



ELSEVIER

The American Journal of  
**PATHOLOGY**

ajp.amjpathol.org

**Artificial Intelligence and Deep Learning in Pathology Theme Issue****MINI-REVIEW****Machine-Learning—Based Evaluation of Intratumoral Heterogeneity and Tumor-Stroma Interface for Clinical Guidance**Arvydas Laurinavicius,<sup>\*†</sup> Allan Rasmuson,<sup>\*†</sup> Benoît Plancoulaine,<sup>\*‡</sup> Michael Shribak,<sup>\*§</sup> and Richard Levenson<sup>\*¶</sup>

From the Department of Pathology, Forensic Medicine and Pharmacology,\* Institute of Biomedical Sciences of the Faculty of Medicine of Vilnius University, Vilnius, Lithuania; the National Center of Pathology,<sup>†</sup> Affiliate of Vilnius University Hospital Santaros Klinikos, Vilnius, Lithuania; the ANTICIPE,<sup>‡</sup> Inserm (UMR 1086), Cancer Center F. Baclesse, Normandy University, Caen, France; the Marine Biological Laboratory of University of Chicago,<sup>§</sup> Woods Hole, Massachusetts; and the Department of Pathology and Laboratory Medicine,<sup>¶</sup> University of California Davis Health, Sacramento, California

Accepted for publication  
April 15, 2021.

Address correspondence to  
Arvydas Laurinavicius, M.D.,  
Ph.D., National Center of Pa-  
thology, Affiliate of Vilnius  
University Hospital Santaros  
Klinikos, P. Baublio Str. 5,  
LT-08406 Vilnius, Lith-  
uania. E-mail: [arvydas.laurinavicius@vpc.lt](mailto:arvydas.laurinavicius@vpc.lt).

Assessment of intratumoral heterogeneity and tumor-host interaction within the tumor microenvironment is becoming increasingly important for innovative cancer therapy decisions because of the unique information it can generate about the state of the disease. However, its assessment and quantification are limited by ambiguous definitions of the tumor-host interface and by human cognitive capacity in current pathology practice. Advances in machine learning and artificial intelligence have opened the field of digital pathology to novel tissue image analytics and feature extraction for generation of high-capacity computational disease management models. A particular benefit is expected from machine-learning applications that can perform extraction and quantification of subvisual features of both intratumoral heterogeneity and tumor microenvironment aspects. These methods generate information about cancer cell subpopulation heterogeneity, potential tumor-host interactions, and tissue microarchitecture, derived from morphologically resolved content using both explicit and implicit features. Several studies have achieved promising diagnostic, prognostic, and predictive artificial intelligence models that often outperform current clinical and pathology criteria. However, further effort is needed for clinical adoption of such methods through development of standardizable high-capacity workflows and proper validation studies. (*Am J Pathol* 2021, ■: 1–8; <https://doi.org/10.1016/j.ajpath.2021.04.008>)

The concepts of intratumoral heterogeneity (ITH) and tumor-host interaction within the tumor microenvironment (TME) have been used in oncology and pathology to designate essential aspects of tumor biology that are becoming increasingly important for selecting personalized cancer therapies. The evaluation of ITH aims to reflect the variability and evolutionary advantage of cancer cell subpopulations in the tumor tissue ecosystem that lead to selection of the most aggressive and adaptive cell clones driving the progression of the disease. Stanta and Bonin<sup>1</sup> outlined several ITH aspects from morphologic [histologic patterns, tissue response, and tumor topography (center versus borders)] to molecular clonal (genetic and epigenetic)

and nonclonal (phenotype and functional). Although ITH of tumor cell subpopulations has been demonstrated in various molecular studies, including multiple sampling from various tumor sites,<sup>2</sup> *in situ* testing techniques are still needed to assess the ITH properties and variance within the spatial context of the TME. The reason is that cell clones' varying sensitivities to different anticancer therapies (toxic agents or

Supported by the European Social Fund grant 09.3.3-LMT-K-712.

Disclosures: None declared.

This article is part of a mini-review series on the applications of artificial intelligence and deep learning in advancing research and diagnosis in pathology.

targeting molecular pathways) are major contributors to variable therapy efficacy and pose challenges for many targeted therapy modes.<sup>3</sup>

The concept of ITH is evolving from an initial focus on tumor cell properties to considering a more complex situation reflecting tumor-host interactions and stroma physicochemical properties within the TME.<sup>4</sup> In part, this is related to an increasing appreciation of the centrality of the anti-tumor response of the host and its relevance for novel immunotherapy modes. For example, different populations of tumor-infiltrating lymphocytes (TILs) and other inflammatory cells within the TME reflect differences in tumor progression. This has led to the concept of immune contexture.<sup>5</sup> Furthermore, research attention is now directed toward the extracellular matrix per se, including tumor-associated collagen architecture,<sup>6</sup> although this has not yet been translated into clinical practice because of technical and affordability challenges.

Recent progress in digital pathology, with its burst of computational power for machine-learning methods, brings opportunities to discover and quantify morphologically resolved features that can be subvisual or even invisible by routine microscopy assessment. A wide spectrum of computer-aided methods, ranging from hand-crafted feature engineering to deep-learning neural networks, have been proposed for diverse tasks in tumor pathology.<sup>7</sup> The field of machine learning and artificial intelligence encompasses complex and sometimes overlapping definitions that makes it difficult to categorize the methods involved. Broadly speaking, there are two general approaches: explicit extraction and exploitation of hand-crafted morphology features; and application of implicit computationally extracted, often multiresolution image primitives. Both of these generate input variables that can be further analyzed by machine-learning tools or more classic statistical engines to generate desired outputs. Common to all the methods is that they extract some features from the image content, and subsequently use these features to complete higher-level tasks. With supervised methods, both the task and the validation data originate from a humanly defined model, whereas unsupervised methods aim to differentiate input elements without linkage to a concrete hypothesis and do so based on implicit features residing in some computer model that is not instantly understood by humans.

The most promising benefits seem to come from artificial intelligence methods that enable quantification of subvisual features of a tumor phenotype in the spatial context of the tissue, which is an imperative task for both ITH and tumor-host interaction studies. Often, there is a considerable overlap between the methods and tasks in these rapidly emerging applications, which, for this review of the latest developments, have been categorized into the following categories: cancer cell ITH; tumor-host interaction; tumor-stroma architecture; and data-driven deep-learning models that can embrace multiple aspects of tumor tissue pathology. [Figure 1](#) illustrates main steps and

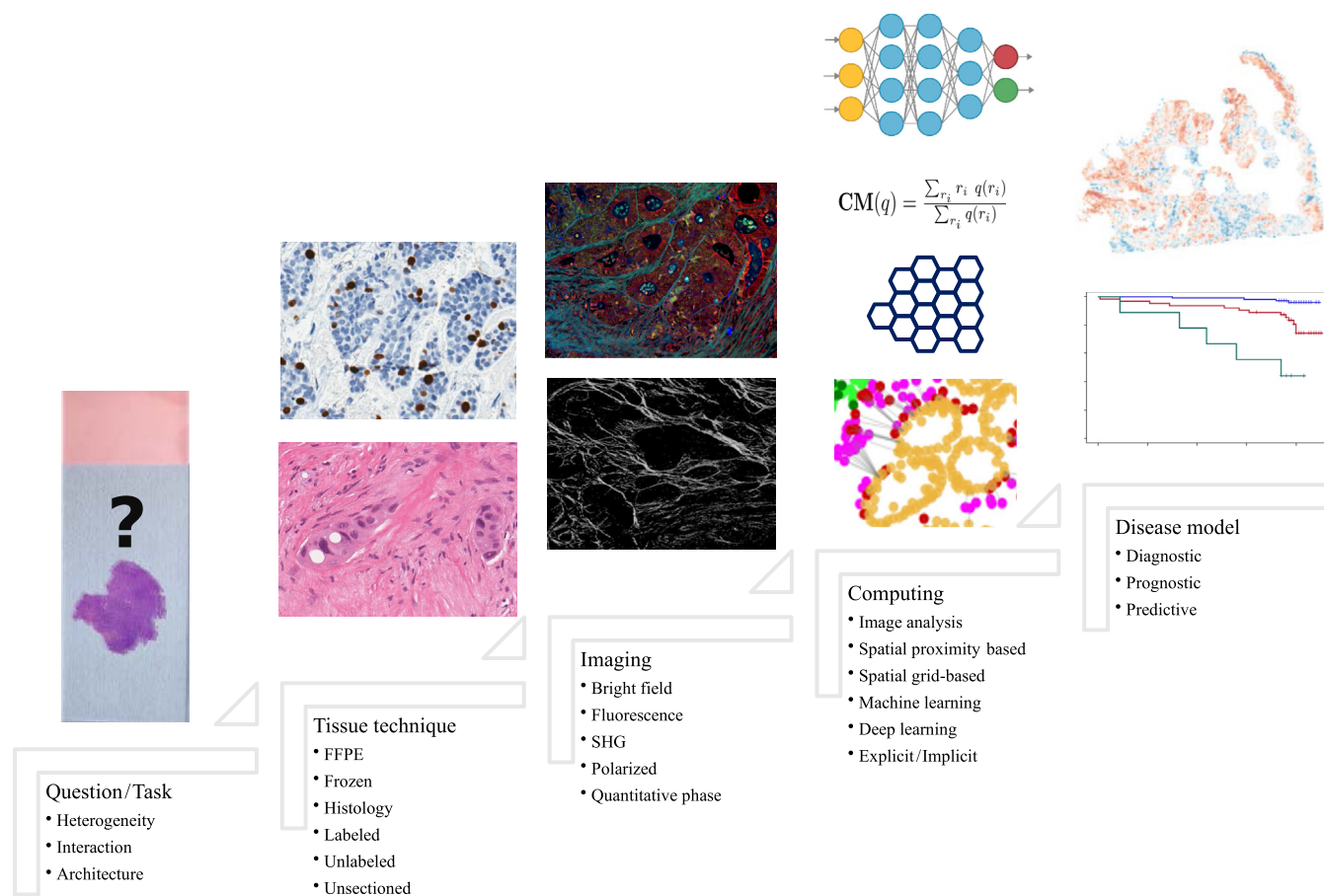
concepts from formulation of hypothesis to final disease model.

## Cancer Cell Intratumoral Heterogeneity

ITH of cancer cell subpopulations became a critical issue for tissue pathology with the introduction of target therapies, first introduced for breast cancer patients. The intratumoral distributions of hormone receptor–, human epidermal growth factor receptor 2 (HER2)–, and Ki-67–positive cancer cells, assessed by immunohistochemistry (IHC), have been characterized by many studies and visual scoring schemes. An illustrative example is set by the efforts to assess proliferative activity of breast cancer by taking the ITH of Ki-67 IHC into account. This led to requirements for pathologists to visually detect the most proliferative areas of the tumor (hotspots) and to enumerate the proliferation rate within the hotspots by counting at least 500 cells. However, the lack of common definitions and objective criteria for a hotspot (size by area or number of cells, density, and contrast) made it difficult to agree on cutoffs for international clinical guidelines; instead of a binary cutoff, the use of continuous quantitative metrics of proliferation rates in various tumors was suggested.<sup>8</sup>

Since the early days of digital pathology, digital image analysis (DIA) studies have demonstrated benefits in terms of capacity, precision, and accuracy for IHC and other tissue biomarker assessment. More importantly, these studies also addressed the issue with ITH because DIA-generated spatial coordinates of biomarker expression within a tissue context enabled application of spatial analytics for various ITH aspects, from automated cell-proximity and density-based hotspot detection to more sophisticated spatial metrics of the ITH.<sup>9</sup> Also, heat maps of feature expressions within the tissue provide augmented visualization for decision support and quality assurance. However, the issue of agreed-on hotspot definitions remains. For example, it has been demonstrated that increasing the size of the evaluated hotspots decreased the resulting Ki-67 scores.<sup>10</sup> An international multicenter study on automated Ki-67 assessment in breast cancer acknowledged that further studies are needed to optimize hotspot sampling and measurements, potentially including clinical outcome data to determine the most predictive analytics.<sup>11</sup>

Grid-based approaches, based on square or hexagonal tiling, are widely used for spatial analyses in geography, ecology, and other areas and form the basis for stereological analysis of biological specimens. Plancoulaine et al<sup>12</sup> proposed a comprehensive quantification of ITH based on hexagonal grid analytics: cell coordinates from DIA of breast cancer Ki-67 IHC whole slide images (WSIs) were systematically subsampled into a grid. This allowed not only computation of local proliferation rates, but also the production of a co-occurrence matrix to generate Haralick texture indicators, which enable quantification of



**Figure 1** Principle steps for machine-learning applications to extract tumor microenvironment features from digital pathology images. Multiple tissue processing, imaging, and computational methods are to be combined into a standardized workflow that generates relevant disease models for specific clinical tasks. An abstraction of the methods is presented in the steps. FFPE, formalin fixed, paraffin embedded; HR, hazard ratio; SHG, second harmonic generation.

proliferation ITH in the form of spatial entropy and uniformity measures. Furthermore, the grid data enable assessment of intratumoral variance by SD, percentiles, and bimodality indicators, which provide alternative metrics to density-based hotspot sampling approaches. For example, 20% of the most proliferative tumor area (Pareto hotspot) corresponding to the upper quintile of the local proliferation rate distribution and represented by the 90<sup>th</sup> percentile (a median for the upper quintile) is one approach that allows for consistent representation of the hottest zone of a biomarker expression.

A common issue for ITH assessment methods is the availability of a ground truth (standard criterion). However, most methods seeking to detect and quantify subvisual features cannot rely on ground truth data obtained by visual assessment. For example, three pathologists independently annotating hotspots in 50 breast cancer surgical excision images had only moderate agreement; on average, only 26% of the hotspot areas coincided for all three observers (ranging from 2% to 71%).<sup>12</sup> The limitations in validation of analytic methods using visually assessed ground truth

indicate the importance of including clinical (diagnostic, prognostic, and predictive) end points to guide the development and testing of applications. Surprisingly, subsequent studies in two independent breast cancer patient cohorts revealed that intratumoral Ki-67 bimodality (heterogeneity) was an independent prognostic factor for overall patient survival, outperforming other Ki-67 indicators (mean, median, or any percentile) of proliferation rate, including those in the manually annotated hotspots.<sup>13,14</sup>

## Tumor-Host Interaction in the Immune Microenvironment

Tumor-host interactions, in particular, antitumor immune responses represented by TIL and inflammatory cells, have been investigated in many studies using both visual and computational means.<sup>15,16</sup> The phenomenon of TIL has been shown to be a prognostic feature in various tumors and is becoming of great clinical importance as a prognostic and, potentially, predictive biomarker for novel immunotherapies.

Computational assessment of TIL poses particular challenges, from TIL staining and digital image segmentation to wide spectrum of immune cell subtypes and TME compartment definitions, as recently reviewed in detail.<sup>17,18</sup>

TIL distribution assessment in the context of the TME is a common task for many of the methods proposed. Cell-to-cell proximities within the compartments of the TME may reflect the dynamics of the host responses and present a unique contribution of tissue-based testing. In consequence, the choice of the spatial analytics methods and data sampling criteria will affect the nature and granularity of information obtained. Also, as with hotspots, clear definitions of the TME compartments are needed. Broadly, the computational approaches can be categorized into those based on cell densities in various TME compartments and those measuring explicit spatial relationships between individual cells and cell types.

DIA and machine-learning applications enable high-capacity automated tools for assessing global as well as intratumoral and stromal TIL densities, and can overcome many of the limitations of visual assessment. DIA was<sup>19</sup> applied to measure CD3<sup>+</sup> and CD8<sup>+</sup> cell densities in the core tumor and at the invasive margin (IM), and combine into the immunoscore for prognostic stratification of colorectal cancer patients, outperforming the utility of conventional clinical and pathologic criteria. Other studies reported prognostic significance of high TIL density in the IM using a variety of IM definitions and widths in different tumors. For example, the IM was defined as 0.5 mm extending into the tumor core and 1.0 mm beyond the tumor for colorectal cancer<sup>20</sup>; 50  $\mu$ m within the tumor to 300  $\mu$ m outside the tumor border for head and neck squamous cell carcinoma<sup>21</sup>; a tumor region of about 400- $\mu$ m width between the tumor and the reticular dermis in primary melanoma<sup>22</sup>; and within 0 to 10  $\mu$ m of the tumor border in breast cancer.<sup>23</sup> Although it is not unreasonable to use a different width of the IM in different tumors, another source of variance between the methods is introduced by different (manual and automated) approaches to detect the tumor edge. Altogether, these studies reveal that even with variable IM definitions, the TME compartment adjacent to the tumor edge in two dimensions, subsampling the tumor-host interface, does nevertheless represent the most informative area to measure tissue antitumor response.

Recently, Rasmusson et al<sup>24</sup> have proposed a concept of interface zone with immunogradient indicators to represent directional TIL density profile across the IM. Instead of sampling the IM as a ribbon of fixed width, the tumor-stroma interface is determined using hexagonal grid processing of tumor location established by an artificial intelligence tissue classifier. The computations are a set of explicit rules to assess the probability of interfaces-ness for each tissue region (hexagon). Each hexagon also maintains its rank in the direction from stroma to tumor, which allows quantification of center of mass, representing the directional TIL density profile (immunogradient) across the interface

zone. More importantly, these statistically defined indicators add new dimensionality to the assessment of TIL densities within the IM and other TME compartments and have demonstrated strong independent prognostic value in breast and colorectal cancer patients.<sup>13,24,25</sup>

A set of indicators were reported<sup>26</sup> based on spatial distributions and proximities between lymphocytes and/or cancer cells in breast cancer hematoxylin and eosin (H&E) WSIs. The method allowed quantification of the immune abundance (lymphocyte fractions to other cells) and immune spatial scores (clusters/hotspots) formed by lymphocytes and/or cancer cells within tumor regions. Remarkably, the spatial scores—not the abundance scores—served as independent prognostic factors. A method based on TIL<sup>27</sup> was identified by deep learning from H&E pathology images from 13 cancer types. It connected regions (clusters) of TIL image patches to explore spatial patterns of local TIL distributions and established their correlations with overall survival. Interestingly, the various indexes were significant across different tumor types. These global structural patterns represent distinct immune responses that can be either specific to cancer subtypes or shared across multiple tumor types, and play a role in determining the nature of the immune responses. Gong et al<sup>28</sup> applied spatial point pattern and morphometric analyses to quantify spatial heterogeneity of CD8<sup>+</sup> cell clusters in colorectal cancer excision samples and found that the number of high-density T-cell clusters of both circular and elongated shapes were higher in patients who responded to programmed cell death protein 1 (PD-1) blockade treatment. The number of clusters in the invasive front was higher in the responders.

Multiple immunofluorescence and other multiplexed imaging techniques combined with spatial analytics enable discovery of complex interactions between immune and other cell subtypes in the TME that are evidently predictive of clinical outcomes. A *Tissue Phenomics* method<sup>29</sup> based on coregistration of immunofluorescence images of prostate cancer was proposed to aggregate the data into a rich set of multiplexed image-based biomarker candidates (*phenes*). Disease recurrence prediction based on selected phenes yielded accuracies clearly outperforming predictions based on the Gleason score. For example, a high ratio of areas of CD8<sup>+</sup> cells and CD34<sup>+</sup> microvessels in the tight tumor border correlated with a good prognosis and long-term disease-free survival. Nearchou et al<sup>30,31</sup> proposed and validated a prognostic signature integrating tumor budding (see the next chapter below), lymphocytic infiltration in proximity to the tumor buds, and CD68<sup>+</sup>/CD163<sup>+</sup> macrophage ratio. This computed score provided more clinically significant cohort stratification for stage II colorectal cancer than studies based on tumor budding, immunoscore, and pT stage.

A recent study<sup>32</sup> based on highly multiplexed tissue imaging using CODEX technology employed DNA-conjugated antibodies and iterative polymerase extension with fluorescent nucleotides. It explored 56 markers in

tissue microarrays representing the IM of colorectal cancer from 35 patients and established nine cellular neighborhoods (regions of the tissue with a specific local density of various cell types) that were conserved across patient groups with different types of lymphoid infiltrates (Crohn-like reaction versus diffuse inflammatory infiltration). In particular, it found that enrichment of PD-1<sup>+</sup>CD4<sup>+</sup> cells only within a granulocyte neighborhood positively correlated with survival in a high-risk patient subset. Another study,<sup>33</sup> using high-plex digital spatial profiling in tissue microarray samples of non-small cell lung cancer from 67 immunotherapy-treated patients uncovered 12 spatially resolved protein biomarkers independently associated with benefit from single-agent PD-1 checkpoint blockade. High expression of CD56 and CD4 in the CD45 compartment was significantly associated with favorable outcomes, whereas high levels of VISTA and CD127 in the tumor compartment were markers of immunotherapy resistance.

In summary, immune-context TME can be interrogated with an expanding spectrum of machine-learning techniques. The discovery process can be aided using high-end spatial multiplexing methods, but as these currently have somewhat limited testing capacity and are susceptible to tissue processing preanalytical variables, the knowledge obtained will help inform more affordable and robust multiple-analyte applications.

## Tumor-Stroma Architecture

Tumor growth patterns form the basis for many histologic classifications and grading systems used in pathology because of their established associations with clinical outcomes. A set of descriptive terms and definitions have been developed; however, the features are based on visual pattern recognition and are prone to interobserver variability and further obscured by the spatial variance of tumor architectures. Furthermore, pathologists have typically focused on tumor cell features (foreground) while the accompanying stroma is in most cases perceived as background (only weakly stained by eosin) with lack of consistent evaluation criteria. On the contrary, the concept of TME considers the tumor and stroma as equal players, where the spatial configuration of the interface as well as structural properties of both components may influence the tumor-host interaction.

The most broadly used concept of tumor-stroma interface in pathology is exemplified by pushing and infiltrative growth, descriptively defined by the irregularity of tumor border. A more specific feature of invasive growth, tumor budding, internationally defined in 2016 “as single tumor cells or small clusters of 4 or fewer tumor cells,”<sup>34</sup> has been carefully examined in many studies.<sup>35</sup> Tumor budding appears to signal occurrence of dissociated tumor spread and is increasingly recognized as a feature of aggressive behavior, providing independent prognostic

information for colorectal and other cancers. Regardless of the agreed on definition and assessment guidelines,<sup>34</sup> intraobserver and interobserver variability and quantification of the features make budding studies subject to the same challenges encountered by DIA and machine-learning techniques in general.<sup>35</sup> The applications were of different levels of automation, mostly based on IHC or immunofluorescence for cytokeratin (the guidelines require quantification based on H&E) and require further standardization for clinical adoption. Not surprisingly, computer-generated data help define more specific morphology and spatial characteristics of the budding, but raise new questions about tumor cell properties, TME compartment sampling criteria, and ITH of the budding. Remarkably, ITH of the tumor budding could be more informative than simple quantification of this feature, as demonstrated in a study<sup>36</sup> showing that only spatial clustering of the tumor buds in hotspots (and especially the number of hotspots), but not the absolute number of tumor buds, correlated with patient outcomes in colorectal cancer.

Interestingly, several computational pathology studies have presented combined predictive models based on both tumor growth pattern and host response indicators, that outperformed standard clinical and pathology parameters. The Spatial Immuno-Oncology Index,<sup>30,31</sup> which integrates tumor budding and lymphocytic infiltration within 50  $\mu\text{m}$  of the tumor buds with the ratio of CD68<sup>+</sup>/CD163<sup>+</sup> macrophages, was able to stratify stage II colorectal cancer and outperformed previous scores that incorporated tumor budding, immunoscore, and pT stage metrics. Farchouk et al<sup>37</sup> reported CD8<sup>+</sup> T-cell density by DIA and intratumoral budding by visual assessment in pretreatment biopsies of rectal adenocarcinoma as independent predictive biomarkers of response to neoadjuvant therapy. Nestarenkaite et al<sup>25</sup> proposed a prognostic Immuno-Interface Score for colorectal cancer survival probability based on independent contributions of CD8 and CD20 immunogradients along with the invasive growth pattern determined by pathologists' visual assessment; this approach outperformed evaluation of tumor budding alone.

The extracellular matrix forms a major component of the TME and represents a dynamic infrastructure mainly supported by fibrillar collagen. The role of collagen fiber organization for cancer progression and its potential for clinical applications have been recently reviewed by Ouellette et al.<sup>6</sup> Studies in various tumor types have shown correlation between fiber features with disease progression and patient outcomes. Although some studies successfully employed special stains, such as Masson trichrome or picrosirius red, there is an ongoing search for label-free collagen imaging methods. Second harmonic generation (SHG) imaging of unstained tissue sections is frequently deployed in these studies and is considered a gold standard collagen imaging technique; however, it is costly, is relatively slow, and requires specialized equipment. Polarization microscopy,<sup>38</sup> which can be enabled using liquid

crystal-based optics, is another label-free imaging modality for collagen. It is less expensive and more feasible for clinical implementation; however, it requires acquisition of five images to compute a birefringence map, which is not suitable for WSIs. Also, this method is insensitive to non-birefringent structures. Polychromatic polarization imaging provides snap-shot orientation-independent birefringence intensity and shows promise for further progress in label-free-imaging.<sup>39</sup>

Apart from imaging techniques, computational assessment of collagen distribution presents another challenge for clinical adoption of image-based collagen fiber biomarkers.<sup>6</sup> An abundance of explicit collagen fiber metrics (eg, density, orientation, alignment, individual fiber properties, texture-based fiber patterns, fiber network branching, and relationships of the fibers and tumor cells) have been demonstrated to provide diagnostic and prognostic value in various tumors.<sup>40</sup> However, despite the accumulated evidence of the informative value of both explicit and implicit collagen microarchitecture, clinical adoption of the models is delayed by the lack of affordable high-capacity workflows of histology processing, imaging, and computation. Recently, computational methods for extracting collagen characteristics from routine H&E sections were proposed by Keikhosravi et al,<sup>41</sup> who trained a convolutional neural network model on SHG data to synthesize SHG images from H&E images. Results were consistent with the SHG ground truth, and the signal generation method can potentially compensate for some of the limitations (orientation dependence and sensitivity) of SHG imaging itself. Fereidouni et al<sup>42</sup> proposed a combined bright-field and fluorescence microscopy method (dual emission transmission, or DUET) to generate trichrome-like images directly from H&E sections. These have been validated against SHG and shown to provide equivalent or superior data on spatially resolved and quantifiable collagen signals. Although these approaches support the potential to use standard histology techniques routinely, their clinical value remains to be investigated.

## Implicit Feature-Based Models

Deep-learning algorithms have been demonstrated as powerful tools to extract novel tumor histology features that can predict molecular biomarker expression in the TME and clinical patient outcomes. Interpretability of these deep-learning-based features is limited in most cases. Although they may take advantage of detailed tumor-cell morphology information, it is likely that they also reflect ITH of tumor cell and/or stroma properties. Rawat et al<sup>43</sup> proposed the concept of tissue fingerprints, which are the features stored in an inner layer of a deep-learning method trained to distinguish different patients using unannotated breast cancer H&E images. Subsequently, a second biomarker neural network was trained to predict estrogen receptor status based on these fingerprints. *ERBB2* amplification-associated morphology

extracted from H&E images of breast cancer also correlated with the efficacy of adjuvant trastuzumab therapy and had a favorable effect on distant disease-free survival in CISH *ERBB2*-positive patients.<sup>44</sup> Another deep-learning model exceeded the performance of experienced gastrointestinal pathologists predicting microsatellite instability from H&E-stained WSIs.<sup>45</sup> More importantly, these methods also allowed localization of the tumor areas responsible for this classification, which help in highlighting biologically and clinically relevant ITH aspects of the tumor morphology. Another aspect emerging from deep-learning studies is the discriminative power of tumor-associated stroma, previously underestimated in diagnostic pathology,<sup>46</sup> although identified previously using conventionally calculated features.<sup>47</sup>

The informative potential of deep-learning-based intratumoral phenotype heterogeneity data is further emphasized by studies aimed at prognostic stratification of patients. In particular, a machine-learning-based approach outperformed a visual prognostic assessment in breast and colorectal cancer tissue microarrays.<sup>48,49</sup> A deep-learning model for predicting postsurgical recurrence of hepatocellular carcinoma from H&E-stained WSIs of liver resections was recently developed and validated.<sup>50</sup> The model stratified patients into low- and high-risk subgroups and exceeded the performance of the standard TNM classification system.

Although these models, based on implicit features, are powerful and can help answer complex questions, a major drawback is their use of pure computerized models that do not add to the knowledge in the human models. There is a lack of understanding of *why* they work. Ongoing studies are expected to help bridge the implicit computerized models with human understanding of disease.<sup>7</sup>

## Summary

The abundance of tissue processing, imaging, and computational processing methods, multiplied by the broad spectrum of tumor types and patient management tasks, presents a vast field of opportunities and challenges for image- and artificial intelligence-derived biomarker analysis. Many studies have demonstrated strong informative value of image-based methods, often outperforming current clinical and histopathologic predictive systems. ITH and tumor-host interaction aspects are of particular interest because they are difficult to include in routine pathology assessment but appear critical for refining and applying innovative cancer therapies. For clinical adoption, standard workflows still need to be established, along with proper analytic and clinical validation steps. Because machine-learning applications, both explicit and implicit, extract quantifiable features of the pathology, the subvisual nature of the features makes analytic validation challenging because of scarcity of robust direct reference values (ground truth). Therefore, validation against clinical end points seems the

main option for implementation decisions. It is likely that the most powerful applications and computational models will integrate quantification of specific molecular (predictive) biomarkers within the rich spatial context of the tissue.

## Author Contributions

A.L., in collaboration with A.R., B.P., M.S., and R.L., drafted the manuscript. All authors critically revised and approved the final manuscript.

## References

- Stanta G, Bonin S: A practical approach to tumor heterogeneity in clinical research and diagnostics. *Pathobiology* 2018, 85:7–17
- Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, Martinez P, Matthews N, Stewart A, Tarpey P, Varela I, Phillimore B, Begum S, McDonald NQ, Butler A, Jones D, Raine K, Latimer C, Santos CR, Nohadani M, Eklund AC, Spencer-Dene B, Clark G, Pickering L, Stamp G, Gore M, Szallasi Z, Downward J, Futreal PA, Swanton C: Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 2012, 366:883–892
- Fisher R, Pusztai L, Swanton C: Cancer heterogeneity: implications for targeted therapeutics. *Br J Cancer* 2013, 108:479–485
- Ramon YCS, Sese M, Capdevila C, Aasen T, De Mattos-Arruda L, Diaz-Cano SJ, Hernandez-Losa J, Castellvi J: Clinical implications of intratumor heterogeneity: challenges and opportunities. *J Mol Med (Berl)* 2020, 98:161–177
- Fridman WH, Pages F, Sautes-Fridman C, Galon J: The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer* 2012, 12:298–306
- Ouellette JN, Drifka CR, Pointer KB, Liu Y, Lieberthal TJ, Kao WJ, Kuo JS, Loeffler AG, Eliceiri KW: Navigating the collagen jungle: the biomedical potential of fiber organization in cancer. *Bioengineering (Basel)* 2021, 8:17
- Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A: Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol* 2019, 16:703–715
- Esposito A, Criscitiello C, Curigliano G: Highlights from the 14(th) St Gallen International Breast Cancer Conference 2015 in Vienna: dealing with classification, prognostication, and prediction refinement to personalize the treatment of patients with early breast cancer. *Ecancermedalscience* 2015, 9:518
- Heindl A, Nawaz S, Yuan Y: Mapping spatial heterogeneity in the tumor microenvironment: a new era for digital pathology. *Lab Invest* 2015, 95:377–384
- Christgen M, von Ahsen S, Christgen H, Länger F, Kreipe H: The region-of-interest size impacts on Ki67 quantification by computer-assisted image analysis in breast cancer. *Hum Pathol* 2015, 46:1341–1349
- Rimm DL, Leung SCY, McShane LM, Bai Y, Bane AL, Bartlett JMS, Bayani J, Chang MC, Dean M, Denkert C, Enwere EK, Galderisi C, Gholap A, Hugh JC, Jadhav A, Kornaga EN, Laurinavicius A, Levenson R, Lima J, Miller K, Pantanowitz L, Piper T, Ruan J, Srinivasan M, Virk S, Wu Y, Yang H, Hayes DF, Nielsen TO, Dowsett M: An international multicenter study to evaluate reproducibility of automated scoring for assessment of Ki67 in breast cancer. *Mod Pathol* 2019, 32:59–69
- Plancoulaine B, Laurinaviciene A, Herlin P, Besusparis J, Meskauskas R, Baltrusaityte I, Iqbal Y, Laurinavicius A: A methodology for comprehensive breast cancer Ki67 labeling index with intra-tumor heterogeneity appraisal based on hexagonal tiling of digital image analysis data. *Virchows Arch* 2015, 467:711–722
- Zilenaite D, Rasmusson A, Augulis R, Besusparis J, Laurinaviciene A, Plancoulaine B, Ostapenko V, Laurinavicius A: Independent prognostic value of intratumoral heterogeneity and immune response features by automated digital immunohistochemistry analysis in early hormone receptor-positive breast carcinoma. *Front Oncol* 2020, 10:950
- Laurinavicius A, Plancoulaine B, Rasmusson A, Besusparis J, Augulis R, Meskauskas R, Herlin P, Laurinaviciene A, Muftah AAA, Miligy I, Aleskandarany M, Rakha EA, Green AR, Ellis IO: Bimodality of intratumor Ki67 expression is an independent prognostic factor of overall survival in patients with invasive breast carcinoma. *Virchows Arch* 2016, 468:493–502
- Hendry S, Salgado R, Gevaert T, Russell PA, John T, Thapa B, et al: Assessing tumor-infiltrating lymphocytes in solid tumors: a practical review for pathologists and proposal for a standardized method from the International Immuno-Oncology Biomarkers Working Group: part 2: TILs in melanoma, gastrointestinal tract carcinomas, non-small cell lung carcinoma and mesothelioma, endometrial and ovarian carcinomas, squamous cell carcinoma of the head and neck, genitourinary carcinomas, and primary brain tumors. *Adv Anat Pathol* 2017, 24:311–335
- Klauschen F, Muller KR, Binder A, Bockmayr M, Hagele M, Seegerer P, Wienert S, Pruneri G, de Maria S, Badve S, Michiels S, Nielsen TO, Adams S, Savas P, Symmans F, Willis S, Grusos T, Park M, Haibe-Kains B, Gallas B, Thompson AM, Cree I, Sotiriou C, Solinas C, Preusser M, Hewitt SM, Rimm D, Viale G, Loi S, Loibl S, Salgado R, Denkert C: International Immuno-Oncology Biomarker Working Group: Scoring of tumor-infiltrating lymphocytes: from visual estimation to machine learning. *Semin Cancer Biol* 2018, 52:151–157
- Amgad M, Stovgaard ES, Balslev E, Thagaard J, Chen W, Dudgeon S, et al: Report on computational assessment of tumor infiltrating lymphocytes from the International Immuno-Oncology Biomarker Working Group. *NPJ Breast Cancer* 2020, 6:16
- Kos Z, Roblin E, Kim RS, Michiels S, Gallas BD, Chen W, et al: Pitfalls in assessing stromal tumor infiltrating lymphocytes (sTILs) in breast cancer. *NPJ Breast Cancer* 2020, 6:17
- Galon J, Pages F, Marincola FM, Angell HK, Thurin M, Lugli A, et al: Cancer classification using the Immunoscore: a worldwide task force. *J Transl Med* 2012, 10:205
- Yoon HH, Shi Q, Heying EN, Muranyi A, Bredno J, Ough F, Djalilvand A, Clements J, Bowermaster R, Liu WW, Barnes M, Alberts SR, Shanmugam K, Sinicrope FA: Intertumoral heterogeneity of CD3(+) and CD8(+) T-cell densities in the microenvironment of DNA mismatch-repair-deficient colon cancers: implications for prognosis. *Clin Cancer Res* 2019, 25:125–133
- Lechner A, Schlosser H, Rothschild SI, Thelen M, Reuter S, Zentis P, Shimabukuro-Vornhagen A, Theurich S, Wennhold K, Garcia-Marquez M, Tharun L, Quaas A, Schauss A, Isensee J, Hucho T, Huebbers C, von Bergwelt-Baildon M, Beutner D: Characterization of tumor-associated T-lymphocyte subsets and immune checkpoint molecules in head and neck squamous cell carcinoma. *Oncotarget* 2017, 8:44418–44433
- Bordry N, Broggi MAS, de Jonge K, Schaeuble K, Gannon PO, Foukas PG, Danenberg E, Romano E, Baumgaertner P, Fankhauser M, Wald N, Cagnon L, Abed-Maillard S, Hajjami HM, Murray T, Ioannidou K, Letovanec I, Yan P, Michielin O, Matter M, Swartz MA, Speiser DE: Lymphatic vessel density is associated with CD8(+) T cell infiltration and immunosuppressive factors in human melanoma. *Oncoimmunology* 2018, 7:e1462878
- Berthel A, Zoernig I, Valous NA, Kahlert C, Klupp F, Ulrich A, Weitz J, Jaeger D, Halama N: Detailed resolution analysis reveals spatial T cell heterogeneity in the invasive margin of colorectal cancer liver metastases associated with improved survival. *Oncoimmunology* 2017, 6:e1286436
- Rasmusson A, Zilenaite D, Nestarenkaite A, Augulis R, Laurinaviciene A, Ostapenko V, Poskus T, Laurinavicius A:

- Immunogradient indicators for anti-tumor response assessment by automated tumor-stroma interface zone detection. *Am J Pathol* 2020, 6:1309–1322
25. Nestarenkaite A, Fadhil W, Rasmusson A, Susanti S, Hadjimichael E, Laurinaviciene A, Ilyas M, Laurinavicius A: Immuno-Interface Score to predict outcome in colorectal cancer independent of microsatellite instability status. *Cancers (Basel)* 2020, 12:2902
  26. Heindl A, Sestak I, Naidoo K, Cuzick J, Dowsett M, Yuan Y: Relevance of spatial heterogeneity of immune infiltration for predicting risk of recurrence after endocrine therapy of ER+ breast cancer. *J Natl Cancer Inst* 2018, 110
  27. Saltz J, Gupta R, Hou L, Kurc T, Singh P, Nguyen V, Samaras D, Shroyer KR, Zhao T, Batiste R, Van Arnam J, Cancer Genome Atlas Research N, Shmulevich I, Rao AUK, Lazar AJ, Sharma A, Thorsson V: Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Rep* 2018, 23:181–193.e7
  28. Gong C, Anders RA, Zhu QF, Taube JM, Green B, Cheng WT, Bartelink IH, Vicini P, Wang B, Popel AS: Quantitative characterization of CD8+ T cell clustering and spatial heterogeneity in solid tumors. *Front Oncol* 2019, 8:649
  29. Harder N, Athelougou M, Hessel H, Brieu N, Yigitsoy M, Zimmermann J, Baatz M, Buchner A, Stief CG, Kirchner T, Binnig G, Schmidt G, Huss R: Tissue phenomics for prognostic biomarker discovery in low- and intermediate-risk prostate cancer. *Sci Rep* 2018, 8:4470
  30. Nearchou IP, Lillard K, Gavriel CG, Ueno H, Harrison DJ, Caie PD: Automated analysis of lymphocytic infiltration, tumor budding, and their spatial relationship improves prognostic accuracy in colorectal cancer. *Cancer Immunol Res* 2019, 7:609–620
  31. Nearchou IP, Gwyther BM, Georgiakakis ECT, Gavriel CG, Lillard K, Kajiwaru Y, Ueno H, Harrison DJ, Caie PD: Spatial immune profiling of the colorectal tumor microenvironment predicts good outcome in stage II patients. *NPJ Digit Med* 2020, 3:71
  32. Schurch CM, Bhate SS, Barlow GL, Phillips DJ, Noti L, Zlobec I, Chu P, Black S, Demeter J, McIlwain DR, Kinoshita S, Samusik N, Goltsev Y, Nolan GP: Coordinated cellular neighborhoods orchestrate antitumoral immunity at the colorectal cancer invasive front. *Cell* 2020, 183:838
  33. Zugazagoitia J, Gupta S, Liu Y, Fuhrman K, Gettinger S, Herbst RS, Schalper KA, Rimm DL: Biomarkers associated with beneficial PD-1 checkpoint blockade in non-small cell lung cancer (NSCLC) identified using high-plex digital spatial profiling. *Clin Cancer Res* 2020, 26:4360–4368
  34. Lugli A, Kirsch R, Ajioka Y, Bosman F, Cathomas G, Dawson H, El Zimaity H, Flejou JF, Hansen TP, Hartmann A, Kakar S, Langner C, Nagtegaal I, Pappa G, Riddell R, Ristimaki A, Sheahan K, Smyrk T, Sugihara K, Terris B, Ueno H, Vieth M, Zlobec I, Quirke P: Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. *Mod Pathol* 2017, 30:1299–1311
  35. Studer L, Blank A, Bokhorst JM, Nagtegaal ID, Zlobec I, Lugli A, Fischer A, Dawson H: Taking tumour budding to the next frontier - a post International Tumour Budding Consensus Conference (ITBCC) 2016 review. *Histopathology* 2021, 78:476–484
  36. Weis CA, Kather JN, Melchers S, Al-ahmdi H, Pollheimer MJ, Langner C, Gaiser T: Automatic evaluation of tumor budding in immunohistochemically stained colorectal carcinomas and correlation to clinical outcome. *Diagn Pathol* 2018, 13:64
  37. Farchoukh L, Hartman DJ, Ma C, Celebrezze J, Medich D, Bahary N, Frank M, Pantanowitz L, Pai RK: Intratumoral budding and automated CD8-positive T-cell density in pretreatment biopsies can predict response to neoadjuvant therapy in rectal adenocarcinoma. *Mod Pathol* 2020, 34:171–183
  38. Keikhosravi A, Liu Y, Drifka C, Woo KM, Verma A, Oldenbourg R, Eliceiri KW: Quantification of collagen organization in histopathology samples using liquid crystal based polarization microscopy. *Biomed Opt Express* 2017, 8:4243–4256
  39. Shribak M: Polychromatic polarization microscope: bringing colors to a colorless world. *Sci Rep* 2015, 5:17340
  40. Xi G, Guo W, Kang D, Ma J, Fu F, Qiu L, Zheng L, He J, Fang N, Chen J, Li J, Zhuo S, Liao X, Tu H, Li L, Zhang Q, Wang C, Boppart SA, Chen J: Large-scale tumor-associated collagen signatures identify high-risk breast cancer patients. *Theranostics* 2021, 11:3229–3243
  41. Keikhosravi A, Li B, Liu Y, Conklin MW, Loeffler AG, Eliceiri KW: Non-disruptive collagen characterization in clinical histopathology using cross-modality image synthesis. *Commun Biol* 2020, 3:414
  42. Fereidouni F, Todd A, Li Y, Chang CW, Luong K, Rosenberg A, Lee YJ, Chan JW, Borowsky A, Matsukuma K, Jen KY, Levenson R: Dual-mode emission and transmission microscopy for virtual histochemistry using hematoxylin- and eosin-stained tissue sections. *Biomed Opt Express* 2019, 10:6516–6530
  43. Rawat RR, Ortega I, Roy P, Sha F, Shibata D, Ruderman D, Agus DB: Deep learned tissue "fingerprints" classify breast cancers by ER/PR/Her2 status from H&E images. *Sci Rep* 2020, 10:7275
  44. Bychkov D, Linder N, Tiulpin A, Kucukel H, Lundin M, Nordling S, Sihto H, Isola J, Lehtimäki T, Kellokumpu-Lehtinen PL, von Smitten K, Joensuu H, Lundin J: Deep learning identifies morphological features in breast cancer predictive of cancer ERBB2 status and trastuzumab treatment efficacy. *Sci Rep* 2021, 11:4037
  45. Yamashita R, Long J, Longacre T, Peng L, Berry G, Martin B, Higgins J, Rubin DL, Shen J: Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. *Lancet Oncol* 2021, 22:132–141
  46. Bejnordi BE, Lin J, Glass B, Mullooly M, Gierach GL, Sherman ME, Karssemeijer N, van der Laak J, Beck AH: Deep learning-based assessment of tumor-associated stroma for diagnosing breast cancer in histopathology images. *Proc/IEEE Int Symp Biomed Imag* 2017, 2017:929–932
  47. Beck AH, Sangoi AR, Leung S, Marinelli RJ, Nielsen TO, van de Vijver MJ, West RB, van de Rijn M, Koller D: Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Sci Transl Med* 2011, 3:108ra13
  48. Turkki R, Bychkov D, Lundin M, Isola J, Nordling S, Kovanen PE, Verrill C, von Smitten K, Joensuu H, Lundin J, Linder N: Breast cancer outcome prediction with tumour tissue images and machine learning. *Breast Cancer Res Treat* 2019, 177:41–52
  49. Bychkov D, Linder N, Turkki R, Nordling S, Kovanen PE, Verrill C, Walliander M, Lundin M, Haglund C, Lundin J: Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci Rep* 2018, 8:3395
  50. Yamashita R, Long J, Saleem A, Rubin DL, Shen J: Deep learning predicts postsurgical recurrence of hepatocellular carcinoma from digital histopathologic images. *Sci Rep* 2021, 11:2047