer tissue morphology and architecture. QH measurements can be divided broadly into three groups: architectural, shape, and texture based.

Architectural features capture the arrangement and spatial topology of histologic primitives such as individual nuclei, tubules, mitoses, and lymphocytes. The spatial location of a particular primitive is considered to be a node in a graph. The nodes are then connected using graph construction algorithms [e.g., Voronoi (53), Delaunay (16), minimum spanning tree (54)]. Quantitative measurements (e.g., internode distance, clustering coefficient of the nodes) can quantitatively characterize the graph and, hence, the image. The spatial architecture and orientation of all nuclei and glands on DP images have recently been expanded to encompass the local architecture of the primitives through the use of algorithms such as cell cluster graphs (55, 56), which successfully predicted progression in p16⁺ oropharyngeal tumors (57). **Figure 2** shows an example of global and cell cluster graphs from a digitized image of a radical prostatectomy specimen.

The shape of individual histologic primitives can indicate the presence of disease. Shape features such as fractal dimension (59), angularity, size, and smoothness of the boundary (47) differ between nuclei and glands in high and low grades of prostate and breast cancers. Finally, disorder (or entropy) in the orientation of nuclei and glands in prostate tissue predicts biochemical recurrence (BCR) postsurgery in patients with prostate cancer (58, 60).

Texture refers to quantitative measures of spatial neighborhood interactions between pixel intensities within local neighborhoods in an image. These could include first-order spatial intensity interactions (e.g., mean, standard deviation, median, variance) within local neighborhoods and second-order interactions (e.g., co-occurrence features). More complex textural features can also be extracted; these include steerable and multiscale gradient features via mathematical operators such as Gabor filters (61), local binary patterns (62), and Laws filters (63). The shape and texture of nuclei within the stroma are significantly correlated with disease recurrence and patient outcome in breast (64), prostate (65), and oropharyngeal cancers (55). **Figure 3** shows the digital stain representation of a routine H&E image, with overlays of nuclear architecture networks and capture of stromal and epithelial textural variations.

2.3. Machine Learning–Based Classification for Disease Presence, Grading, and Outcome Prediction

A typical analysis pipeline involves a machine learning classifier that takes as input a series of manually or computer-extracted features and employs those features to render a prediction. In

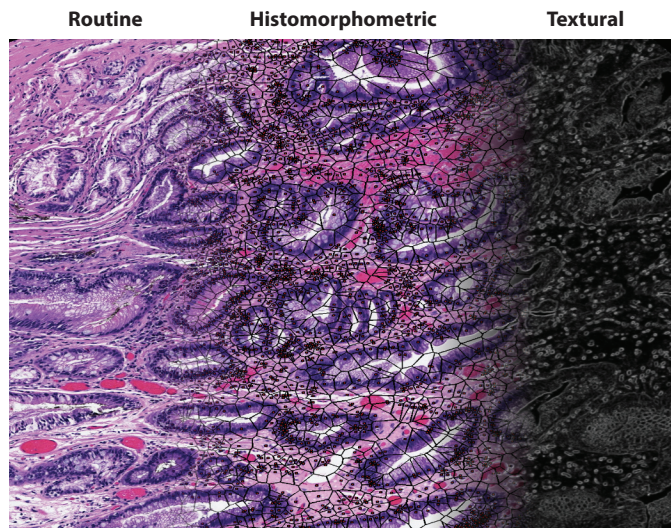


Figure 3

A digital stain. A routine hematoxylin and eosin tissue image (*left*) can be converted into a histomorphometric representation comprising nuclear architecture (*middle*) and textural measurements (*right*).

the context of DP, predictions might involve a low-level recognition (e.g., Is the primitive a nucleus or not?), a diagnostic decision (e.g., Is the tissue region of interest cancerous or not?), or a prognostication (e.g., Will the patient have early or distant disease recurrence?). Machine learning classifiers can be broadly categorized into supervised and unsupervised approaches. Whereas unsupervised learning approaches try to identify natural groupings directly from the feature space, supervised approaches rely on labeled training exemplars in order to learn the best classifier that optimally distinguishes instances from the classes of interest. However, the challenges involved in training supervised classifiers in the context of WSI data are data density and, hence, the need to address issues pertaining to computational complexity. To address these issues in processing WSI data, one can use multiresolution classifiers that initially perform predictions at lower resolutions and map results to the next level of image resolution. A new, higher-resolution classifier is then employed at the next level in the image pyramid and is more efficient because the interrogation is limited to areas identified as suspicious at the lower resolution. This approach, modeled on the way pathologists tend to review slides under a microscope, is employed for both diagnosis (66) and grading (45).

Both supervised and unsupervised classification approaches tend to be susceptible to the curse of dimensionality (67), an issue when the size of the feature space is significantly larger than the number of data samples. To address this issue, especially for multispectral images, investigators are increasingly using data dimensionality reduction methods. These methods can be broadly categorized into linear and nonlinear methods (68) on the basis of the approach used to map the higher-dimensional feature space onto a set of eigenvectors. Supervised and unsupervised classifiers can then be constructed in the lower-dimensional eigenvector space instead of the original high-dimensional feature vector space. A problem posed by dimensionality reduction methods is that they tend to obfuscate the original feature space, leading to a lack of transparency in the resultant classifiers. A method called FINE (feature importance in nonlinear methods) (69) has recently been proposed to reverse map eigenvector representations to the original set of image features.

2.4. Applications in Digital Pathology

Applications of image analysis for DP include not only computer-assisted diagnosis of disease but also telepathology for remote consulting, education, and precision medicine. In this section, we describe some key problems in health care and use selected examples to show how DP solutions are emerging to address them.

2.4.1. Precision medicine. There is a broad need for predictive and prognostic assays to distinguish aggressive from less-aggressive phenotypes of cancer so as to identify optimal therapies in individual patients and guide clinical trials (70). Most prognostic tests in the United States and Europe are based on gene expression assays. Recent studies have shown extensive genetic heterogeneity among cancer cells between tumors and even within the same tumor. In fact, molecular signatures for both good and bad prognoses can be found in the same tumor. Interestingly, tumor morphology observed on standard H&E staining is still remarkably useful for tumor characterization. In reality, tumor morphology reflects the sum of all temporal genetic and epigenetic changes and alterations in tumor cells, thereby providing incredible utility for predicting tumor biology, clinical behavior, and treatment response. Consequently, there has been recent interest in identifying computer-assisted image features from routine H&E images that can help predict disease progression (71) and disease recurrence (58, 64). The successful validation of these H&E-based tests could have a significant impact in low- and middle-income countries that do not have access to the more expensive, molecular-based prognostic tests and assays. **Figure 4** shows an example of the use of image informatics in the context of prediction outcome in oropharyngeal cancers. Specifically, Lewis and colleagues (71) showed how the combination of nuclear graphs within both the stromal and epithelial compartments can predict progression in p16⁺ oropharyngeal cancers.

2.4.2. Telepathology. Telepathology (5, 72, 73)—the remote analysis of biopsy samples using digital imaging technology—has been steadily increasing in popularity, not only in the United States (74) and Europe but also in low- and middle-income countries, for two reasons. First, the technology sector and Internet infrastructure are rapidly growing. Second, WSI scanners are becoming increasingly available, enabling rapid digitization in routine pathology. Telepathology can be highly effective. For example, a recent two-center study showed that a clinically important diagnosis was achieved via telepathology in 93% of 100 patient cases (75). Most cases (79%) were analyzed within 3 days, demonstrating the strong potential of telepathology for servicing low-resource areas.

2.4.3. Education. Digitized slides can be shared and accessed from any location with an Internet connection (76, 77), which can prove beneficial for medical practice as well as education. For example, Ventana Medical Systems, Inc. has developed PathXchange (78), an educational image sharing system that has a global professional community and case library for networking, telepathology, sharing, and archival. A recent study (79) showed that second- and third-year medical students easily adapted to virtual microscopy, found it user friendly, and thought that the opportunity to view slides remotely was a huge advantage. The study also found that virtual microscopy increased collaboration and interactions among students. Recently, researchers introduced an Internet-based training tool called Score the Core, which employs tissue microarrays to train pathologists to visually score estrogen receptor, progesterone receptor (percent positive), and Ki-67 (percent positive) (80). A software infrastructure for tracking viewed regions in WSI and, using the regions tracked during a practical exam in oral pathology (81), analyzed the collected data to discover students' viewing behavior.

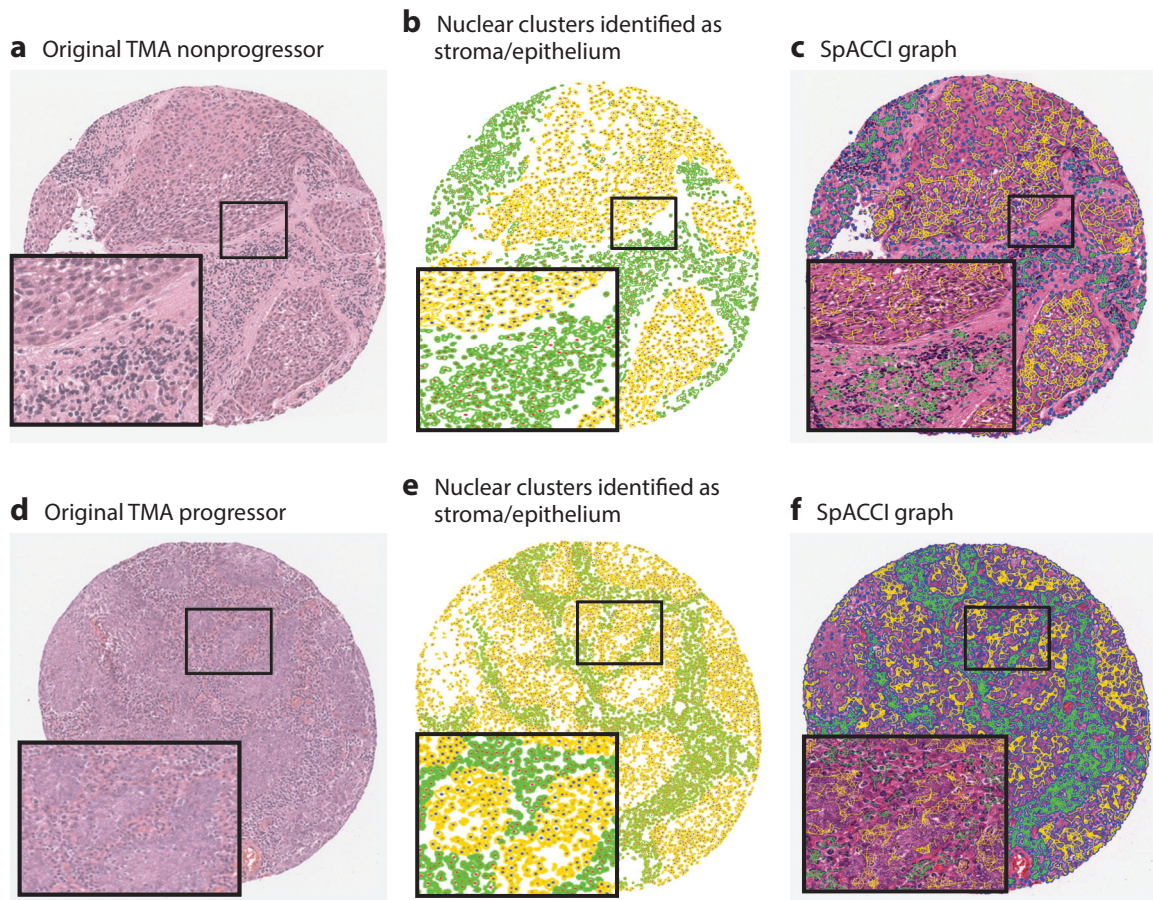


Figure 4

Nuclear architecture features can be extracted within the stromal and epithelial compartments within hematoxylin and eosin tissue sections. The combination of stromal and epithelial nuclear architecture features, referred to as a SpACCI graph, enables improved prediction of which p16⁺ oropharyngeal cancers will and will not progress. Abbreviation: TMA, tissue microarray.

3. MOLECULAR DIGITAL PATHOLOGY

3.1. Instrumentation and Computing Resources for Label-Free Molecular Digital Pathology

Because the acquisition of IHC microscopic images is similar to the use of current instrumentation in pathology, we focus on the state of the art of spectroscopic imaging, or chemical imaging (CI). IR imaging (82) is one of the major modalities that promises excellent image contrast, fast data recording, and exceptional molecular sensitivity; therefore, we consider it an example of the progress and potential of using CI for DP.

Given that vibrational frequencies within molecules directly resonate with optical frequencies in the mid-IR spectral region, light absorption provides a quantitative molecular fingerprint of the material, providing ample molecular biomarkers. No dyes or stains are needed to visualize molecular content, so data can be recorded from a variety of samples without prior knowledge of

the type or composition of the sample (**Figure 5a–e**). Thereafter, informatics techniques are used to extract the desired information or to discover new information (83).

3.2. Recent Trends in Infrared Chemical Imaging for Molecular Digital Pathology

In state-of-the-art instrumentation, an interferometer is combined with a microscope to yield both spectral multiplexing and spatial multichannel advantages. An image is produced with a full spectrum acquired for every pixel. In this Fourier transform (FT) approach, entire spectra (typically, 1,024–2,048 points for 4–8 cm^{-1} resolution and 2^n points for fast FT) from many pixels (typically, 16–16,384) are rapidly and simultaneously acquired (84). The small format of these detectors remains a major impediment to data recording speed and will likely be improved in order to enable WSI.

Interestingly, classification methods have shown (85) that collection of entire spectra is unnecessary. This idea drove the development of discrete frequency IR (DFIR) spectroscopy and imaging (82). Although DFIR imaging was initially achieved with the use of inexpensive micro-fabricated filters (86, 87), the throughput was low, and advances in tunable lasers can considerably enhance the DFIR approach and make it practical for use in mDP. In addition to hardware, another exciting development involves increasing the quality and speed of data from CI.

3.2.1. High-resolution imaging. Akin to optical microscopes, the vast majority of experiments in IR imaging have generally relied on transmission sampling measurements. In attenuated total reflection imaging, the evanescent electric field generated from a solid immersion lens (SIL) probes a limited sample volume that is in contact with the lens. First, this technique can allow imaging of fresh-frozen samples, potentially enabling intraoperative mDP by reducing water absorption (88). Second, subwavelength SIL imaging provides $\sim 1.25\text{-}\mu\text{m}$ spatial resolution (of the surface layer) with no need for sectioning, fixation, or staining. Numerous biomedical applications have been reported (89, 90); however, the unique optical configuration, loss of throughput, and need for contact complicate the experiment. Investigators have also used a subwavelength tip probe (91) for nanoscale measurements. Use of nanoscale imaging systems (92) in research to identify molecular changes is likely imminent, but clinical applications may not be practical in the near future.

3.2.2. New, fast, and high-performance instrumentation. Guided by new instrumentation (93) and theoretical developments (94), the information content of IR images has been shown to be much higher than previously thought, resulting in high-definition (HD) IR imaging. **Figure 6** depicts an example of the improvement in image quality using HD optics compared with the conventional design. This approach, either with or without computational extensions (95), implies that, despite the longer wavelengths in the IR region, image quality similar to low-power optical microscopy can be obtained. HD instrumentation, with rationally designed optics (96), also poses challenges (e.g., lower throughput, need for greater speed), which can be addressed with signal processing approaches (97) and/or higher throughput, new sources (98, 99) such as quantum cascade lasers (QCLs) (100–102), and DFIR imaging. Recent studies on HD tissue imaging (103, 104), which has subcellular sensitivity, present an additional challenge to informatics methods. Broadband QCL microscopy was only recently shown to be faster than FT-IR imaging (25). Although initial attempts to use the new laser technology (105) did not provide data of sufficiently high quality, new commercial instruments (106) and laboratory prototypes have demonstrated great potential (107). In particular, through the coupling of spectral variable selection and imaging, a three-orders-of-magnitude-shorter time was recently achieved in comparison to the fastest

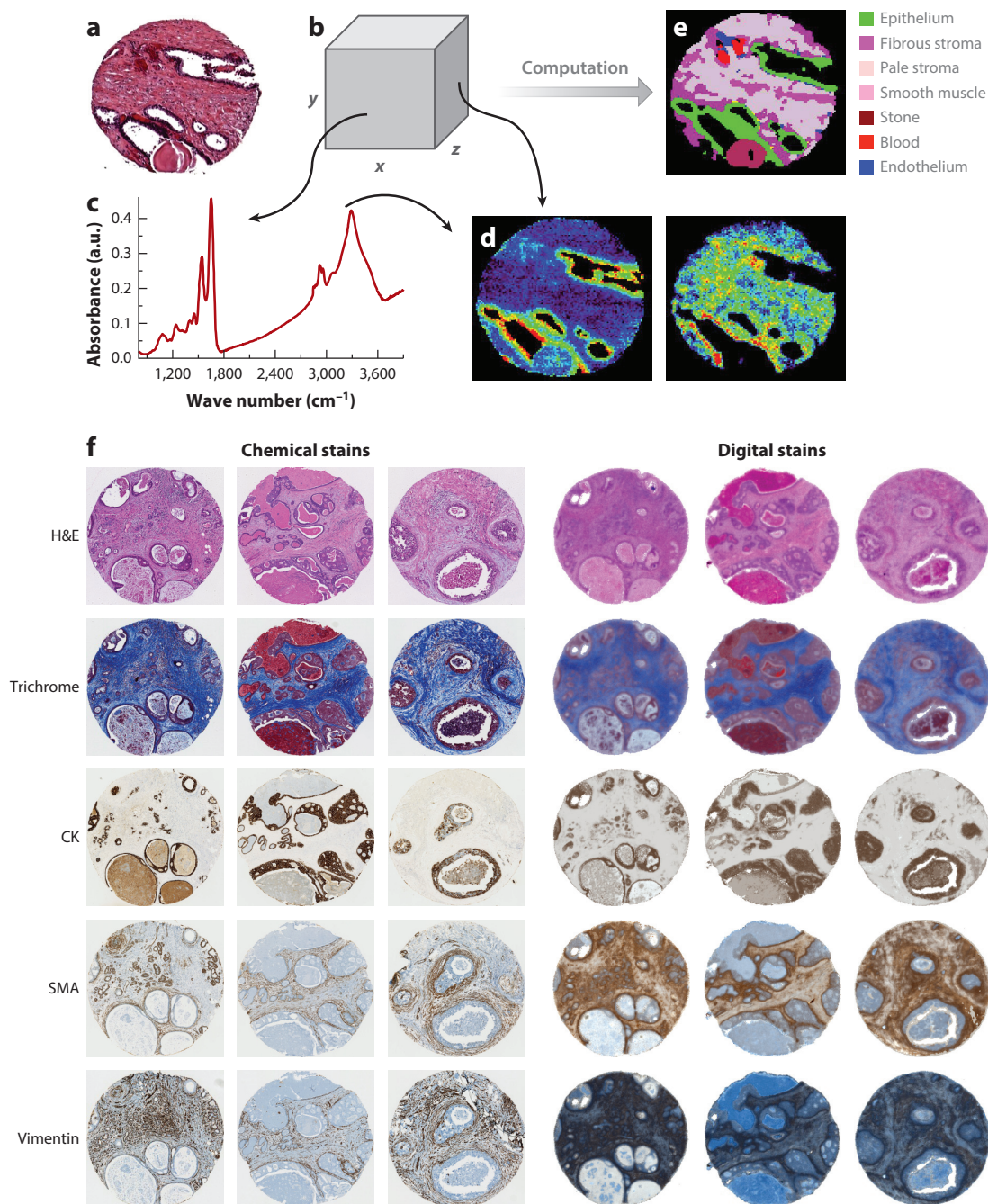


Figure 5

(a) Conventional imaging in pathology requires dyes and a human to recognize cells. (b) In chemical imaging data, both (c) a spectrum at any pixel and the spatial distribution of any spectral feature can be observed, as in (d, left) nucleic acids (at $\sim 1,080 \text{ cm}^{-1}$) and (right) collagen (at $\sim 1,245 \text{ cm}^{-1}$). (e) Computational tools can then translate the chemical imaging data into knowledge used in pathology.

(f) In addition to the color-coded image in panel e, molecular imaging (left) can be reproduced by chemical imaging (right).

Abbreviations: CK, cytokeratin; H&E, hematoxylin and eosin; SMA, smooth muscle α -actin.

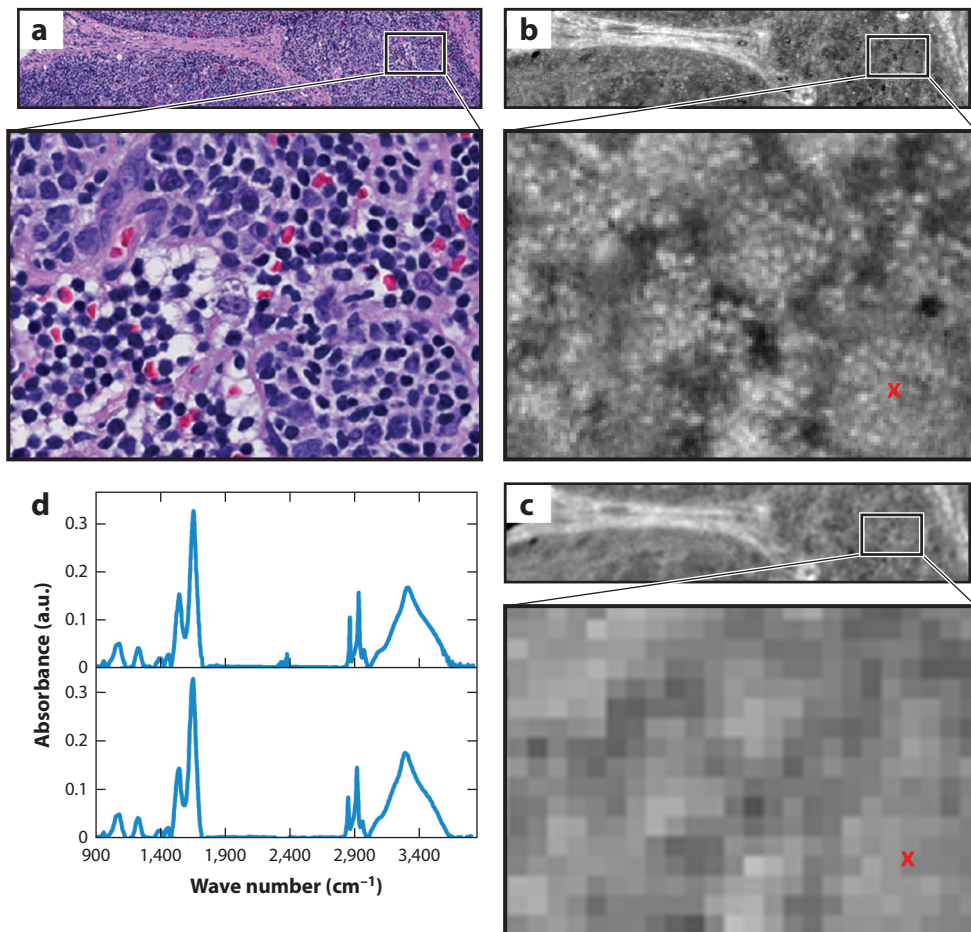


Figure 6

Comparison of hematoxylin and eosin (H&E)-stained optical microscopy and infrared (IR) images of lymph node tissue. (a) An H&E-stained image from a healthy submandibular lymph node biopsy. (b) A high-definition IR image of a serial section of the lymphoid tissue. (c) The same region imaged with a lower-resolution Fourier transform IR (FT-IR) imaging spectrometer. The IR images show the absorbance at $3,075\text{ cm}^{-1}$ after baseline correction. (d) Sample spectra plotted from the pixel marked with a red *x* in panel c. There is a slight discordance between the H&E and IR images because they are on different tissue sections.

available HD FT-IR imaging system. We anticipate that the new technology will allow imaging of biopsy samples in only a few minutes.

3.2.3. Fundamentals of spectroscopic imaging: midinfrared theory and bottom-up biology experiments. The straightforward hardware coupling of an interferometer with a microscope and array detector in FT-IR imaging incorrectly signaled that the theoretical models used to understand recorded data were also straightforward. Investigators now realize that a sample's geometry, morphology, microenvironment, and imaging optics also affect the spectrum (96, 108). Although one way to find features that are relatively insensitive to these effects is to use ANOVA (analysis of variance) (109) or bar coding (110), another is to use models (111, 112) of scattering prior

to employing informatics approaches. Accurate modeling of the image formation process (113) to recover structural, optical, and morphologic properties of the sample requires more research. Complementary to informatics experiments, many studies seek to understand the biological origins of the signal itself. The important transformations and factors relevant to pathology that have been studied (114–117) include the cell cycle, sample preparation, and the response of biological systems to various stimuli. Although it is impossible to review the large body of research here, fundamental studies remain necessary to provide input for theory-based reconstruction efforts as well as to correlate pathologic identities and optical readouts.

3.3. Informatics Methods for Chemical Imaging

Because CI requires no prior sample knowledge, whole slides can typically be imaged. Pixel-wise computer algorithms typically identify cell types or the presence of disease, as shown on the left side of **Figure 3** and described in, for example, References 118 and 119. Whereas informatics methods mirror those for DP, the inputs are spectral and the domain of operation is spectral and not usually spatial. In two schools of thought, either specific spectral features or full spectra (raw or derived quantities such as principal components) serve as input. The advantages of both discrete data points and continuous spectra can be exploited through the use of sparse sampling methods (120). More commonly, dimensionality reduction and classification are combined to provide a classifier that is robust and fast. To this end, a recent trend has been to use optimized hardware–software approaches for real-time processing (121).

3.4. Computational and Big Data Analytic Methods for Molecular Digital Pathology

Because many frequencies are imaged, data sets are considerably larger for spectroscopic imaging. As a result, many more features become available for classification and data sizes are often more extensive than in optical microscopy. The information obtained is also multifaceted (molecular, morphologic, and diagnostic), presenting unique challenges that go beyond those for DP. Below is a brief review of informatics and computational approaches to the quantification of single and multiplexed biomarkers, as well as to spectroscopic imaging and stainless staining modalities.

3.4.1. Biomarker quantification methods. IHC methods are frequently used to assess biomarkers in order to aid diagnoses. Because IHC staining is variable and nonlinear and the manual interpretation is subjective, several approaches have been developed for automated biomarker quantification (see **Table 1** for a list of examples). Most studies use supervised learning to identify pixels stained for the specific biomarker, typically through the use of a low-level color or texture feature.

3.4.2. Feature-based approaches for quantifying multiplexed biomarkers. Apart from simply quantifying the expression of the individual biomarkers, understanding the spatial interaction between the differently expressing biomarkers (122) has attracted researchers' interest; an example is the three epitopes (123) involved in human tonsil pathology. Through quantitation of the fluorescence intensity for each marker, a colocalization pattern demonstrated that 12% of the total CD34 was colocalized with CK18, whereas only 1% of the CK18 was colocalized with CD34. Similarly, software in conjunction with a multispectral imaging system has been used to analyze CK18, α -methylacyl-CoA racemase, and androgen receptor expression in prostate cancer (124). A multiplexed biomarker imaging approach (8) used to study the tissue systems of Barrett's

Table 1 Open-source and commercial image analysis platforms for image analysis and biomarker quantification from digital pathology and multiplexed immunohistochemical (IHC) images

Digital pathology software	Capabilities and features
Precision (Aperio)	Stain separation, cell counting, separation of broad tissue classes
inForm (Caliper)	Separation of broad tissue classes
Virtuoso (Ventana)	Quantification of IHC staining
Halo (Indica)	Segmentation, analysis
Visiopharm	Segmentation, feature analysis of angiogenesis, bone growth, neural structure; cancer detection
TissueStudio	Image analysis of IHC staining
Developer XD (Definiens)	Separation of broad classes and features on the basis of shape, morphology, color of segmented objects
FARSIGHT	Optical microscopy applications, focusing on fluorescence
ImageJ	Java image analysis package designed with an open architecture to provide extensibility via plugins; recent push to integrate digital pathology images

esophagus showed 14 epithelial and stromal biomarkers, providing significant differences between high-grade dysplasia and reactive atypia.

3.4.3. Molecular digital pathology from spectroscopic imaging data. Although mass spectrometry and Raman imaging show great promise, the discussion in this review is limited to IR imaging. An overwhelming majority of studies utilize standard statistical pattern recognition techniques, including linear discriminant analysis, Bayesian methods, neural networks, and random forests, to find cell types or signatures of disease (reviewed in Reference 125). Investigators are now focusing on patient issues and outcomes (126). Notably, recent studies show that it is possible to classify tissue into carcinoma and benign at human-competitive accuracy (127), relate the cancer detected to the most probable grade (128), and to prognosticate (129).

Initial studies typically involved hundreds of spectra, a few cell types, and tens of samples. A paradigm change has occurred with the common use of tissue microarrays (TMAs) in high-throughput sampling, fast imaging, and robust informatics to provide statistically validated (130) protocols (131, 132). Whereas early techniques focused on reproducing histopathology, newer approaches are exploring data in finer detail in terms of progression in time (133), involvement of the spatial microenvironment (134), secretions (135), variants of tumors (136), and engineered model systems (137). Both supervised and unsupervised methods can be used for these purposes.

3.4.4. Stainless staining. Although staining is used for routine pathology (138), molecular techniques often require multiple markers for disease stratification (139) to be imaged on the same platform. The label-free methods described above can be useful in this regard, but because of their color-coded images they do not easily translate to these established practices in pathology. In a new approach (140), CI data were used to generate computationally stained images resembling molecular stains that are common in mDP. Because the tissue is not actually stained and the histopathologic information is algorithmically obtained, this approach has been termed stainless staining. **Figure 5f** shows a metrics approach using a neural network model to relate the biochemical input to molecular or dye parameters in order to evaluate various stains. Eliminating the need to stain can enable faster mDP in time-limited (intraoperative) settings, precious or limited samples can be imaged without perturbation, and exploratory mDP for many epitopes can be performed.

4. MULTIMODAL DIGITAL PATHOLOGY

4.1. Informatics Methods for Combining Functional and Morphologic Parameters

Tumors with similar morphologic phenotypes may have significantly different behaviors and outcomes (141, 142). There is a general consensus that the intelligent combination of multiple, independent sources of clinical, molecular, and pathological data (both visual and subvisual) can provide more predictive power (143–150), but combining these diverse channels of information is challenging because the dimensionalities of the individual features differ. Multikernel learning (MKL) and canonical correlation analysis (CCA) (151, 152) enable combinations of these attributes to create a unified predictor. MKL strategies (153) allow one to mathematically project an individual set of measurements into new eigenvector representations (which have a lower dimensionality than the original features), which can then be combined into a fused representation (150). CCA strategies (151, 152) aim to find a common representation for multimodal data in which class separation is maximized while noise is minimized. Supervised classifiers can then be constructed in the MKL and CCA reduced spaces. For example, in glioblastoma multiforme (GBM) (154), a multidimensional representation of the computer-extracted nuclear shape and architecture features from whole-mount tissue sections enabled identification of subtypes on the basis of morphometric indices and the predictive capability of each subtype, and the molecular correlates of the subtypes were consistent with previous findings.

The use of CI data for mmDP is less well established but is becoming more common (**Figure 7**). For example, cell type information (IR histology) has been combined with morphologic variables from H&E-stained images (41). The morphological features thus extracted, optimized by a two-stage feature selection method using a minimum-redundancy maximal-relevance criterion and sequential floating forward selection, were used to classify tissue samples as either cancer or noncancer. This study achieved high accuracy [area under the receiver operating characteristic (ROC) curve (AUC) > 0.97] in cross-validation trials on each of two data sets that were stained under different conditions. In the absence of IR data, the performance of the same classification system decreased both within data sets and between data sets. Only through a combination of IR and optical data could high-accuracy classification be achieved. Other investigators have combined IR with mass spectrometric data and overlain IHC (155) and IR data (156).

4.2. Applications in Precision Medicine

Recently, there has been interest in using DP informatics to predict disease recurrence. Research (157) has shown that the combination of nuclear architecture features from H&E- and Feulgen-stained images enabled prediction of BCR in prostate cancer patients; the combination of Feulgen and H&E resulted in better prognosis prediction than did H&E- or Feulgen-derived measurements alone. The results of this study are shown in **Figure 8**. The 10 top-ranked computer-extracted features from among nuclear shape, nuclear architecture (described via the Voronoi and Delaunay tessellations and cell graphs), orientation, and texture for each stain type were identified via feature selection and used to train a random forest classifier to predict BCR. The features derived from Feulgen, H&E, and both Feulgen and H&E measurements resulted in statistically significant separation of the predicted BCR and nonrecurrence groups; the combination yielded the best discrimination between BCR and nonrecurrence populations (**Figure 8**).

In a recent study (152), machine learning and data fusion methods (i.e., CCA) were used to computationally combine computer-extracted histomorphometric features and protein expression

features (via mass spectrometry) from prostate tissue specimens. The fused predictor predicted the risk of BCR within 5 years of surgery with significantly greater accuracy (AUC = 0.92) than morphometric or proteomic features alone.

The use of CI demonstrated both epithelial and stromal changes associated with cancer recurrence without the use of dyes or destructive analysis. Furthermore, because no stains were required, the stromal changes could be probed on a single sample, providing an effective means of adding microenvironment data. In a midgrade dominant prostate cancer patient cohort for whom existing prognostic tools were ineffective, the CI approach outperformed two widely used tools, the Kattan nomogram and the CAPRA-S (Cancer of the Prostate Risk Assessment, postsurgical) score, while providing a histologic explanation for the prediction (**Figure 9**). An interesting aspect of this study was the use of a frequent pattern mining approach, which enabled similar groups of

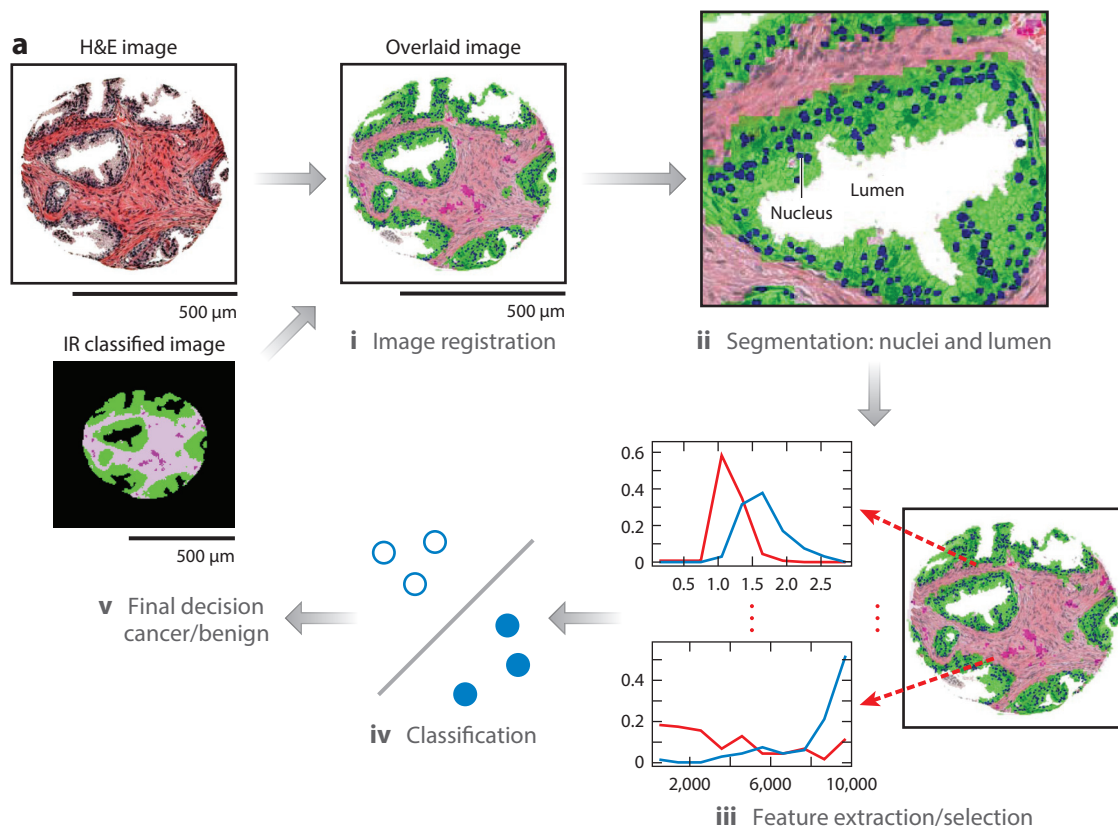


Figure 7a

(a) Overview of a multimodal digital pathology system. (i) A Fourier transform infrared (FT-IR) spectroscopy data-based cell type classification (IR classified image) is overlain on a hematoxylin & eosin (H&E)-stained image, leading to (ii) segmentation of nuclei and lumens in a tissue sample. (iii) Features are extracted and selected, then (iv) used by the classifier to (v) predict whether the sample is cancerous or benign. (b) Example features. Each panel shows one feature, along with the distributions of the feature's values for cancer (red) and benign (blue) classes. (c) IR and H&E images can be overlaid with our automated alignment algorithm, and the features allow better classification of cancer than does H&E staining alone. Abbreviations: AUC, area under the curve; AVG, average; STD, standard deviation.

cells in both the tumor and its microenvironment to be probed simultaneously. The IR spectral information was found to be independent of currently used variables in nomograms in a logistic regression model, providing a new avenue for investigation both of other aspects of the disease, such as therapy response, and of other tumors. Currently, it is not possible to make these numerous stromal observations in a laboratory setting practical for a clinical test. The challenges of measuring multiple cell types via multiple cumbersome steps and of manually combining the data from multiple sections and with clinical management have limited the potential of using the microenvironment. mmDP addresses this need directly.

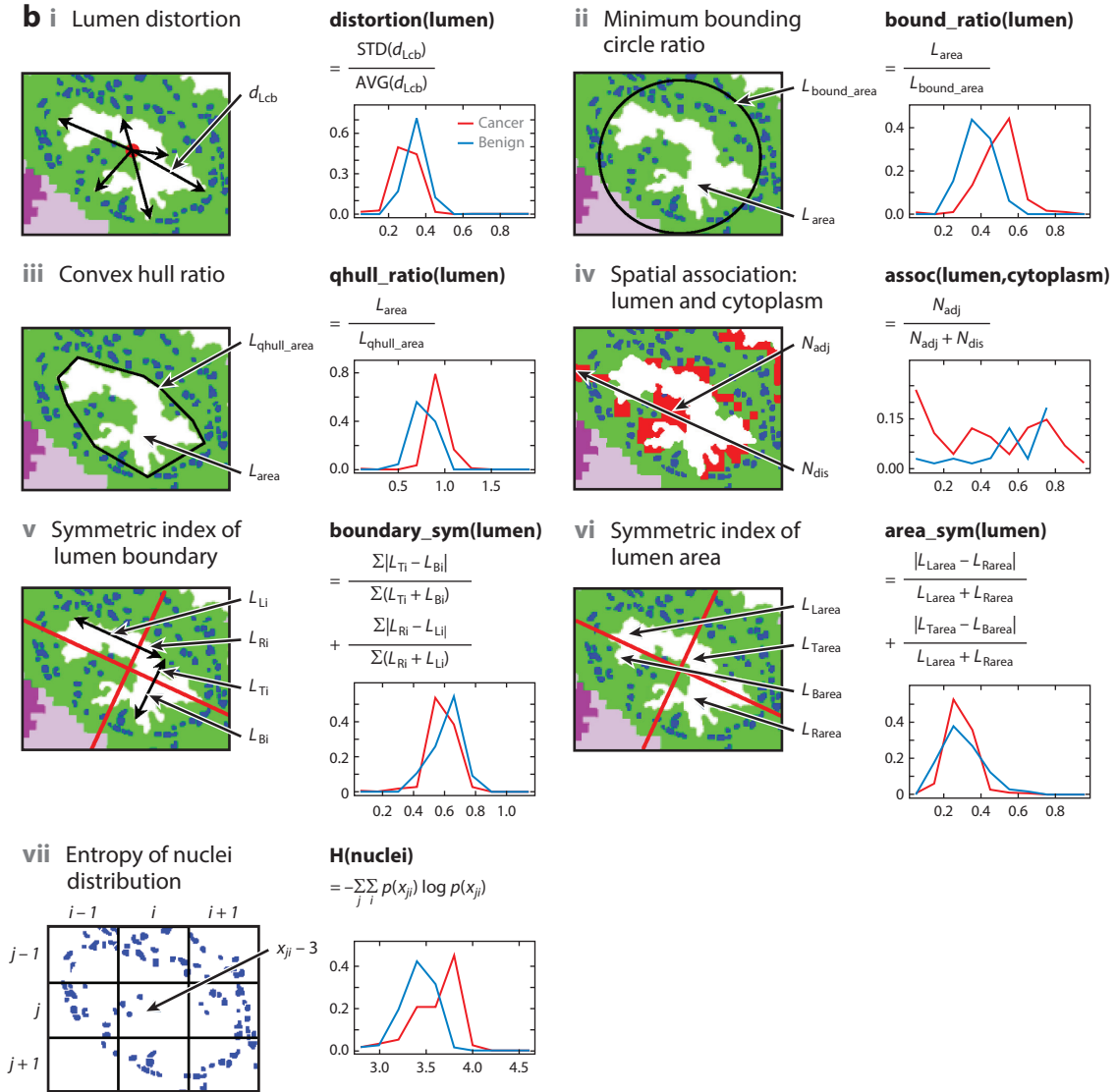


Figure 7b

(Continued)

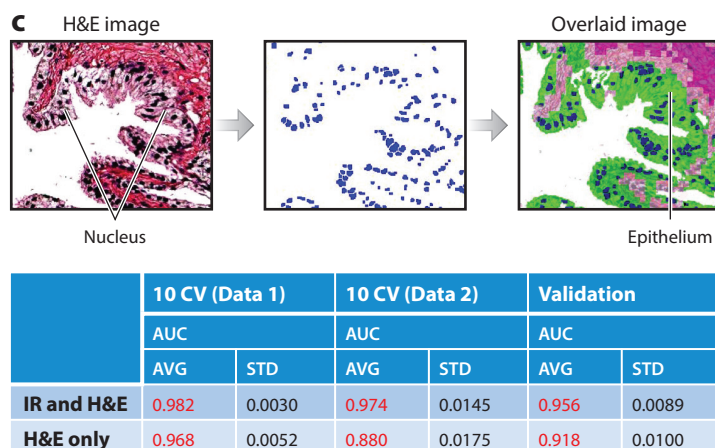


Figure 7c

(Continued)

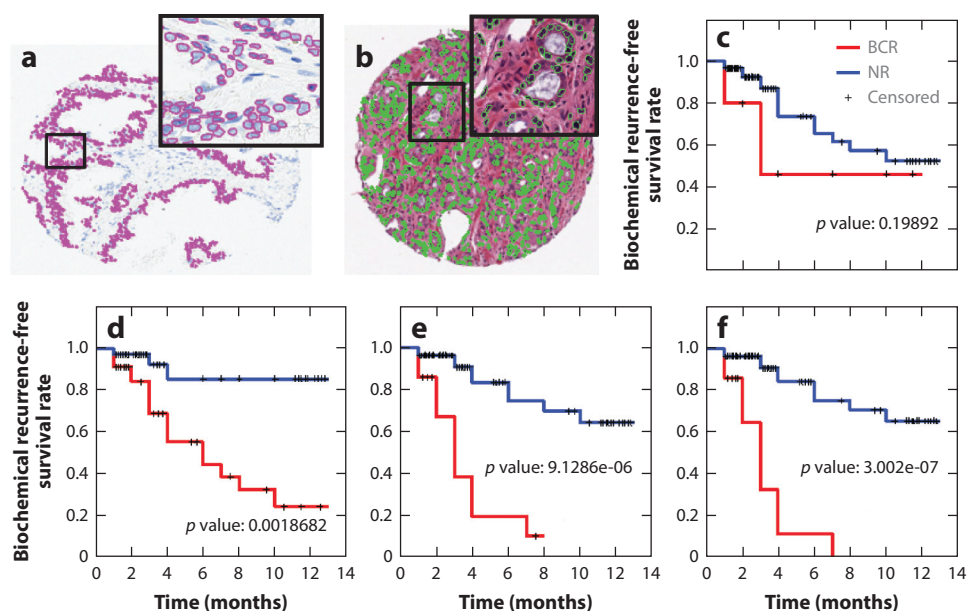


Figure 8

Combination of histomorphometric and histochemical features is more predictive of biochemical recurrence (BCR) in prostate cancer patients than either modality alone. Data are from 69 patients. Nuclei were first automatically detected in (a) Feulgen and (b) hematoxylin and eosin (H&E) images. (c–f) Kaplan–Meier curves for (c) Gleason staining, (d) the top 10 features from H&E staining, (e) a combination of H&E and Feulgen staining, and (f) a combination of Gleason-, Feulgen-, and H&E-derived image markers. Abbreviation: NR, no recurrence.

Quartile	1 (low)	2	3	4 (high)
IR OR (95% CI)	1	6.5 (1.3–46.8)	10.9 (2.0–86.8)	21.3 (4.2–168.4)
CAPRA-S OR (95% CI)	1	3.2 (0.76–15.0)	1.3 (0.28–5.3)	6.1 (1.1–40.3)
Kattan OR (95% CI)	1	0.81 (0.20–3.1)	1.5 (0.35–6.5)	1.9 (0.38–10.2)

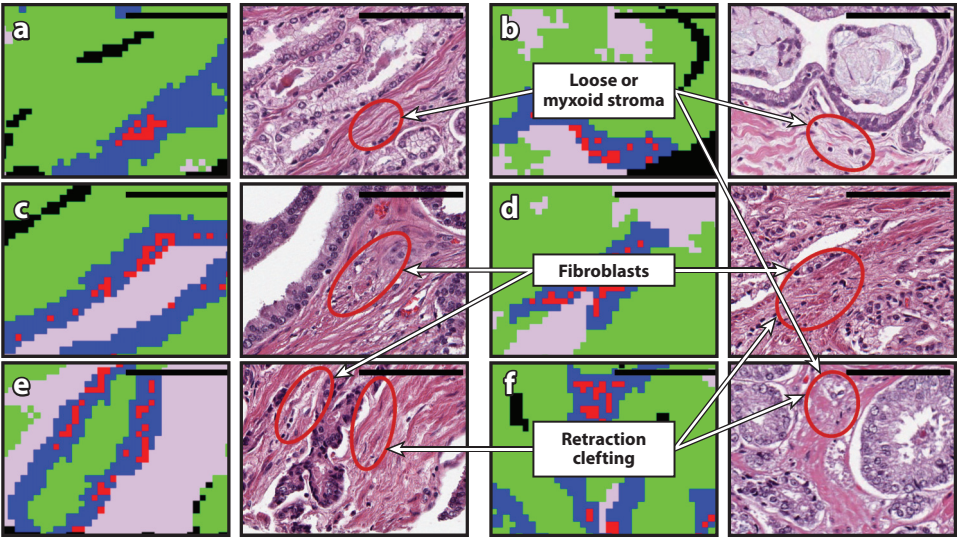
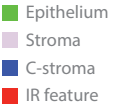


Figure 9
(Top, table) Odds ratios (ORs) of recurrence in prostate cancer shows that the infrared (IR) score is better than the clinical standards, the Kattan nomogram and the Cancer of the Prostate Risk Assessment, postsurgical (CAPRA-S) score. (a–f) The power of the IR score comes from chemical changes in the tumor microenvironment, as shown for six cases. The three most important changes are labeled.

5. CONCLUDING REMARKS

Technological advances and innovations, in both instrumentation and informatics tools, are driving several exciting new directions in DP. We have summarized the state of the art with selected examples that introduce the reader to this emerging area. In particular, we note that the field of pathology is moving toward more informative multimodal tissue imaging, wherein integration of computer-derived morphologic and functional tissue-based measurements is becoming a central theme. In addition to conventional decision making, these transitions in DP will lead to a greater impact in the areas of telepathology, education, and precision medicine. The present challenges extend beyond the need for progress in hardware or software to encompass engineering pipelines for robust decision making and improving patient outcomes. This progress will require the integration of engineering and pathology at an unprecedented scale.

DP has not yet been approved for primary diagnosis in the United States. The slow pace of regulatory approval is changing (158), but further impetus is needed to drive clinical adoption

(159). Pathology has been practiced with a traditional microscope for more than 100 years; understandably, mass adoption is unlikely to happen spontaneously. Lessons from radiology are pertinent: Following the introduction of digital mammography, radiologists soon gave up looking at mammographic films and made the transition to digital techniques. DP involves additional barriers, including the fact that (unlike radiology) it is not implicitly digital—that is, the tissue slide still needs to be constructed prior to digitization—but the feature sets (both morphologic and molecular) are much richer. The use of advanced computational image analytic tools will offer substantial help to pathologists in efficiently navigating and interpreting the digitized imagery. This issue will become more important as multiplexed tissue biomarkers and multimodal tissue imaging becomes more routine, requiring analytic and informatics methods to scale up to the impending digital data deluge (160). To be accessible to a practicing pathologist, these tools must be able to run on a stand-alone desktop computer. Clearly, in order for DP to be broadly adopted in the clinic, it will be critical to optimize these methods.

The rapid pace of progress in mDP and multimodal DP has been driven by advances in instrumentation, but emerging measurement techniques now need to be standardized (161) and broadly validated. These two advances can offer a wealth of new, independent information that not only allows clinicians to make more effective clinical decisions but also helps them discover underlying biological changes. Such data can also inspire research and aid in the formulation of new hypotheses. In addition to conventional clinical diagnoses, the additional information obtained with mDP and mmDP may enable more molecular-specific diagnoses and rapid intraoperative assessment (162).

Recent progress has led to a transformation in the practice of pathology—one that may rival the introduction of the microscope by Antonie van Leeuwenhoek in the 1600s. The impact and spread of this revolution, however, will hinge on mass adoption following regulatory approvals and represent a rich area of investigation for biomedical engineers and data scientists.

DISCLOSURE STATEMENT

A.M. is a scientific advisory board member/consultant and has equity in Inspirata Inc. A.M. also has equity in Elucid Bioimaging and is working with PathCore Inc. on National Institutes of Health grant 1U24CA199374-01.

ACKNOWLEDGMENTS

Research reported by R.B. in this review is currently supported by the National Institutes of Health (NIH) (grants R01CA197516, R01GM117594, UO1MH109062, R21CA190120, and R01EB009745). Research reported by A.M. in this review was supported by the National Cancer Institute of the NIH (awards 1U24CA199374-01, R21CA179327-01, and R21CA195152-01); the National Institute of Diabetes and Digestive and Kidney Diseases (award R01DK098503-02); a US Department of Defense (DOD) Prostate Cancer Synergistic Idea Development Award (PC120857); a DOD Lung Cancer Idea Development New Investigator Award (LC130463); a DOD Prostate Cancer Idea Development Award; a Case Comprehensive Cancer Center Pilot Grant; and a VeloSano Grant from the Cleveland Clinic and the Wallace H. Coulter Foundation Program in the Department of Biomedical Engineering at Case Western Reserve University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

LITERATURE CITED

1. Ghaznavi F, Evans A, Madabhushi A, Feldman M. 2013. Digital imaging in pathology: whole-slide imaging and beyond. *Annu. Rev. Pathol. Mech. Dis.* 8:331–59
2. Madabhushi A. 2009. Digital pathology image analysis: opportunities and challenges. *Imaging Med.* 1:7–10
3. Pantanowitz L, Sinard JH, Henricks WH, Fatheree LA, Carter AB, et al. 2013. Validating whole slide imaging for diagnostic purposes in pathology: guideline from the College of American Pathologists Pathology and Laboratory Quality Center. *Arch. Pathol. Lab. Med.* 137:1710–22
4. Snead DR, Tsang Y-W, Meskiri A, Kimani PK, Crossman R, et al. 2015. Validation of digital pathology imaging for primary histopathological diagnosis. *Histopathology*. In press. doi: 10.1111/his.12879
5. Farahani N, Pantanowitz L. 2015. Overview of telepathology. *Surg. Pathol. Clin.* 8:223–31
6. Pantanowitz L, Dickinson K, Evans AJ, Hassell LA, Henricks WH, et al. 2014. American Telemedicine Association clinical guidelines for telepathology. *J. Pathol. Inform.* 5:39
7. Pantanowitz L, McHugh J, Cable W, Zhao C, Parwani AV. 2015. Imaging file management to support international telepathology. *J. Pathol. Inform.* 6:17
8. Prichard JW, Davison JM, Campbell BB, Repa KA, Reese LM, et al. 2015. TissueCypherTM: a systems biology approach to anatomic pathology. *J. Pathol. Inform.* 6:48
9. Ali S, Veltri R, Epstein JI, Christudass C, Madabhushi A. 2015. Selective invocation of shape priors for deformable segmentation and morphologic classification of prostate cancer tissue microarrays. *Comput. Med. Imaging Graph.* 41:3–13
10. Ali S, Madabhushi A. 2012. An integrated region-, boundary-, shape-based active contour for multiple object overlap resolution in histological imagery. *IEEE Trans. Med. Imaging* 31:1448–60
11. Ali S, Madabhushi A. 2011. Graphical processing unit implementation of an integrated shape-based active contour: application to digital pathology. *J. Pathol. Inform.* 2:S13
12. Fatakdwala H, Xu J, Basavanahally A, Bhanot G, Ganesan S, et al. 2010. Expectation-maximization-driven geodesic active contour with overlap resolution (EMaGACOR): application to lymphocyte segmentation on breast cancer histopathology. *IEEE Trans. Biomed. Eng.* 57:1676–89
13. Monaco J, Hipp J, Lucas D, Smith S, Balis U, Madabhushi A. 2012. Image segmentation with implicit color standardization using spatially constrained expectation maximization: detection of nuclei. *Med. Image Comput. Comput. Assist. Interv.* 15:365–72
14. Veta M, van Diest PJ, Willems SM, Wang H, Madabhushi A, et al. 2015. Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Med. Image Anal.* 20:237–48
15. Wang AH, Cruz A, Basavanahally A, Gilmore H, Shih N, et al. 2014. Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. *J. Med. Imaging* 1:034003
16. Basavanahally AN, Ganesan S, Agner S, Monaco JP, Feldman MD, et al. 2010. Computerized image-based detection and grading of lymphocytic infiltration in *HER2*⁺ breast cancer histopathology. *IEEE Trans. Biomed. Eng.* 57:642–53
17. Chen J-M, Qu A-P, Wang L-W, Yuan J-P, Yang F, et al. 2015. New breast cancer prognostic factors identified by computer-aided image analysis of HE stained histopathology images. *Sci. Rep.* 5:10690
18. Chu P, Weiss L. 2014. *Modern Immunohistochemistry*. Cambridge, UK: Cambridge Univ. Press. 2nd ed.
19. Camp RL, Chung GG, Rimm DL. 2002. Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nat. Med.* 8:1323–27
20. Angelo M, Bendall SC, Finck R, Hale MB, Hitzman C, et al. 2014. Multiplexed ion beam imaging of human breast tumors. *Nat. Med.* 20:436–42
21. Xing Y, Chaudry Q, Shen C, Kong KY, Zhau HE, et al. 2007. Bioconjugated quantum dots for multiplexed and quantitative immunohistochemistry. *Nat. Protoc.* 2:1152–65
22. Evans CL, Potma EO, Puoris'haag M, Côté D, Lin CP, Xie XS. 2005. Chemical imaging of tissue in vivo with video-rate coherent anti-Stokes Raman scattering microscopy. *PNAS* 102:16807–12
23. Freudiger CW, Min W, Saar BG, Lu S, Holtom GR, et al. 2008. Label-free biomedical imaging with high sensitivity by stimulated Raman scattering microscopy. *Science* 322:1857–61

24. Nandakumar P, Kovalev A, Volkmer A. 2009. Vibrational imaging based on stimulated Raman scattering microscopy. *New J. Phys.* 11:033026
25. Yeh K, Kenkel S, Liu J-N, Bhargava R. 2015. Fast infrared chemical imaging with a quantum cascade laser. *Anal. Chem.* 87:485–93
26. Lasch P, Haensch W, Naumann D, Diem M. 2004. Imaging of colorectal adenocarcinoma using FT-IR microspectroscopy and cluster analysis. *Biochim. Biophys. Acta* 1688:176–86
27. Al-Janabi S, Huisman A, Van Diest PJ. 2012. Digital pathology: current status and future perspectives. *Histopathology* 61:1–9
28. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. 2009. Histopathological image analysis: a review. *IEEE Rev. Biomed. Eng.* 2:147–71
29. Irshad H, Veillard A, Roux L, Racoceanu D. 2014. Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential. *IEEE Rev. Biomed. Eng.* 7:97–114
30. Pantanowitz L, Valenstein PN, Evans AJ, Kaplan KJ, Pfeifer JD, et al. 2011. Review of the current state of whole slide imaging in pathology. *J. Pathol. Inform.* 2:36
31. Rocha R, Vassallo J, Soares F, Miller K, Gobbi H. 2009. Digital slides: present status of a tool for consultation, teaching, and quality control in pathology. *Pathol. Res. Pract.* 205:735–41
32. Kanade T. 1973. *Picture processing system by computer complex and recognition of human faces*. PhD thesis, Dep. Inf. Sci., Kyoto Univ., Kyoto, Jpn.
33. Nishikawa RM, Giger ML, Doi K, Vyborny CJ, Schmidt RA. 1993. Computer-aided detection of clustered microcalcifications: an improved method for grouping detected signals. *Med. Phys.* 20:1661–66
34. Prewitt JM, Mendelsohn ML. 1966. The analysis of cell images. *Ann. N.Y. Acad. Sci.* 128:1035–53
35. Jack CR Jr., Bernstein MA, Fox NC, Thompson P, Alexander G, et al. 2008. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* 27:685–91
36. Mincione G, Di Nicola M, Di Marcantonio MC, Muraro R, Piattelli A, et al. 2015. Nuclear fractal dimension in oral squamous cell carcinoma: a novel method for the evaluation of grading, staging, and survival. *J. Oral Pathol. Med.* 44:680–84
37. Khan AM, El-Daly H, Rajpoot N. 2012. RanPEC: random projections with ensemble clustering for segmentation of tumor areas in breast histology images. In *Proceedings of the 16th Conference on Medical Image Understanding and Analysis*, ed. X Xie, pp. 17–23. Swansea, UK: Dep. Comput. Sci., Univ. Swansea
38. Hipp JD, Cheng JY, Toner M, Tompkins RG, Balis UJ. 2011. Spatially invariant vector quantization: a pattern matching algorithm for multiple classes of image subject matter including pathology. *J. Pathol. Inform.* 2:13
39. Kong J, Sertel O, Shimada H, Boyer K, Saltz J, Gurcan M. 2007. Computer-aided grading of neuroblastic differentiation: multi-resolution and multi-classifier approach. In *Proceedings of the IEEE International Conference on Image Processing*, 5:525–28. Piscataway, NJ: IEEE
40. Wang Y-Y, Chang S-C, Wu L-W, Tsai S-T, Sun Y-N. 2007. A color-based approach for automated segmentation in tumor tissue classification. In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6577–80. Piscataway, NJ: IEEE
41. Kwak JT, Hewitt SM, Bhargava R. 2011. Multimodal microscopy for automated histologic analysis of prostate cancer. *BMC Cancer* 11:62
42. Guo Y, Xu X, Wang Y, Wang Y, Xia S, Yang Z. 2014. An image processing pipeline to detect and segment nuclei in muscle fiber microscopic images. *Microsc. Res. Tech.* 77:547–59
43. Gurcan MN, Madabhushi A, Rajpoot N. 2010. Pattern recognition in histopathological images: an ICPR 2010 contest. In *Recognizing Patterns in Signals, Speech, Images and Videos*, ed. D Ünay, Z Çataltepe, S Aksoy, pp. 226–34. Berlin: Springer
44. Basavanahally A, Ganesan S, Feldman M, Shih N, Mies C, et al. 2013. Multi-field-of-view framework for distinguishing tumor grade in ER⁺ breast cancer from entire histopathology slides. *IEEE Trans. Biomed. Eng.* 60:2089–99
45. Doyle S, Feldman MD, Shih N, Tomaszewski J, Madabhushi A. 2012. Cascaded discrimination of normal, abnormal, and confounder classes in histopathology: Gleason grading of prostate cancer. *BMC Bioinform.* 13:282

46. Sparks R, Madabhushi A. 2013. Statistical shape model for manifold regularization: Gleason grading of prostate histology. *Comput. Vis. Image Underst.* 117:1138–46
47. Sparks R, Madabhushi A. 2013. Explicit shape descriptors: novel morphologic features for histopathology classification. *Med. Image Anal.* 17:997–1009
48. Kass M, Witkin A, Terzopoulos D. 1988. Snakes: active contour models. *Int. J. Comput. Vis.* 1:321–31
49. Osher S, Sethian JA. 1988. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton–Jacobi formulations. *J. Comput. Phys.* 79:12–49
50. van Ginneken B, Frangi AF, Staal JJ, ter Haar Romeny BM, Viergever MA. 2002. Active shape model segmentation with optimal features. *IEEE Trans. Med. Imaging* 21:924–33
51. Cireşan DC, Giusti A, Gambardella LM, Schmidhuber J. 2013. Mitosis detection in breast cancer histology images with deep neural networks. *Med. Image Comput. Comput. Assist. Interv.* 16:411–18
52. Cruz-Roa A, Basavanthally A, González F, Gilmore H, Feldman M, et al. 2014. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. *Proc. SPIE* 9041:904103
53. Bilgin C, Demir C, Nagi C, Yener B. 2007. Cell-graph mining for breast tissue modeling and classification. In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5311–14. Piscataway, NJ: IEEE
54. Marcelpoil R. 1993. Normalization of the minimum spanning tree. *Anal. Cell Patol.* 5:177–86
55. Ali S, Lewis J, Madabhushi A. 2013. Spatially aware cell cluster (spACC1) graphs: predicting outcome in oropharyngeal p16⁺ tumors. *Med. Image Comput. Comput. Assist. Interv.* 16:412–19
56. Ali S, Veltri R, Epstein JA, Christudass C, Madabhushi A. 2013. Cell cluster graph for prediction of biochemical recurrence in prostate cancer patients from tissue microarrays. *Proc. SPIE* 8676:H86760
57. Lewis JS Jr., Ali S, Luo J, Thorstad WL, Madabhushi A. 2014. A quantitative histomorphometric classifier (QuHbIC) identifies aggressive versus indolent p16-positive oropharyngeal squamous cell carcinoma. *Am. J. Surg. Patol.* 38:128–37
58. Lee G, Sparks R, Ali S, Shih NNC, Feldman MD, et al. 2014. Co-occurring gland angularity in localized subgraphs: predicting biochemical recurrence in intermediate-risk prostate cancer patients. *PLOS ONE* 9:e97954
59. Yu E, Monaco JP, Tomaszewski J, Shih N, Feldman M, Madabhushi A. 2011. Detection of prostate cancer on histopathology using color fractals and probabilistic pairwise Markov models. In *Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3427–30. Piscataway, NJ: IEEE
60. Lee G, Ali S, Veltri R, Epstein JJ, Christudass C, Madabhushi A. 2013. Cell orientation entropy (CoRE): predicting biochemical recurrence from prostate cancer tissue microarrays. *Med. Image Comput. Comput. Assist. Interv.* 16:396–403
61. Jain AK, Farrokhnia F. 1991. Unsupervised texture segmentation using Gabor filters. *Pattern Recognit.* 24:1167–86
62. Ojala T, Pietikainen M, Harwood D. 1994. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In *Proceedings of the 12th LAPR International Conference on Pattern Recognition*, 1:582–85. Los Alamitos, CA: IEEE Comput. Soc. Press
63. Laws KI. 1980. *Textured image segmentation*. Univ. South. Calif. Image Proc. Inst. (USCIPI) rep. 940, Dep. Electr. Eng., USCIPI, Los Angeles
64. Beck AH, Sangoi AR, Leung S, Marinelli RJ, Nielsen TO, et al. 2011. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Sci. Transl. Med.* 3:108ra113
65. Gertych A, Ing N, Ma Z, Fuchs TJ, Salman S, et al. 2015. Machine learning approaches to analyze histological images of tissues from radical prostatectomies. *Comput. Med. Imaging Graph.* 46:197–208
66. Doyle S, Feldman M, Tomaszewski J, Madabhushi A. 2012. A boosted Bayesian multiresolution classifier for prostate cancer detection from digitized needle biopsies. *IEEE Trans. Biomed. Eng.* 59:1205–18
67. Bellman R, 1957. *Dynamic Programming*. Princeton, NJ: Princeton Univ. Press
68. Lee G, Rodriguez C, Madabhushi A. 2008. Investigating the efficacy of nonlinear dimensionality reduction schemes in classifying gene and protein expression studies. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 5:368–84

69. Ginsburg S, Lee G, Ali S, Madabhushi A. 2015. Feature importance in nonlinear embeddings (FINE): applications in digital pathology. *IEEE Trans. Med. Imaging* 99:1
70. Paik S, Shak S, Tang G, Kim C, Baker J, et al. 2004. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* 351:2817–26
71. Ali S, Lewis JS, Madabhushi A. 2012. Use of quantitative histomorphometrics to classify disease progression in HPV-positive squamous cell carcinoma. *J. Clin. Oncol.* 30:73 (abstr.)
72. Boyce BF. 2015. Whole slide imaging: uses and limitations for surgical pathology and teaching. *Biotech. Histochem.* 90:321–30
73. Hartman DJ. 2015. Mobile technologies for the surgical pathologist. *Surg. Pathol. Clin.* 8:233–38
74. Jones NC, Nazarian RM, Duncan LM, Kamionek M, Lauwers GY, et al. 2015. Interinstitutional whole slide imaging teleconsultation service development: assessment using internal training and clinical consultation cases. *Arch. Pathol. Lab. Med.* 139:627–35
75. Desai S, Patil R, Kothari A, Shet T, Kane S, et al. 2004. Static telepathology consultation service between Tata Memorial Centre, Mumbai and Nargis Dutt Memorial Charitable Hospital, Barshi, Solapur, Maharashtra: an analysis of the first 100 cases. *Indian J. Pathol. Microbiol.* 47:480–85
76. Leifer Z. 2015. The use of virtual microscopy and a wiki in pathology education: tracking student use, involvement, and response. *J. Pathol. Inform.* 6:30
77. Van Es SL, Kumar RK, Pryor WM, Salisbury EL, Velan GM. 2015. Cytopathology whole slide images and adaptive tutorials for postgraduate pathology trainees: a randomized crossover trial. *Hum. Pathol.* 46:1297–305
78. Kaplan KJ. 2012. PathXchange case of the week. *Digital Pathology Blog*, June 20. <http://tissuepathology.com/2012/06/20/pathxchange-case-of-the-week-1#axzz3pyCcAMuX>
79. Sağol Ö, Yörükoğlu K, Lebe B, Durak MG, Ulukuş Ç, et al. 2015. Transition to virtual microscopy in medical undergraduate pathology education: first experience of Turkey in Dokuz Eylül University Hospital. *Türk Patol. Derg.* 31:175–80
80. Engelberg JA, Retallack H, Balassanian R, Dowsett M, Zabaglo L, et al. 2015. “Score the Core” web-based pathologist training tool improves the accuracy of breast cancer IHC4 scoring. *Hum. Pathol.* 46:1694–704
81. Walkowski S, Lundin M, Szymas J, Lundin J. 2014. Students’ performance during practical examination on whole slide images using view path tracking. *Diagn. Pathol.* 9:208
82. Bhargava R. 2012. Infrared spectroscopic imaging: the next generation. *Appl. Spectrosc.* 66:1091–120
83. Diem M, Mazur A, Lenau K, Schubert J, Bird B, et al. 2013. Molecular pathology via IR and Raman spectral imaging. *J. Biophotonics* 6:855–86
84. Levin IW, Bhargava R. 2005. Fourier transform infrared vibrational spectroscopic imaging: integrating microscopy and molecular recognition. *Annu. Rev. Phys. Chem.* 56:429–74
85. Bhargava R, Fernandez DC, Hewitt SM, Levin IW. 2006. High throughput assessment of cells and tissues: Bayesian classification of spectral metrics from infrared vibrational spectroscopic imaging data. *Bioclim. Biophys. Acta* 1758:830–45
86. Kodali AK, Schulmerich M, Ip J, Yen G, Cunningham BT, Bhargava R. 2010. Narrowband midinfrared reflectance filters using guided mode resonance. *Anal. Chem.* 82:5697–706
87. Liu J-N, Schulmerich MV, Bhargava R, Cunningham BT. 2011. Optimally designed narrowband guided-mode resonance reflectance filters for mid-infrared spectroscopy. *Opt. Express* 19:24182–97
88. Geiger FB, Koerdell M, Schick A, Heimann A, Matiassek K, Herkommer AM. 2015. Concept and setup for intraoperative imaging of tumorous tissue via attenuated total reflection spectroscopy with quantum cascade lasers. *Proc. SPIE* 9412:F94125
89. Chan KLA, Kazarian SG. 2007. Chemical imaging of the stratum corneum under controlled humidity with the attenuated total reflection Fourier transform infrared spectroscopy method. *J. Biomed. Opt.* 12:044010
90. Kazarian SG, Chan KLA. 2006. Applications of ATR-FTIR spectroscopic imaging to biomedical samples. *Bioclim. Biophys. Acta* 1758:858–67
91. Marcott C, Lo M, Kjoller K, Domanov Y, Balooch G, Luengo GS. 2013. Nanoscale infrared (IR) spectroscopy and imaging of structural lipids in human stratum corneum using an atomic force microscope to directly detect absorbed light from a tunable IR laser source. *Exp. Dermatol.* 22:419–21

92. Centrone A. 2015. Infrared imaging and spectroscopy beyond the diffraction limit. *Annu. Rev. Anal. Chem.* 8:101–26
93. Nasse MJ, Walsh MJ, Mattson EC, Reininger R, Kajdacsy-Balla A, et al. 2011. High-resolution Fourier-transform infrared chemical imaging with multiple synchrotron beams. *Nat. Methods* 8:413–16
94. Reddy RK, Walsh MJ, Schulmerich MV, Carney PS, Bhargava R. 2013. High-definition infrared spectroscopic imaging. *Appl. Spectrosc.* 67:93–105
95. van Dijk T, Mayerich D, Bhargava R, Carney PS. 2013. Rapid spectral-domain localization. *Opt. Express* 21:12822–30
96. van Dijk T, Mayerich D, Carney PS, Bhargava R. 2016. Understanding the optics-sample interaction in infrared spectroscopic imaging. *Analyst*. In press
97. Reddy RK, Bhargava R. 2010. Accurate histopathology from low signal-to-noise ratio spectroscopic imaging data. *Analyst* 135:2818–25
98. Phillips MC, Hô N. 2008. Infrared hyperspectral imaging using a broadly tunable external cavity quantum cascade laser and microbolometer focal plane array. *Opt. Express* 16:1836–45
99. Serrano AL, Ghosh A, Ostrander JS, Zanni MT. 2015. Wide-field FTIR microscopy using mid-IR pulse shaping. *Opt. Express* 23:17815–27
100. Capasso F. 2010. High-performance midinfrared quantum cascade lasers. *Opt. Eng.* 49:111102
101. Faist J, Capasso F, Sivco DL, Sirtori C, Hutchinson AL, Cho AY. 1994. Quantum cascade laser. *Science* 264:553–56
102. Luo G, Peng C, Le H, Pei S, Hwang W-Y, et al. 2001. Grating-tuned external-cavity quantum-cascade semiconductor lasers. *Appl. Phys. Lett.* 78:2834–36
103. Leslie LS, Wrobel TP, Mayerich D, Bindra S, Emmadi R, Bhargava R. 2015. High definition infrared spectroscopic imaging for lymph node histopathology. *PLOS ONE* 10:e0127238
104. Sreedhar H, Varma VK, Nguyen PL, Davidson B, Akkina S, et al. 2015. High-definition Fourier transform infrared (FT-IR) spectroscopic imaging of human tissue sections towards improving pathology. *J. Vis. Exp.* 95:52332
105. Kole MR, Reddy RK, Schulmerich MV, Gelber MK, Bhargava R. 2012. Discrete frequency infrared microspectroscopy and imaging with a tunable quantum cascade laser. *Anal. Chem.* 84:10366–72
106. Bassan P, Weida MJ, Rowlette J, Gardner P. 2014. Large scale infrared imaging of tissue micro arrays (TMAs) using a tunable quantum cascade laser (QCL) based microscope. *Analyst* 139:3856–59
107. Kröger N, Egl A, Engel M, Gretz N, Haase K, et al. 2014. Quantum cascade laser-based hyperspectral imaging of biological tissue. *J. Biomed. Opt.* 19:111607
108. Mohlenhoff B, Romeo M, Diem M, Wood BR. 2005. Mie-type scattering and non-Beer-Lambert absorption behavior of human cells in infrared microspectroscopy. *Biophys. J.* 88:3635–40
109. Kwak JT, Reddy R, Sinha S, Bhargava R. 2012. Analysis of variance in spectroscopic imaging data from human tissues. *Anal. Chem.* 84:1063–69
110. Nallala J, Piot O, Diebold M-D, Gobinet C, Bouché O, et al. 2013. Infrared imaging as a cancer diagnostic tool: introducing a new concept of spectral barcodes for identifying molecular changes in colon tumors. *Cytometry A* 83:294–300
111. Bassan P, Kohler A, Martens H, Lee J, Byrne HJ, et al. 2010. Resonant Mie scattering (RMieS) correction of infrared spectra from highly scattering biological samples. *Analyst* 135:268–77
112. Bassan P, Kohler A, Martens H, Lee J, Jackson E, et al. 2010. RMieS-EMSC correction for infrared spectra of biological cells: extension using full Mie theory and GPU computing. *J. Biophotonics* 3:609–20
113. van Dijk T, Mayerich D, Carney PS, Bhargava R. 2013. Recovery of absorption spectra from Fourier transform infrared (FT-IR) microspectroscopic measurements of intact spheres. *Appl. Spectrosc.* 67:546–52
114. Holton SE, Bergamaschi A, Katzenellenbogen BS, Bhargava R. 2014. Integration of molecular profiling and chemical imaging to elucidate fibroblast-microenvironment impact on cancer cell phenotype and endocrine resistance in breast cancer. *PLOS ONE* 9:e96878
115. Holton SE, Walsh MJ, Kajdacsy-Balla A, Bhargava R. 2011. Label-free characterization of cancer-activated fibroblasts using infrared spectroscopic imaging. *Biophys. J.* 101:1513–21
116. Kumar S, Shabi TS, Goormaghtigh E. 2014. A FTIR imaging characterization of fibroblasts stimulated by various breast cancer cell lines. *PLOS ONE* 9:e111137

117. Wald N, Goormaghtigh E. 2015. Infrared imaging of primary melanomas reveals hints of regional and distant metastases. *Analyst* 140:2144–55
118. Baker MJ, Trevisan J, Bassan P, Bhargava R, Butler HJ, et al. 2014. Using Fourier transform IR spectroscopy to analyze biological materials. *Nat. Protoc.* 9:1771–91
119. Tiwari S, Bhargava R. 2015. Extracting knowledge from chemical imaging data using computational algorithms for digital cancer diagnosis. *Yale J. Biol. Med.* 88:131–43
120. Brady SP, Do MN, Bhargava R. 2009. Reconstructing FT-IR spectroscopic imaging data with a sparse prior. In *Proceedings of the 16th International IEEE Conference on Image Processing*, pp. 829–32. Piscataway, NJ: IEEE
121. Mayerich D, RB Walsh MJ, Schulmerich MV. 2013. Real-time interactive data mining for chemical imaging information: application to automated histopathology. *BMC Bioinform.* 14:156
122. Stack EC, Wang C, Roman KA, Hoyt CC. 2014. Multiplexed immunohistochemistry, imaging, and quantitation: a review, with an assessment of Tyramide signal amplification, multispectral imaging and multiplex analysis. *Methods* 70:46–58
123. Sweeney E, Ward TH, Gray N, Womack C, Jayson G, et al. 2008. Quantitative multiplexed quantum dot immunohistochemistry. *Biochem. Biophys. Res. Commun.* 374:181–86
124. Teverovskiy M, Vengrenyuk Y, Tabesh A, Sapir M, Fogarasi S, et al. 2008. Automated localization and quantification of protein multiplexes via multispectral fluorescence imaging. In *Proceedings of the 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 300–3. Piscataway, NJ: IEEE
125. Diem M. 2015. *Modern Vibrational Spectroscopy and Micro-Spectroscopy: Theory, Instrumentation and Biomedical Applications*. New York: Wiley
126. Hughes C, Baker MJ. 2016. Can mid-infrared biomedical spectroscopy of cells, fluids and tissue aid improvements in cancer survival? A patient paradigm. *Analyst* 141:467–75
127. Keith FN, Reddy RK, Bhargava R. 2016. Development of a practical spatial-spectral analysis protocol for breast histopathology using Fourier transform infrared spectroscopic imaging. *Faraday Discuss.* doi: 10.1039/C5FD00199D. 27 pp.
128. Akalin A, Mu X, Kon MA, Ergin A, Remiszewski SH, et al. 2015. Corrigendum: Classification of malignant and benign tumors of the lung by infrared spectral histopathology (SHP). *Lab. Investig.* 95:697
129. Kwak JT, Kajdacsy-Balla A, Macias V, Walsh M, Sinha S, Bhargava R. 2015. Improving prediction of prostate cancer recurrence using chemical imaging. *Sci. Rep.* 5:8758
130. Beleites C, Salzer R. 2008. Assessing and improving the stability of chemometric models in small sample size situations. *Anal. Bioanal. Chem.* 390:1261–71
131. Bhargava R. 2007. Towards a practical Fourier transform infrared chemical imaging protocol for cancer histopathology. *Anal. Bioanal. Chem.* 389:1155–69
132. Fernandez DC, Bhargava R, Hewitt SM, Levin IW. 2005. Infrared spectroscopic imaging for histopathologic recognition. *Nat. Biotechnol.* 23:469–74
133. Petibois C, Délérès G. 2006. Chemical mapping of tumor progression by FT-IR imaging: towards molecular histopathology. *Trends Biotechnol.* 24:455–62
134. Kumar S, Desmedt C, Larsimont D, Sotiriou C, Goormaghtigh E. 2013. Change in the microenvironment of breast cancer studied by FTIR imaging. *Analyst* 138:4058–65
135. Travo A, Piot O, Wolthuis R, Gobinet C, Manfait M, et al. 2010. IR spectral imaging of secreted mucus: a promising new tool for the histopathological recognition of human colonic adenocarcinomas. *Histopathology* 56:921–31
136. Hughes C, Iqbal-Wahid J, Brown M, Shanks JH, Eustace A, et al. 2013. FTIR microspectroscopy of selected rare diverse sub-variants of carcinoma of the urinary bladder. *J. Biophotonics* 6:73–87
137. Kong R, Reddy RK, Bhargava R. 2010. Characterization of tumor progression in engineered tissue using infrared spectroscopic imaging. *Analyst* 135:1569–78
138. Kumar V, Abbas AK, Fausto N, Aster JC. 2014. *Robbins and Cotran Pathologic Basis of Disease, Professional Edition*. Philadelphia: Saunders Elsevier. 8th ed.
139. Bratthauer GL, Moynier F, Stamatakis MD, Mezzetti TP, Shekitka KM, et al. 2002. Combined E-cadherin and high molecular weight cytokeratin immunoprofile differentiates lobular, ductal, and hybrid mammary intraepithelial neoplasias. *Hum. Pathol.* 33:620–27

140. Mayerich D, Walsh MJ, Kadjacsy-Balla A, Ray PS, Hewitt SM, Bhargava R. 2015. Stain-less staining for computed histopathology. *Technology* 3:27–31
141. Boutros PC, Fraser M, Harding NJ, de Borja R, Trudel D, et al. 2015. Spatial genomic heterogeneity within localized, multifocal prostate cancer. *Nat. Genet.* 47:736–45
142. Gallee MPW, Ten Kate FJW, Mulder PGH, Blom JHM, Van der Heul RO. 1990. Histological grading of prostatic carcinoma in prostatectomy specimens. Comparison of prognostic accuracy of five grading systems. *Br. J. Urol.* 65:368–75
143. Albertsen PC. 2008. Predicting survival for men with clinically localized prostate cancer: What do we need in contemporary practice? *Cancer* 112:1–3
144. Albertsen PC. 2010. Treatment of localized prostate cancer: When is active surveillance appropriate? *Nat. Rev. Clin. Oncol.* 7:394–400
145. Albertsen PC, Hanley JA, Fine J. 2005. 20-year outcomes following conservative management of clinically localized prostate cancer. *JAMA* 293:2095–101
146. Albertsen PC, Hanley JA, Penson DF, Barrows G, Fine J. 2007. 13-year outcomes following treatment for clinically localized prostate cancer in a population based cohort. *J. Urol.* 177:932–36
147. Epstein JI. 2010. An update of the Gleason grading system. *J. Urol.* 183:433–40
148. Fine SW, Epstein JI. 2008. A contemporary study correlating prostate needle biopsy and radical prostatectomy Gleason score. *J. Urol.* 179:1335–38
149. Singanamalli A, Wang H, Lee G, Shih N, Ziober A, et al. 2014. Supervised multi-view canonical correlation analysis: fused multimodal prediction of disease prognosis. *Proc. SPIE* 9038:903805
150. Viswanath S, Madabhushi A. 2012. Consensus embedding: theory, algorithms and application to segmentation and classification of biomedical data. *BMC Bioinform.* 13:26
151. Golugula A, Lee G, Master SR, Feldman MD, Tomaszewski JE, et al. 2011. Supervised regularized canonical correlation analysis: integrating histologic and proteomic measurements for predicting biochemical recurrence following prostate surgery. *BMC Bioinform.* 12:483
152. Lee G, Singanamalli A, Wang H, Feldman M, Master S, et al. 2014. Supervised multi-view canonical correlation analysis (SMVCCA): integrating histologic and proteomic features for predicting recurrent prostate cancer. *IEEE Trans. Med. Imaging* 34:284–97
153. Tiwari P, Kurhanewicz J, Madabhushi A. 2013. Multi-kernel graph embedding for detection, Gleason grading of prostate cancer via MRI/MRS. *Med. Image Anal.* 17:219–35
154. Chang H, Han J, Borowsky A, Loss L, Gray JW, et al. 2013. Invariant delineation of nuclear architecture in glioblastoma multiforme for clinical and molecular association. *IEEE Trans. Med. Imaging* 32:670–82
155. Großerueschkamp F, Kallenbach-Thieltges A, Behrens T, Brüning T, Altmayer M, et al. 2015. Marker-free automated histopathological annotation of lung tumour subtypes by FTIR imaging. *Analyst* 140:2114–20
156. Kallenbach-Thieltges A, Großerueschkamp F, Mosig A, Diem M, Tannapfel A, Gerwert K. 2013. Immunohistochemistry, histopathology and infrared spectral histopathology of colon cancer tissue sections. *J. Biophotonics* 6:88–100
157. Gawlik A, Lee G, Whitney J, Epstein J, Veltri R, Madabhushi A. 2015. *Computer extracted nuclear features from Feulgen and H&E images predict prostate cancer outcomes*. Presented at Annu. Meet. Biomed. Eng. Soc., Tampa, Oct. 7–10
158. US Food Drug Admin. 2015. *Draft guidance for industry and Food and Drug Administration staff: technical performance assessment of digital pathology whole slide imaging devices*. Draft rep., US Food Drug Admin., Rockville, MD
159. Bellis M, Metias S, Naugler C, Pollett A, Jothy S, Yousef GM. 2013. Digital pathology: attitudes and practices in the Canadian pathology community. *J. Pathol. Inform.* 4:3
160. Webster JD, Dunstan RW. 2014. Whole-slide imaging and automated image analysis: considerations and opportunities in the practice of pathology. *Vet. Pathol.* 51:211–23
161. Ergin A, Großerueschkamp F, Theisen O, Gerwert K, Remiszewski S, et al. 2015. A method for the comparison of multi-platform spectral histopathology (SHP) data sets. *Analyst* 140:2465–72
162. Steiner G, Kirsch M. 2014. Optical spectroscopic methods for intraoperative diagnosis. *Anal. Bioanal. Chem.* 406:21–25