Artificial intelligence and digital pathology: Opportunities and implications for immuno-oncology

Faranak Sobhani, Ruth Robinson, Azam Hamidinekoo, Ioannis Roxanis, Navita Somaiah, Yinyin Yuan

PII: S0304-419X(21)00019-6

DOI: https://doi.org/10.1016/j.bbcan.2021.188520

Reference: BBACAN 188520

To appear in: *BBA - Reviews on Cancer*

Received date: 10 March 2020

Revised date: 4 January 2021

Accepted date: 30 January 2021

Please cite this article as: F. Sobhani, R. Robinson, A. Hamidinekoo, et al., Artificial intelligence and digital pathology: Opportunities and implications for immuno-oncology, *BBA - Reviews on Cancer* (2021), https://doi.org/10.1016/j.bbcan.2021.188520

Artificial Intelligence and Digital Pathology: Opportunities and Implications for Immuno-Oncology

Faranak Sobhani[a1] faranak.sobhani@icr.ac.uk, Ruth Robinson[b1] ruthrobinson2@nhs.net, Azam Hamidinekoo[a] azam.nekoo@icr.ac.uk, Ioannis Roxanis[b] ioannis.roxanis@icr.ac.uk, Navita Somaiah[b] navita.somaiah@icr.ac.uk, Yinyin Yuan[a,*] yinyin.yuan@icr.ac.uk

[a]Division of Molecular Pathology, Institute of Cancer Research (ICR), London, UK

[b]Division of Radiotherapy and Imaging, Institute of Cancer Research (ICR), The Royal Marsden NHS Foundation Trust, London, UK

[*]Corresponding author.

Abstract

The field of immuno-oncology has expanded rapidly over the past decade, but key questions remain. How does tumour-immune interaction regulate disease progression? How can we prospectively identify patients who will benefit from immunotherapy? Identifying measurable features of the tumour immune-microenvironment which have prognostic or predictive value will be key to making meaningful gains in these areas. Recent developments in deep learning enable a big-data analysis of pathological samples. Digital approaches allow data to be acquired, integrated and analysed far beyond what is possible with conventional techniques, and to do so efficiently and at scale. This has the potential to reshape what can be achieved in terms of volume, precision and reliability of output, enabling data for large cohorts to be summarised and compared. This review examines applications of Artificial intelligence (AI) to important questions in Immuno-oncology (IO). We discuss general considerations that need to be taken into account before AI can be applied in any clinical setting. We describe AI methods that have been applied to the field of IO to date and present several examples of their use.

Keywords

Artificial intelligence (AI), Deep Learning (DL), Digital Pathology (DP), Immuno-Oncology (IO)

[1] Joint authorship

1. Introduction

The ability to evade immune destruction is a seminal feature of cancer (Hanahan & Weinberg, 2011). Agents designed to ramp up the anti-tumour immune response have had therapeutic traction across a range of tumour sites and histologies (Tang et al., 2018) with some patients experiencing durable disease control. Aside from this, traditional cytotoxic therapies have been shown to mediate some of their anti-tumour effects through immune mechanisms (Bertin-Ciftci et al., 2013). Clinical success from immunotherapy is far from universal and the majority of unselected patients have a poor objective response. Besides, these agents have a significant toxicity profile (Gibney et al., 2016). To maximise the clinical gains - and minimise harm - it is essential that we have robust predictive biomarkers that are able to prospectively discriminate between those more or less likely to benefit from IC.

1.1. Predictive Assays in Current Use

IHC markers - PD-L1 expression by tumour and/or local immune cells, as assessed by single marker immunohistochemistry is used across a spectrum of solid tumours to select for a benefit from immune checkpoint inhibitors. However, its utility as a biomarker is limited by intra-tumoural heterogeneity and dynamic changes in expression. We lack a standardised approach to scoring and significance thresholds. Reliability of scoring is affected by inter-observer variation as well as technical differences between the various assays in use (Balar & Weber, 2017).

Genomic tools - Genomic tools including targeted panels to estimate tumour mutational burden are also used to select for likely responders. Tumour mutational burden (TMB) correlates with neoantigen load and has been shown to predict response to IO in lung, bladder and head and neck tumours (Chan et al., 2019b). Cancers with defective mismatch repair (dMMR) tend to have high TMB as consequence, and IO is therefore of particular benefit in this subgroup. dMMR is most commonly seen in cancers associated with the inherited Lynch syndrome (colorectal, endometrial, small intestine, urothelial, central nervous system and sebaceous gland cancers) and can be detected through the use of antibodies against nuclear MMR proteins, plus or minus PCR to identify microsatellite instability - a downstream manifestation of dMMR (Luchini et al., 2019). Although both are predictive biomarkers for sensitivity to immune checkpoint blockade, TMB and PDL1 do not necessarily select for the same patients as illustrated by the fact that dual checkpoint blockade for NSCLC was beneficial with high TMB, irrespective of PDL1 status (Hellmann et al.,

2018). This underlines the fact that clinical response to IO is determined by multiple factors. A recent meta-analysis showed that composite biomarkers incorporating PD-L1, TMB and simultaneous quantification of multiple proteins via multiplex IHC/immunofluorescence performed better than either PD-L1 or TMB in isolation (Lu et al., 2019). However, the increased cost and complexity of these techniques need to be considered if aiming to implement more widely.

Assays of immune reaction - The density of tumour-infiltrating immune effector cells also shows promise as a clinically useful biomarker. In colorectal cancer, the Immunoscore has been shown to be a better predictor of outcome than traditional TNM staging. This score is based on the density of CD3 and CD8-positive cells at the invasive margin and the centre of the tumour. Notably, patients who experienced disease relapse had low immune reaction irrespective of the T stage of the primary tumour (Mlecnik et al., 2011). A standardised system exists for the manual scoring of stromal tumour infiltrating lymphocytes (TILs) on H&E slides in breast cancer (Salgado et al., 2015). The score is a semi-quantitative assessment, expressed as an average across all assessable tumour stroma. The intensity of the baseline immune infiltrate has prognostic and predictive significance in HER2-positive and triple-negative subtypes (Denkert et al., 2018). In triple-negative breast cancer, TILs score predicts pathological and clinical response to checkpoint inhibitors in the neoadjuvant and metastatic settings respectively (Chan et al., 2020). Predictive power may be further increased by combining TILs scores with PDL1 assessment (Gonzalez-Ericsson et al., 2020). The consensus TILs scoring methodology represents a pragmatic approach that has shown good rates of inter-user reproducibility. However, its granularity is limited and it does not attempt to capture detail about how immune cells may be distributed within a specimen. Additionally, even a straightforward manual scoring system is time-consuming to implement at scale, for example, to analyse a trial cohort with thousands of samples.

1.2. Opportunities

A host of clinical trials are currently evaluating novel IO therapies and treatment combinations (Tang et al., 2018). Longitudinal tissue specimens collected from patients undergoing treatment with IO are a valuable source of potential information. Studying changes in the distribution and activity of immune cells with therapeutic intervention, and correlating these with clinical outcomes can provide mechanistic insights into treatment resistance and identify

candidates for predictive biomarkers. In particular, pathological analyses have the advantage of using material such as H&E stained tissue sections, which are widely available and retain information around tissue architecture and spatial organisation. Direct visual assessment of a prepared glass slide using a microscope remains the gold standard in the pathological assessment. However, these traditional manual methods are time-consuming and require a highly trained workforce, which is already under pressure from increasing volume and complexity of histopathology requests (Bainbridge et al., 2016). Use of minimally invasive procedures has expanded at the same time as our interest in tissue biomarkers. Therefore pathologists are being asked to report on ever more complex continuous variables, but with less available tissue. Even for an experienced practitioner, manual techniques are inherently vulnerable to inter-and intra-observer variability. There are natural upper limits on precision and limited scope to describe complex topographical features in an objective and quantifiable manner. Digital approaches offer a potential solution to these issues.

2. Digital pathology and AI: General Principles

In digital pathology (DP), glass-mounted specimens are captured as a whole-slide image (WSI) for downstream computer-based analysis. AI techniques applied to the digitised specimen can utilise various features to perform segmentation and classification tasks. By far the most common AI technique used in these papers and IO research to date is supervised classification. Classification is the task of predicting an output label for each input data point.

Supervised method refers to the fact that the training model is shown example pairs of inputs and labels, and thereby learns the relationship between the two. The model attempts to draw boundaries – implicitly or explicitly – in the input space, separating data points which belong to different classes. Whilst being considerably easier to train than unsupervised techniques, the drawback of supervised methods is their reliance upon the input of large amounts of labelled 'ground truth' data – information collected from the real world, for example, annotations by a pathologist. However, it is worth noting that considerable amounts of annotated data are already in existence within the public domain as well as open-source models and easy-to-use software packages.

Unsupervised methods, on the other hand, usually bypass the need for labelled data (Yamamoto et al., 2019; Cheerla & Gevaert, 2019; Liu et al., 2020b; Ren et al., 2019; Liu et al.,

2020a). Instead, they rely upon the machine being able to discover relevant features for tasks such as grouping together unlabelled data points with high similarity. There are four major types of unsupervised methods (Gentleman & Carey, 2008): (i) exclusive (ii) agglomerative (iii) overlapping and (iv) probabilistic. These models discover unknown patterns in the data, however, in the main, they remain experimental and computationally complex. In specific problems, it can be difficult for the network to converge on a globally optimal solution due to redundant feature representations (Chang et al., 2013) and it is likely to perform less well than supervised training approaches (Zhou et al., 2014). However, such methods may be the best approach for truly novel insights. These techniques involve a diverse set of models and algorithms but all centre around the concept that computers can learn from data as humans learn from experience, and can make decisions about novel data without the need for ongoing instruction. Of particular interest in our setting are deep learning (DL) models. These consist of cascades of trainable, multi-stage layers inspired by the organisation of neurons. A signal input into the model is propagated and modified in a layer-by-layer fashion along these networks to produce an output. DL models have a wide range of architectures themselves, the choice of which depends on the particular task being solved; for example, in image analysis convolutional neural networks (CNNs) (Krizhevsky et al., 2012), generative adversarial networks (GANs) (Goodfellow et al., 2014), fully convolutional neural networks (FCNNs) (Long et al., 2015) and recurrent convolutional neural networks (RCNNs) (Liang & Hu, 2015) are a popular choice.

Histopathological image analysis methods can be broadly categorized into cell-level (identifying/segmenting single cells) or semantic region-based (patch-based; larger extracted patches from whole-slide images, i.e. 512pix×512pix) analysis. Cell-level analysis methods identify structures known as histologic primitives (e.g. nuclei). These features can be correlated with clinical characteristics, such as response to a specific treatment. Early studies applied DL approaches using small patches of manually selected regions of interest extracted from the slides (Raza et al., 2019). For example, object detection can be performed by training a deep CNN on patches centred on the objects of interest such as nuclei. These approaches consider only the information within these size-limited patches, which encompass the object and its immediate neighbourhood, and are mostly suitable for identifying small histologic primitives. Accurate detection of these histologic primitives serves as the basis for a larger number of tasks such as morphological grading, molecular profiling and IO assays. Table 1 gives an overview of small size

level analysis approaches.

The semantic region-based analysis seeks certain special regions inside the whole section like glands, tubules, ducts, etc. These methods are most suitable for identifying meaningful connectives inside an image. Cell level analysis classifies the patches (often small, i.e 56x56 pixels) of an image into different defined classes while semantic region-based analysis can be regarded as semantic identification of objects in a larger image (i.e. 256x256 pixels) in which a pixel-level classification has resulted, i.e. it classifies the pixels into its corresponding classes. Both approaches (cell/semantic region-based methods) can be used for different tasks including segmentation, detection and classification based on the type of annotation and ground truth being used in the methodology set-up. Table 2 gives an overview of region based analysis approaches.

Many reviews of digital analysis of histopathological images exist in the literature and address the various problems associated with the use of different types of histopathology images (Doyle et al., 2008; Gurcan et al., 2009; Irshad et al., 2014; Xing et al., 2017; Pichat et al., 2018; Hamidinekoo et al., 2018; Komura & Ishikawa, 2018; Bera et al., 2019; Niazi et al., 2019). In their recent review (Shimizu et al., 2020) have described numerous recent examples of the applications of AI in oncology and highlight resources and datasets that can help utilise AI tools in cancer research. Table 3 gives an overview of the variety of problems being tackled with DL techniques that are demonstrating promising results.

3. Considerations for Use of AI in Clinical Settings

The backbone of any effective digital pathology service includes (but is not limited to): capturing images using WSI; storing, analysing and archiving the digital images; performing quality control checks; sharing images with other institutions and integrating outputs into clinical decision making. Regulatory requirements and financial viability need to be considered throughout. Workflows require continuous adaptation to evolving demands. In this review, we focus on three main challenges concerning the application of AI algorithms to DP data: (i) generalizability of the model (ii) explainability of the model (iii) limitations on quantity or quality of the data which can be used by the designed model.

Generalizability - This is a measure of how well the complexity of the model matches the complexity of the data. Problems arise when the model has merely memorised training samples but fails to form a general understanding - a problem known as over-fitting. In this case, the model will

perform well with training data but fail to identify relevant information in the novel data. The primary goal, and greatest challenge, for any ML practitioner is for the model to correctly apply what it has learned when unleashed on entirely new data. This is crucial for the deployment of AI in DP across hospitals and laboratories. Table 4 and 5, gives a summary of recent studies in the IO that have evaluated the generalizability of the AI-based models using a large number of internal and external cases. Generalizability may be improved by (i) adjusting network parameters based on the complexity of target data (the greater the number of parameters, the greater the chance of over-fitting); (ii) using dropout neurons (training multiple possible configurations of a network, then calculating the average of all the corresponding subset network weights, which promotes accumulation of independent learning); (iii) weight regularization (to avoid focusing on certain features in the training data, which leads to a continuous increase of weights); (iv) ensuring similar distribution between the training and the upcoming data when deploying the model; (v) frequent re-training rounds (also called fine-tuning) in order to keep up with the change in cohorts.

Explainability - Also known as interpretability, this refers to how well we understand the factors influencing the model's decision making. It is crucial that a model is explainable when used for healthcare purposes, in order to ensure that predictions are being made in an ethical, reliable and transparent manner. Inability to detect bias could have potentially dangerous consequences. Traditional 'bottom-up' ML approaches focus their analysis on specific fundamental characteristics and micro-attributes of a histology image. Deconvoluting the decision-making processes in this scenario is more intuitive and can be approached in several different ways including activation maps (and its derivatives) (Chan et al., 2019a), as well as attention methods (Fraz et al., 2019) and compensating dataset bias and scarcity (Ye et al., 2020).

By contrast, it can be very difficult to identify the salient features being used by the model when using an end-to-end DL approach. For example, Courtiol et al. (2019) identified strongly associated features with either progression/survival; however, some of these features were unexpected (i.e. stromal regions with inflammation and other histological features that were not within the tumour microenvironment). However, progress has been made in this area and there are examples in the literature where DL has yielded biologically interpretable results. For example, Beck et al. (2011) developed a prognostic model incorporating morphometric descriptors and higher-level contextual image features and implicated stromal morphologic structure as a prognostic determinant for breast cancer. Ali et al. (2013) designed spatially aware cell cluster

graphs to predicting tumour outcome in Oropharyngeal p16+ and showed that combining stromal and epithelial nuclear architectural contributions yield superior prognostic performances. Yamamoto et al. (2019) extracted explainable features from histopathology images and several studies have addressed patient stratification by DL methods using H&E images through identifying specific areas of tissue strongly associated with either progression or survival (Steiner et al., 2018; Mobadersany et al., 2018; Liu et al., 2019). As pathologists will retain overall clinical supervision for conclusions drawn from patient samples, transparency is needed in order for them to understand when algorithms should be applied and under what circumstances the output should be used with caution (Huss & Coupland, 2020).

Quantity and quality of data - Digital techniques require the pathology specimens to be scanned at high resolution. Investment in infrastructure is required to cope with this additional step in the pre-diagnostic pipeline, and also to store the colossal amounts of data (e.x, one H&E slide with 20x magnification has a file size of 473,869,300 bytes) with appropriate security considerations and inventory management capabilities. The advent of a graphics processing unit (GPU) based processing, in which vast amounts of data is handled in a parallel fashion has enabled up-scaling to extremely large neural networks which allow huge training sets to be loaded and processed. The quality of the acquired digital images needs to be certified and accepted both by pathologists and the Computer-Assisted Diagnosis system. Presence of artefacts or unintentional loss of information during data acquisition can have a significant influence on down-stream processing. Digital image artefacts may be introduced at any point along the pathway of histopathology slide preparation, from surgical removal through to fixation, tissue processing, embedding, microtomy, staining, mounting, as well as the final digitisation step (Taqi et al., 2018). It is important to be able to identify commonly occurring artefacts such as blurriness, over-straining, air bubbles and colour variation which would adversely affect the interpretation and cause the sample to be diagnostically useless. To address these issues, various preprocessing methods have been proposed to reduce noise: conversion to grayscale, colour normalization (Ehteshami Bejnordi et al., 2015; Ciompi et al., 2017; Khan et al., 2014; Cho et al., 2017) or colouraugmentation (Lafarge et al., 2017; Lin et al., 2017).

Alternatively, Janowczyk et al. (2019) proposed an automated quality control approach to precisely localise artefacts on slides to be avoided during computational analysis. Steiner et al. (2018) have developed a novel convolutional neural network (DeepFocus) to automatically

identify out-of-focus regions in histopathological images. In addition, results of medical interest such as survival prediction are sensitively influenced by the accuracy of the designed algorithm. Most of these medical approaches are supervised methods relying on 'ground truth' information i.e. data collected from the real world. For most problems, the expert opinion of histopathologists and other medical doctors provide the gold standard for training automated decision support systems. However, in many settings, it may be impossible for clinicians to provide this training information with absolute certainty. In summary, although the performance of an algorithm is often measured by accuracy this is not the only feature that is required if the tool is to be of use in everyday applications, including in the field of IO. Training a model on diverse and noisy clinical cohorts will cause accuracy to decrease, but is of pivotal importance in achieving a generalizable algorithm. It is crucial that any model undergoes careful and rigorous validation, preferably within the context of a multicentre prospective trial (Banna et al., 2019). Once applied in real-world scenarios, a clinical team will still be required to make a final judgement on the utility of the output for any individual, bearing in mind the additional context and influencing factors.

4. AI Methodology in the Field of IO

In Table 4, we present some of the DP approaches that have been used to facilitate different pathology workflows for various immune biomarkers, some of which have characterised the TME through spatial analysis and multiplexing. In Table 5, we present non-comprehensive collections of DP approaches that have been used to facilitate different pathology and data integration workflows for IO. These work have characterised the TME through cell analysis, spatial analysis, multiplexing, and omics data integration, which will be divided into 4 sections for in-depth discussions.

4.1. Applications in IO Research

- **Evaluating TME topography** - The functionality of individual cells within the TME is influenced by their precise location, including proximity to other cell types and features of the supporting stroma. Macrophages, for example, display location-dependent phenotypic plasticity; behaviour varies according to whether they are located in the invasive, stromal or hypoxic zones of the tumour (Yang et al., 2018). Single-cell RNA-seq has also contributed to the discovery of functionally distinct cell subsets in the TME, which hold

independent prognostic and predictive value in determining response to immunotherapy (Bartoschek et al., 2018). Tissue sections preserve spatial information and are therefore an ideal substrate for computational analysis of topographical patterns. DL-based image analysis has been used extensively to study the spatial organisation of the immune infiltrate across cancer types, revealing rich and diverse patterns from routine clinical H&E (Failmezger et al., 2019). Effland et al. (2019) demonstrate the use of a machine learning algorithm which can detect immune cells in the immediate neighbourhood of tumour cells. The model could also be used to identify immune cells proximate to other immune cells, and thereby define immune-rich zones. One interesting aspect of this work was the use of an artificial training dataset, generated stochastically from a handful of real-life images. This approach avoids the requirement for extensive numbers of annotations by pathologists but may threaten generalizability. Fibroblasts may provide growth factors and extracellular matrix components providing an extrinsic mechanism of immune-escape. Using a combination of flow cytometry and spatial histology assessment, studies in both breast and pancreatic cancer independently identified specific immunosuppressive fibroblast subsets that localize to the boundary of tumour nests (Costa et al., 2018). The observations of specific spatial compartmentalization of these cell subsets are intriguing, and automated spatial histology analysis could help accelerate and standardize such studies. For example, Failmezger et al. (2019) have recently demonstrated the use of network topological analysis to define a physical barrier of lymphocytic infiltration formed by stromal cells within the TME of metastatic melanoma. In lung cancer, the fractal complexity of the cancer-stromal cell interface has been used to characterize the spatial arrangement of immune cells (AbdulJabbar et al., 2020). The box-counting algorithm, also known as the Minkowski–Bouligand dimension, was modified in order to capture coarse-to-fine geometric details of the cancer-stroma interface over a range of spatial scales determined by cell distributions. Using this method complex morphological patterns dictating cancer-stromal cell contact emerged, which were preserved over varying spatial scales. Fractal dimension was significantly higher in immune-cold tumour regions, and this could not be explained by stromal cell abundance. This supports the conclusion that stroma-based inhibition associated with immune cold phenotypes is a specific morphological pattern. Spatial measures of the immune response such as these have been

shown to correlate with resistance to immunotherapy and with patient outcomes, and therefore have the potential for clinical application as predictive biomarkers.

- **Optimisation of immune scoring** - The availability of AI tools in DP has renewed interests in the development of immune scores for predicting prognosis and response to immunotherapy. Koelzer et al. (2018) demonstrated an example of computational quantitation of membranous PDL1 expression using multiplexed IHC and the HALO™ digital image analysis software. The authors then employed a supervised machine learning algorithm (random forest model) to classify and exclude immune cells from analysis. By restricting PD-L1 scoring to melanoma cells, the authors aimed to reduce apparent heterogeneity which would otherwise lead to artificially high scores. The checkpoint inhibitor ipilumimab is an antibody directed against cytotoxic T-lymphocyte antigen (CTLA-4). There is an unmet need for biomarkers predicting response to CTLA blockade. Harder et al. (2019) used an AI approach to discover novel immune-based signatures associated with clinical response. WSI were generated from melanoma biopsies taken prior to exposure to ipilumimab, slides had been stained for CD3, CD8, and FoxP. Objects of interest (CD4 and CD8 positive cells) stained in a similar way to melanin and therefore a DL classification step was used to identify the immune cells. Image-based features from regions of interest were then extracted and mined for correlation with patient outcomes, although the small sample size was limiting in this study with respect to clinically translatable conclusions. Successful digital approaches to TILs scoring not only enhance speed and precision but also permit the integration of spatial information (Amgad et al., 2020). For example, in early-stage lung cancer, a set of spatial descriptors of co-localisation patterns of TILs and tumour cells were associated with recurrence (Corredor et al., 2019). In bronchoscopic biopsies from pre-invasive lesions, regressive carcinoma-in-situ lesions harbour more infiltrating immune cells, measured by AI and DP, than those that progress to cancer, suggesting that host immune surveillance is strongly implicated in regression of such lesions (Pennycuick et al., 2019). Conversely, the presence of a poorly-infiltrated tumour is a negative prognostic indicator in solid tumours. For example, in one of the first studies to investigate the immune landscape across multiple metastases using pathological samples, the immunoscore for the least immune-infiltrated metastases was found to be the strongest prognosticator in colorectal cancer (Mlecnik et

al., 2018). Similarly, multi-region sampling in lung cancer found a strong association between the number of tumour regions with diminished lymphocytic infiltration and the risk of disease relapse. Prognostic value was independent of tumour size and stage and further validated in an independent cohort of 970 patients with 4,324 multi-region tumour samples, representing the largest multi-region fully automated computational pathology analysis to date (AbdulJabbar et al., 2020). Thus, even if there is above-average immune infiltration across the tumour(s) as a whole, it is the presence of immune-cold regions which appears to drive the clinical outcome and is, therefore, the more significant feature. Automated techniques can enhance our ability to detect such regions. Neural networks enable the integration of heterogeneous data. Reiman and colleagues demonstrated a model which incorporated bulk RNA sequencing data and morphological features from H&E specimens to estimate abundance of immune cell subtypes. This enabled the identification of key effector immune cells without the need for more specialised laboratory techniques such as multiplexed immunofluorescence or single-cell RNA sequencing (Reiman et al., 2019). The approach was flexible and the authors envisioned that additional clinical or molecular information could be incorporated, such as radiological features or data from methylation assays. Thus DP and AI could be applied to the measurement of composite, multi-modality biomarkers.

- **Accounting for intra-tumoural heterogeneity in biomarker development** - When assessing the immunogenicity of a given tissue sample, pathological and molecular approaches may produce discordant results. Spatial heterogeneity may also account, at least in part, for the lack of reproducibility in molecular testing on diagnostic tumour samples, due to sampling bias. Indeed, up to 50% of patients from a multi-region dataset were vulnerable to this issue when using published prognostic signatures (Biswas et al., 2019). Identifying genes expressed uniformly (‘clonally’) across different regions within the same tumour, and deriving a molecular read-out on this basis is likely to be more robust to which part of the tumour is biopsied than conventional methods. The ORACLE signature was significantly associated with mortality in a meta-analysis of 904 lung cancer patients sourced from five separate cohorts. In a study using multi-region sampling, DL pathological image analysis and RNA-sequencing data were derived from the same frozen tissue samples in non-small cell lung cancer (AbdulJabbar et al., 2020). Immune

assessment based on these two data types were in agreement in the majority of samples, with the exception of patients that exhibited high intra-tumoural heterogeneity of immune cell distribution as based on RNA-seq and exome-sequencing data. Moreover, in the discordant tumour regions, pathological images showed a high level of spatial heterogeneity in TIL distribution, measured by immune spatial clustering. Thus, spatial heterogeneity of lymphocyte distribution is likely to be the explanatory factor for the discrepancy between data types generated from adjacent tumour sections. Approaches such as this that consider intra-tumoural heterogeneity may help overcome the reproducibility problem for tumour molecular biomarkers.

- **Deciphering cancer evolution towards immune escape** - The TME can be considered as an ecosystem made up of interacting populations of cancer cells and stroma (Merlo et al., 2006; Weinberg, 2008). Intra-tumoural genetic diversity of cancer cells provides a substrate for evolution according to Darwinian principles (Greaves, 2015). The anti-cancer host immune response, enhanced by IO therapeutics, exerts a selective force which favours expansion of clonal populations that are able to resist this pressure – this is known as immunoediting (Rooney et al., 2015). Immune-escape may be mediated by cancer-cell intrinsic adaptations, such as modulation of immune checkpoint pathways, or through selection advantages conferred by the cancer-associated stroma (Dong et al., 2002; Vinay et al., 2015). By combining pathological immune scoring with sequencing efforts, it has been shown that immune edited tumour clones of colorectal cancer were eliminated while progressing clones were immune-privileged, such that branched evolution across space and time could be traced back to immune-escaping clones (Angelova et al., 2018). In high-grade serous ovarian cancer, a negative association between epithelial CD8+ TILs scored using AI and DP and cancer genetic diversity was found, providing evidence of immunological pruning of tumour clones (Zhang et al., 2018). Thus, DP coupled with omics data will allow the expanded application of these techniques to discover unique spatial signatures that signify immune regulation and evasion.

5. Conclusion

AI and DP tools, tailored for use with routine clinical samples and cutting-edge multiplex tissue imaging techniques have the potential to enable precise descriptions of the complex spatial

organization of the tumour ecosystem to emerge. Integrating this information with genomic and transcriptomic data could unveil mechanisms of immune escape evolving with and without treatment. AI could therefore drive the discovery of novel biomarkers of immune sensitivity and resistance, and identify novel therapeutic targets DL approaches have been popular in early computational pathology efforts. However, there are unavoidable challenges in their application to clinical data. Many current DL algorithms are regarded as 'black box' models, for which it is difficult to produce an explanation for a particular predictive outcome or identify the salient features upon which a decision was made. This is one reason why it has not yet yielded validated, comprehensive, high-level systems. A collaborative approach between data scientists and clinical pathologists in this field will provide the optimal conditions for the development of robust solutions that are sufficiently interpretable to cross into clinical use.

Acknowledgements

F.S. acknowledges funding from NIH U54 CA217376 and R01 CA185138, CDMRP Breast Cancer Research Program Award BC132057. A.H. acknowledges support by Children's Cancer and Leukaemia Group (CCLGA2019 6). N.S. and R.R. acknowledge NHS funding to the NIHR Biomedical Research Centre at The Royal Marsden NHS Foundation Trust and the Institute of Cancer Research.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

The funders had no role in the design of the study; the collection, analysis, or interpretation of the data; the writing of the manuscript; or the decision to submit the manuscript for publication.

Yinyin: Y.Y. has received speakers bureau honoraria from Roche and is a consultant for Merck and Co Inc.

Table 1: Overview of papers using deep learning for digital pathology at cell level for various tasks including detection, segmentation, and classification.

Reference	Topic	Staining	Method
(Sahiner et al., 1996)	Mitosis detection	H&E	CNN-based pixel classifier
(Malon & Cosatto, 2013)	Mitosis detection	H&E	Combines shape based features with CNN
(Wang et al., 2014)	Mitosis detection	H&E	CNN and handcrafted features
(Shkolyar et al., 2015)	Mitosis detection	-	CNN-based patch classifier
(Veta et al., 2016a)	Mitosis detection	-	-
(Chen et al., 2016a)	Mitosis detection	-	convolutional neural network,transfer learning
(Albarqouni et al., 2016)	Mitosis detection	H&E	fCNN, CNN for segmentation
(Mao & Yin, 2016)	Mitosis detection	-	Hierarchical CNNs for patch sequence classificatio
(Litjens et al., 2017)	Mitosis detection	-	survey on nuclei analysis
(Irshad et al., 2014)	nuclei detection	IHC	review on nuclei detection
(Sirinukunwattana et al., 2015)	nuclei detection	-	spatially constrained network
(Xie et al., 2015b)	Nucleus detection	H&E, Ki-67	CNN-based structured regression model
(Xie et al., 2015a)	Nucleus detection	Ki-67	CNN model
(Akram et al., 2016)	cell detection	-	employed the bounding box for cell (nucleus) detection
(Sirinukunwattana et al., 2016)	Nucleus detection	H&E	CNN with
(Kashif et al., 2016)	Nucleus detection	H&E	Combination of CNN and hand-crafted features
(Xu & Huang, 2016)	Nucleus detection	-	General deep learning framework
(Xie et al., 2016a)	Nucleus detection	FL, H&E	fully convolutional regression networks
(Romo-Bucheli et al., 2016)	Tubule nuclei	H&E	CNN-based classification

	detection		
(Jacobs et al.)	Nucleus detection	-	transfer learning
(Turkki et al., 2016)	Nucleus detection	H&E	CNN-based classification of superpixels
(Xu et al., 2016a)	Nucleus detection	H&E	Stacked sparse auto-encoders (SSAE)
(Veta et al., 2016b)	Nuclear area measurement	H&E	CNN
(Chen et al., 2016c)	Nucleus classification	IFL	Deep regression network (DRN)
(Han et al., 2016)	Nucleus classification IFL	H&E	Classification with CNN
(Mishra et al., 2016)	Classification of mitochondria EM	EM	CNN-based patch classifier
(Phan et al., 2016)	Nucleus classification FL	H&E	transfer learning (pre-trained CNN)
(Albarqouni et al., 2016)	Nucleus classification	IHC	CNN framework
(Yao et al., 2016)	Nucleus classification H&E	H&E	-
(Wang et al., 2016b)	Subtype cell detection	H&E	Combination of two CNNs
(Xing et al., 2016)	Nucleus segmentation	H&E, IHC	CNN and selection-based sparse shape model
(Gao et al., 2017)	Nucleus classification IFL	IFL	CNN
(Zhao et al., 2017)	Classification of leukocytes RM	RM	CNN-based patch classifier
(Song et al., 2015)	Nuclei segmentation	H&E	Multi-scale CNN and graph-partitioning-based method
(Ronneberger et al., 2015)	Cell segmentation	-	U-Net with deformation augmentation

(Janowczyk et al., 2016)	Nucleus segmentation H&E	H&E	deep hierarchical learning scheme
(Akram et al., 2016)	Nuclei segmentation	-	extracted bounding box information
(Yang et al., 2016)	Glial cell segmentation TPM	TPM	fCNN with an iterative k-terminal cut algorithm
(Song et al., 2017)	Cell segmentation H&E	H&E	Multi-scale CNN
(Pennycuick et al., 2019)	Cell detection	H&E,IHC	Deconvolving convolutional neural network
(Hagos et al., 2019)	Cell detection	H&E,IHC	Cell Detection
(Zhou et al., 2014)	Tissue classification	-	multispectral unsupervised feature learning

Table 2: Overview of papers using deep learning at tissue level for various tasks including detection, segmentation, and classification.

Reference	Topic	Staining	Method
(Ciresan et al., 2012)	Segmentation of neuronal membranes	EM	Ensemble of several CNNs with different architectures
(Kainz et al., 2015)	Segmentation of colon glands	H&E	Used two CNNs to segment glands
(Apou et al., 2016)	Detection of lobular structures in breast	IHC	CNN and a texture classification
(BenTaieb & Hamarneh, 2016)	Segmentation of colon glands	H&E	fCNN with a loss accounting
(BenTaieb et al., 2016)	Segmentation of colon glands	H&E	A multi-loss fCNN
(Chen et al., 2016d)	Neuronal membrane, fungus segmentation	EM	Combination of bi-directional LSTM-RNNs and kU-Nets
(Chen et al., 2016b)	Segmentation of colon glands	H&E	Deep contour-aware CNN

(Çiçek et al., 2016)	Segmentation of xenopus kidney	CM	3D U-Net
(Drozdzal et al., 2016)	Segmentation of neuronal structures	EM	fCNN with skip connections
(Li et al., 2016)	Segmentation of colon glands	H&E	Compares CNN with an SVM using hand-crafted features
(Wang et al., 2016a)	Segmentation of messy, muscle regions	H&E	Conditional random field jointly trained with an fCNN
(Xie et al., 2016b)	Perimysium segmentation	H&E	2D spatial clockwork RNN
(Xu et al., 2016b)	Segmentation of colon glands	H&E	Used three CNNs to predict gland and contour pixels
(Xie et al., 2015a)	Segmenting epithelium & stroma	H&E. IHC	CNNs applied to over-segmented image regions
(Gecer et al., 2018)	Detection and classification of cancer in whole slide breast	H&E	detection, classification and pixel-wise labeling of WSI
(Rodner et al., 2019)	Pixel-wise classification	H&E,IHC	semantic segmentation using a FCN

Table 3: Overview of held challenges in the field of digital pathology.

Name	Aims	tissue	Dataset released			Year	Provided ground-truth
			Staining	Training	Testing		
ICPR https://mitos-atypia-14.grand-challenge.org/	mitosis detection,	breast	H&E	32 WSIs		2014	centroids of mitosis,

	nuclear atypia score						nuclear atypia score
GlaS https://warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/	gland segmentation	colon	H&E	85 images	80 images	2015	binary masks
BioImaging http://www.bioimaging2015.ineb.up.pt/challenge_overview.html	ccancer classification	breast	H&E	140 images of 2048×1536	20 images	2015	labels
TUMAC http://tupac.tue-image.nl	tumor detection	breast	H&E	573 WSIs	321 WSIs	2016	tumor proliferation score, molecular proliferation score.
CAMELYON'16 https://camelyon16.grand-challenge.org/	detection of cancer metastasis	breast	H&E	270 WSIs	130 WSIs	2016	annotated contours, binary masks
HER2 Scoring https://warwick.ac.uk/fac/sci/dcs/research/	HER2 scoring	breast	IHC	100 WSIs		2020	HER2 and

arch/tia/her2contest	g			16	%age scores
TMA analysis in thyroid cancer diagnosis http://www-o.ntust.edu.tw/~cvmi/ISBI2017/	cancer diagnosis	thyroid	H&E, IHC	28 TMAs, 616 tissue cores	2017	-
CAMELYON'17 https://camelyon17.grand-challenge.org	detection of cancer metastasis	breast	H&E	1399 WSIs	2017	metastases annotations in WSI, patient pN-stage label
BACH https://iciar2018-challenge.grand-challenge.org/	classification and pixel-wise labelling of WSIs	breast	H&E	400+ images, 20 WSIs, 10 WSIs	2018	pixel-wise labels
PatchCamelyon https://patchcamelyon.grand-challenge.org	metastasis detection	lymph node		327,680 images	2018	binary label indicating presence of

						metastatic tissue
ACDC-LungHP https://acdc-lunghp.grand-challenge.org/	cancer detection, classification	lung	H&E	150 WSIs	50 WSIs	2019	annotation of cancer regions
ANHIR https://anhir.grand-challenge.org/Intro/	image registration	lesions, lung-lobes, mammary-gland	H&E, IHC	50+ WSIs		2019	-
LYSTO https://lysto.grand-challenge.org/LYSTO	assessment of lymphocytes	breast, colon and prostate	IHC	20,000 patches of size 299×299	12,000 patches	2019	number of lymphocytes for each patch
DigestPath https://digestpath2019.grand-challenge.org/Home/		mucus-secreting glands	H&E	99 WSIs	56 WSIs	2019	cell bounding boxes
PAIP https://paip2019.grand-challenge.org/Home/	liver cancer Segmentati	liver	H&E	60 WSIs	40 WSIs	2019	tumor area segmentatio

		on						n, viable tumor area
CodaLab https://competitions.codalab.org/competitions/20395#learn_the_details-overview	classification	blood normal cells	-	73 cases	45 cases	2019		lables
LYON https://lyon19.grand-challenge.org/Home/	lymphocyte detection	breast, colon and prostate	IHC	no training data	441 region of interests	2019		-
ECDP2020 https://ecdp2020.grand-challenge.org/Home/	identify HER2+ from HER2-	breast	H&E	360 WSIs		2020		-
Gleason https://gleason2019.grand-challenge.org	gleason grading	prostate	TMA (H&E)	245 cores	88 cores	2019		maps and labels

Table 4: Overview of different pathology workflows for various immune biomarkers that have been addressed by deep learning approaches.

Reference	Aims	Methodology		Dataset used		Results
		Task	Approach	Tissu	Modality	

				e		
(Turkki et al., 2016)	quantification of tumor-infiltrating immune cells	supervised classification of immune cell-rich/poor regions	1-features extraction by CNN 2- binary classification by SVM	breast	H&E, CD45	F-score=0.94; $K_{DL=0.79}$ vs $K_{manual=078}$
(Saltz et al., 2018)	spatial organization and molecular correlation of TIL maps with survival, tumor subtypes, and immune profiles	1-supervised classification of patches with low/high lymphocyte by CNN; 2-Supervised segmentation of necrosis regions by CNN	lymphocyte and necrosis semi-supervised CNN	various	H&E, molecular data	-
(Mezheyeuski et al., 2018)	quantification of immune infiltrates in situ in the environment of epithelial and stromal compartments	-	-	lung	TMA (including: CD8, CD20, CD4, FOXP3, CD45RO, and pan-cytokeratin)	correlation of DL vs manual lymphocytes quantification for: CD45RO (R = 0.52), FOXP3 (R= 0.87), CD4 (R= 0.79), CD20 (R= 0.81),CD8 (R= 0.90)
(Turkki et al., 2019)	patient outcome prediction	supervised classification of samples	1-feature extraction with a deep CNN; 2-	breast	TMA	$ACC_{automated}$=0.60 (95% CI 0.55-0.65) vs

		into low/high digital risk score	feature pooling with IFV; 3-PCA; 4-classification with SVM			ACC_{manual}=0.58 (95% CI 0.53-0.63)
(Amgad et al., 2019)	region and nucleus segmentation for characterization of TILs	1-supervised classification of histologic compartments; 2-segmentation of nucleus; 3- calculate TIL scores;	1-FCN to output a combined mask. 2-decomposing output for region and nucleus segmentation; 3-seed classifications from the cell segmentation.	breast	H&E	Dice=0.78, ROC-AUC=0.89, R=0.73, p<0.001
(Aprupe et al., 2019)	quantification of biomarkers of immune cells	Supervised binary classification	1-features extraction by a CNN; 2- binary classification by softmax	lung	CD3, CD8, CD20	cell count difference to humans=0.033 cells on average
(Koelzer et al., 2019)	precision immunoprofiling, digital scoring of PD-L1 expression		characterization of the tumor microenvironment through spatial analysis and multiplexing; spatial analysis of T-cell	colon	TMA	-

			infiltrationn			
(Bug et al., 2019)	checkpoint inhibitor response prediction using patient derived xenografts in humanized mice	tissue classification using HistoNet model with eight distinct classes	automatic extraction of meta-features for the characterization of the tumor	H&E	lung (mouse-trial)	F1-score of 83%; $ACC_{tumor\text{-}response}$=84%

Table 5: Overview of different collections of DP approaches that have been used to facilitate data integration work-flows for IO

Reference	Topics	Aim	Summary
(Schmauch et al., 2020)	A deep learning model to predict RNA-Seq expression of tumor from Whole slide images	Predict RNA-Seq profiles form Whole-slide images	The developed model (HE2RNA) could predict subsets of genes expressed in different cancer types and the expression of a subset of protein-coding genes. It could also quantify immune infiltration, including genes involved in immune cell activation status and immune cell signalling
(Gamper et al., 2020)	PanNuke Dataset Extension, Insights and Baselines	release of PanNuke dataset for nucleus segmentation and classification; eliminate the process of	Comparing instance segmentation performance of several models using the

		verification and quality control by the clinical professionals and incorporating	prepared PanNuke dataset. The models trained on PanNuke generalise to other inseen tissues.
(Fu et al., 2020)	Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis	pan-cancer computational histopathology (PC-CHiP) study associations between computational histopathological features and genomic driver alterations, whole transcriptomes and survival	Pan-cancer computational histopathology analysis with deep learning extracts histopathological patterns and accurately discriminates 28 cancer and 14 normal tissue types, Computational histopathology predicts whole-genome duplications, focal amplifications and deletions, as well as driver gene mutations
(Kather et al., 2020)	Pan-cancer image-based detection of clinically actionable genetic alterations	deep learning can predict point mutations, molecular tumor subtypes and immune-related gene expression signatures directly from routine histological images of tumor tissue	deep learning can predict point mutations, molecular tumor subtypes and immune-related gene expression signatures directly from routine histological images of tumor tissue
(Mobadersany et	Predicting cancer	developed a computational	Approach surpasses the

al., 2018)	outcomes from histology and genomics using convolutional networks	approach based on DL to predict the overall survival of patients diagnosed with brain tumours from microscopic images of tissue biopsies and genomic biomarkers, present an approach called survival convolutional neural networks (SCNNs), which provide a highly accurate prediction of time-to-event outcomes from histology images	prognostic accuracy of human experts using the current clinical standard for classifying brain tumours and presents an innovative approach for the objective, accurate, and integrated prediction of patient outcomes.
(Narayanan et al., 2019)	Unmasking the tissue microecology of ductal carcinoma in situ with deep learning	Automate the identification of DCIS; quantify the spatial relationship of DCIS with TILs, providing a new way to study immune response and identify new markers of progression improving clinical management.	Developed a deep learning pipeline that integrates tissue segmentation, DCIS segmentation, single cell classification and spatial analysis in routine H&E histology images.
(Pantanowitz et al., 2020)	An artificial intelligence algorithm for prostate cancer diagnosis in WSI of core needle biopsies: a blinded clinical validation and deployment study	Predicting slide-level scores for probability of cancer, Gleason score, Gleason pattern, and perineural invasion and calculation of cancer percentage present in CNB material.	the trained model was tested on internal and external datasets elucitating generalisibity of the algorithms.

References

AbdulJabbar, K., Raza, S. E. A., Rosenthal, R., Jamal-Hanjani, M., Veeriah, S., Akarca, A., Lund, T., Moore, D. A., Salgado, R., Al Bakir, M., Zapata, L., Hiley, C. T., Officer, L., Sereno, M., Smith, C., Loi, S., Marafioti, T., Quezada, S. A., McGranahan, N., Quesne, J. L., & Yuan, Y. (2020). Geospatial immune variability illuminates differential evolution of lung adenocarcinoma. *Nature Medicine*,.

Akram, S. U., Kannala, J., Eklund, L., & Heikkilä, J. (2016). Cell segmentation proposal network for microscopy image analysis. In *International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis* (pp. 21–29). Springer.

Albarqouni, S., Baur, C., Achilles, F., Belagiannis, V., Demirci, S., & Navab, N. (2016). Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images. *IEEE transactions on medical imaging*, *35*, 1313–1321.

Ali, S., Lewis, J., & Madabhushi, A. (2013). Spatially aware cell cluster (spaccl) graphs: predicting outcome in oropharyngeal p16+ tumors. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 412–419). Springer.

Amgad, M., Sarkar, A., Srinivas, C., Redman, R., Ratra, S., Bechert, C. J., Calhoun, B. C., Mrazeck, K., Kurkure, U., Cooper, L. A. et al. (2019). Joint region and nucleus segmentation for characterization of tumor infiltrating lymphocytes in breast cancer. In *Medical Imaging 2019: Digital Pathology* (p. 109560M). International Society for Optics and Photonics volume 10956.

Amgad, M., Stovgaard, E. S., Balslev, E., Thagaard, J., Chen, W., Dudgeon, S., Sharma, A., Kerner, J. K., Denkert, C., Yuan, Y. et al. (2020). Report on computational assessment of tumor infiltrating lymphocytes from the international immuno-oncology biomarker working group. *NPJ breast cancer*, *6*, 1–13.

Angelova, M., Mlecnik, B., Vasaturo, A., Bindea, G., Fredriksen, T., Lafontaine, L., Buttard, B., Morgand, E., Bruni, D., Jouret-Mourin, A. et al. (2018). Evolution of metastases in space and time under immune selection. *Cell*, *175*, 751–765.

Apou, G., Schaadt, N. S., Naegel, B., Forestier, G., Schönmeyer, R., Feuerhake, F., Wemmert, C., & Grote, A. (2016). Detection of lobular structures in normal breast tissue. *Computers in*

biology and medicine, *74*, 91–102.

Aprupe, L., Litjens, G., Brinker, T. J., van der Laak, J., & Grabe, N. (2019). Robust and accurate quantification of biomarkers of immune cells in lung cancer micro-environment using deep convolutional neural networks. *PeerJ*, *7*, e6335.

Bainbridge, S., Cake, R., Meredith, M., Furness, P., & Gordon, B. (2016). Testing times to come? an evaluation of digital pathology capacity across the uk. *Cancer Research UK*,.

Balar, A. V., & Weber, J. S. (2017). Pd-1 and pd-l1 antibodies in cancer: current status and future directions. *Cancer Immunology, Immunotherapy*, *66*, 551–564.

Banna, G. L., Oliver, T., Rundo, F., Malapelle, U., Fraggetta, F., Libra, M. L., & Addeo, A. (2019). The promise of digital biopsy for the prediction of tumor molecular features and clinical outcomes associated with immunotherapy. *Frontiers in medicine*, *6*, 172.

Bartoschek, M., Oskolkov, N., Bocci, M., Lövrot, J., Larsson, C., Sommarin, M., Madsen, C. D., Lindgren, D., Pekar, G., Karlsson, G. et al. (2016). Spatially and functionally distinct subclasses of breast cancer-associated fibroblasts revealed by single cell rna sequencing. *Nature Communications*, *9*, 1–13.

Beck, A. H., Sangoi, A. R., Leung, S., Marinelli, R. J., Nielsen, T. O., Van De Vijver, M. J., West, R. B., Van De Rijn, M., & Koller, D. (2011). Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Science translational medicine*, *3*, 108ra113–108ra113.

BenTaieb, A., & Hamarneh, G. (2016). Topology aware fully convolutional networks for histology gland segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 460–468). Springer.

BenTaieb, A., Kawahara, J., & Hamarneh, G. (2016). Multi-loss convolutional networks for gland analysis in microscopy. In *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on* (pp. 642–645). IEEE.

Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V., & Madabhushi, A. (2019). Artificial intelligence in digital pathologyâ€"new tools for diagnosis and precision oncology. *Nature Reviews Clinical Oncology*, *16*, 703–715.

Bertin-Ciftci, J., Barré, B., Le Pen, J., Maillet, L., Couriaud, C., Juin, P., & Braun, F. (2013). prb/e2f-1-mediated caspase-dependent induction of noxa amplifies the apoptotic effects of the bcl-2/bcl-xl inhibitor abt-737. *Cell Death & Differentiation*, *20*, 755–764.

Biswas, D., Birkbak, N. J., Rosenthal, R., Hiley, C. T., Lim, E. L., Papp, K., Boeing, S., Krzystanek, M., Djureinovic, D., La Fleur, L. et al. (2019). A clonal expression biomarker associates with lung cancer mortality. *Nature Medicine*, *25*, 1540–1548.

Bug, D., Feuerhake, F., Oswald, E., Schüler, J., & Merhof, D. (2019). Semi-automated analysis of digital whole slides from humanized lung-cancer xenograft models for checkpoint inhibitor response prediction. *Oncotarget*, *10*, 4587.

Chan, J. J., Tan, T. J., & Dent, R. A. (2020). Integrating immunotherapy in the (neo) adjuvant setting of early breast cancer. *Current Opinion in Oncology*, *32*, 575–584.

Chan, L., Hosseini, M. S., Rowsell, C., Plataniotis, K. N., & Damaskinos, S. (2019a). Histosegnet: Semantic segmentation of histological tissue type in whole slide images. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 10662–10671).

Chan, T. A., Yarchoan, M., Jaffee, E., Swanton, C., Quezada, S. A., Stenzinger, A., & Peters, S. (2019b). Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Annals of Oncology*, *30*, 44–56.

Chang, H., Nayak, N., Spellman, P. T., & Parvin, B. (2013). Characterization of tissue histopathology via predictive sparse decomposition and spatial pyramid matching. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 91–98). Springer.

Cheerla, A., & Gevaert, O. (2019). Deep learning with multimodal representation for pancancer prognosis prediction. *Bioinformatics*, *35*, i446–i454.

Chen, H., Dou, Q., Wang, X., Qin, J., Heng, P.-A. et al. (2016a). Mitosis detection in breast cancer histology images via deep cascaded networks. In *AAAI* (pp. 1160–1166).

Chen, H., Qi, X., Yu, L., & Heng, P.-A. (2016b). Dcan: Deep contour-aware networks for accurate gland segmentation. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 2487–2496).

Chen, H., Wang, X., & Heng, P. A. (2016c). Automated mitosis detection with deep regression networks. In *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on* (pp. 1204–1207). IEEE.

Chen, J., Yang, L., Zhang, Y., Alber, M., & Chen, D. Z. (2016d). Combining fully convolutional and recurrent neural networks for 3d biomedical image segmentation. In *Advances in Neural Information Processing Systems* (pp. 3036–3044).

Cho, H., Lim, S., Choi, G., & Min, H. (2017). Neural stain-style transfer learning using gan for histopathological images. *arXiv preprint arXiv:1710.08543,*.

Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 424–432). Springer.

Ciompi, F., Geessink, O., Bejnordi, B. E., de Souza, G. S., Baidoshvili, A., Litjens, G., van Ginneken, B., Nagtegaal, I., & van der Laak, J. (2017). The importance of stain normalization in colorectal tissue classification with convolutional networks. In *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on* (pp. 160–163). IEEE.

Ciresan, D., Giusti, A., Gambardella, L. M., & Schmidhuber, J. (2012). Deep neural networks segment neuronal membranes in electron microscopy images. In *Advances in neural information processing systems* (pp. 2843–2851).

Corredor, G., Wang, X., Zhou, Y., Lu, C., Fu, P., Syrigos, K., Rimm, D. L., Yang, M., Romero, E., Schalper, K. A. et al. (2019). Spatial architecture and arrangement of tumor-infiltrating lymphocytes for predicting likelihood of recurrence in early-stage non–small cell lung cancer. *Clinical Cancer Research*, *25*, 1526–1534.

Costa, A., Kieffer, Y., Scholer-Dahirel, A., Pelon, F., Bourachot, B., Cardon, M., Sirven, P., Magagna, I., Fuhrmann, L., Bernard, C. et al. (2018). Fibroblast heterogeneity and immunosuppressive environment in human breast cancer. *Cancer Cell*, *33*, 463–479.

Courtiol, P., Maussion, C., Moarii, M., Pronier, E., Pilcer, S., Sefta, M., Manceron, P., Toldo, S., Zaslavskiy, M., Le Stang, N. et al. (2019). Deep learning-based classification of mesothelioma improves prediction of patient outcome. *Nature Medicine*, *25*, 1519–1525.

Denkert, C., von Minckwitz, G., Darb-Esfahani, S., Lederer, B., Heppner, B. I., Weber, K. E., Budczies, J., Huober, J., Klauschen, F., Furlanetto, J. et al. (2018). Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *The lancet oncology*, *19*, 40–50.

Dong, H., Strome, S. E., Salomao, D. R., Tamura, H., Hirano, F., Flies, D. B., Roche, P. C., Lu, J., Zhu, G., Tamada, K. et al. (2002). Tumor-associated b7-h1 promotes t-cell apoptosis: a potential mechanism of immune evasion. *Nature medicine*, *8*, 793–800.

Doyle, S., Agner, S., Madabhushi, A., Feldman, M., & Tomaszewski, J. (2008). Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features. In *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on* (pp. 496–499). IEEE.

Drozdzal, M., Vorontsov, E., Chartrand, G., Kadoury, S., & Pal, C. (2016). The importance of skip connections in biomedical image segmentation. In *Deep Learning and Data Labeling for Medical Applications* (pp. 179–187). Springer.

Effland, A., Kobler, E., Brandenburg, A., Klatzer, T., Neuhäuser, L., Hölzel, M., Landsberg, J., Pock, T., & Rumpf, M. (2019). Joint reconstruction and classification of tumor cells and cell interactions in melanoma tissue sections with synthesized training data. *International journal of computer assisted radiology and surgery*, *14*, 587–599.

Ehteshami Bejnordi, B., Litjens, G., Timofeeva, N., Otte-Höller, I., Homeyer, A., Karssemeijer, N., & van der Laak, J. (2015). Stain specific standardization of whole-slide histopathological images,.

Failmezger, H., Muralidhar, S., Rullan, A., de Andrea, C. E., Sahai, E., & Yuan, Y. (2019). Topological tumor graphs: a graph-based spatial model to infer stromal recruitment for immunosuppression in melanoma histology. *Cancer Research*,.

Fraz, M., Khurram, S., Graham, S., Shaban, M., Hassan, M., Loya, A., & Rajpoot, N. (2019). Fabnet: feature attention-based network for simultaneous segmentation of microvessels and nerves in routine histology images of oral cancer. *Neural Computing and Applications*, (pp. 1–14).

Fu, Y., Jung, A. W., Torne, R. V., Gonzalez, S., Vöhringer, H., Shmatko, A., Yates, L. R., Jimenez-Linan, M., Moore, L., & Gerstung, M. (2020). Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. *Nature Cancer*, (pp. 1–11).

Gamper, J., Koohbanani, N. A., Graham, S., Jahanifar, M., Khurram, S. A., Azam, A., Hewitt, K., & Rajpoot, N. (2020). Pannuke dataset extension, insights and baselines. *arXiv preprint arXiv:2003.10778*,.

Gao, Z., Wang, L., Zhou, L., & Zhang, J. (2017). Hep-2 cell image classification with deep convolutional neural networks. *IEEE journal of biomedical and health informatics*, *21*, 416–428.

Gecer, B., Aksoy, S., Mercan, E., Shapiro, L. G., Weaver, D. L., & Elmore, J. G. (2018). Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks. *Pattern recognition*, *84*, 345–356.

Gentleman, R., & Carey, V. J. (2008). Unsupervised machine learning. In *Bioconductor case studies* (pp. 137–157). Springer.

Gibney, G. T., Weiner, L. M., & Atkins, M. B. (2016). Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *The Lancet Oncology*, *17*, e542–e551.

Gonzalez-Ericsson, P. I., Stovgaard, E. S., Sua, L. F., Reisenbichler, E., Kos, Z., Carter, J. M., Michiels, S., Le Quesne, J., Nielsen, T. O., Lænkholm, A.-V. et al. (2020). The path to a better biomarker: application of a risk management framework for the implementation of pd-l1 and tils as immuno-oncology biomarkers in breast cancer clinical trials and daily practice. *The Journal of pathology*, *250*, 667–684.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672–2680)

Greaves, M. (2015). Evolutionary determinants of cancer. *Cancer Discovery*, *5*, 806–820.

Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., & Yener, B. (2009). Histopathological image analysis: A review. *IEEE reviews in biomedical engineering*, *2*, 147–171.

Hagos, Y. B., Narayanan, P. L., Akarca, A. U., Marafioti, T., & Yuan, Y. (2019). Concorde-net: Cell count regularized convolutional neural network for cell detection in multiplex immunohistochemistry images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 667–675). Springer.

Hamidinekoo, A., Denton, E., Rampun, A., Honnor, K., & Zwiggelaar, R. (2018). Deep learning in mammography and breast histology, an overview and future trends. *Medical Image Analysis*, *47*, 45–67.

Han, X.-H., Lei, J., & Chen, Y.-W. (2016). Hep-2 cell classification using k-support spatial pooling in deep cnns. In *Deep Learning and Data Labeling for Medical Applications* (pp. 3–11). Springer.

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *cell*, *144*, 646–674.

Harder, N., Schönmeyer, R., Nekolla, K., Meier, A., Brieu, N., Vanegas, C., Madonna, G., Capone, M., Botti, G., Ascierto, P. A. et al. (2019). Automatic discovery of image-based signatures for ipilimumab response prediction in malignant melanoma. *Scientific reports*, *9*, 1–19.

Hellmann, M. D., Ciuleanu, T.-E., Pluzanski, A., Lee, J. S., Otterson, G. A., Audigier-Valette, C., Minenza, E., Linardou, H., Burgers, S., Salman, P. et al. (2018). Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. *New England Journal of Medicine*, *378*, 2093–2104.

Huss, R., & Coupland, S. E. (2020). Software-assisted decision support in digital histopathology. *The Journal of Pathology*, *250*, 685–692.

Irshad, H., Veillard, A., Roux, L., & Racoceanu, D. (2014). Methods for nuclei detection, segmentation, and classification in digital histopathology: a reviewâ€”current status and future potential. *IEEE reviews in biomedical engineering*, *7*, 97–114.

Jacobs, J. G., Brostow, G. J., Freeman, A., Alexander, D. C., & Panagiotaki, E. (). Detecting and classifying nuclei on a budget,.

Janowczyk, A., Doyle, S., Gilmore, H., & Madabhushi, A. (2016). A resolution adaptive deep hierarchical (radhical) learning scheme applied to nuclear segmentation of digital pathology images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, (pp. 1–7).

Janowczyk, A., Zuo, R., Gilmore, H., Feldman, M., & Madabhushi, A. (2019). Histoqc: an open-source quality control tool for digital pathology slides. *Journal of Clinical Oncology, Clinical Cancer Informatics*, *3*, 1–7.

Kainz, P., Pfeiffer, M., & Urschler, M. (2015). Semantic segmentation of colon glands with deep convolutional neural networks and total variation segmentation. *arXiv preprint arXiv:1511.06919*,.

Kashif, M. N., Raza, S. E. A., Sirinukunwattana, K., Arif, M., & Rajpoot, N. (2016). Handcrafted features with convolutional neural networks for detection of tumor cells in histology images. In *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on* (pp. 1029–1032). IEEE.

Kather, J. N., Heij, L. R., Grabsch, H. I., Loeffler, C., Echle, A., Muti, H. S., Krause, J., Niehues, J. M., Sommer, K. A., Bankhead, P. et al. (2020). Pan-cancer image-based detection of

clinically actionable genetic alterations. *Nature Cancer*, (pp. 1–11).

Khan, A. M., Rajpoot, N., Treanor, D., & Magee, D. (2014). A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution. *Biomedical Engineering, IEEE Transactions on*, *61*, 1729–1738.

Koelzer, V. H., Gisler, A., Hanhart, J. C., Griss, J., Wagner, S. N., Willi, N., Cathomas, G., Sachs, M., Kempf, W., Thommen, D. S. et al. (2018). Digital image analysis improves precision of pd-l1 scoring in cutaneous melanoma. *Histopathology*, *73*, 397–406.

Koelzer, V. H., Sirinukunwattana, K., Rittscher, J., & Mertz, K. D. (2019). Precision immunoprofiling by image analysis and artificial intelligence. *Virchows Archiv*, *474*, 511–522.

Komura, D., & Ishikawa, S. (2018). Machine learning methods for histopathological image analysis. *Computational and Structural Biotechnology Journal*, *16*, 34–42.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097–1105).

Lafarge, M. W., Pluim, J. P., Eppenhof, K. A., Moeskops, P., & Veta, M. (2017). Domain-adversarial neural networks to address the appearance variability of histopathology images. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (pp. 83–91). Springer.

Li, W., Manivannan, S., Akbar, S., Zhang, J., Trucco, E., & McKenna, S. J. (2016). Gland segmentation in colon histology images using hand-crafted features and convolutional neural networks. In *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on* (pp. 1405–1408). IEEE.

Liang, M., & Hu, X. (2015). Recurrent convolutional neural network for object recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3367–3375).

Lin, H., Chen, H., Dou, Q., Wang, L., Qin, J., & Heng, P.-A. (2017). Scannet: A fast and dense scanning framework for metastatic breast cancer detection from whole-slide images. *arXiv preprint arXiv:1707.09597*,.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Van Der Laak, J. A., Van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical

image analysis. *Medical Image Analysis*, *42*, 60–88.

Liu, D., Zhang, D., Song, Y., Zhang, F., O'Donnell, L., Huang, H., Chen, M., & Cai, W. (2020a). Pdam: A panoptic-level feature alignment framework for unsupervised domain adaptive instance segmentation in microscopy images. *IEEE Transactions on Medical Imaging,*.

Liu, D., Zhang, D., Song, Y., Zhang, F., O'Donnell, L., Huang, H., Chen, M., & Cai, W. (2020b). Unsupervised instance segmentation in microscopy images via panoptic domain adaptation and task re-weighting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4243–4252).

Liu, Y., Kohlberger, T., Norouzi, M., Dahl, G. E., Smith, J. L., Mohtashamian, A., Olson, N., Peng, L. H., Hipp, J. D., & Stumpe, M. C. (2019). Artificial intelligence–based breast cancer nodal metastasis detection: Insights into the black box for pathologists. *Archives of pathology & laboratory medicine*, *143*, 859–868

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431–3440).

Lu, S., Stein, J. E., Rimm, D. L., Wang, D. W., Bell, J. M., Johnson, D. B., Sosman, J. A., Schalper, K. A., Anders, R. A., Wang, H. et al. (2019). Comparison of biomarker modalities for predicting response to pd-1/pd-l1 checkpoint blockade: a systematic review and meta-analysis. *JAMA oncology*, *5*, 1195–1204.

Luchini, C., Bibeau, F., Ligtenberg, M., Singh, N., Nottegar, A., Bosse, T., Miller, R., Riaz, N., Douillard, J.-Y., Andre, F. et al. (2019). Esmo recommendations on microsatellite instability testing for immunotherapy in cancer, and its relationship with pd-1/pd-l1 expression and tumour mutational burden: a systematic review-based approach. *Annals of Oncology*, *30*, 1232–1243.

Malon, C. D., & Cosatto, E. (2013). Classification of mitotic figures with convolutional neural networks and seeded blob features. *Journal of pathology informatics*, *4*.

Mao, Y., & Yin, Z. (2016). A hierarchical convolutional neural network for mitosis detection in phase-contrast microscopy images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 685–692). Springer.

Merlo, L. M., Pepper, J. W., Reid, B. J., & Maley, C. C. (2006). Cancer as an evolutionary and ecological process. *Nature Reviews Cancer*, *6*, 924–935.

Mezheyeuski, A., Bergsland, C. H., Backman, M., Djureinovic, D., Sjöblom, T., Bruun, J., & Micke, P. (2018). Multispectral imaging for quantitative and compartment-specific immune infiltrates reveals distinct immune profiles that classify lung cancer patients. *The Journal of Pathology*, *244*, 421–431.

Mishra, M., Schmitt, S., Wang, L., Strasser, M. K., Marr, C., Navab, N., Zischka, H., & Peng, T. (2016). Structure-based assessment of cancerous mitochondria using deep networks. In *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on* (pp. 545–548). IEEE.

Mlecnik, B., Van den Eynde, M., Bindea, G., Church, S. E., Vasaturo, A., Fredriksen, T., Lafontaine, L., Haicheur, N., Marliot, F., Debetancourt, D. et al. (2018). Comprehensive intrametastatic immune quantification and major impact of immunoscore on survival. *JNCI: Journal of the National Cancer Institute*, *110*, 97–108.

Mlecnik, B., Tosolini, M., Kirilovsky, A., Berger, A., Bindea, G., Meatchi, T., Bruneval, P., Trajanoski, Z., Fridman, W.-H., Pages, F. et al. (2011). Histopathologic-based prognostic factors of colorectal cancers are associated with the state of the local immune reaction. *Journal of clinical oncology*, *29*, 610–618.

Mobadersany, P., Yousefi, S., Amgad, M., Gutman, D. A., Barnholtz-Sloan, J. S., Vega, J. E. V., Brat, D. J., & Cooper, L. A. (2018). Predicting cancer outcomes from histology and genomics using convolutional networks. *Proceedings of the National Academy of Sciences*, *115*, E2970–E2979.

Narayanan, P. L., Raza, S. E. A., Hall, A. H., Marks, J. R., King, L., West, R. B., Hernandez, L., Dowsett, M., Gusterson, B., Maley, C. et al. (2019). Unmasking the tissue microecology of ductal carcinoma in situ with deep learning. *BioRxiv*, (p. 812735).

Niazi, K. A. K., Akhtar, W., Khan, H. A., Yang, Y., & Athar, S. (2019). Hotspot diagnosis for solar photovoltaic modules using a naive bayes classifier. *Solar Energy*, *190*, 34–43.

Pantanowitz, L., Quiroga-Garza, G. M., Bien, L., Heled, R., Laifenfeld, D., Linhart, C., Sandbank, J., Shach, A. A., Shalev, V., Vecsler, M. et al. (2020). An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: a blinded clinical validation and deployment study. *The Lancet Digital Health*, *2*, e407–e416.

Pennycuick, A., Teixeira, V. H., AbdulJabbar, K., Raza, S. E. A., Lund, T., Akarca, A., Rosenthal, R., Pipinikas, C. P., Lee-Six, H., Chandrasekharan, D. P. et al. (2019). Immune

surveillance in clinical regression of pre-invasive squamous cell lung cancer. *bioRxiv*, (p. 833004).

Phan, H. T. H., Kumar, A., Kim, J., & Feng, D. (2016). Transfer learning of a convolutional neural network for hep-2 cell image classification. In *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on* (pp. 1208–1211). IEEE.

Pichat, J., Iglesias, J. E., Yousry, T., Ourselin, S., & Modat, M. (2018). A survey of methods for 3d histology reconstruction. *Medical Image Analysis*, *46*, 73–105.

Raza, S. E. A., AbdulJabbar, K., Jamal-Hanjani, M., Veeriah, S., Le Quesne, J., Swanton, C., & Yuan, Y. (2019). Deconvolving convolutional neural network for cell detection. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* (pp. 891–894). IEEE.

Reiman, D., Sha, L., Ho, I., Tan, T., Lau, D., & Khan, A. A. (2019). Integrating rna expression and visual features for immune infiltrate prediction. In *PSB* (pp. 284–295). World Scientific.

Ren, J., Hacihaliloglu, I., Singer, E. A., Foran, D. J., & Qi, X. (2019). Unsupervised domain adaptation for classification of histopathology whole-slide images. *Frontiers in bioengineering and biotechnology*, *7*, 102.

Rodner, E., Bocklitz, T., von Eggeling, F., Ernst, G., Chernavskaia, O., Popp, J., Denzler, J., & Guntinas-Lichius, O. (2019). Fully convolutional networks in multimodal nonlinear microscopy images for automated detection of head and neck carcinoma: Pilot study. *Head & Neck*, *41*, 116–121.

Romo-Bucheli, D., Janowczyk, A., Gilmore, H., Romero, E., & Madabhushi, A. (2016). Automated tubule nuclei quantification and correlation with oncotype dx risk categories in er+ breast cancer whole slide images. *Scientific reports*, *6*, 32706.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 234–241). Springer.

Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G., & Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*, *160*, 48–61.

Sahiner, B., Chan, H.-P., Petrick, N., Wei, D., Helvie, M. A., Adler, D. D., & Goodsitt, M. M. (1996). Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. *IEEE transactions on Medical Imaging*,

15, 598–610.

Salgado, R., Denkert, C., Demaria, S., Sirtaine, N., Klauschen, F., Pruneri, G., Wienert, S., Van den Eynden, G., Baehner, F. L., Pénault-Llorca, F. et al. (2015). The evaluation of tumor-infiltrating lymphocytes (tils) in breast cancer: recommendations by an international tils working group 2014. *Annals of oncology*, *26*, 259–271.

Saltz, J., Gupta, R., Hou, L., Kurc, T., Singh, P., Nguyen, V., Samaras, D., Shroyer, K. R., Zhao, T., Batiste, R. et al. (2018). Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Reports*, *23*, 181–193.

Schmauch, B., Romagnoni, A., Pronier, E., Saillard, C., Maillé, P., Calderaro, J., Kamoun, A., Sefta, M., Toldo, S., Zaslavskiy, M. et al. (2020). A deep learning model to predict rna-seq expression of tumours from whole slide images. *Nature Communications*, *11*, 1–15.

Shimizu, H., Okada, M., Toh, Y., Doki, Y., Endo, S., Fukuda, H., Hirata, Y., Iwata, H., Kobayashi, J., Kumamaru, H. et al. (2020). Thoracic and cardiovascular surgeries in japan during 2018. *General thoracic and cardiovascular surgery*, (pp. 1–34).

Shkolyar, A., Gefen, A., Benayahu, D., & Greenspan, H. (2015). Automatic detection of cell divisions (mitosis) in live-imaging microscopy images using convolutional neural networks. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE* (pp. 743–746). IEEE.

Sirinukunwattana, K., Raza, S. E. A., Tsang, Y.-W., Snead, D., Cree, I., & Rajpoot, N. (2015). A spatially constrained deep learning framework for detection of epithelial tumor nuclei in cancer histology images. In *International Workshop on Patch-based Techniques in Medical Imaging* (pp. 154–162). Springer.

Sirinukunwattana, K., Raza, S. E. A., Tsang, Y.-W., Snead, D. R., Cree, I. A., & Rajpoot, N. M. (2016). Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE transactions on medical imaging*, *35*, 1196–1206.

Song, Y., Tan, E.-L., Jiang, X., Cheng, J.-Z., Ni, D., Chen, S., Lei, B., & Wang, T. (2017). Accurate cervical cell segmentation from overlapping clumps in pap smear images. *IEEE transactions on medical imaging*, *36*, 288–300.

Song, Y., Zhang, L., Chen, S., Ni, D., Lei, B., & Wang, T. (2015). Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph

partitioning. *IEEE Transactions on Biomedical Engineering*, *62*, 2421–2433.

Steiner, D. F., MacDonald, R., Liu, Y., Truszkowski, P., Hipp, J. D., Gammage, C., Thng, F., Peng, L., & Stumpe, M. C. (2018). Impact of deep learning assistance on the histopathologic review of lymph nodes for metastatic breast cancer. *The American journal of surgical pathology*, *42*, 1636.

Tang, J., Shalabi, A., & Hubbard-Lucey, V. (2018). Comprehensive analysis of the clinical immuno-oncology landscape. *Annals of Oncology*, *29*, 84–91.

Taqi, S. A., Sami, S. A., Sami, L. B., & Zaki, S. A. (2018). A review of artifacts in histopathology. *Journal of Oral and Maxillofacial Pathology: JOMFP*, *22*, 279.

Turkki, R., Byckhov, D., Lundin, M., Isola, J., Nordling, S., Kovanen, P. E., Verrill, C., von Smitten, K., Joensuu, H., Lundin, J. et al. (2019). Breast cancer outcome prediction with tumour tissue images and machine learning. *Breast Cancer Research and Treatment*, *177*, 41–52.

Turkki, R., Linder, N., Kovanen, P. E., Pellinen, T., & Lundin, J. (2016). Antibody-supervised deep learning for quantification of tumor-infiltrating immune cells in hematoxylin and eosin stained breast cancer samples. *Journal of Pathology Informatics*, *7*.

Veta, M., Van Diest, P. J., Jiwa, M., Al-Janabi, S., & Pluim, J. P. (2016a). Mitosis counting in breast cancer: Object-level interobserver agreement and comparison to an automatic method. *PloS one*, *11*, e0161286.

Veta, M., Van Diest, P. J., & Pluim, J. P. (2016b). Cutting out the middleman: measuring nuclear area in histopathology slides without segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 632–639). Springer.

Vinay, D. S., Ryan, E. P., Pawelec, G., Talib, W. H., Stagg, J., Elkord, E., Lichtor, T., Decker, W. K., Whelan, R. L., Kumara, H. S. et al. (2015). Immune evasion in cancer: Mechanistic basis and therapeutic strategies. In *Seminars in cancer biology* (pp. S185–S198). Elsevier volume 35.

Wang, H., Roa, A. C., Basavanhally, A. N., Gilmore, H. L., Shih, N., Feldman, M., Tomaszewski, J., Gonzalez, F., & Madabhushi, A. (2014). Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. *Journal of Medical Imaging*, *1*, 034003.

Wang, J., MacKenzie, J. D., Ramachandran, R., & Chen, D. Z. (2016a). A deep learning approach

for semantic segmentation in histology tissue images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 176–184). Springer.

Wang, S., Yao, J., Xu, Z., & Huang, J. (2016b). Subtype cell detection with an accelerated deep convolution neural network. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 640–648). Springer.

Weinberg, R. A. (2008). Coevolution in the tumor microenvironment. *Nature Genetics*, *40*, 494–495.

Xie, W., Noble, J. A., & Zisserman, A. (2016a). Microscopy cell counting and detection with fully convolutional regression networks. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, (pp. 1–10).

Xie, Y., Kong, X., Xing, F., Liu, F., Su, H., & Yang, L. (2015a). Deep voting: A robust approach toward nucleus localization in microscopy images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 374–382). Springer.

Xie, Y., Xing, F., Kong, X., Su, H., & Yang, L. (2015b). Beyond classification: structured regression for robust cell detection using convolutional neural network. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 358–365). Springer.

Xie, Y., Zhang, Z., Sapkota, M., & Yang, L. (2016b). Spatial clockwork recurrent neural network for muscle perimysium segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 185–193). Springer.

Xing, F., Xie, Y., Su, H., Liu, F., & Yang, L. (2017). Deep learning in microscopy image analysis: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, *29*, 4550–4568.

Xing, F., Xie, Y., & Yang, L. (2016). An automatic learning-based framework for robust nucleus segmentation. *IEEE transactions on medical imaging*, *35*, 550–566.

Xu, J., Xiang, L., Liu, Q., Gilmore, H., Wu, J., Tang, J., & Madabhushi, A. (2016a). Stacked sparse autoencoder (ssae) for nuclei detection on breast cancer histopathology images. *IEEE transactions on medical imaging*, *35*, 119–130.

Xu, Y., Li, Y., Liu, M., Wang, Y., Lai, M., Eric, I., & Chang, C. (2016b). Gland instance segmentation by deep multichannel side supervision. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 496–504). Springer.

Xu, Z., & Huang, J. (2016). Detecting 10,000 cells in one second. In *International Conference on*

Medical Image Computing and Computer-Assisted Intervention (pp. 676–684). Springer.

Yamamoto, Y., Tsuzuki, T., Akatsuka, J., Ueki, M., Morikawa, H., Numata, Y., Takahara, T., Tsuyuki, T., Tsutsumi, K., Nakazawa, R. et al. (2019). Automated acquisition of explainable knowledge from unannotated histopathology images. *Nature communications*, *10*, 1–9.

Yang, L., Zhang, Y., Guldner, I. H., Zhang, S., & Chen, D. Z. (2016). 3d segmentation of glial cells using fully convolutional networks and k-terminal cut. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 658–666). Springer.

Yang, M., McKay, D., Pollard, J. W., & Lewis, C. E. (2018). Diverse functions of macrophages in different tumor microenvironments. *Cancer Research*, *78*, 5492–5503.

Yao, J., Wang, S., Zhu, X., & Huang, J. (2016). Imaging biomarker discovery for lung cancer survival prediction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 649–657). Springer.

Ye, H.-J., Chen, H.-Y., Zhan, D.-C., & Chao, W.-L. (2020). Identifying and compensating for feature deviation in imbalanced deep learning. *arXiv preprint arXiv:2001.01385*,.

Zhang, A. W., McPherson, A., Milne, K., Kroeger, D. R., Hamilton, P. T., Miranda, A., Funnell, T., Little, N., de Souza, C. P., Laan, S. et al. (2018). Interfaces of malignant and immunologic clonal dynamics in ovarian cancer. *Cell*, *173*, 1755–1769.

Zhao, J., Zhang, M., Zhou, Z., Chu, J., & Cao, F. (2017). Automatic detection and classification of leukocytes using convolutional neural networks. *Medical & biological engineering & computing*, *55*, 1287–1301.

Zhou, Y., Chang, H., Barner, K., Spellman, P., & Parvin, B. (2014). Classification of histology sections via multispectral convolutional sparse coding. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3081–3088).