


Future of biomarker evaluation in the realm of artificial intelligence algorithms: application in improved therapeutic stratification of patients with breast and prostate cancer

Jenny Fitzgerald ¹, Debra Higgins,² Claudia Mazo Vargas,³ William Watson,³ Catherine Mooney,³ Arman Rahman,³ Niamh Aspell,¹ Amy Connolly,¹ Claudia Aura Gonzalez,³ William Gallagher³

¹Invent Building, Deciphex Ltd, Dublin City University, Dublin, Ireland

²OncoAssure, Nova UCD, Belfield Innovation Park, Dublin, Ireland

³School of Biomolecular and Biomedical Science, Conway Institute, University College Dublin, Dublin, Ireland

Correspondence to

Dr Jenny Fitzgerald, Deciphex Ltd, Dublin City University, Dublin, Ireland; Jenny.fitzgerald@deciphex.com

Received 23 April 2021

Accepted 25 April 2021

Published Online First

10 June 2021

ABSTRACT

Clinical workflows in oncology depend on predictive and prognostic biomarkers. However, the growing number of complex biomarkers contributes to costly and delayed decision-making in routine oncology care and treatment. As cancer is expected to rank as the leading cause of death and the single most important barrier to increasing life expectancy in the 21st century, there is a major emphasis on precision medicine, particularly individualisation of treatment through better prediction of patient outcome. Over the past few years, both surgical and pathology specialties have suffered cutbacks and a low uptake of pathology specialists means a solution is required to enable high-throughput screening and personalised treatment in this area to alleviate bottlenecks. Digital imaging in pathology has undergone an exponential period of growth. Deep-learning (DL) platforms for hematoxylin and eosin (H&E) image analysis, with preliminary artificial intelligence (AI)-based grading capabilities of specimens, can evaluate image characteristics which may not be visually apparent to a pathologist and offer new possibilities for better modelling of disease appearance and possibly improve the prediction of disease stage and patient outcome. Although digital pathology and AI are still emerging areas, they are the critical components for advancing personalised medicine. Integration of transcriptomic analysis, clinical information and AI-based image analysis is yet an uncultivated field by which healthcare professionals can make improved treatment decisions in cancer. This short review describes the potential application of integrative AI in offering better detection, quantification, classification, prognosis and prediction of breast and prostate cancer and also highlights the utilisation of machine learning systems in biomarker evaluation.

INTRODUCTION

Developments in artificial intelligence (AI) technology have allowed the mining of previously hidden data from routine histology images of cancer, providing potentially clinically meaningful information. The considerable degree of uncertainty in traditional pathology analysis when determining whether patients have indolent or aggressive disease leads to overtreatment of patients and subsequent secondary complications (eg, surgical complications including sepsis, chemotherapy-associated toxicity), significantly impacting patients' quality of life. Routinely available

tumour tissue contains an abundance of clinically relevant information that is currently not fully exploited. Timely and accurate investigation of tumour histomorphology is critical, and determining pertinent prognostic markers is the key to personalised cancer management. Machine learning (ML) techniques are commonly used in biomarker development and increases in labelled/annotated data and images are enabling deep neural networks (DNNs).¹ Recently, studies have shown that DNN models based on digitised slides have potential in several cancer types in areas including tumour diagnosis, prognostic prediction and identification of pathological features such as biomarkers.^{2–4}

CLINICAL NEED

Breast cancer is the most common form of cancer diagnosis in women (excluding nonmelanoma skin cancer), with over 2 million women diagnosed worldwide annually.⁵ Breast cancer is a heterogeneous disease that still presents challenges for clinicians in predicting the likelihood of disease progression, particularly in patients where the disease is detected in the early stages. To manage the increasing volume of breast cancer cases, we need to meaningfully interpret breast cancer progression to establish prognostic factors and limit the number of patients going for unnecessary treatments. Approximately 30% of patients develop a recurrence of the disease within 10 years and, therefore, require aggressive chemotherapy.⁶ However, it is quite difficult to differentiate between those whose disease will or will not recur. Most early stage patients with breast cancer are treated with chemotherapy, despite many not benefiting from such treatment, thereby exposing these individuals to severe side effects. Hence, unwarranted treatment burdens healthcare systems with additional patients and with huge associated costs. The need for carefully collected clinical tissues being made available to investigators is key to facilitate advancement of diagnostics, thus making initiatives such as the Innovative Medicine Initiative's (IMI) Big Picture Consortium pivotal in allowing greater wide-spread access to data.⁷ Recent legislative changes have made this more difficult and both surgical and pathology support are needed at a senior level to enable high-throughput screening and personalised treatment in this area to alleviate



© Author(s) (or their employer(s)) 2021. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Fitzgerald J, Higgins D, Mazo Vargas C, et al. *J Clin Pathol* 2021;**74**:429–434.

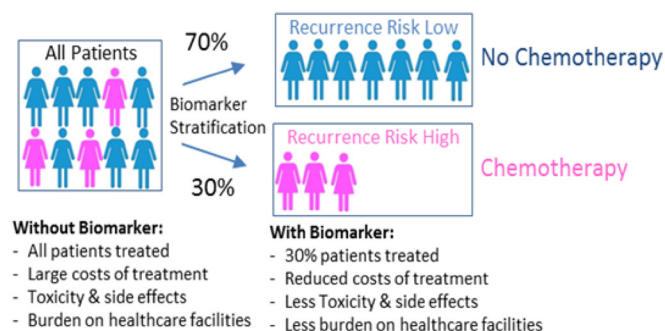


Figure 1 Biomarker stratification can determine risk of recurrence and will identify patients who will benefit from chemotherapy treatment and who will not benefit. This prognostic test will also result in reduced costs of treatment and less strain on healthcare systems.

bottlenecks.⁸ There is a definitive clinical need for the development of a highly sensitive and specific prognostic assay (figure 1). The ideal test should give a low/high risk (binary) output to allow for easier decision-making, while also being highly accurate in terms of its ability to differentiate between patient classes that may benefit from aggressive treatment.

Similarly, prostate cancer is the second most common cancer in men (after nonmelanoma skin cancer) and approximately one in six men will be diagnosed with prostate cancer during their lifetime.⁵ There are approximately 1.2 million men diagnosed worldwide annually with prostate cancer representing 7.1% of all cancers in men.⁹ According to the American Cancer Society, 248 530 men will be diagnosed with prostate cancer in the USA in 2021.¹⁰ As a result of prostate-specific antigen (PSA) screening and digital rectal examination testing, many prostate cancers are now detected at an early stage. Approximately 80% of men have slow growing or indolent prostate cancer and can leave their disease untreated while undergoing frequent monitoring known as active surveillance. The remaining 20% of patients have aggressive prostate cancer, which requires aggressive treatment.¹¹

Clinical decision-making for prostate cancer is set out in the European Association of Urology (EAU) guidelines from localised to advanced therapy.¹² These are informed by clinical features with increasing inclusion of genomics and molecular biomarkers.¹³ Potential biomarkers should provide additional independent information from already established clinical and pathological variables, to improve the predictive accuracy for prostate cancer diagnosis, prognosis and treatment response. Current clinical tools cannot determine the true risk of metastases, thus accurately identifying patients with aggressive prostate cancer (ie, predicting metastatic potential), represents a major clinical unmet need and is crucially important for selecting the appropriate treatment options for patients with prostate cancer. A prognostic test that accurately identifies patients with aggressive prostate cancers prior to radical prostatectomy will help identify the 20% of high-risk patients who will benefit from aggressive treatment of their disease and give the 80% of low-risk patients, who will not benefit from treatment, peace of mind that active surveillance is the correct treatment option for them.

CURRENT APPROACHES

Currently, patients are diagnosed with breast or prostate cancer based on traditional (manual) pathological analysis of H&E-stained patient tumour samples obtained either through biopsy or surgical procedures. The clinician uses the pathological assessment in combination with other clinical features (eg, urine or blood-based biochemistry results and physical parameters), to determine whether

the patient has indolent or aggressive disease. However, traditional pathology analysis comes with a considerable degree of uncertainty. Histopathological assessment is essential in the diagnosis of cancer; however, individual evaluation of histopathology slides cannot accurately predict patient prognosis. Advances in technology, particularly computing speed, have rapidly increased the growth of digital imaging in pathology.^{14 15} Deep-learning (DL) AI platforms, such as H&E image analysis with AI-based grading capabilities of specimens, allows pathologists to better assess and determine diagnostic and prognostic indicators and are the key to improving the prediction of disease stage and providing personalised medicine. DNN models based on digitised slides have proven to show potential in predicting the prognosis of patients with cancer, thus advancing precision oncology.¹⁶ Although digital pathology and AI are still rapidly evolving areas, they are vital for facilitating personalised medicine for patient treatment. The full utility of the integration of transcriptomic analysis, clinical information and AI-based image analysis has yet to be realised by healthcare professionals who can make better treatment decisions in cancer (figure 2). Integrative AI assessment offers better detection, quantification, classification, prognosis and prediction and allows for easier collaboration between pathologists, clinicians, scientists and industry, which is vital to move the field forward in a meaningful way.

Solutions

The development of prognostic assays has created new opportunities for improving both breast and prostate cancer treatment decisions; however, these assays are commonly run in a centralised lab setting, thereby limiting accessibility. The development of decentralised tests, which combines both pathology image and molecular data, can be integrated into the standard clinical workflow of the hospital setting and could prove advantageous. The use of AI approaches to predict patient prognosis will better inform appropriate clinical decision-making. AI can be leveraged to identify histological features from digital images of a tumour and can associate these features with molecular data from the same tumour tissue to predict patient outcome. AI-based diagnostic methods can harness computational and mathematical vigour previously unrealised by traditional approaches. The integrated analysis of molecular data with AI-mined image data can provide greater context for clinicians, which can in turn provide greater stratification for appropriate treatment approaches.

Prognostics can play a vital role in patient management and decision-making. Currently, approaches to predict status or progression of cancers are primarily based on immunohistochemical screening for specific biomarkers or the use of gene expression signatures (ie, transcriptomic data). Research surrounding the integration of omic and clinical data for cancer prognostics is very limited and not widely reported. Studies that have taken advantage of singular integration approaches, such as combination of genomic and image data, have found great utility in predicting patient outcome.^{17–19} The advent of digital pathology has allowed for the investigation of other forms of prognostic data from digital tissue images, including quantitative pathology. Quantitative pathology can involve the enumeration of specific cells or morphometric/densitometric analysis to classify tissue in a metric form; this is seen as a more objective approach compared with subjective assessment performed by pathologists. In computer-aided diagnostic systems, ML techniques are widely employed for cancer detection and diagnosis.^{20 21} Recently, AI-based techniques applied for automated classification of whole slide images have advanced significantly via DL.^{22 23} A primary advantage of DL is that it can generate high-level feature representation directly from the raw images. In addition, with the

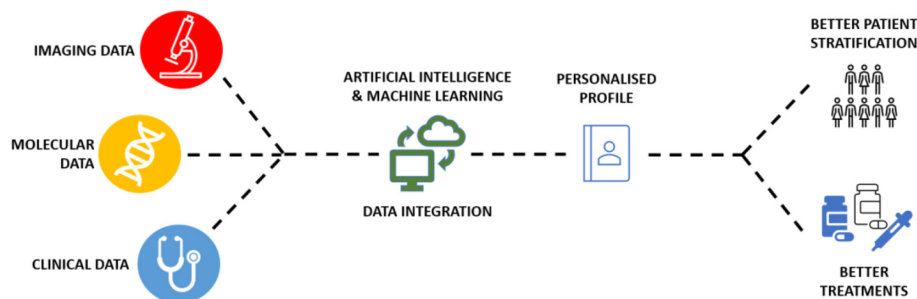


Figure 2 Integrating imaging, molecular and clinical data with artificial intelligence and deep learning to develop more sophisticated personalised patient profiles and accordingly use this information for improved patient stratification to guide appropriate treatment strategies.

support of massive parallel architecture, graphic processing units and DL techniques have gained enormous success in many fields in recent years including applications in cancer detection and diagnosis to predict patient survival time directly from cancer pathological images.²⁴ AI approaches offer the opportunity to make better use of data-driven clinical diagnosis and can pave the way for a revolution in prognostic stratifiers.

Impact

Advances in AI technology are revolutionising patient care, through monitoring, diagnosis and even prognosis.²⁵ In medical imaging, the evolution of digital pathology has led to the incorporation of AI and DL approaches to analyse whole slide images in pathology. Predictive and prognostic ancillary testing in cancer have increasingly placed the pathologist in a new role as a 'diagnostic oncologist' who performs, interprets and integrates pathology data on each patient's tissue such that individualised treatment decisions can be made.²⁶ With the evolution of whole slide imaging (WSI) and digital pathology, there has been an increase in the demand for their use as diagnostic aids supporting contemporary clinical care and to also aid in research.²⁷ The Food and Drug Administration (FDA) has released numerous reports on guidance for the use of digital pathology, with the most recent being a report on guidance for exploiting digital pathology in 2020 due to the COVID-19 (SARS-CoV-2) pandemic.²⁸ Additionally, in 2019, the FDA released a discussion paper proposing a fast-track regulatory framework for AI/ML-based software as a medical device.²⁹ The recent pandemic has highlighted the need for digital pathology services and, at present, is of utmost importance as by using remote digital pathology devices, it may aid in pathology and diagnostic services, while reducing physical contact between healthcare personnel. A previous barrier to digital pathology was clinical adoption; however, it is envisaged after the pandemic, this obstacle will be considerably mitigated by greater widespread acceptance of digital pathology platforms.^{30–32}

Clinical workflows in oncology are increasingly relying on predictive and prognostic molecular biomarkers. The acceptance and adoption of commercially available prognostic assays were initially slow; however, established assays based on diverse biomarkers have impacted clinical practice and forged a route to market, which should facilitate the introduction of new and improved assays. As of early February 2021, there were 45 approved companion diagnostic assays³³ (in vitro diagnostic and image analysis) approved by the FDA for cancer diagnosis. Several publications have highlighted the requirement for integrative technologies to provide a better understanding of disease progression for proper patient stratification.³⁴ Current prognostic assays for breast cancer are extremely costly and suffer from ambiguity, a key aspect that needs to be addressed. Similarly, for prostate cancer, risk stratification has traditionally relied on clinicopathologic features, such as PSA, grade group, clinical stage

and percentage of positive biopsy cores, to define prognostic risk groups.^{35,36} Only recently have molecular tests been developed that may better determine the aggressiveness of prostate cancer based on general features of malignancy (namely, proliferation indices).^{37–39} A validated DL algorithm to improve Gleason scoring of patients with prostate cancer was recently described,⁴⁰ whereby the proposed solution significantly improved the accuracy of the scoring to 70% when compared with an average 61% accuracy achieved by 29 pathologists. This substantiates the requirement for digital triaging of pathology slides.

AI-based image analysis systems enable the extraction of quantitative features such as standard morphometric descriptors of image objects and higher level contextual, relational and global image features from H&E images. These can then be used independently to construct prognostic models.^{41–44} This has the potential to create several highly novel technologies, which will offer healthcare solutions, transform patient treatments, reduce physician workload and also reduce the impact of healthcare industries on the environment too. Despite improvements in diagnostics and treatments, the response rates of patients to treatments remains very low, that is only 4%–25% for the top 10 US best selling drugs.⁴⁵ Specifically, in cancer the overall response rate to FDA-approved drugs is 41%, but only 6% of patients achieve a complete regression of the tumour.⁴⁶ A major reason for highly variable drug responses is that current clinical tools are based on population studies and measure the average response instead of the responses of individual patients. However, these statistics also give hope that improved patient stratification and personalised therapies can shift many patients to complete response. Early intervention through a robust and validated assay can strengthen this shift.

In recent years, there has been a rapid increase in both the breast and prostate cancer diagnostics market and as a result of this, several assays have been developed that aim to stratify patients with cancer into those at low or high risk of recurrence. Early stage breast cancer, particularly those that are oestrogen receptor-positive (OR+), represents approximately 75% of newly diagnosed breast cancer cases each year.⁴⁷ For OR+ patients, several tests are available on the market that attempt to stratify based on likelihood of future recurrence of the disease, such as the leading breast cancer prognostic assays; Oncotype DX (Exact Sciences)⁴⁸ and MammaPrint (Agendia),⁴⁹ with a novel test in this space, OncoMasTR, recently developed also.^{50,51} Oncotype DX is a 21 gene RT-qPCR assay that stratifies patients into three groups: low, intermediate and high risk; whereas MammaPrint is a 70 gene microarray assay that stratifies patients into two groups: low and high risk. However, both Oncotype Dx and MammaPrint can only be performed in a centralised laboratory setting, thereby limiting their utility and adding to their high cost (approx. US\$ 4000 per patient). There is a clear critical need for a cost-effective, decentralised test in this space.

The global prostate cancer diagnostics market size was valued at US\$ 2.83 billion in 2019 and is anticipated to expand at a growth rate of 13.2% over the next 7 years.⁵² Approximately 4 million biopsies are performed worldwide each year with 25% of positive biopsies classified as high-risk and result in patients going directly for treatment. The US is the leading global market with the largest revenue share in 2017 accounting for nearly 40% of the prostate cancer diagnostics market⁵³ from the two market-leading prostate cancer prognostic signatures tests; Oncotype Dx, Genomic Prostate Score (GPS) (Exact Sciences) and Prolaris (Myriad), with combined revenue of US\$48 million. Oncotype Dx GPS is a 17 gene quantitative reverse transcription PCR (RT-qPCR) assay that provides a GPS score (0–100) and compares a patient's score to the average score across the National Comprehensive Cancer Network (NCCN) clinical categories. Prolaris is a 46 gene RT-qPCR assay that provides a cell cycle progression score (0–10), which is combined with a clinical CAPRA score to determine a 10-year risk of death. Again, the need for a decentralised test is seen in this arena also.

Moreover, there is concern among the medical community that there is no solution that completely addresses the market needs—for example, the market-leading test in the breast cancer prognostic arena, namely, Oncotype DX, classifies more than half of patients into an 'intermediate risk' category and was found to have some limitations.⁵⁴ MammaPrint classifies approximately 50% of patients as 'high-risk', resulting in significant over treatment. Other tests on the market also introduce an intermediate group and/or lack accuracy. A comprehensive gap analysis⁵⁵ surrounding breast cancer research noted these issues and pinpointed critical gaps in the analysis of variant patient groups stratified with biomarkers. The gap analysis stressed the requirements for research infrastructure that is multidisciplinary and will provide access to new technologies pursuing innovative research avenues.

Innovative Infrastructures

Recently, the rapidly increasing number and clinical importance of molecular biomarkers in routine clinical practice allow cancer treatments to be tailored more specifically according to the genetic make-up of a particular tumour; consequently, however, the cost, turnaround time and tissue requirements in routine workflows also increase.⁵⁶ Although most new biomarkers in oncology are based on molecular diagnostic assays, advances in AI-based DL are facilitating the extraction of otherwise hidden information directly from routinely available image data. H&E slides are available for almost every patient with cancer, making them an easy-to-obtain, information-rich data source for assessment by image analysis methods using DL. Each patient is unique and, therefore, the 'one size fits all' approach typically used within healthcare is not always effective. In oncology, some of the goals facilitated by personalised medicine are individualisation of patient treatments, predicting risk of recurrence and predicting survival time.⁵⁷

Current state-of-the-art

Breast cancer prognostics

There are several breast cancer studies that have aimed to bring together images, genomic signatures, molecular subtype characterisation and clinically used recurrence scores in different ways.^{58–60} Such studies have tried to integrate different data sources to improve results obtained individually. Natrajan *et al* proposed a histology–genomic integration analysis for diagnosis of patients at high risk of relapse using tumour microenvironment heterogeneity, clinical annotation and DNA/RNA data matching.⁶¹ The K-means algorithm, tumour microenvironment heterogeneity by Shannon diversity index, Gaussian mixture

clustering and a Cox proportional hazards regression model were used in this study. They concluded that microenvironment heterogeneity, together with key genomic alterations, could be used to identify those patients at high risk of relapse and facilitate treatment decisions. Other studies correlating image characteristics with clinically available prognostic genomic assays have explored prediction of risk of recurrence.^{62–64} Verma *et al* used morphometric information from H&E samples (mitosis, architectural patterns, nuclei) and Oncotype DX scores as inputs for a Cox proportional hazard analysis in OR-positive patients with breast cancer,⁶⁵ showing outperformance of either assay alone.

Heindl *et al* evaluated the relevance of spatial heterogeneity of immune infiltration for predicting risk of recurrence. Different scores such as immune cell abundance—intratumour lymphocytes, adjacent-to-tumour lymphocytes and distal-to-tumour lymphocytes, spatial scores—immune, cancer and immune-cancer hotspots, prognostic scores—OR, Progesterone Receptor (PgR), Human Epidermal Growth Factor Receptor 2 (HER2), Ki67 expression, clinical treatment score, node status, size, grade, age and treatment, OncoType DX recurrence and PAM50 risk of recurrence were used, where they were able to provide an association between spatial scores and late recurrence in OR+ disease using unsupervised clustering to score immune cell abundance and spatial heterogeneity.⁶² Another example using integrated spatial assessment is described by Yuan *et al* who used microenvironmental heterogeneity quantification via digital image analysis integrated with RNA gene expression and DNA copy number profiling data to identify molecular changes and subsequently develop an integrated predictor of survival for patients with OR-negative disease.⁶⁶ For this work, morphological image information, image-based lymphocyte proportion, pathological lymphocyte infiltration (LI) score, expression LI signature, integrated LI and stromal spatial pattern were used with a support vector machine (SVM) and kernel smoother. Similarly, Sun *et al* integrated genomic data and pathological images to effectively predict the survival time of patients with breast cancer.⁶⁷ Genomic data (gene expression, copy number alteration, gene methylation and protein expression) and pathological images were considered as input of an SVM based on multiple kernel learning. Established validated methods similar to the ones described with high-throughput capacity will have great potential for early intervention and treatment support of patients with breast cancer.

Prostate cancer prognostics

The histologic spectrum of prostate cancer ranges from its precursor lesion, high-grade prostatic intraepithelial neoplasia, to dedifferentiated cancer and displays a wide spectrum of morphological patterns.^{68–71} The Gleason score is currently the strongest predictor of prostate cancer recurrence with Gleason 4 and Gleason 5 patterns strongly associated with poorer outcome.^{72–73} More recently, the presence of cribriform architecture in pathology H&E sections is considered to provide additional prognostic information as Gleason grading suffers from substantial interobserver variation, limiting its usefulness.^{74–75} Digital pathology coupled with automated DL systems is being explored to see whether prostate cancer diagnostics can benefit from a robust and reproducible Gleason grading system.^{76–77}

At the molecular level, a number of molecular subtypes have been identified.⁷⁸ Most assays for prostate cancer are focused on these pre-diagnostic markers to consider their ability to identify clinically significant prostate cancer (PCa), while avoiding unnecessary biopsies and also to decide whom to biopsy (PHI, 4Kscore, SelectMDx, MiPS) and when to rebiopsy (PCA3 and ConfirmMDx). Several image-based prostate analysis methods have been described. Bulten

Take home messages

- ▶ AI-based diagnostic and prognostic approaches are becoming for ubiquitous patient treatment decisions.
- ▶ Molecular signatures have been developed to assist clinicians in stratifying patients based on risk of recurrence; while these have shown promise, there is significant room for improving prognostic capacity which can be supplemented using AI approaches.

et al and Ström *et al* independently developed an automated DL system for Gleason grading of prostate cancer biopsies at Radboud University Medical Center and Karolinska Institute, respectively, designed to (1) identify individual glands, (2) assign Gleason growth patterns and (3) determine biopsy-level grade.^{76–77} They found that their DL systems performed as well as pathologists and could potentially provide assistance through screening biopsies, providing second opinions on grade groups and presenting qualitative measurements of tumour volume percentages.

For prognostics, Oncotype DX GPS and ProMark might help decide who to treat, whereas Prolaris tries to identify patients who might be at risk for distant metastases and, therefore, need further treatment. The real challenge is to find the best combination, which includes biomarkers and clinical data, without increasing the cost and the risk of overdiagnosis.⁷⁹ The combined clinical and molecular heterogeneity of prostate cancer necessitates the use of prognostic, predictive and diagnostic biomarkers to assist the clinician with treatment selection. The pathologist plays a critical role in guiding molecular biomarker testing in prostate cancer and requires a thorough knowledge of the current testing options.

Multiplexed Immunohistochemistry (IHC) enables identification and quantification of multiple biomarkers and reveals spatial context within a digital workflow, resulting in rapid generation of data.⁸⁰ The spatial architecture of tissue samples can strongly influence disease pathology, progression and treatment response. To our knowledge, there are currently no product solutions commercially available fusing imaging, molecular and clinical data, which would provide greater information to clinicians than any currently available methods. With the advancement of digital AI-based solutions, the current gap in available offerings in the diagnostics and prognostics space will undoubtedly be populated quite rapidly over the next few years.

Conclusion

AI and ML approaches provide significant new opportunities in precision oncology, particularly in relation to lower costs, ergonomic healthcare settings for pathologists and improved patient stratification.

Pathologists spend an inordinate amount of time at a microscope reviewing boxes of glass slides, this is neither efficient (costing time and money) nor ergonomic. Physicians are often 'over-treating' as a precautionary measure without full prognostic awareness'. Removing low value processes (normal tissue) from an overloaded pathologist can increase their capacity to review high priority cases and allows for faster review. Finally, personalised medicine has been greatly improved due to digital pathology and AI, which provide greater certainty regarding prognosis.

Handling editor Runjan Chetty.

Contributors All authors contributed to the design and implementation of the review and to the writing of the manuscript.

Funding Funding is acknowledged from the Irish Cancer Society Collaborative Cancer Research Centre BREASTPREDICT [grant number CCRC13GAL], as well as Science Foundation Ireland (SFI) under the Investigator Programme OPTI-PREDICT [grant number 15/IA/3104] and the Strategic Research Programme Precision Oncology Ireland [grant number 18/SPP/3522]. Funding is also acknowledged from Enterprise Ireland (EI) and from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement number 713654. Funding is acknowledged from Enterprise Ireland's Disruptive Technologies Innovation Fund 2021.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Commissioned; internally peer reviewed.

ORCID iD

Jenny Fitzgerald <http://orcid.org/0000-0003-3963-197X>

REFERENCES

- Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. *Nat Rev Genet* 2015;16:321–32.
- Jiang Y, Yang M, Wang S, *et al*. Emerging role of deep learning-based artificial intelligence in tumor pathology. *Cancer Commun* 2020;40:154–66.
- Harder N, Athelogou M, Hessel H, *et al*. Tissue phenomics for prognostic biomarker discovery in low- and intermediate-risk prostate cancer. *Sci Rep* 2018;8:4470.
- Chen RJ, Lu MY, Wang J, *et al*. Pathomic fusion: an integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. *IEEE Trans Med Imaging* 2020;PP. doi:10.1109/TMI.2020.3021387. [Epub ahead of print: 03 Sep 2020] <http://arxiv.org/abs/1912.08937>
- Bray F, Ferlay J, Soerjomataram I, *et al*. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
- Redig AJ, McAllister SS. Breast cancer as a systemic disease: a view of metastasis. *J Intern Med* 2013;274:113–26.
- Moulin P, Grünberg K, Barale-Thomas E, *et al*. IMI-Bigpicture: a central repository for digital pathology. *Toxicol Pathol* 2021;49:711–3.
- Metter DM, Colgan TJ, Leung ST, *et al*. Trends in the US and Canadian pathologist workforces from 2007 to 2017. *JAMA Netw Open* 2019;2:e194337.
- Rawla P. Epidemiology of prostate cancer. *World J Oncol* 2019;10:63–89.
- Siegel RL, Miller KD, Fuchs HE, *et al*. Cancer statistics, 2021. *CA Cancer J Clin* 2021;71:7–33.
- Choyke PL, Loeb S. Active surveillance of prostate cancer. *Oncology* 2017;31:67–70.
- EAU Guidelines: Prostate Cancer. Uroweb. [cited 2021 Mar 1]. Available: <https://uroweb.org/guideline/prostate-cancer/>
- Oon SF, Pennington SR, Fitzpatrick JM, *et al*. Biomarker research in prostate cancer--towards utility, not futility. *Nat Rev Urol* 2011;8:131–8.
- Bera K, Schalper KA, Rimm DL, *et al*. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol* 2019;16:703–15.
- Parwani AV. Next generation diagnostic pathology: use of digital pathology and artificial intelligence tools to augment a pathological diagnosis. *Diagn Pathol* 2019;14:s13000-019-0921-2:138.
- Yu K-H, Zhang C, Berry GJ, *et al*. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nat Commun* 2016;7:12474.
- Murphy K, Murphy BT, Boyce S, *et al*. Integrating biomarkers across omic platforms: an approach to improve stratification of patients with indolent and aggressive prostate cancer. *Mol Oncol* 2018;12:1513–25.
- Bychkov D, Linder N, Turkki R, *et al*. Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci Rep* 2018;8:3395.
- Yousefi S, Amrollahi F, Amgad M, *et al*. Predicting clinical outcomes from large scale cancer genomic profiles with deep survival models. *Sci Rep* 2017;7:11707.
- Santos MK, Ferreira Júnior JR, Wada DT, *et al*. Artificial intelligence, machine learning, computer-aided diagnosis, and radiomics: advances in imaging towards to precision medicine. *Radiol Bras* 2019;52:387–96.
- Ramadan SZ. Methods used in computer-aided diagnosis for breast cancer detection using mammograms: a review. *J Healthc Eng* 2020;2020:e9162464:2020–1.
- Zhu W, Xie L, Han J, *et al*. The application of deep learning in cancer prognosis prediction. *Cancers* 2020;12:603.
- Ahuja AS. The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ* 2019;7:e7702.
- Acs B, Rantalainen M, Hartman J. Artificial intelligence as the next step towards precision pathology. *J Intern Med* 2020;288:62–81.
- Barisoni L, Lafata KJ, Hewitt SM, *et al*. Digital pathology and computational image analysis in nephropathology. *Nat Rev Nephrol* 2020;16:669–85.
- Allison KH. Prognostic and predictive parameters in breast pathology: a pathologist's primer. *Mod Pathol* 2021;34:94–106.
- Dimitriou N, Arandjelović O, Caie PD. Deep learning for whole slide image analysis: an overview. *Front Med* 2019;6:264 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6882930/>

- 28 Enforcement Policy for Remote Digital Pathology Devices During the Coronavirus Disease. COVID-19 Public Health Emergency - Guidance for Industry, Clinical Laboratories, Healthcare Facilities, Pathologists and Food and Drug Administration Staff 2019;9.
- 29 Health C for D and R. artificial intelligence and machine learning in software as a medical device, 2021. Fda. Available: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>
- 30 A Turning Point for Digital Pathology. The pathologist. Available: <https://thepathologist.com/outside-the-lab/a-turning-point-for-digital-pathology> [Accessed 2 Mar 2021].
- 31 The Future of Digital Pathology. The pathologist.. Available: <https://thepathologist.com/inside-the-lab/the-future-of-digital-pathology> [Accessed 2 Mar 2021].
- 32 Telepathology in the era of COVID-19 | Charles river laboratories. Available: <https://www.criver.com/eureka/telepathology-in-the-era-of-covid-19> [Accessed cited 2021 Mar 2].
- 33 Health C for D and R. list of cleared or Approved companion diagnostic devices (in vitro and imaging tools), 2021. FDA. Available: <https://www.fda.gov/medical-devices/vitro-diagnostics/list-cleared-or-approved-companion-diagnostic-devices-vitro-and-imaging-tools>
- 34 Mazo C, Kearns C, Mooney C, *et al.* Clinical decision support systems in breast cancer: a systematic review. *Cancers* 2020;12:369.
- 35 Cooperberg MR, Pasta DJ, Elkin EP, *et al.* The University of California, San Francisco cancer of the prostate risk assessment score: a straightforward and reliable preoperative predictor of disease recurrence after radical prostatectomy. *J Urol* 2005;173:1938–42.
- 36 Reiter RE. Risk stratification of prostate cancer 2016. *Scand J Clin Lab Invest Suppl* 2016;245:S54–9.
- 37 Klein EA, Maddala T, Millward C, *et al.* Development of a needle biopsy-based genomic test to improve discrimination of clinically aggressive from indolent prostate cancer. *Journal of Clinical Oncology* 2012;30:4560.
- 38 Erho N, Crisan A, Vergara IA, *et al.* Discovery and validation of a prostate cancer genomic classifier that predicts early metastasis following radical prostatectomy. *PLoS One* 2013;8:e66855.
- 39 Czucik J, Swanson GP, Fisher G. Prognostic value of an RNA expression signature derived from cell cycle proliferation genes for recurrence and death from prostate cancer: a retrospective study in two cohorts. *Lancet Oncol* 2011;12:245–55.
- 40 Nagpal K, Foote D, Liu Y, *et al.* Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer. *NPI Digit Med* 2019;2:1–10.
- 41 Pati P, Jaume G, Foncubierta A. Hierarchical Cell-to-Tissue graph representations for breast cancer subtyping in digital pathology. 2021 arXiv:2102.11057v1
- 42 Xie C, Sigel C, Doukas M. EPIC-Survival: end-to-end part inferred clustering for survival analysis. *Featuring Prognostic Stratification Boosting* 2021 arXiv:2101.11085v1
- 43 Saltz J, Gupta R, Hou L, *et al.* Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Rep* 2018;23:181–93.
- 44 Jin YW, Jia S, Ashraf AB, *et al.* Integrative data augmentation with U-Net segmentation masks improves detection of lymph node metastases in breast cancer patients. *Cancers* 2020;12:2934.
- 45 Schork NJ. Personalized medicine: time for one-person trials. *Nature* 2015;520:609–11.
- 46 Chen EY, Raghunathan V, Prasad V. An overview of cancer drugs Approved by the US food and drug administration based on the surrogate end point of response rate. *JAMA Intern Med* 2019;179:915.
- 47 Dai X, Xiang L, Li T, *et al.* Cancer hallmarks, biomarkers and breast cancer molecular subtypes. *J Cancer* 2016;7:1281–94.
- 48 Paik S, Tang G, Shak S, *et al.* Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* 2006;24:3726–34.
- 49 Soliman H, Shah V, Srkalovic G, *et al.* Mammprint guides treatment decisions in breast cancer: results of the impact trial. *BMC Cancer* 2020;20:81.
- 50 Buus R, Sestak I, Barron S, *et al.* Validation of the OncoMasTR risk score in estrogen Receptor-Positive/HER2-Negative patients: a TransATAC study. *Clin Cancer Res* 2020;26:623–31.
- 51 Lanigan F, Brien GL, Fan Y, *et al.* Delineating transcriptional networks of prognostic gene signatures refines treatment recommendations for lymph node-negative breast cancer patients. *Febs J* 2015;282:3455–73.
- 52 Prostate Cancer Diagnostics Market Size | Industry Report, 2027. Available: <https://www.grandviewresearch.com/industry-analysis/prostate-cancer-diagnostics-market> [Accessed 1 Mar 2021].
- 53 Cancer biomarker market size | global industry research report, 2025. Available: <https://www.grandviewresearch.com/industry-analysis/cancer-biomarker-market> [Accessed 1 Mar 2021].
- 54 Tiberi D, Masucci L, Shedid D, *et al.* Limitations of personalized medicine and gene assays for breast cancer. *Cureus* 2017;9:e1100.
- 55 Eccles SA, Aboagye EO, Ali S, *et al.* Critical research gaps and translational priorities for the successful prevention and treatment of breast cancer. *Breast Cancer Res* 2013;15:R92.
- 56 Echle A, Rindtorff NT, Brinker TJ, *et al.* Deep learning in cancer pathology: a new generation of clinical biomarkers. *Br J Cancer* 2021;124:686–96.
- 57 El Naqa I, Ruan D, Valdes G, *et al.* Machine learning and modeling: data, validation, communication challenges. *Med Phys* 2018;45:e834–40.
- 58 Cho N. Molecular subtypes and imaging phenotypes of breast cancer. *Ultrasonography* 2016;35:281–8.
- 59 Falato C, Tobin NP, Lorent J, *et al.* Intrinsic subtypes and genomic signatures of primary breast cancer and prognosis after systemic relapse. *Mol Oncol* 2016;10:517–25.
- 60 Bustamante Eduardo M, Popovici V, Imboden S, *et al.* Characterization of molecular scores and gene expression signatures in primary breast cancer, local recurrences and brain metastases. *BMC Cancer* 2019;19:549.
- 61 Natrajan R, Saleem H, Mardakheh FK, *et al.* Microenvironmental heterogeneity parallels breast cancer progression: a Histology-Genomic integration analysis. *PLoS Med* 2016;13:e1001961.
- 62 Heindl A, Sestak I, Naidoo K, *et al.* Relevance of spatial heterogeneity of immune infiltration for predicting risk of recurrence after endocrine therapy of ER+ breast cancer. *J Natl Cancer Inst* 2018;110:166–75.
- 63 Woodard GA, Ray KM, Joe BN, *et al.* Qualitative Radiogenomics: association between Oncotype DX test recurrence score and BI-RADS mammographic and breast MR imaging features. *Radiology* 2018;286:60–70.
- 64 Veta M, van Diest PJ, Kornegoor R, *et al.* Automatic nuclei segmentation in H&E stained breast cancer histopathology images. *PLoS One* 2013;8:e70221.
- 65 Verma V, Sprave T, Haque W, *et al.* A systematic review of the cost and cost-effectiveness studies of immune checkpoint inhibitors. *J Immunother Cancer* 2018;6:128.
- 66 Yuan Y, Failmezger H, Rueda OM, *et al.* Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Sci Transl Med* 2012;4:157ra143–57.
- 67 Sun Q, Jankovic MV, Budzinski J, *et al.* A dual mode adaptive Basal-Bolus advisor based on reinforcement learning. *IEEE J Biomed Health Inform* 2019;23:2633–41.
- 68 Bostwick DG, Liu L, Brawer MK, *et al.* High-Grade prostatic intraepithelial neoplasia. *Rev Urol* 2004;6:360–79.
- 69 Moch H, Cubilla AL, Humphrey PA, *et al.* The 2016 who classification of tumours of the urinary system and male genital Organs-Part A: renal, penile, and testicular tumours. *Eur Urol* 2016;70:93–105.
- 70 Shoag J, Barbieri CE. Clinical variability and molecular heterogeneity in prostate cancer. *Asian J Androl* 2016;18:543.
- 71 Tolkach Y, Kristiansen G. The heterogeneity of prostate cancer: a practical approach. *Pathobiology* 2018;85:108–16.
- 72 Sharma M, Miyamoto H. Percent Gleason pattern 4 in stratifying the prognosis of patients with intermediate-risk prostate cancer. *Transl Androl Urol* 2018;7:S484–9.
- 73 Andrén O, Fall K, Franzén L, *et al.* How well does the Gleason score predict prostate cancer death? A 20-year followup of a population based cohort in Sweden. *J Urol* 2006;175:1337–40.
- 74 van der Slot MA, Hollemans E, den Bakker MA, *et al.* Inter-Observer variability of cribriform architecture and percent Gleason pattern 4 in prostate cancer: relation to clinical outcome. *Virchows Arch* 2021;478:249–256.
- 75 van Leenders GJLH, Verhoef EI, Hollemans E. Prostate cancer growth patterns beyond the Gleason score: entering a new era of comprehensive tumour grading. *Histopathology* 2020;77:850–61.
- 76 Bulten W, Pinckaers H, van Boven H, *et al.* Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study. *Lancet Oncol* 2020;21:233–41.
- 77 Ström P, Kartasalo K, Olsson H, *et al.* Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet Oncol* 2020;21:222–32.
- 78 Gerlinger M, Catto JW, Orntoft TF, *et al.* Intratumour heterogeneity in urologic cancers: from molecular evidence to clinical implications. *Eur Urol* 2015;67:729–37.
- 79 Brönimann S, Pradere B, Karakiewicz P, *et al.* An overview of current and emerging diagnostic, staging and prognostic markers for prostate cancer. *Expert Rev Mol Diagn* 2020;20:841–50.
- 80 Tan WCC, Nerurkar SN, Cai HY, *et al.* Overview of multiplex immunohistochemistry/immunofluorescence techniques in the era of cancer immunotherapy. *Cancer Commun* 2020;40:135–53.