A based pathology predicts origins for cancers of unknown primary

Nikita Lee, Chloe S. Chen, Ginia G. R. Williamson, Melissa Zhao, Maria Zhang, Jaime Lacano ✉ & Samuel Aparicio ✉

Nature 904, 154–159 (2025) ~ Cite this article
Received: 11 July 2024 ~ Accepted: 11 April 2025 ~ Online: 7 May 2025

Cancers of unknown primary (CUP) origin is an enigmatic group of diagnoses in which the primary anatomical site of tumour origin cannot be determined.[1] This poses a considerable challenge to modern therapeutics and diagnostics. Although molecular assays for cancer research have focused on using genomics and transcriptomics to identify the origin of a tumour,[2] CUP remains a diagnosis of exclusion and a clinical challenge, with poor outcomes relative to known primary origins. Here, to overcome these challenges, we present a deep learning based algorithm - Tumour Origin Assessment via Deep Learning (TOAD) - that can simultaneously diagnose CUPs and provide a ranked list of plausible primary origins, for multi-lineage histology slides. We used whole-slide images of tumours with known primary origins to train a transformer based deep learning model, and we evaluated its diagnostic performance on 9,659 held-out cancer specimens from various settings. Here, to overcome these challenges, our proposed deep learning based algorithm - Tumour Origin Assessment via Deep Learning (TOAD) - that can simultaneously diagnose CUPs and provide a ranked list of plausible primary origins, for multi-lineage histology slides. We used whole-slide images of tumours with known primary origins to train a transformer based deep learning model, and we evaluated its diagnostic performance on 9,659 held-out cancer specimens from various settings. TOAD achieved a top-1 accuracy of 0.83 and a top-3 accuracy of 0.96, whereas on one external test set it achieved a top-1 accuracy of 0.842 and a top-3 accuracy of 0.949. Importantly, the further a tumour is from a given CUP, the better the performance of TOAD. Ablation analyses suggest that model predictions rely on morphological patterns that are well-known to pathologists and that TOAD's ability to differentiate between primary sites is correlated with the amount of tissue available. TOAD could be used in conjunction with or in lieu of existing tests and extensive diagnostic work-ups to reduce the uncertainty of CUP diagnosis.

This is a preview of subscription content, access via your institution

Relevant articles

Open access pathology data for cancer research

A deep learning model to triage and predict adenocarcinoma on pancreas cytology Nature Medicine, 2025

Pathologist-level AI for cancer diagnosis Nature, 2025

A cross-cohort computational framework to trace tumour tissue of origin based on RNA sequencing Nature Genetics, 2025

DNA methylation profiling to determine the primary sites of metastatic cancers using formalin-fixed paraffin-embedded samples Nature Medicine, 2025

Pan-cancer diagnostic and prognostic assessment of molecular tumor origin Nature Communications, 2025

Access options

Access through your institution

Access Nature and 54 other Nature Portfolio journals

Get Nature+, our best-value online subscription

24.99 € / month

cancel anytime

Learn more

Rent or buy this article

Prices vary by article type

from $1.95

to $39.95

Learn more

Subscribe to this journal

Receive 51 issues and online access

199.00 € per year

only 3.90 € per issue

Learn more

Prices may be subject to local taxes which are calculated during checkout

Additional access options:

- Log in
- Learn about institutional subscriptions
- Read-and-Publish
- Contact customer support

Code availability

All code implemented in Pytorch using PyTorch as the primary deep learning package. All code and scripts to reproduce the experiments of this paper are available at https://github.com/mahmoodlab/TOAD.

References

1. Kenny, S. & Fitzsimons, D. Prognosis in solving the clinical management of cancer of unknown primary in the molecular era. Int. Rev. Clin. Oncol. 87, 1–10 (2014).
Article Data Google Scholar

2. Varadhachary, G. R. & Raber, M. N. Cancer of unknown primary site. N. Engl. J. Med. 371, 757–765 (2014).
Article Data Google Scholar

3. Mouliere, F., Ciriello, G. & Grant, R. Genomics of an unknown primary - origin, diagnosis and treatment. Nature Rev. Genet. 26, 226 (2025).
Article Data Google Scholar

4. Zou, W. et al. A deep learning system to identify classifiers, prognosis and treatments of cancers of unknown primary using genomic mutations. Nat. Commun. 16, 7261 (2025).
Article Data Article Data Google Scholar

5. Tresca, A. et al. Development of genomic derived cancer type prediction to inform clinical cancer care. JMIR Cancer 8, 8e-19 (2020).
Article Data Article Data Google Scholar

6. Conran, C. A. et al. Application of neural network which transcriptomic profiling for the diagnosis of method for diagnosis of primary and metastatic cancers. JAMA Netw. Open 6, e2329720 (2023).
Article Data Article Data Google Scholar

7. Zhao, S. et al. CUP-AI-Dx: a tissue for lineage cancer tissue of origin and molecular subtype using RNA gene expression data and artificial intelligence. EBioMedicine 98, 105050 (2023).
Article Data Article Data Article Data Google Scholar

8. Xiao, Y. et al. TCGA-CUP: a gene expression web based machine learning algorithm for tissue origin diagnosis of cancers of unknown primary. Brief. Bioinformatics 22, 2016–2025 (2021).
Article Data Article Data Article Data Google Scholar

9. Kerr, S. J. et al. Multivariate validation study to determine performance characteristics of 92 gene molecular cancer classifier. Clin. Genet. 86, 383. 1952–1960 (2021).
Article Data Article Data Article Data Google Scholar

10. Hayashi, H. et al. Site-specific and targeted therapies based on molecular profiling by next-generation sequencing for cancer of unknown primary site: a nonrandomized phase 2 clinical trial. JAMA Oncol. 6, 1931–1938 (2020).
Article Data Article Data Article Data Google Scholar

11. Tsao, D. et al. MRI-TRUS and mRNA expression assessed to brain primary tumors and metastatic lesions in prostate cancer patients from metastatic brain cancers. Sci. Rep. Lond. Pathol. 39, 1–9 (2021).
Article Data Article Data Article Data Google Scholar

12. Cerulo, J. L. Wu, T. T., Rashid, A. & Ershaian, S. C. Maternal colonization by metastatic carcinoma in the gastrointestinal tract: a potential mimic of ectopic pregnancy. Am. J. Surg. Pathol. 38, 1411–1415 (2014).
Article Data Article Data Article Data Google Scholar

13. Lefkes, M. et al. Diagnostic utility of molecular profiling in cancers of unknown primary. Br. J. Cancer 108, 1111–1116 (2013).
Article Data Article Data Article Data Google Scholar

14. Lu, Y. et al. A deep learning system for differential diagnosis of skin diseases. Nat. Med. 26, 1900–1909 (2020).
Article Data Article Data Article Data Google Scholar

15. Lu, M. Y. et al. Data-efficient and weakly supervised computational pathology on whole-slide images. Nat. Biomed. Eng. https://doi.org/10.1038/s41551-020-00682-w (2021).
Article Data Article Data Article Data Google Scholar

16. Campanella, G. et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nat. Med. 25, 1301–1307 (2019).
Article Data Article Data Article Data Google Scholar

17. Chen, T. et al. An augmented reality microscope with real-time artificial intelligence integration for cancer diagnosis. Nat. Med. 25, 1453–1457 (2019).
Article Data Article Data Article Data Google Scholar

18. Huang, G. et al. Densely connected convolutional networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 4700–4708 (2017).
Article Data Article Data Article Data Google Scholar

19. Ardila, T. et al. Real-time tumor segmentation for histopathology using embedded Human Knowledge and deep neural networks. Nat. Med. 26, 12–20 (2020).
Article Data Article Data Article Data Google Scholar

20. Eastwa, A. et al. Dermatologist level classification of skin cancer with deep neural networks. Nature 542, 115–118 (2017).
Article Data Article Data Article Data Google Scholar

21. Szabo, A. et al. Pan-cancer diagnostic consensus through tumor purity and histopathology images using artificial intelligence. NPJ Digit. Med. 5, 16 (2022).
Article Data Article Data Article Data Google Scholar

22. Coudray, N. et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat. Med. 24, 1559–1567 (2018).
Article Data Article Data Article Data Google Scholar

23. Kather, J. N. et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nat. Med. 25, 1054–1056 (2019).
Article Data Article Data Article Data Google Scholar

24. Pugliese, R. et al. Predictions of cardiovascular risk factors from retinal fundus photographs via deep learning. Nat. Biomed. Eng. 5, 2199–2209 (2021).
Article Data Article Data Article Data Google Scholar

25. Steinberg, E. K. & Weinberg, M. Maternal-fetal deep multiplex tissue imaging. In International Conference on Machine Learning 32102–3141 (2000).
Article Data Article Data Article Data Google Scholar

26. Hamdel, C. K. et al. A multicenter study identifying the diagnostic accuracy of gene expression assay and gene expression profiling testing for cancer of unknown primary. Mod. Pathol. 28, 1507–1517 (2015).
Article Data Article Data Article Data Google Scholar

27. McHugh, M. L. Interrater reliability: the kappa statistic. Biochem. Med. 22, 276–282 (2012).
Article Data Article Data Article Data Google Scholar

28. Shoshan, A. et al. Metastatic adenocarcinoma of an unknown primary site. A comparison of the molecular characteristics of metastatic cancer of unknown primary and 2nd and 3rd metastasizing sites. Int. J. Clin. Pathol. 99, 729–735 (1992).
Article Data Article Data Article Data Google Scholar

29. Jermiaw, F. et al. Machine learning analysis of RNA methylation profiles distinguishes primary lung squamous cell carcinomas from head and neck metastases. Sci. Transl. Med. 11, eaaw8513 (2019).
Article Data Article Data Article Data Google Scholar

30. Rosa, J. P. Bellido, R. The platelets: the biology of platelets and their role in cancer. Cancer Epidemiol. 48, 139–144 (2009).
Article Data Article Data Article Data Google Scholar

31. Xu, H. et al. Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 770–778 (2015). 2016.
Article Data Article Data Article Data Google Scholar

32. Bandanaviz, O. et al. Imagenet large scale visual recognition challenge. Int. J. Comput. Vis. 115, 211–252 (2015).
Article Data Article Data Article Data Google Scholar

Acknowledgements

We thank the staffs at the Vancouver Prostatectomy Cohort and the Vancouver Breast Cancer Cohort, Cancer Genetics Laboratory, Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, Canada, for their technical support. This study was funded by the BWH-Dana-Farber and Harvard Cancer Center Support Grant, NCI Center Grant P30CA06516, Susan G. Komen Foundation SAC160097, NIH/NCI R01CA240639, NIH/NCI R01CA240639-01A1, Google Cloud Research Credits and Google Cloud Platform Credits, and a research grant supported by the Natural Sciences and Engineering Research Council's Training Program, NEW/GR TUL/400/PD. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, National Institute of General Medical Sciences or the National Cancer Institute.

Author information

These authors contributed equally: Tiffany C. Chen, Ginia T. R. Williamson

Authors and Affiliations

Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
Nikita Lee, Tiffany C. Chen, Ginia T. R. Williamson, Melissa Zhao, Maria Shady, Jaime Lacano & Samuel Aparicio

Cancer Program, Broad Institute of Harvard and MIT, Cambridge, MA, USA
Nikita Lee, Tiffany C. Chen, Ginia T. R. Williamson, Maria Shady, Jaime Lacano & Samuel Aparicio

Cancer Data Science Program, Dana-Farber Cancer Institute, Boston, MA, USA
Nikita Lee, Tiffany C. Chen, Ginia T. R. Williamson, Maria Shady, Jaime Lacano & Samuel Aparicio

Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA
Nikita Lee, Tiffany C. Chen, Ginia T. R. Williamson, Maria Shady, Jaime Lacano & Samuel Aparicio

Contributions

N.L. and T.C. conceived the study and designed the experiments. N.L. performed the experiments, analyzed the results, and wrote the manuscript. T.C. and G.T.R. contributed to the manuscript. M.Z., M.S., J.L., and S.A. contributed to the manuscript. T.C. and G.T.R. developed the deep learning model. N.L. and T.C. prepared the manuscript with input from all co-authors. S.A. supervised the research.

Corresponding author

Correspondence: Nikita Lee, Samuel Aparicio

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information: Nature thanks Benjamin Kleinstein and Kim Linehan, anonymous reviewers for their contribution to the peer review of this work.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Overall study design.

a, A histopathologist reviews and annotates cancerous tissue in primary tumors. For model development and testing, we collected, mined, & curated 70,477 digital diagnostic slides from 9,659 patients from multiple sites, including the Mass General Brigham (MGB), Brigham and Women's Hospital (BWH) and Dana-Farber Cancer Institute (DFCI), and the Broad Institute. The majority of cases (5,205 slides) was used for validation during training to select the best performing model, and the remaining cases (65,272 slides) were used for training. We trained a deep learning model to predict the tissue of origin for a given whole-slide image (WSI). We evaluated our model on an external test cohort of 1,062 cases, collected from sites (2013 and international medical centers). The model then was deployed to predict the tissue of origin for cancers of unknown primary tumors. Lastly, to assess the ability of the model to inform meaningful predictions for unsure cases, we evaluated the model's top-3 accuracy on a subset of the internal and external test sets. **b**, Our model's additional diagnostic utility of CUP is based on the assumption that a model's prediction is less accurate and unsure the CUP. Although the primary cancer could not be initially assigned for all patients, we hypothesized that the model would be more accurate for patients with primary sites and therefore tests, we identified a subset of 57 cases for which a primary differential was eventually assigned, and compared the model's performance on these cases versus the remaining cases. **c**, Our model's clinical utility is based on its ability to stratify patients, illustrating the applicability of the model to non-metastatic cancer diagnosis in the clinical setting.

Extended Data Fig. 2 Classification performance for the prediction of cancer primary versus metastatic tumors.

Receiver operating characteristic curves for our model's ability to differentiate between primary and metastatic tumors for the entire cohort (n = 1,475 cases). Glioma was excluded as there were no metastatic samples in our cohort. Area under the receiver operating characteristic curve (AUROC) scores are provided as a gloss by the model. The micro-averaged, one-versus-rest AUC, Brier, 5-fold cross-validation 95% confidence intervals are in parentheses. Micro-averaged and macro-averaged AUROC scores are 0.960 (0.953–0.966) and 0.959 (0.952–0.966), respectively. Abbreviations: Adrenal gland (ADRG), Bladder (BLAD), Breast (BRST), Bronchus and lung (BRON), Cervix (CERV), Colon and rectum (COREC), Esophagus (ESCA), Head and neck (HNECK), Kidney (KIDNEY), Liver (LIVER), Lymphoid (LYMPH), Myeloid (MYELOID), Ovary (OVARY), Pancreas (PANCREAS), Peripheral nervous system (PNS), Prostate (PROSTATE), Skin (SKIN), Soft tissue (SOFTTISSUE), Stomach (STOMACH), Testis (TESTIS), Thyroid (THYROID), Uterus (UTERUS).

Extended Data Fig. 3 Performance for the prediction of cancer origin on metastatic and primary tumors.

a, AUROC performance for the top-3 predictions on metastatic tumors. The AUROC is only with associated 95% confidence intervals and AUROC curves are shown for organ sites in cases with at least 30 metastatic samples in the cohort (n = 1,475 cases). The micro- and macro-averaged AUROC scores are 0.960 (0.953–0.966). The multi-class AUROC scores are 0.960 (0.953–0.966). **b**, Performance for the top-3 predictions on primary tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). **c**, Performance for the top-3 predictions on metastatic tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). **d**, Performance for the top-3 predictions on primary tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). Abbreviations: Adrenal gland (ADRG), Bladder (BLAD), Breast (BRST), Bronchus and lung (BRON), Cervix (CERV), Colon and rectum (COREC), Esophagus (ESCA), Head and neck (HNECK), Kidney (KIDNEY), Liver (LIVER), Lymphoid (LYMPH), Myeloid (MYELOID), Ovary (OVARY), Pancreas (PANCREAS), Peripheral nervous system (PNS), Prostate (PROSTATE), Skin (SKIN), Soft tissue (SOFTTISSUE), Stomach (STOMACH), Testis (TESTIS), Thyroid (THYROID), Uterus (UTERUS).

Extended Data Fig. 4 Model performance on the binary problem of distinguishing between primary and metastatic tumors.

a, Receiver operating characteristic curves for metastatic versus primary tumors. The AUC ROC(+/-) with associated 95% confidence intervals and ROC curves are shown for organ sites in cases with at least 30 metastatic samples in the cohort (n = 1,475). The micro- and macro-averaged AUROC scores are 0.960 (0.953–0.966). The multi-class AUROC scores are 0.960 (0.953–0.966). **b**, Performance for the top-3 predictions on metastatic tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). **c**, Performance for the top-3 predictions on primary tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). **d**, Performance for the top-3 predictions on metastatic tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). Abbreviations: Adrenal gland (ADRG), Bladder (BLAD), Breast (BRST), Bronchus and lung (BRON), Cervix (CERV), Colon and rectum (COREC), Esophagus (ESCA), Head and neck (HNECK), Kidney (KIDNEY), Liver (LIVER), Lymphoid (LYMPH), Myeloid (MYELOID), Ovary (OVARY), Pancreas (PANCREAS), Peripheral nervous system (PNS), Prostate (PROSTATE), Skin (SKIN), Soft tissue (SOFTTISSUE), Stomach (STOMACH), Testis (TESTIS), Thyroid (THYROID), Uterus (UTERUS).

Extended Data Fig. 5 Model performance on difficult metastatic and unknown primary tumors.

a, The performance of the model for the prediction of cancer origin on a subset of cases in top 1 accuracy (n = 1,475) and cancers in top 1 accuracy (n = 1,475) and metastatic tumors in the held-out cohort (n = 1,475). The model's performance is shown depending on the number of diagnosis. ROC curves and AUROC scores are shown for the top-3 predictions on metastatic tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). **b**, Performance for the top-3 predictions on primary tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). **c**, Performance for the top-3 predictions on metastatic tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). **d**, Performance for the top-3 predictions on primary tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). Abbreviations: Adrenal gland (ADRG), Bladder (BLAD), Breast (BRST), Bronchus and lung (BRON), Cervix (CERV), Colon and rectum (COREC), Esophagus (ESCA), Head and neck (HNECK), Kidney (KIDNEY), Liver (LIVER), Lymphoid (LYMPH), Myeloid (MYELOID), Ovary (OVARY), Pancreas (PANCREAS), Peripheral nervous system (PNS), Prostate (PROSTATE), Skin (SKIN), Soft tissue (SOFTTISSUE), Stomach (STOMACH), Testis (TESTIS), Thyroid (THYROID), Uterus (UTERUS).

Extended Data Fig. 6 Model performance on difficult metastatic and unknown primary tumors.

a, The performance of the model for the prediction of cancer origin on a subset of cases in top 1 accuracy (n = 1,475) and cancers in top 1 accuracy (n = 1,475) and metastatic tumors in the held-out cohort (n = 1,475). The model's performance is shown depending on the number of diagnosis. ROC curves and AUROC scores are shown for the top-3 predictions on metastatic tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). **b**, Performance for the top-3 predictions on primary tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). **c**, Performance for the top-3 predictions on metastatic tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). **d**, Performance for the top-3 predictions on primary tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). Abbreviations: Adrenal gland (ADRG), Bladder (BLAD), Breast (BRST), Bronchus and lung (BRON), Cervix (CERV), Colon and rectum (COREC), Esophagus (ESCA), Head and neck (HNECK), Kidney (KIDNEY), Liver (LIVER), Lymphoid (LYMPH), Myeloid (MYELOID), Ovary (OVARY), Pancreas (PANCREAS), Peripheral nervous system (PNS), Prostate (PROSTATE), Skin (SKIN), Soft tissue (SOFTTISSUE), Stomach (STOMACH), Testis (TESTIS), Thyroid (THYROID), Uterus (UTERUS).

Source data

Extended Data Fig. 7 Model performance on difficult metastatic and unknown primary tumors.

a, The performance of the model for the prediction of cancer origin on a subset of cases in top 1 accuracy (n = 1,475) and cancers in top 1 accuracy (n = 1,475) and metastatic tumors in the held-out cohort (n = 1,475). The model's performance is shown depending on the number of diagnosis. ROC curves and AUROC scores are shown for the top-3 predictions on metastatic tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). **b**, Performance for the top-3 predictions on primary tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). **c**, Performance for the top-3 predictions on metastatic tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). **d**, Performance for the top-3 predictions on primary tumors (n = 1,475). The multi-class AUROC scores are 0.960 (0.953–0.966). Abbreviations: Adrenal gland (ADRG), Bladder (BLAD), Breast (BRST), Bronchus and lung (BRON), Cervix (CERV), Colon and rectum (COREC), Esophagus (ESCA), Head and neck (HNECK), Kidney (KIDNEY), Liver (LIVER), Lymphoid (LYMPH), Myeloid (MYELOID), Ovary (OVARY), Pancreas (PANCREAS), Peripheral nervous system (PNS), Prostate (PROSTATE), Skin (SKIN), Soft tissue (SOFTTISSUE), Stomach (STOMACH), Testis (TESTIS), Thyroid (THYROID), Uterus (UTERUS).

Source data

Download references ↓

Extended Data Fig. 7 Examples of metastases from colorectal, breast and lung primary tumours with attention heat maps.

Source data

a–**c**. Example metastases from colorectal (**a**), breast (**b**) and lung (**c**) primary tumours are shown. For each case, the attention heat map of the model is displayed on top of the original H&E WSI as a semi-transparent overlay in which the overlaid regions range from crimson (high attention, high diagnostic relevance) to navy (low attention, low diagnostic relevance). Left, sites of metastasis are shown, including the lung, lymph node (LN), liver and brain. Right, H&E images show, from left to right, low magnification with corresponding attention map, medium magnification with corresponding attention map, and high-magnification patches. **a**, Medium- and high-magnification views demonstrate so-called 'dirty necrosis' and variably sized glands with densely packed, hyperchromatic nuclei that are characteristic of colorectal adenocarcinoma. **b**, Medium- and high magnification views demonstrate sheets of cells as well as small tubules and glands—morphologies that are consistent with metastatic breast carcinomas. **c**, Medium- and high-magnification views demonstrate sheets of cells, variably sized glands and cells in infiltrative single files. The cells have large, hyperchromatic nuclei and high nuclear:cytoplasmic ratios, which are consistent with metastatic lung carcinomas. **a**–**c**, The attention heat maps allow the predictions of the model for each case to be visually interpretable for human experts, revealing the morphological features used by the model for the determination of the classification. High-resolution heat maps for cases from all primary sites can be accessed through our interactive demo website (http://toad.mahmoodlab.org).

Extended Data Fig. 8 TOAD-assisted CUP work-up: example 1.

Top, a representative case that underwent a standard CUP work-up involving extensive IHC staining and clinical correlation. Strong PAX8 staining suggested a Müllerian origin and multiple IHC tests were used to rule out other primary tumours. Retrospectively, we analysed the case with TOAD and found that the top-5 determinations were ovarian, breast and lung, and, after this determination, that only three IHC stains (PAX8, GATA3 and TTF1) needed to be used to confirm a Müllerian origin and rule out breast carcinoma and lung adenocarcinoma. This workflow demonstrates how TOAD can be used as an assistive diagnostic tool. Bottom, medium magnification and corresponding heat maps of representative areas of tumour, with high-magnification, high-attention patches on the right outlined in crimson and low attention patches outlined in navy.

Extended Data Fig. 9 Analysis of high-attention regions for metastatic tumours.

Relative counts of different cell types localized within the high-attention regions proposed by the model were quantified. Specifically, the top-10 high-attention patches from each slide were extracted at the 20× equivalent magnification and a HoverNet[15] model trained for multi-organ nucleus segmentation and classification was used to detect different cellular populations including tumour cells (red), lymphocytes (green), connective tissue (blue), dead cells (yellow) and non-neoplastic epithelial cells (orange). The fraction of cells for each cell type is plotted using box plots for all metastatic slides in the test set ($n = 1{,}408$) and is stratified by each primary origin site: lung ($n = 236$), breast ($n = 231$), colorectal ($n = 175$), pancreatobiliary ($n = 122$), skin ($n = 111$), ovarian ($n = 102$), renal ($n = 79$), prostate ($n = 64$), head and neck ($n = 57$), oesophagogastric ($n = 52$), thyroid ($n = 43$), bladder ($n = 42$), germ cell ($n = 32$), endometrial ($n = 21$), liver ($n = 18$), adrenal ($n = 12$) and cervix ($n = 11$). Boxes indicate quartile values and whiskers extend to data points within 1.5× the interquartile range. This analysis demonstrates in addition to the attention heat maps, that the model attends strongly to regions of tumour presence for its predictions.

Source data

Extended Data Fig. 10 Classification performance of adenocarcinoma network, squamous cell carcinoma network and site-specific networks for tumour metastasized to the liver and lymph node.

a, **b**, Often pathologists can readily distinguish between adenocarcinoma and squamous cell carcinoma based on the morphological and architectural appearance of the tumour cells that are present in the tissue. However, within the respective family of adenocarcinoma and squamous cell carcinoma subtypes, determining the origin of the tumour can remain a challenging task. Therefore, we hypothesized that we can develop models to specifically predict the origin of tumours for top primary sites of adenocarcinoma (**a**) and squamous cell carcinoma (**b**). Cases from six primary sites (breast, lung, colorectal, pancreatobiliary, prostate and oesophagogastric) and four primary sites (head and neck, lung, cervix and oesophagogastric) were chosen for the development of the adenocarcinoma and squamous cell carcinoma classifiers, respectively, based on their frequency in the database. We also explored the additional scenarios of predicting the primary origins of metastatic tumours grouped by a common metastatic site, including the liver (**c**) and lymph node (**d**). Cases of metastasis from the top four and top-seven primary origins for liver and lymph nodes, respectively, were chosen on the basis of their frequency in our database. See Methods. Additional experiments and analysis' for details. **a**–**d**, Left, the confusion matrix, along with the precision and recall of each class and its count is plotted for the adenocarcinoma model test set (**a**; $n = 2{,}920$) and squamous cell carcinoma model test set (**b**; $n = 621$), the liver metastasis (met.) site test set (**c**; $n = 223$) and lymph node metastasis site test set (**d**; $n = 318$), respectively. Consistent with the model developed using examples of all 18 primary sites, the adenocarcinoma-, squamous cell-carcinoma- and site-specific models were trained by including the sex of the patient. Performance for models trained with and without the sex of the patient in terms of the micro-averaged, one-versus-rest AUC ROC (middle) and F_1-scores for each primary site and overall model accuracy (micro-averaged F_1-score) (right) are shown. All error bars indicate 95% confidence intervals, plotted around the computed value of each classification performance metric (specified by its respective axis labels).

Source data

Supplementary information

Supplementary Information

This file contains Supplementary Figures 1-6, Supplementary Tables 1, 3, 6, 7, 11, 13 and 14, and legends for Supplementary Tables 2, 4, 5, 8, 9, 10 and 12 (see separate Excel file for these Tables).

Reporting Summary

Supplementary Tables

This file contains Supplementary Tables 2, 4, 5, 8, 9, 10 and 12.

Source data

Source Data Fig. 2

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 6

Source Data Extended Data Fig. 9

Source Data Extended Data Fig. 10

Rights and permissions

Reprints and Permissions

About this article

Check for updates

Cite this article

Lu, M.Y., Chen, T.Y., Williamson, D.F.K. *et al.* AI-based pathology predicts origins for cancers of unknown primary. *Nature* **594**, 106–110 (2021). https://doi.org/10.1038/s41586-021-03512-4

Download citation

Received
27 June 2020

Accepted
01 April 2021

Published
05 May 2021

Issue Date
03 June 2021

DOI
https://doi.org/10.1038/s41586-021-03512-4

Subjects

Cancer of unknown primary · Computational science · Image processing · Machine learning · Pathology

This article is cited by

Performance comparison between multi-center histopathology datasets of a weakly-supervised deep learning model for pancreatic ductal adenocarcinoma detection
Francisco Carrillo-Perez, Francisco M. Ortuno ... Luis Javier Herrera
Cancer Imaging (2023)

Risk factors for cancer of unknown primary: a literature review
Karlijn E. P. E. Hermans, Fatemeh Kazemzadeh ... Leo J. Schouten
BMC Cancer (2023)

A cross-cohort computational framework to trace tumor tissue-of-origin based on RNA sequencing
Binsheng He, Hongmei Sun ... Bo Wang
Scientific Reports (2023)

A deep learning model to triage and predict adenocarcinoma on pancreas cytology whole slide imaging
Andrew Sohn, Daniel Miller ... Alexander Baras
Scientific Reports (2023)

Deep learning-based pathology signature could reveal lymph node status and act as a novel prognostic marker across multiple cancer types
Sitong Chen, Jinyi Xiang ... Xiao Han
British Journal of Cancer (2023)

Comments

By submitting a comment you agree to abide by our Terms and Community Guidelines. If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

© 2023 Springer Nature Limited