nature portfolio

Corresponding author(s): Shenghong Ju

Last updated by author(s): Mar 1, 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	We used ITK-SNAP 3.8.0 for data visualization.
Data analysis	All statistical analyses were conducted using SPSS version 22.0 (IBM), GraphPad Prism version 8, R version 4.0.2, and Python version 3.9.7. The model's code is publicly available on GitHub https://github.com/GFizzz/Deep-Learning-CT-Signature-for-Predicting-Early-Liver-Metastases).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The in-house data used in this study, obtained with institutional permission, contain sensitive patient information and are therefore not publicly available. Access requests may be made to the corresponding author with a detailed research proposal and are subject to approval and a data use agreement. Data use is restricted

to non-commercial academic research purposes only. However, source data supporting the findings of this study are provided within the paper and its supplementary information files.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	In this study, the biological attribute of sex (male/female) was used as a covariate in the analysis. Gender identity or presentation was not collected or analyzed, as the research focus is on biological and clinical factors related to PDAC, which are primarily associated with biological sex rather than social gender constructs.
Reporting on race, ethnicity, or other socially relevant groupings	The study cohort was predominantly composed of individuals of Han Chinese ethnicity, which is the largest ethnic group in China. Given the high degree of homogeneity within this population, race and ethnicity were not used as categorization variables in the primary analyses of this study.
Population characteristics	A total of 905 patients who underwent upfront surgery [mean (SD) age, 64.58 (9.41) years; 538 (59.4%) males] and 158 patients who received NAT followed by surgery [mean (SD) age, 62.01 (8.54) years; 78 (49.4%) males] were included in the study.
Recruitment	Patients diagnosed with localized PDAC and who underwent upfront surgery with pathological confirmation were retrospectively enrolled from five institutions. the retrospective study design inherently introduces potential selection bias and unmeasured confounding.
Ethics oversight	This study was approved by the Institutional Review Board (IRB) with protocol number 2021ZDSYLL029-Y01.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample size was determined by the availability of eligible patients within the institutional database who met the predefined inclusion criteria for the study. No formal statistical power calculation was performed prior to data collection, as this is a retrospective study where the available cohort size is constrained by historical data.
Data exclusions	Surgery-related death: Patients who died within 30 days after surgery due to complications directly attributable to the surgical procedure were excluded from the analysis. Lost to follow-up: Patients who did not have complete follow-up data available. Image artifact: Patients whose preoperative CT images contained significant artifacts that impaired the quality of image interpretation and potentially affected the accuracy of model training or evaluation were excluded.
Replication	The Mamba-based model was first trained on a training cohort (n=450), then validated on an internal validation cohort (n=113) derived from the same institution, and finally tested on an external validation cohort (n=342) from a different institution.
Randomization	Randomization was not applicable to this retrospective study design.
Blinding	Blinding was implemented during the analysis phase. The investigators responsible for evaluating the model's performance on the validation cohorts were blinded to the patients' treatment (upfront surgery vs. NAT) and clinical outcomes at the time of analysis.

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	
Research sample	
Sampling strategy	

Data collection	
Timing	
Data exclusions	
Non-participation	
Randomization	

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	
Research sample	
Sampling strategy	
Data collection	
Timing and spatial scale	
Data exclusions	
Reproducibility	
Randomization	
Blinding	

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	
Location	
Access & import/export	
Disturbance	

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Validation

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

Authentication

Mycoplasma contamination

Commonly misidentified lines
(See ICLAC register)

Palaeontology and Archaeology

Specimen provenance

Specimen deposition

Dating methods

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals

Wild animals

Reporting on sex

Field-collected samples

Ethics oversight

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	NA
Study protocol	NA
Data collection	Patients with localized pancreatic ductal adenocarcinoma (PDAC) who underwent upfront surgery with pathological confirmation were retrospectively enrolled from four centers: The First Affiliated Hospital of the University of Science and Technology of China (Center 1), The First Affiliated Hospital of Wenzhou Medical University (Center 2), Zhongda Hospital of Southeast University (Center 3), and Northern Jiangsu People's Hospital (Center 4). For external validation, a separate cohort was prospectively collected from The First Affiliated Hospital, Zhejiang University School of Medicine (Center 5), using identical inclusion criteria. To assess the Mamba-based model's utility in personalized neoadjuvant therapy (NAT) management, an additional cohort of patients receiving NAT at Center 5 between June 2019 and December 2022 was included, along with a contemporaneous comparator group undergoing

upfront surgery during the same period. Data collection spanned from January 2015 to December 2022 across all centers, with standardized electronic medical record review and pathology review conducted by designated investigators.

Outcomes

Patients who developed liver metastasis within six months postoperatively, confirmed by radiological or pathological evidence, were classified as having ELM.

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☒	☐	Public health
☒	☐	National security
☒	☐	Crops and/or livestock
☒	☐	Ecosystems
☒	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☒	☐	Demonstrate how to render a vaccine ineffective
☒	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☒	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☒	☐	Increase transmissibility of a pathogen
☒	☐	Alter the host range of a pathogen
☒	☐	Enable evasion of diagnostic/detection modalities
☒	☐	Enable the weaponization of a biological agent or toxin
☒	☐	Any other potentially harmful combination of experiments and agents

Plants

Seed stocks

NA

Novel plant genotypes

NA

Authentication

NA

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

Files in database submission

Genome browser session
(e.g. UCSC)

Methodology

Replicates

Sequencing depth

Antibodies

Peak calling parameters

Data quality

Software

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Instrument

Software

Cell population abundance

Gating strategy

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type

Design specifications

Behavioral performance measures

Acquisition

Imaging type(s)

Field strength

Sequence & imaging parameters

Area of acquisition

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software

Normalization

Normalization template

Noise and artifact removal

Volume censoring

Statistical modeling & inference

Model type and settings

Effect(s) tested

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference

(See Eklund et al. 2016)

Correction

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Graph analysis

Multivariate modeling and predictive analysis