Policy briefing

making science work for health | PHG FOUNDATION

Pathomics, genomics and AI in cancer care

Summary

- Pathomics, a new term in the omics field, involves quantitative analysis of digital pathology images
- In cancer, pathomics can provide spatial and morphological information about a tumour, whereas genomics provides the missing molecular information
- Combined analysis of pathomic and genomic data has been enabled by artificial intelligence (AI) models that can infer complex patterns from large and varied data types
- Such AI models are being developed across the globe to provide novel insights into cancer pathophysiology with promising applications in cancer diagnostics, prognostics, and therapeutics
- Barriers around standardising and scaling-up of pathomic analysis, and optimising sample processing for combined pathomic and genomic analysis, remain to be addressed

Pathology workflows are increasingly being digitised. This involves scanning and converting pathology glass slides into high-resolution digital images known as whole-slide images (WSIs), followed by management, sharing, and interpretation of WSIs. Such digital workflows offer benefits including remote pathology analysis and increased collaboration between pathologists. However, these workflows involve high set-up costs and are an additional step to the traditional pathology workflow, which has limited its clinical uptake.

Whole-slide images allow quantitative features to be extracted and analysed from cell and tissue images, which is now the focus of a new omics field called 'pathomics' [1]. Pathomics allows objective detection, segmentation, labelling, and classification of tumour samples, which could help reduce errors in traditional pathology assessments. Furthermore, quantification of pathological data can help create large-scale pathology databases, which can be used to train new pathologists, enable novel biomarker research, and conduct multiomic analyses, for example, combined pathomic and genomic analyses.

Combining pathomics with genomics is enabled by AI algorithms that are capable of handling large and heterogeneous data types and deriving novel insights from them. While pathomics provides an 'outside view' of the cell (the spatial and morphological information about a cancer tumour), genomics provides an 'inside view' (the molecular details of cellular activity within the tumour), improving our understanding of cancer pathophysiology [2].

WYNG FOUNDATION

We thank the WYNG Foundation for supporting the research behind this briefing paper. Views expressed in this briefing note do not necessarily reflect those of the WYNG Foundation.

Improving cancer care

Cancer is the leading cause of death in Hong Kong, with an average yearly increase of 3.6% between 2011-2021. Early intervention is essential to tackling this growing burden. The comprehensive view provided by combining pathomic and genomic data could help with many segments of the cancer pathway including [2,3]:

Early detection and improved diagnosis: AI models trained on pathomic and genomic data can identify novel pathophysiological features that may otherwise be missed by clinicians, supporting earlier and more accurate cancer diagnosis. Novel tools such as 'Pathomic Fusion' have been developed for this purpose but await use in clinics.

Personalised treatment decisions based on:

- **Tumour grading or classification:** traditionally, tumour grading is based on subjective judgements about tumour morphology and can be inconclusive. Combined knowledge of underlying genetic alterations (genomics) and quantitative image features (pathomics) could provide better information for tumour classification [2].
- **Patient stratification and prognostication:** combining pathomics and genomics means more data that improves the power of AI models to stratify patients into risk-based groups or predict overall patient survival.
- **Drug response:** access to a patient's genomic data can help clinicians make individualised drug recommendations. Pathomics further enables understanding of drug effectiveness via assessments of immune cell recruitment to tumour samples. Together, these omics data could allow clinicians to prescribe the most effective cancer drug for a patient.

Streamlining genomic testing for faster diagnosis and treatment: AI-based models capable of predicting disease-specific mutations from digital pathology images are being developed. For example, 'Panoptes', a deep learning model trained on pathomic and genomic data, can predict the mutation status of endometrial cancer from digital pathology images without the need for genomic sequencing. Model optimisation and clinical utility assessments are pending, but such software could enable pathologists to make faster and better clinical decisions. Such models also limit the need for genetic testing to cases where the model results have low confidence. In a 2023 case report on lung cancer, an algorithm which uses pathology images was used to predict disease-causing variants which could be confirmed via a rapid polymerase chain reaction [4]. The algorithm had mitigated the need for whole exome or whole genome sequencing to identify disease-causing variants, cutting time to treatment by 10-14 days [4].

Global efforts to support cancer care

Research efforts across the globe are assessing the value of combined pathomic and genomic analyses. Although several genomic data repositories are available, wider adoption of digital pathology and large-scale pathomic data repositories are needed to enable integrated pathomic and genomic analyses. In China, while complete digital pathology workflows remain to be implemented in the clinic, a cost-effectiveness study conducted in 2024 supported its wider uptake.

In the UK, the Genomics Pathology Imaging Collection – a joint initiative between Genomics England and the UK National Pathology Imaging Co-operative (NPIC) – plans to curate whole-slide images for the entire 100,000 Genomes Project cancer cohort. There are also

plans to set up Cellular Pathology Genomic Centres across the UK where pathology is used to triage patients according to clinical urgency and genomics is used to assess sample suitability for cancer vaccine studies. In British Columbia, Canada, the Digital Pathology project aims to collect 10,000 pathology slides of 30 different cancer types and build the necessary infrastructure to automate the tumour identification process for cancer diagnosis. It interfaces with the Personalised OncoGenomics (POG) program, which comprises a multidisciplinary team, including pathologists, who are evaluating the potential of genomics and transcriptomics to advance cancer care. How these two projects plan on conducting integrated pathomic and genomic analyses remains to be seen.

Barriers to clinical implementation

Despite the advantages of combining pathomics and genomics, clinical implementation is challenged by:

Sampling issues: often only a limited amount of tumour sample is available, and the choice between freezing these samples, which provides better nucleic acid quality for genomics, or using formalin-fixed, paraffin-embedded (FFPE) tissue samples, commonly used for pathology analysis, hinders the integration of pathomics and genomics. Moreover, immune cells sometimes found in tumours and normal tissue within the tumour sample can add noise to genomic analysis.

Tumour heterogeneity: pathomic and genomic data from a single biopsy may not capture the varying degrees of malignancy across the whole tumour. Multiple biopsies of a tumour may not always be possible due to location or size of the tumour. Effectively addressing tumour heterogeneity remains an open question in the field.

Challenge of integrating pathomic and genomic data: different omics modalities reflect different types of biological data that need to be extensively pre-processed - data normalisation and noise filtration need to be conducted to minimise any biases in the data. These processes are resource-intensive, require specialised knowledge, and are still being developed.

Lack of data standardisation in pathomics: whole-slide images are sensitive to the type of scanners used for image capture and inter-scanner image differences can be heterogenous enough to prevent data integration and pathomic analyses. Efforts are underway to improve standardisation, for example, via the adoption of the Digital Imaging and Communications in Medicine (DICOM) - a widely used international standard for medical images.

Amount of data to manage and store: a single whole-slide images can be four gigabytes - equivalent to a two-hour high definition (HD) movie, and approximately 5000 WSIs translates to several petabytes of data or 64-years' worth of HD movies [5]. Clinical systems currently struggle with terabytes of genomic data. Therefore, storing and managing both genomic and pathomic data is likely to present significant challenges. Cloud-based archiving options offer a potential solution although file retrieval comes at additional cost.

AI bias and explainability: AI models can inherit biases from training datasets that can lead to inaccurate diagnoses, and these models often lack transparency which makes it difficult for pathologists to understand and trust AI diagnoses. Improving the explainability of AI-based models, training these models on genetically diverse data, and ensuring that these models are properly regulated is necessary to support their clinical use.

Infrastructure and staff limitations: pathomics requires digital pathology workflows to be implemented, and combining pathomics and genomics requires a multidisciplinary team,

including pathologists, data scientists, geneticists, and bioinformaticians, to effectively manage and interpret these data types. These requirements mean high set-up costs and the need for extensive staff training.

Looking ahead

As previously mentioned, a promising area of research is training AI models on pathomic and genomic data to predict genomic features, such as mutational status, from whole-slide images. However, wider clinical adoption of pathomics is hindered by the limited uptake of digital pathology workflows, which still rely on pathology glass slides.

Slide-free direct-to-digital imaging techniques being developed such as MUSE – Microscopy with Ultraviolet Surface Excitation – could revolutionise this space. This technology uses ultraviolet light to generate tissue images without having to pre-slice tissue samples. It also uses non-damaging water-soluble dyes to stain tissue samples, which help preserve these samples for molecular assessments such as genomic analyses. This technique, compared to traditional pathology workflows, was found to reduce sample preparation time from hours to minutes.

Conclusion

Integrating pathomics and genomics for cancer care is being actively and widely researched. Several barriers remain to be overcome around sample processing and management of pathomic and genomic data, ensuring clinical settings have access to the necessary infrastructure and staff to support large-scale AI analysis and ensuring AI models in clinics are properly regulated. However, slide-free direct-to-digital imaging technologies hold the potential to transform the field of pathomics, and AI-based models that can infer molecular features from pathology slides are likely to push pathomics into clinical settings. Although the full clinical potential of combined pathomics and genomics approach remains to be realised, the future looks promising.

References

1. Gupta, R., Kurc, T., Sharma, A., Almeida, J. S. & Saltz, J. The emergence of pathomics. Curr. Pathobiol. Rep. 7, 73–84 (2019)
2. Lu, C., Shiradkar, R. & Liu, Z. Integrating pathomics with radiomics and genomics for cancer prognosis: A brief review. Chin. J. Cancer Res. 33, 563–573 (2021)
3. Yu, K.-H. et al. Association of omics features with histopathology patterns in lung adenocarcinoma. Cell Syst. 5, 620-627.e3 (2017)
4. Waissengrin, B. et al. Artificial intelligence (AI) molecular analysis tool assists in rapid treatment decision in lung cancer: a case report. J. Clin. Pathol. 76, 790–792 (2023)
5. Zuraw, A. Digital Pathology 101: All you need to know to start and continue your digital pathology journey

With sincere thanks to Dr Aleksandra Zuraw for her time and expertise in providing valuable information that has helped shape this briefing.

Author: Chaitanya Erady **Date:** December 2024