Precision immunoprofiling by image analysis and artificial intelligence

Viktor H. Koelzer1,2, Korsuk Sirinukunwattana3, Jens Rittscher3,4,5, Kirsten D. Mertz6

Abstract
Clinical success of immunotherapy is driving the need for new prognostic and predictive assays to inform patient selection and stratification. This requirement can be met by a combination of computational pathology and artificial intelligence. Here, we critically assess computational approaches supporting the development of a standardized methodology in the assessment of immune-oncology biomarkers, such as PD-L1 and immune cell infiltrates. We examine immunoprofiling through spatial analysis of tumor-immune cell interactions and multiplexing technologies as a predictor of patient response to cancer treatment. Further, we discuss how integrated bioinformatics can enable the amalgamation of complex morphological phenotypes with the multiomics datasets that drive precision medicine. We provide an outline to machine learning (ML) and artificial intelligence tools and illustrate fields of application in immuno-oncology, such as pattern-recognition in large and complex datasets and deep learning approaches for survival analysis. Synergies of surgical pathology and computational analyses are expected to improve patient stratification in immuno-oncology. We propose that future clinical demands will be best met by (1) dedicated research at the interface of pathology and bioinformatics, supported by professional societies, and (2) the integration of data sciences and digital image analysis in the professional education of pathologists.

Keywords
Personalized medicine · Immuno-oncology · Immunotherapy · Digital pathology · Image analysis · Machine learning · Artificial intelligence

Introduction
A picture is worth a thousand words. This is the essence of the technological transition from macropathology to micropathology as captured in Virchow’s core principle “omnis cellula et cellulae” first popularized in 1858 [81]. One hundred sixty years later, pathologists are in a technological transition phase of similar importance [61]: We are beginning to recognize that images of cells contain more information than what can be extracted by the human eye [20, 67, 83]. Computer-aided image analysis has the potential to make complex morphological information more accessible in daily diagnostic practice, improving prognostic and predictive patient stratification. Artificial intelligence has already been successfully employed in the setting of computational pathology to categorize diseases based on their molecular features [15, 21, 59]. In combination, it seems likely that image-based digital pathology in combination with artificial intelligence will become part of a pathologist’s tool repertoire in the near future.

Immuno-oncology requires a detailed understanding of the tumor microenvironment, including the identification and quantification of different immune cell subsets, their spatial context, and the expression of immune checkpoint markers. Changes in immune cell infiltration and biomarker expression before and after therapeutic intervention are critical parameters for clinical development [80]. Image analysis tools can
carry out complex and repetitive biomarker analyses with high precision and excellent reproducibility and can thus greatly assist pathologists in their key role to integrate clinical, morphologic, and molecular information for personalized treatment [20, 83, 84]. Actively engaging in image analysis methods is therefore becoming increasingly important for pathologists to maintain their role as leaders in precision medicine and diagnostics.

In recent decades, electronic data processing has transformed medical radiology, molecular diagnostics, and genetic testing. At the same time, the practice of pathology which forms a crucial link between these clinical disciplines has not changed significantly. The current convergence of new imaging technologies with tissue-based multiplexed immunohistochemistry (IHC) (reviewed in “Multiplexing”) and molecular phenotyping, our ability to digitize and process large collections of histology slides, and the promise of supporting human interpretation through automated analysis and artificial intelligence will have a dramatic impact on the field [15, 59].

While traditional medical device companies, such as Philips, GE, and Leica advance new platforms for digital pathology and commercialize new imaging technologies, the major IT companies, including Google, IBM, and Microsoft, as well as numerous start-up companies (e.g., PathAI) enter the space by applying their expertise in big data science and artificial intelligence to data analysis and integrated decision making. Pharmaceutical companies have also recognized the importance of advanced pathology to their own work. Roche has already received FDA clearance for the VENTANA MMR IHC Panel for patients diagnosed with colorectal cancer (CRC) [2]. Developing and maintaining an understanding of digital imaging and data mining will therefore be beneficial skills for pathologists practicing in the twenty-first century [67, 83].

This paper provides a comprehensive outline of image analysis and machine learning (ML) applications for precision immunoprofiling. We will demonstrate how digital tools can facilitate pathology workflows in the assessment of established immune biomarkers and enable the deep characterization of the tumor microenvironment through spatial analysis and multiplexing. Further, we will identify computational methods driving morpho-molecular integration and deep learning methodologies for the discovery of novel therapeutic targets.

Computational pathology to assist in the assessment of established biomarkers

Immuo-oncology has been revolutionized by the introduction of immune checkpoint inhibitors (ICI). ICI are monoclonal antibodies targeting immuno-regulatory molecules on the surface of T cells, antigen-presenting cells, and neoplastic cell populations [39]. Clinical success of reagents blocking the CTLA-4 (cytotoxic T lymphocyte-associated protein 4, CD152) and PD-1/PD-L1 checkpoints (programmed cell death protein 1, CD279; programmed death-ligand 1, CD274) has driven rapid regulatory approval for treatment of patients with both solid and hematologic malignancies [33]. Assessment of PD-L1 expression by IHC has emerged as an important predictive biomarker for patients with non-small cell lung cancer (NSCLC) [12, 13], urothelial carcinoma [34], and renal cell cancer [60]. However, assessment of PD-L1 is inherently difficult due to expression in both neoplastic and non-neoplastic cell populations, considerable marker heterogeneity and non-intuitive cutoffs. To complicate matters, PD-L1 positivity thresholds vary [80]. Inter-observer variability of PD-L1 assessment by pathologists is a known problem [17, 77]. This may contribute to inaccurate patient stratification and the misinterpretation of the impact of PD-L1 expression on clinical outcome: Subjective decisions can lead to radically different therapeutic stratification when scoring around cutoffs.

Digital scoring of PD-L1 can assist the pathologist to overcome these barriers by providing standardized metrics for biomarker assessment at single cell resolution across whole tissue sections [38, 44, 55] (Fig. 1a, b). A single tumor section can contain up to 10^5–10^7 cells, which can be rapidly and reproducibly scored using purpose-built image analysis algorithms [55]. Assessment of full tissue sections is a robust approach to account for tissue heterogeneity and to reliably measure PD-L1 expression at low levels [44, 55]. Fine-scale differences in biomarker compartmentalization can be captured at single cell resolution. ML technologies are beginning to be utilized in digital image analysis [40]. Pathologist input is critical for the creation of homogenous artificial intelligence training datasets that incorporate a sufficiently large number of true positive and negative staining examples as well as cases with unspecific staining and common technical artifacts. Expertly designed ML applications then have the potential to aid the human observer in assigning biomarker scores to defined cell populations based on morphological criteria and staining characteristics. For PD-L1 scoring, this approach can be particularly helpful to include or exclude PD-L1 expression in tumor infiltrating immune cells and tumor regions with non-specific, e.g., cytoplasmic staining (Fig. 1c–f) [44].

Standardized biomarker assessment protocols are essential to accelerate clinical development of immuno-oncology therapeutics [27]. Co-development of purpose-built computational pathology solutions in parallel to clinical trials may help to harmonize immunotherapy companion diagnostics, since image analysis algorithms can be easily shared and standardized across diagnostic labs. Alternatively, standardized biomarker testing may be conducted via telepathology at a central location. In addition, computational pathology applications can also increase the quality of biomarker assessment within an institution by using standardized tissue blocks and scanning
procedures. While computational pathology will become a crucial tool in extracting quantitative information from digitized slides, digitization will also enable new applications. It will be possible to compare current cases with annotated phenotype libraries, hence effectively integrating a vast amount of knowledge into clinical decision-making [52]. Taken together, computational pathology methods are likely to become important tools for making the assessment of immune biomarkers more reproducible, robust, and reliable.

Spatial analysis of tumor immune cell infiltration

The identification, localization, and spatial relationships of specific immune cells—before, during, or after therapy—have significant prognostic and predictive potential [10]. Both commercial and freely available open-source image analysis solutions are available to perform area-based quantification of immune cells by IHC or immunofluorescence (IF) for prognostic and predictive analyses [6, 25, 54]. This includes the enumeration of peri- and intratumoral CD3+ and CD8+ T cell populations per mm² of tumor tissue to form the Immunoscore® (IS) as marketed by HalioDx [63]. This approach has shown independent prognostic value in CRC beyond usual risk factors and has strong potential to aid patient stratification in other solid tumors [41]. A worldwide consortium-based validation study in stage I–III CRC reported significantly longer recurrence free intervals in patients with high numbers of tumor-infiltrating T cells [63]. Recent data from stage III and stage IV CRC patients treated with adjuvant chemotherapy also indicate that the quantification of IS markers by image analysis has the potential to guide decisions on treatment duration and follow-up in advanced disease [57]. However, the predictive potential of the Immunoscore® for immunotherapy response beyond well-known predictors, such as DNA mismatch-repair (MMR) deficiency remains to be investigated [62].

Assessing immune cell infiltration is an important component of the “Cancer Immunogram” for patient stratification in future immunotherapy trials [10]. Initial studies have identified a broad association of T cell location at the tumor invasive margin and in nodal metastasis with response to ICI in melanoma patients [19, 79]. Modern digital image analysis techniques will enable a more detailed analysis: The exact coding of the x-y location of each individual marker positive cell on a
histological slide for precise nearest neighbor and infiltration analyses is already feasible (Fig. 2a–c). This will allow to better understand the pathophysiology of tumor-host interaction, checkpoint molecule expression, and therapy effects in archival samples and preclinical models [8, 31]. Using these modern digital tools to record immune cell infiltrates over the course of therapy will empower data mining with patient characteristics, clinical response profiles, and genomic markers to...
build novel predictive indicators. Spatial profiling and compartmentalization studies can also be used to monitor and better understand immunotherapy associated adverse events which are commonly observed with ICI therapy [47].

**Multiplexing**

To empower more accurate patient stratification for immunotherapy, histological analysis should aim for a simultaneous characterization of both immune and tumor-related pathways in a single tissue sample [10, 29, 73]. Multiplexed immunoprofiling is key to generate comprehensive biomarker datasets for correlation with clinical parameters [11, 80]. Digital image analysis is an effective tool to extract comprehensive information on biomarker expression levels, co-localization, and compartmentalization. Technical approaches include virtual multistaining by computational methods, simultaneous multimarker staining of single slides, and sequential staining and quenching protocols.

Image registration is a powerful approach to generate virtual multistains from serial sections using computational methods: Following sectioning, staining, and scanning, serial sections can be computationally aligned by automated or user guided approaches [82]. Digital image analysis is carried out and the results are used to generate virtual multimarker profiles. Standard manual or automated staining procedures can be used for visualization. Virtual multistaining also offers the unique opportunity of combining distinct histological methods, such as RNA in situ hybridization and IHC for multiparametric tissue profiling. This approach has been recently used to spatially correlate the expression of specific microRNAs, interferon-stimulated target genes, T cell infiltration, and the expression of cytotoxic effector molecules in the microenvironment of CRC [45]. Limitations include the requirement for multiple tissue sections, inevitable sectioning artifacts, and the need to perform complex image transformation procedures for precise image alignment.

Highly optimized multistaining panels in combination with purpose-built digital image analysis software can overcome these drawbacks. However, substantial investment into a multimodal staining platform and analysis software is required to tap into this technology. Examples of this approach include the Vectra® Polaris™ Automated Quantitative Pathology Imaging System introduced by PerkinElmer [35]. Using a single slide, this workflow allows to generate highly multiplexed IF panels to investigate six or more antigens in a strictly quantitative manner. Standardization is facilitated by commercially available marker panels, which may assist future clinical applications. A recent study by Mezheyevskii and colleagues elegantly illustrates the combination of the Vectra® approach with tissue microarray technology for a high-throughput profiling of the immune-environment in NSCLC samples [56]. By combining a tumor-specific marker with a panel of immune-cell associated antigens, the authors analyze both quantitative and spatial information of specific lymphocyte subpopulations with RNA-expression levels and prognosis.

Interesting alternatives to simultaneous multiplexed IF are sequential workflows consistent of staining, digitalization, inactivation of fluorescent dyes, and re-staining. Profiling of up to 100 IF biomarkers on a single tissue section is technically feasible [29, 70]. Exemplary applications include multiplexed fluorescence microscopy method (MxIF) [29, 73] and multipetptide-ligand cartography (MELC) [70]. Similar approaches have been tested and successfully applied for IHC protocols [66, 78]. The single channel outputs from each staining cycle can be easily merged into comprehensive expression maps for characterization of the tumor microenvironment using computational methods [78]. However, individual optimization of the staining protocols, tissue degradation with iterative staining cycles, process time, and standardization may be hurdles in applying this technology to large clinical trial sample sets.

**CO-Detection by indEXing (CODEX)** is among the most recent and innovative approaches for high-dimensional imaging of antibody-tagged epitopes in FFPE tissue [30]. CODEX staining uses antibodies labeled with unique DNA barcodes that are iteratively detected by in-situ polymerization with fluorescently labeled dNTP analogues. Tissue slides are stained in a single incubation step, followed by imaging cycles for the spatially resolved visualization of all antibody binding events. Pilot studies for have demonstrated the technical feasibility and power of this approach for deep profiling of immune tissue architecture [30]. Application to immunoprofiling of cancer tissues is a logical next step for detailed analysis of the microenvironment in correlation with clinical and molecular parameters. A decisive advantage of this approach is the possibility to image CODEX-labels on a standard three-color fluorescence microscope or scanner.
Combining spatial imaging approaches with mass spectrometry may serve to reach yet another dimension in precision profiling of the tumor microenvironment. Multiplexed ion beam imaging (MIBI) uses antibodies labeled with isotopically pure elemental metal reporters (mass tags) to simultaneously detect up to 100 targets on FFPE tissue sections [5]. Conventional histologic stains can be included in this panel, allowing precise reconstruction of virtual histology images with antibody expression data. Decisive technological advantages of MIBI include the stability of mass tags, a sensitivity exceeding chromogenic IHC by up to three log levels, and no spectral overlap between individual labels [5]. Recent studies have demonstrated the technical feasibility to integrate RNAscope-based metal in situ hybridization with MIBI for multiparametric profiling [71]. Purpose-built digital image analysis methods are available to reconstruct virtual histology images and explore cell phenotypes, spatial interaction, and morphological structures [69].

Further technological advances using dissociated tissue samples allow to perform single cell proteomic analysis using a mass spectrometry-flow cytometry hybrid device, the so-called “CyTOF” [9]. However, substantial investment and expertise in mass spectrometry are necessary to access this technology.

**Morpho-molecular integration**

The recently proposed “Cancer Immunogram” suggests seven categories including tumor foreignness, general immune status, immune cell infiltration, absence of checkpoints, absence of soluble inhibitors, absence of inhibitory tumor metabolism, and tumor sensitivity to immune effectors as the most important predictors of immunotherapy response [10]. As such, the “Cancer Immunogram” is highly integrative and includes both tumor- and immune-related parameters assessed with both molecular and image-based methods for individualized prediction of immunotherapy response. By providing continuous data on tissue-based parameters, such as immune cell infiltration and expression of immune checkpoints, computational pathology methods are ideally suited for data integration with molecular parameters.

Tumors with a high mutational load frequently exhibit an immunologically activated phenotype [4]. Recent study has linked defects in DNA mismatch repair to high frequencies of neoantigens and immunotherapy response in advanced solid tumors [49, 62]. However, the correlation between genomic parameters and immunogenicity is far from linear [53]. In some solid tumors, such as melanoma, a lack of association of neoantigen type and frequency with baseline immune activation has been elegantly demonstrated [74]. In these tumors, strong regulatory T cell infiltration may curb the anti-tumoral immune response [75]. Other known mechanisms that may impact tumor antigenicity are defects in antigen-presentation and overexpression of immunosuppressive molecules, such as CD47, PD-L1, and indolamine-2,3-dioxygenase [42, 43, 75]. These image-based features can be reliably captured by computational image analysis methods for in-depth profiling of the tumor microenvironment.

Incorporating digital information on immune cell infiltration with molecular data is a powerful approach to inform the “Cancer Immunogram.” Conde and colleagues recently provided a proof of principle for the association of specific genetic alterations with defined immunophenotypes by combining digital assessment of CD8+ T cell infiltration in lung squamous-cell carcinoma with manual scoring of PD-L1 and targeted next generation sequencing [16]. Joining digital and molecular pathology also facilitates the development of novel assays to predict immunotherapy response. We have recently developed and validated a computational pathology assay that identifies specific PD1-positive subpopulations in NSCLC as a powerful predictive indicator of response to ICI treatment [76]. Methodologically, this assay was driven by morpho-molecular integration of RNA sequencing data with tissue-based methods. Computational image analysis was a key tool in this process, driving the development of a standardized algorithm for the detection of tumor-infiltrating lymphocyte populations with uniquely high expression levels of PD1 protein (PD1⁺ lymphocytes). PD1⁺ cells were reproducibly detected and quantified in pretreatment biopsies of lung cancer patients and correlated strongly with treatment response to ICI in two independent clinical cohorts. This translational approach highlights how digital image analysis can represent a powerful companion diagnostic for cancer immunotherapy applications.

**Machine learning and artificial intelligence**

ML will transform the field of immuno-oncology. We expect that ML will drive a paradigm shift in the way data are collected and analyzed to discover new prognostic markers or to construct more rigorous risk classification to empower stratified medicine. ML is essentially a set of computer algorithms that learn generic rules to perform any given task directly from data, without requirements of predefined knowledge or domain expertise. This, in many ways, is similar to how pathologists have previously gained knowledge and expertise through continuous practice that has led to new diagnostic classifications or prognostic factors in the clinical routine.

Since ML is heavily data-driven, it enables a means to derive unbiased statistics from data. Accurate and continuous variables are more informative and can provide more biologically relevant information than the semi-quantitative scores presently implemented in diagnostic practice. Although several classical regression models exist to facilitate the discovery process of predictive or prognostic factors, the power of ML surpasses that of traditional tools when data is large and complex. ML can crunch through a vast amount of data and learn
in the training of image analysis algorithms is therefore of critical importance. Furthermore, ML tends to “overfit” to discovery data which results in an over-optimistic estimate of the performance of a model, while, in fact, it cannot be generalized well to new unseen datasets. It is, therefore, necessary that ML models are rigorously validated on new datasets that are independent of those used during the model development. Originating from the field of computer vision, deep convolutional neural network (DCNN) is a ML approach that is specialized in image analysis tasks [46, 50, 51]. In recent years, DCNN has been successfully applied to key applications in diagnostic pathology, including cell classification [72], cell enumeration [85], tumor grading [23], cancer diagnosis [22], and cancer prognostication [58].

Application of machine learning in immuno-oncology: pattern recognition

In immuno-oncology, ML as a pattern recognition tool enables an accurate and reproducible means for the unbiased assessment of regularities in the expression of immunohistochemical markers, tumor morphology, and the spatial distribution of tumor infiltrating lymphocytes (TILs). The ability of ML tools to detect key features in complex immunophenotypic datasets underlines their potential importance for the development of novel predictive models in cancer research. Initial studies underline the potential of ML methods for clinical translation:

Yuan et al. computationally profiled cells of triple-negative breast cancer cohorts on H&E-stained sections and were able to uncover three categories of lymphocytes (intra-tumor, adjacent-tumor, and distal-tumor) with an unsupervised clustering method based on the proximities of lymphocytes to tumor cells [87]. Interestingly, the ratio between the total number of intratumoural lymphocytes and the number of cancer cells was strongly associated with disease-specific survival and strongly correlated with the expression of CTLA-4, a known immunotherapy target. Saltz et al. employed a DCNN model to determine the probability of TIL infiltration for every small region of H&E stained sections [68]. This resulted in a map which indicates the degree of local lymphocyte infiltration. Unsupervised clustering methods were employed to group small TIL regions into spatially coherent structures [28]. The resulting clustering patterns were characterized using parameters related to cluster size and shape. Importantly, the covariates summarizing complex characteristics of clusters were found to be associated with overall survival in various cancer groups, illustrating the potential of ML methodologies applied to standard H&E slides for the development of biomarkers in immuno-oncology.

Heindl et al. demonstrated that quantifying the number of spatial clusters or hotspots of immune cells and cancer cells is prognostic in ER+ breast cancer using a fully automated H&E stained image analysis algorithm [32]. The spatial clustering is identified as an area in which the number of cells of interest is greater than expected by chance given the distribution of the cells on the whole tissue section. The increased immune spatial clustering is associated with poor prognosis and has a level of prognostic significance on par with the IHC test [18] and the Oncotype DX 21-gene recurrence score [64]. Interestingly, immune scores which are related to the ratios of distinct categories of lymphocytes to the total number of cancer cells did not provide prognostic information in this study. This illustrates the importance of ML tools to detect novel morphologic features that provide prognostic information beyond established classifications.

Application of machine learning in Immuno-oncology: survival analysis

In medical research, survival analysis is a traditional means to assess the prognostic significance of each candidate covariate. The most widely used survival model is the Cox proportional hazards model. It assumes that the risk of an event of interest to occur (failure) is not time-dependent, is determined by a linear combination of all covariates, and might also include interaction terms between the covariates. This somewhat too simplistic assumption may not be able to capture the precise effects of the observed medical covariates on the risk of failure. This is particularly true in immuno-oncology where variable effects of immune infiltrates and immune-related gene signatures are observed depending on the tumor type and host immune status [7]. DNN applied to survival analysis could provide a better way to model complex nonlinear relationships among prognostic factors that better fit to the survival data.

It has been shown that DNN performs better than the Cox proportional hazard model in various medical applications [24, 58, 86]. Katzman et al. investigated the risk associated with treatment choices and demonstrated that their DeepSurv system could provide treatment recommendations that increase the median survival time of patients [37]. Yousef et al. used their SurvivalNet to analyze large-scale genomic profiles obtained from the TCGA database [14, 86]. They have demonstrated the robustness of the model across different cancer types even if the number of input covariates is
sometimes considerably larger than the sample size. Moreover, the model also allows the interpretation of the prognostic significance of individual covariates based on their contributions to overall risk. The same model has been extended to allow integration between histology image and genomic data in a unified framework [58]. Likewise, specific methylome signatures queried by ML were shown to be suitable for prediction of response to immune-checkpoint inhibitors. In analogy to image analysis, this modality takes both neoplastic and reactive cells into account [21]. These results suggest the potential of deep learning as a discovery tool to provide insight into the biology of immuno-oncology.

Challenges

Regulatory issues in the use of digital pathology devices have been broadly recognized by national and international bodies. Pioneering companies have recently received CE certification for routine pathology applications in the European Union under the In vitro diagnostic medical devices directive [1]. Comprehensive, non-binding recommendations have been issued by the Food and Drug Administration (FDA) and the College of American pathologists (CAP) [3, 65]. These guidelines highlight two important aspects of testing emerging biomarkers for immunotherapy by computational pathology methods: First, the quality and reliability of the imaging system as recommended by the FDA and mandated by the CE certification process, and second, the consistency of diagnoses made by pathologists using digital systems as advised by the CAP [3, 65]. Both are of essential importance in the design of reliable computational pathology workflows to inform clinical decision making.

Concerns have been raised that inaccurate study design, statistical analysis, and reporting of research lead to a significant waste of research funding and inaccurate scientific conclusions [36]. A major challenge inherent to the use of digital image analysis is the substantial expansion in the number of variables. Although it may seem attractive from a scientific perspective to extract as much information as possible from a limited amount of tissue, it is important to recognize that big data mining inherently increases the level of statistical noise. Robust statistical considerations are required to avoid multiple testing problems and misleading conclusions. As with any diagnostic tool, digital image analysis is therefore critically dependent on scientific rigor, reliable documentation, consequent quality control, and adequate training of the pathology workforce [48]. Unsolved problems also concern the standardization of the preanalytical steps before the imaging and analysis of tissue slides. A perfect example is Ki-67 staining, which has proven difficult to harmonize between labs despite being one of the most frequently performed assay of its kind [26]. The same applies to the majority of immunohistochemical stains, in particular those where expression levels matter.

Conclusions

Advanced multiparametric imaging applications and ML have the potential to translate our evolving understanding of tumor-host interaction into better patient stratification and new treatment strategies. Computational pathology will help to derive complete, standardized, and reproducible datasets to facilitate the individualized prediction of immunotherapy response. Upskilling the pathology workforce through professional education and the recognition of computational pathology by professional societies will be essential to meet future clinical demands for optimal patient care.

Author contributions VHK, KS, JR, and KDM jointly conceptualized this work, drafted the manuscript, and approved the final version to be published.

Funding information Viktor H Koelzer is supported by a research fellowship of the Swiss National Science Foundation (P2SKP3_168322). Kirsten D Mertz is supported by a grant of the Krebsliga beider Basel (KLbB-4182-03-2017) and by a grant of the “Freiwillige Akademische Gesellschaft Basel.” KS and JR were supported by the Oxford NIHR Oxford Biomedical Research Centre.

Compliance with ethical standards

No research involving human participants and/or animals was performed for this study.

Conflict of interest KS and JR have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents received or pending. V.H.K. and K.D.M. are participants of a patent application co-owned by the Netherlands Cancer Institute (NKI-AVL) and the University of Basel on the assessment of cancer immunotherapy biomarkers by digital pathology. Based on NKI-AVL and the University of Basel policy on management of intellectual property, V.H.K. and K.D.M. would be entitled to a portion of received royalty income. V.H.K. has served as an invited speaker on behalf of Indica Labs.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices. https://ec.europa.eu/growth/single-market/european-standards/harmonised-standards/iv-diagnostic-medical-devices_en. Accessed March 6th, 2018
2. Roche receives FDA clearance for the VENTANA MMR IHC panel for patients diagnosed with colorectal cancer. http://www.ventana.com/roche-receives-fda-clearance-ventana-mmr-ihc-panel-patients-diagnosed-colorectal-cancer/. Accessed May 14th, 2018
3. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Devices and Radiological Health, Office
26. Focke CM, Burger H, van Diest PJ, Finsterbusch K, Glaser D, Faraggi D, Simon R (1995) A neural network model for survival.

24. Feng Z, Bethmann D, Kappler M, Ballesteros-Merino C, Eckert A, Duruisseaux M, Martinez-Cardus A, Calleja-Cervantes ME, Moran MF, Gut M, Gut I, Teixido C, Vilarino N, Prat A, Reguat N, Benito N, Garrido P, Barragan I, Emile JF, Rosell R, Brambilla E, Esteller M (2018) Epigenetic prediction of response to anti-PD-1 treatment in non-small-cell lung cancer: a multicentre, retrospective analysis. Lancet Respir Med 6:771–781. https://doi.org/10.1016/S2213-2600(18)30284-4

22. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, van Norden D, Bartelsman JF, Bosch J, Karssemeijer N, Litjens G, van der Laak JA, the CC, Hermens M, Mansen QF, Balkenhol M, Geesink S, Stathonkos N, van Dijk MC, Bult P, Beca F, Beck AH, Wang D, Khosla A, Gargeya R, Irshad H, Zhong A, Dou Q, Li Q, Chen H, Lin HJ, Heng PA, Hass C, Bruni E, Wong Q, Halici U, Oner MU, Cetin-Atlay R, Berseth M, Khvatkov V, Vleyezghanin A, Kraus O, Shaban M, Rajpoot N, Awan R, Sirinukunwattana K, Qaiser T, Tsang YW, Tellez D, Annumishe J, Hufnagl P, Valkonen M, Kartasalo K, Latonen L, Ruusuvuori P, Liimatainen K, Albarqouni S, Munagl B, George A, Demirel S, Navab N, Watanabe S, Seno S, Takenaka Y, Matsuda H, Ahmady Phoulady H, Kovalev V, Kalinovsky A, Lianuich V, Bueno G, Fernandez-Camboles MM, Serrano I, Deniz O, Raococeanu D, Venancio R (2017) Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA 318:2199–2210. doi: https://doi.org/10.1001/jama.2017.14585

23. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542:114–118. https://doi.org/10.1038/nature21056

24. Farag D, Simon R (1995) A neural network model for survival data. Stat Med 14:73–82

25. Feng Z, Bethmann D, Kappler M, Ballesteros-Merino C, Eckert A, Bell RB, Cheng A, Bui T, Leidner R, Urba WJ, Johnson K, Hoyt C, Bifutulo CB, Bukar J, Wickenhauer C, Seliger B, Fox BA (2017) Multiparametric immune profiling in HPV-oral squamous cell cancer. JCI Insight 2. https://doi.org/10.1172/jci.insight.93652

26. Frey BJ, Dueck D (2007) Clustering by passing messages between data points. Science 315:972–976. https://doi.org/10.1126/science.1136800

27. von Hoegen P, Wang E, Welters MJ, Winter H, Withington T, Wolchok JD, Xiao W, Zitvogel L, Zwierzina H, Marincola FM, Gajewski TF, Wigginton JM, Disis ML (2011) Defining the critical hurdles in cancer immunotherapy. J Transl Med 9:214. https://doi.org/10.1186/1479-5876-9-214

28. Frey BJ, Dueck D (2007) Clustering by passing messages between data points. Science 315:972–976. https://doi.org/10.1126/science.1136800

29. Gerdes MJ, Sevinsky CJ, Sood A, Adak S, Bello MO, Bordwell A, Can A, Corwin A, Dinn S, Filkins RJ, Hollman D, Kamath V, Kaanumalle S, Kenny L, Larsen M, Li Q, Lowes C, McCulloch CC, McDonough E, Montalto MC, Pang Z, Rittscher J, Santamaria-Pang A, Saranchan BD, Seel ML, Seppo A, Shaiikh K, Sui Y, Zhang J, Ginty F (2013) Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. Proc Natl Acad Sci USA 110:11982–11987. https://doi.org/10.1073/pnas.1300316110

30. Goltsvev Y, Samusik N, Kennedy-Darling J, Bhatre S, Hale M, Vasquez G, Black S, Nolan G (2018) Deep profiling of mouse splenic architecture with CODEX multiplexed imaging. bioRxiv. https://doi.org/10.1101/203166

31. Heindl A, Nawaz S, Yuan Y (2015) Mapping spatial heterogeneity in the tumor microenvironment: a new era for digital pathology. Lab Invest 95:377–384. https://doi.org/10.1038/labinvest.2014.155

32. Heindl A, Sestak I, Naidoo K, Cuzick J, Dowsett M, Yuan Y (2018) Relevance of spatial heterogeneity of immune infiltration for predicting risk of recurrence after endocrine therapy of ER+ breast cancer. J Nat Cancer Inst 110. https://doi.org/10.1093/jnci/djy137

33. Hoos A (2016) Development of immuno-oncology drugs—from CTLA4 to PD1 to the next generations. Nat Rev Drug Discov 15: 235–247. https://doi.org/10.1038/nrd.2015.35

34. Hsu FS, Su CH, Huang KH (2017) A comprehensive review of US FDA-approved immune checkpoint inhibitors in urothelial carcinoma. J Immunol Res 2017:6940546. https://doi.org/10.1155/2017/6940546

35. Huang W, Henrickk C, Drew S (2013) A colorful future of quantitative pathology: validation of Vectra technology using chromogenic multiplexed immunohistochemistry and prostate tissue microarrays. Hum Pathol 44:29–38. https://doi.org/10.1016/j.humpath.2012.05.009

36. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, Schulz KF, Tishbriani R (2014) Increasing value and reducing waste in research design, conduct, and analysis. Lancet 383:166–175. https://doi.org/10.1016/S0140-6736(13)62227-8

37. Katzman JL, Shahu M, Cloninger A, Bates J, Jiang T, Kluger Y (2018) DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. BMC Med Res Methodol 18:24. https://doi.org/10.1186/s12874-018-0482-1

38. Kearney S, Black J, Aeffner F, Black J, Pratte L, Krueger J (2017) Abstract 4582: evaluating benefits of PD-L1 image analysis for the clinical setting. Cancer Res 77:4582–4582. https://doi.org/10.1158/1538-7445.am.2017-4582

39. Khagi Y, Kurzrock R, Patel SP (2017) Next generation predictive biomarkers for immune checkpoint inhibition. Cancer Metastasis Rev 36:179–190. https://doi.org/10.1007/s10555-016-9652-y

40. Khosravi P, Kazemi E, Imielinski M, Elemento O, Hajirasouliha I (2018) Deep convolutional neural networks enable discrimination of heterogeneous digital pathology images. ElBioMedicine 27:317–328. https://doi.org/10.1016/j.ebiom.2017.12.026

41. Kirilovsky A, Marlrot F, El Sissy C, Haicheur N, Galon J, Pages F (2016) Rational bases for the use of the immunoscore in routine clinical settings as a prognostic and predictive biomarker in cancer patients. Int Immunol 28:373–382. https://doi.org/10.1093/intimm/dxw021
42. Koelzer VH, Canonica K, Dawson H, Sokol L, Karamitopoulopou-Diamantis E, Lugli A, Zlobec I (2016) Phenotyping of tumor-associated macrophages in colorectal cancer: impact on single cell invasion (tumor budding) and clinicopathological outcome. Oncoimmunology 5:e1106677. https://doi.org/10.1080/2162402X.2015.1106677

43. Koelzer VH, Dawson H, Anderson E, Karamitopoulopou E, Masucci GV, Lugli A, Zlobec I (2015) Active immunosurveillance in the tumor microenvironment of colorectal cancer is associated with low frequency tumor budding and improved outcome. Transl Res 166:207–217. https://doi.org/10.1016/j.trsl.2015.02.008

44. Koelzer VH, Gisler A, Hanhart JC, Griss J, Wagner SN, Willi N, Cathomas G, Sachs M, Kempf W, Thommen DS, Mertz KD (2018) Digital image analysis improves precision of programmed death ligand 1 (PD-L1) scoring in cutaneous melanoma. Histopathology. https://doi.org/10.1111/his.13528

45. Koelzer VH, Sokol L, Zahir S, Christe L, Dawson H, Berger MD, Inderbitzin D, Zlobec I, Lugli A (2017) Digital analysis and epigenetic regulation of the signature of rejection in colorectal cancer. Oncoimmunology 6:e1288330. https://doi.org/10.1080/2162402X.2017.1288330

46. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst:1097–1105

47. Laubli H, Koelzer VH, Matter MS, Herzig P, Dolder Schlienger B, Wiese MN, Lardinois D, Mertz KD, Zippelius A (2018) The T cell repertoire in tumors overlaps with pulmonary inflammatory lesions in patients treated with checkpoint inhibitors. Oncoimmunology 7:e1386362. https://doi.org/10.1080/2162402X.2017.1386362

48. Lawler M, Alsina D, Adams RA, Anderson AS, Brown G, Fearnhead NS, Fenwick SW, Halloran SP, Hochhauser D, Hull MA, Koelzer VH, McNair AGK, Monahan KJ, Nathke I, Norton C, Novelli MR, Steele RJC, Thomas AL, Wilde LM, Wilson RH, Tomlinson I, Bowel Cancer UKCRCiGCI (2018) Critical research gaps and recommendations to inform research prioritisation for more effective prevention and improved outcomes in colorectal cancer. Gut 67:179–193. https://doi.org/10.1136/gutjnl-2017-315333

49. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, Martin NT, Black JC, Pollack Z, Aetfen F, Krueger J (2017) Abstract 661: evaluating “harmonization” of PD-L1 assays using image analysis. Cancer Res 77:661–661. https://doi.org/10.1158/1538-7445.am2017-661

50. Mezheyuski A, Bergsland CH, Backman M, Djureinovic D, Sjoblom T, Braun J, Mieck P (2017) Multispectral imaging for quantitative and compartment-specific immune infiltrates reveals distinct immune profiles that classify lung cancer patients. J Pathol. https://doi.org/10.1002/path.4052

51. Mclnich B, Van den Eynde M, Bindea G, Church SE, Vastuoro A, Fredriksen T, Lafontaine L, Haicheur N, Marliot F, Debattencourt D, Painet G, Jourret-Mourin A, Gigot JF, Hubert C, Danse E, Dragean C, Carrasco J, Humblet Y, Valge-Archer V, Berger A, Pages F, Machiels JP, Galon J (2018) Comprehensive Intratumatistic immune quantification and major impact of immunoscore on survival. J Natl Cancer Inst 110. https://doi.org/10.1093/jnci/djy123

52. Mobadersany P, Youselli S, Amgad M, Gutman DA, Barmholz-Sloan JS, Velazquez Vega JE, Brat DJ, Cooper LAD (2018) Predicting cancer outcomes from histology and genomics using convolutional networks. Proc Natl Acad Sci U S A 115:E2970–E2979. https://doi.org/10.1073/pnas.1717391195

53. Moran S, Martinez-Carus A, Sayols S, Musalen E, Balana C, Estival-Gonzalez A, Moulinho C, Heyn H, Diaz-Lageses A, de Moura MC, Stella GM, Comoglio PM, Ruiz-Miro M, Matias-Guiu X, Pazo-Cid R, Anton A, Lopez-Lopez R, Soler G, Longo F, Guerra I, Fernandez S, Assenov Y, Plass C, Morales R, Carles J, Bowtell D, Mileskin L, Sia D, Thothill R, Taberner J, Llovet JM, Esteller M (2016) Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. Lancet Oncol 17:1386–1395. https://doi.org/10.1016/S1470-2045(16)30297-2

54. Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, Tykodi SS, Sosman JA, Procopio G, Pili E, Castellano D, Cheoueiri TK, Gurney H, Donskov F, Bono P, Westrich E, Tugler HC, Ueda T, Tomita Y, Schutze FA, Holland SM, Kollmannsberger C, Larkin J, Ravaud A, Simon JS, Xu LA, Waxman IM, Sharma P, CheckMate I (2015) Nivolumab versus everolimus in advanced renal-cell carcinoma. N Engl J Med 373:1803–1813. https://doi.org/10.1056/NEJMoa1510665

55. Obermeyer Z, Emanuel EJ (2016) Predicting the future—big data, machine learning, and clinical medicine. N Engl J Med 375:1216–1219. https://doi.org/10.1056/NEJMsa1606181

56. Overman MJ, McDermott R, Leah J, Lonardi S, Lenh Z, Morse MA, Desai J, Hill A, Axelsson M, Moss RA, Goldberg MV, Cao Z, Ledein JM, Maglinie GA, Kopetz S, Andre T (2017) Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. Lancet Oncol 18:1189–1191. https://doi.org/10.1016/S1470-2045(17)30422-9

57. Pagés F, Mlecnik B, Marliot F, Binadea G, Ou F-S, Bifulco C, Lugli A, Zlobec I, Rau TT, Berger MD, Nagtegaal ID, Vink-Böger E, Hartmann A, Geppert C, Kolweiler J, Merkel S, Grützmann R, Van den Eynde M, Jouret-Mourin A, Kärthuser A, Léonard D, Remue C, Wiese MN, Lardinois D, Mertz KD, Zippelius A (2014) Computational pathology: an emerging definition. Arch Pathol Lab Med 138:1133–1138. https://doi.org/10.5858/arpa.2014-0034-ED

58. Lukasza M, Riaz N, Makarov V, Balachandran VP, Hellmann MD, Solovyov A, Rizvi NA, Merghouth T, Levine AJ, Chan TA, Wolchok JD, Greenbaum BD (2017) A neoantigen fitness model predicts tumor response to checkpoint blockade immunotherapy. Nature 551:517–520. https://doi.org/10.1038/nature24473
H. Okuno K, Torigoe T, Sato N, Furuhata T, Takemasa I, Itoh K, Patel PS, Vora HH, Shah B, Patel JB, Rajvik KN, Pandya SJ, Shukla SN, Wang Y, Zhang G, Kawakami Y, Marincola FM, Ascieto PA, Sargent DJ, Fox BA, Galon J International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. Lancet. https://doi.org/10.1016/S0140-6736(18)30789-X

64. Paik S, Shah S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N (2004) A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 351:2817–2826. https://doi.org/10.1056/NEJMoa041588

65. Pantanowitz L, Sinard JH, Henricks WH, Fatheree LA, Carter AB, Remark R, Merghoub T, Grabe N, Litjens G, Damotte D, Wolchok JD, Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Saltz J, Gupta R, Hou L, Kurc T, Singh P, Nguyen V, Samaras D, Sood A, Miller AM, Brogi E, Sui Y, Hernandez KM, Li Y, Gajewski TF (2016) Density of immunogenic antigens does not explain the presence or absence of the T cell-inflamed tumor microenvironment in melanoma. Proc Natl Acad Sci USA 113:E7759–E7768. https://doi.org/10.1073/pnas.1609376113

75. Spranger S, Spaanem RM, Zha Y, Williams J, Meng Y, Ha TT, Gajewski TF (2013) Up-regulation of PD-L1, IDO, and T(regs) in the melanoma tumor microenvironment is driven by CD8(+) T cells. Sci Transl Med 5:200ra116. https://doi.org/10.1126/scitranslmed.3006504

102. Thrommen DS, Koehler VH, Herzig P, Roller A, Trefny M, Dimeloe S, Kievalainen A, Hanhart J, Schill C, Hess C, Savic Prince S, Wiebe M, Lardinois D, Ho PC, Klein C, Karanikas V, Mertz KD, Schumacher TN, Zipfel A (2018) A transcriptionally and functionally distinct PD-1(+)/CD8(+) T cell pool with predictive potential in non-small-cell lung cancer treated with PD-1 blockade. Nat Med 24:994–1004. https://doi.org/10.1038/s41591-018-0057-z

103. Troncone G, Gridelli C (2017) The reproducibility of PD-L1 scoring in lung cancer: can the pathologists do better? Transl Lung Cancer Res 6:S74–S77. https://doi.org/10.21037/tlcr.2017.10.05

104. Tsujikawa T, Kumar S, Borkar RN, Azimi V, Thibault G, Chang YH, Bailer A, Kawashima R, Choe G, Sauer D, El Rassi E, Clayburgh DR, Kulesz-Martin MF, Lutz ER, Zheng J, Laffee EM, Leyschuck P, Margolin AA, Mori M, Gray JW, Flint PW, Coussens LM (2017) Quantitative multiplex immunohistochemistry reveals myeloid-inflamed tumor-immune complexity associated with poor prognosis. Cell Rep 19:203–217. https://doi.org/10.1016/j.celrep.2017.03.037

105. Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJ, Robert L, Chmielowski B, Spasic M, Henry S, Plisbou V, West AN, Carmona M, Kivork C, Seja E, Cherry G, Gutierrez AJ, Grogan TR, Mateus C, Tomasic G, Glaspy JA, Emerson RO, Robins H, Pierce RH, Elashoff DA, Robert C, Ribas A (2014) PD-1 blockade induces responses by inhibiting adaptive immune resistance. Nature 515:568–571. https://doi.org/10.1038/nature13954

106. Ung C, Kokkx M, Waumans Y (2017) Digital pathology in immuno-oncology—a roadmap for clinical development. Expert Rev Precis Med Drug Dev 2:9–19. https://doi.org/10.1080/21681163.2016.1149104

107. Virchow R (1858) Die Cellularpathologie in ihrer Begründung auf historische, biologische, und pathologische Gewebelehre. August Hirschwald, Berlin

108. Wang CW, Ka SM, Chen A (2014) Robust image registration of cancer tissue samples by mass cytometry. Cell Syst 6:25. https://doi.org/10.1016/j.cels.2017.12.001

109. Watanabe K, Ueno S, Kurihara Y, Yamashita H, Kikuchi S, Tsuchiya K, Itoh M, Hayashi J, Minoura M, Nasu K, Yamanaka H, Nakajima K, Nishimura K, Saeki Y, Yoda K, Morikawa K, Kojima S, Iwane H, Hattori H, Uchida A, Tsuchiya K, Umehara H, Hattori T, Kato I, Matsubara M, Morimoto R, Asato H, Kakinuma H (2014) Multiplexed immunofluorescence delineates proteomic profiles with deep survival models. Sci Rep 7:11707. https://doi.org/10.1038/srep06050

110. Webb S (2018) Deep learning for biology. Nature 554:555–557. https://doi.org/10.1038/d41586-018-02174-z

111. Williams BJ, Bottoms D, Torean D (2017) Future-proofing pathology: the case for clinical adoption of digital pathology. J Clin Pathol 70:1010–1018. https://doi.org/10.1136/jclinpath-2017-204644

112. Xie W, Noble JA, Zisserman A (2018) Microscopy cell counting and detection with fully convolutional regression networks. Comput Meth Biomech Biomed Eng Imaging Visualization 6:283–292. https://doi.org/10.1080/21681163.2016.1149104

113. Yousefi S, Amrollahi F, Amgad M, Dong C, Lewis JE, Song C, Gutman DA, Halani SH, Velazquez Vega JE, Brat DJ, Cooper LAD (2017) Predicting clinical outcomes from large scale cancer genomic profiles with deep survival models. Sci Rep 7:11707. https://doi.org/10.1038/s41598-017-11876-7

114. Yuan Y (2015) Modelling the spatial heterogeneity and molecular correlates of lymphocytic infiltration in triple-negative breast cancer. J R Soc Interface 12. https://doi.org/10.1098/rsif.2014.1153