Artificial intelligence-based pathology as a biomarker of sensitivity to atezolizumab–bevacizumab in patients with hepatocellular carcinoma: a multicentre retrospective study

Qinghe Zeng, Christophe Klein, Stefano Caruso, Pascale Maille, Daniela S Allende, Beatriz Mínguez, Massimo Iavarone, Massih Ningarhari, Andrea Casadei-Gardini, Federica Pedica, Margherita Rimini, Riccardo Perbellini, Camille Boulagnon-Rombi, Alexandra Heurgué, Marco Maggioni, Mohamed Rela, Mukul Vij, Sylvain Baulande, Patricia Legoix, Sonia Lameiras, the HCC-AI study group, Léa Bruges, Viviane Gnemmi, Jean-Charles Nault, Claudia Campani, Hyungjin Rhee, Young Nyun Park, Mercedes Iñarrairaegui, Guillermo Garcia-Porrero, Josepmaria Argemi, Bruno Sangro, Antonio D'Alessio, Bernhard Scheiner, David James Pinato, Matthias Pinter, Valérie Paradis, Aurélie Beaufrère, Simon Peter, Lorenza Rimassa, Luca Di Tommaso, Arndt Vogel, Sophie Michalak, Jérôme Boursier, Nicolas Loménie, Marianne Ziol, Julien Calderaro*

Summary

Background Clinical benefits of atezolizumab plus bevacizumab (atezolizumab–bevacizumab) are observed only in a subset of patients with hepatocellular carcinoma and the development of biomarkers is needed to improve therapeutic strategies. The atezolizumab–bevacizumab response signature (ABRS), assessed by molecular biology profiling techniques, has been shown to be associated with progression-free survival after treatment initiation. The primary objective of our study was to develop an artificial intelligence (AI) model able to estimate ABRS expression directly from histological slides, and to evaluate if model predictions were associated with progression-free survival.

Methods In this multicentre retrospective study, we developed a model (ABRS-prediction; ABRS-P), which was derived from the previously published clustering-constrained attention multiple instance learning (or CLAM) pipeline. We trained the model fit for regression analysis using a multicentre dataset from The Cancer Genome Atlas (patients treated by surgical resection, n=336). The ABRS-P model was externally validated on two independent series of samples from patients with hepatocellular carcinoma (a surgical resection series, n=225; and a biopsy series, n=157). The predictive value of the model was further tested in a series of biopsy samples from a multicentre cohort of patients with hepatocellular carcinoma treated with atezolizumab–bevacizumab (n=122). All samples in the study were from adults (aged ≥18 years). The validation sets were sampled between Jan 1, 2008, to Jan 1, 2023. For the multicentre validation set, the primary objective was to assess the association of high versus low ABRS-P values, defined relative to cross-validation median split thresholds in the first biopsy series, with progression-free survival after treatment initiation. Additionally, we performed spatial transcriptomics and matched prediction heatmaps with in situ expression profiles.

Findings Of the 840 patients sampled, 641 (76%) were male and 199 (24%) were female. Across the development and validation datasets, hepatocellular carcinoma risk factors included alcohol intake, hepatitis B and C virus infections, and non-alcoholic steatohepatitis. Using cross-validation in the development series, the mean Pearson's correlation between ABRS-P values and ABRS score (mean expression of ABRS genes) was 0·62 (SD 0·09; mean p<0·0001, SD<0·0001). The ABRS-P generalised well on the external validation series (surgical resection series, r=0·60 [95% CI 0·51–0·68], p<0·0001; biopsy series, r=0·53 [0·40–0·63], p<0·0001). In the 122 patients treated with atezolizumab–bevacizumab, those with ABRS-P-high tumours (n=74) showed significantly longer median progression-free survival than those with ABRS-P-low tumours (n=48) after treatment initiation (12 months [95% CI 7–not reached] *vs* 7 months [4–9]; p=0·014). Spatial transcriptomics showed significantly higher ABRS score, along with upregulation of various other immune effectors, in tumour areas with high ABRS-P values versus areas with low ABRS-P values.

Interpretation Our study indicates that AI applied on hepatocellular carcinoma digital slides is able to serve as a biomarker for progression-free survival in patients treated with atezolizumab–bevacizumab. This approach could be used in the development of inexpensive and fast biomarkers for targeted therapies. The combination of AI heatmaps with spatial transcriptomics provides insight on the molecular features associated with predictions. This methodology could be applied to other cancers or diseases and improve understanding of the biological mechanisms that drive responses to treatments.

Funding Institut National du Cancer, Fondation ARC, China Scholarship Council, Ligue Contre le Cancer du Val de Marne, Fondation de l'Avenir, Ipsen, and Fondation Bristol Myers Squibb Pour la Recherche en Immuno-Oncologie.

Lancet Oncol 2023; 24: 1411–22

Published **Online** November 8, 2023 https://doi.org/10.1016/S1470-2045(23)00468-0

*Members listed in appendix 1 (pp 16–17)

Centre d'Histologie, d'Imagerie et de Cytométrie, Centre de Recherche des Cordeliers, INSERM, Sorbonne Université, Université Paris Cité, Paris, France (Q Zeng MSc, C Klein MSc); **Laboratoire d'Informatique Paris Descartes, Université Paris Cité, Paris, France** (Q Zeng, N Loménie PhD); **Université Paris Est Créteil, INSERM, IMRB, Créteil, France** (S Caruso PhD, P Maille MSc, Prof J Calderaro MD); **Department of Pathology, Henri Mondor-Albert Chenevier University Hospital, AP-HP, Créteil, France** (S Caruso, P Maille, Prof J Calderaro); **Pathology Department and Robert J Tomsich Pathology and Laboratory Medicine Institute, Cleveland Clinic, Cleveland, OH, USA** (D S Allende MD); **Liver Unit, Hospital Universitario Vall d´Hebron, Barcelona, Spain** (B Mínguez MD); **Liver Cancer Research Group, Liver Diseases, Vall d'Hebron Institut de Recerca, Barcelona, Spain** (B Mínguez); **Department of Medicine, Universitat Autònoma de Barcelona, Bellaterra, Spain** (B Mínguez); **Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas, Instituto de Salud Carlos III, Madrid, Spain** (B Mínguez); **Division of Gastroenterology and Hepatology** (M Iavarone MD, R Perbellini MD) **and**

Department of Pathology (M Maggioni MD), Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy; Centre Hospitalier Universitaire de Lille, Hôpital Huriez, Maladies de l'Appareil Digestif, Lille, France (M Ningarhari PhD, L Bruges MD); Université de Lille, INSERM, INFINITE, Lille, France (M Ningarhari, L Bruges); Department of Oncology, Vita-Salute San Raffaele University, IRCCS San Raffaele Scientific Institute Hospital, Milan, Italy (A Casadei-Gardini MD, M Rimini MD); Pathology Unit, IRCCS Ospedale San Raffaele, Milan, Italy (Prof F Pedica MD); Reims University Hospital, Department of Pathology, Reims, France (C Boulagnon-Rombi PhD); Research Unit CNRS UMR 7369 MEDyC, Université de Reims Champagne-Ardenne, Faculté de Médecine de Reims, Reims, France (C Boulagnon-Rombi); Reims University Hospital, Department of Hepatology, Reims, France (A Heurgué MD); The Institute of Liver Disease and Transplantation (Prof M Rela MD) and Department of Pathology (M Vij MD), Dr Rela Institute and Medical Centre, Bharath Institute of Higher Education and Research, Chennai, India; Institut Curie Genomics of Excellence NGS Platform, Institut Curie, Paris, France (S Baulande PhD, P Legoix MSc, S Lameiras MSc); Service d'Anatomie Pathologique, Centre de Biologie Pathologique, CHU Lille, Lille, France (V Gnemmi PhD); JPARC-Jean-Pierre Aubert Research Center, Lille, France (V Gnemmi); AP-HP Paris Nord, Hôpital Universitaire Avicenne, Service d'Hépatologie, Paris, France (Prof J-C Nault MD); Centre de Recherche des Cordeliers, Sorbonne Université, Paris, France (Prof J-C Nault MD, C Campani MD); INSERM, Université de Paris Cité, Functional Genomics of Solid Tumors, Paris, France (Prof J-C Nault, C Campani); Department of Radiology, Research Institute of Radiological Science, Center for Clinical Imaging Data Science, Severance Hospital, Yonsei University College of Medicine, Seoul, South Korea

Research in context

Evidence before this study

We searched PubMed on May 15, 2023, for relevant articles published from database inception using the search terms "artificial intelligence" OR "deep learning" OR "machine learning" AND "treatment" AND "slides", without language restrictions. Few artificial intelligence (AI) studies have been performed with the aim to directly assess the sensitivity of patients to targeted and immune therapies from whole-slide images. No research had investigated whether gene signatures associated with response to anticancer drugs could be estimated with use of AI-based pathology.

Added value of this study

Our multicentre study shows that AI-based pathology can be used as a biomarker for patients with cancers treated with systemic therapies. Our findings also indicate the robustness of such models, with validation on two datasets characterised by clinical and technical differences compared with the development set. Additionally, we developed a pipeline that matches deep learning prediction heatmaps with spatial transcriptomics, and showed, in situ, that areas predicted to have high atezolizumab–bevacizumab response signature (ABRS) values displayed higher ABRS expression than areas predicted to have low ABRS values.

Implications of all the available evidence

Pending prospective validation, this novel type of AI-based biomarker could offer a rapid and lower cost alternative to current techniques for the identification of patients who are likely to benefit from specific treatments, including immunotherapy. Combination of AI heatmaps with spatial profiling could be applied to other cancers or diseases and improve understanding of the biological mechanisms that drive response or resistance to treatments.

Introduction

Hepatocellular carcinoma is the sixth most common cancer and the third cause of cancer-related deaths worldwide.[1] Combination therapy with the immune checkpoint inhibitor atezolizumab and the anti-angiogenic drug bevacizumab (atezolizumab–bevacizumab) is now one of the standards of care for patients with advanced disease.[2,3] However, responses are observed only in a minority of patients and the development of biomarkers is needed to improve clinical management.[3]

Zhu and colleagues recently identified a gene signature, the atezolizumab–bevacizumab response signature (ABRS), associated with progression-free survival after atezolizumab–bevacizumab initiation.[4] The genes in this signature are *CXCR2P1*, *ICOS*, *TIMD4*, *CTLA4*, *PAX5*, *KLRC3*, *FCRL3*, *AIM2*, *GBP5*, and *CCL4*, with some of these genes involved in the regulation of T-cell activation (*CTLA4* and *ICOS*) and innate immunity (*KLRC3*, *FCRL3*, and *AIM2*). However, the implementation of gene signatures as biomarkers in clinical practice is challenging because their assessment is costly and requires expertise in molecular biology. Standardisation issues might also substantially hamper robustness. These disadvantages probably explain why no gene signature is currently used for decision making except in breast cancer (in which assays are used for prediction of recurrence after surgery).[5]

Histological slides are readily accessible from pathology departments, and we and others have shown that the processing of histological slides by artificial intelligence (AI) can be used to predict several relevant molecular alterations.[6–11] However, the use of histological slides to identify subsets of patients who are likely to benefit from targeted therapy and immunotherapy is far less studied.[12,13] Therefore, in this proof-of-concept study, we investigated whether histology in combination with an AI-based approach could serve as a biomarker for systemic therapy in patients with advanced hepatocellular carcinoma. Training AI models with use of a clinical outcome as the direct endpoint might be the most efficient method toward this aim. However, it is challenging to obtain substantial numbers of hepatocellular carcinoma histological samples from treated patients, given the existence of non-invasive diagnostic criteria, and we thus opted to create an algorithm to predict the expression of ABRS as a surrogate biomarker.

Methods

Study design and participants

The workflow of the study is presented in figure 1. For all patient cohorts, one histological sample per patient was used. The Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA-LIHC) public dataset was used to train the model. This dataset includes adult patients (aged ≥18 years) with primary hepatocellular carcinoma treated by surgical resection in more than 20 different centres from the USA. We extracted data in March, 2020. Inclusion criteria were: (1) unequivocal diagnosis of hepatocellular carcinoma by two liver pathologists (MZ and JC), (2) available digital histological slides (haematein-eosin stained) from formalin-fixed paraffin-embedded (FFPE) materials, and (3) available gene expression profiling (RNA sequencing). We screened the whole series (n=379) according to our inclusion criteria and 43 cases were excluded due to absence of whole-slide images from FFPE material, absence of RNA sequencing data, or existence of equivocal histological features (morphological areas suggestive of a diagnosis of combined hepatocellular-cholangiocarcinoma). Thus 336 patients were included.

The first external validation series included samples from adults (aged ≥18 years) with hepatocellular

carcinoma who underwent surgical resection (n=225 resection samples; 240 samples were investigated of which 15 were excluded due to poor mRNA or sequencing quality) at Henri Mondor University Hospital (Créteil, France) between Jan 1, 2008 and Jan 1, 2019. Inclusion criteria were: (1) unequivocal histological diagnosis of hepatocellular carcinoma by two liver pathologists (MZ and JC), (2) available histological slides, (3) available baseline clinical data, and (4) available FFPE tumour blocks. Preoperative treatments and comorbidities were permitted.

The second external validation set consisted of samples obtained from adults (aged ≥18 years) with hepatocellular carcinoma who underwent liver biopsy (n=157 biopsy samples; 160 samples were investigated of which three were excluded due to poor mRNA or sequencing quality) at two French liver centres (Henri Mondor University Hospital, Créteil, and Avicenne University Hospital, Bobigny) between Jan 1, 2008 and Jan 1, 2019. Inclusion criteria were the same as those used for resected samples.

For both validation sets, slides were stained with haematein-eosin-saffron in each department and scanned with a NanoZoomer S360 or S60 device (Hamamatsu Photonics, Hamamatsu, Japan). Both of these validation sets differed from the TCGA-LIHC series in terms of sampling methods, staining protocols, and RNA sequencing technologies.

A fourth series of patients with hepatocellular carcinoma who underwent liver biopsy was used to examine the predictive value of the model in providing a biomarker for atezolizumab–bevacizumab treatment outcome. This cohort comprised 122 adults (aged ≥18 years) from 20 clinical centres in Asia, Europe, and the USA (appendix 1 p 1), who received 1200 mg atezolizumab plus 15 mg/kg bodyweight bevacizumab intravenously every 3 weeks. Patients in this set were sampled between June 1, 2020, and Jan 1, 2023. Inclusion criteria were: (1) treatment with atezolizumab–bevacizumab, (2) available tumour biopsy sample from before treatment initiation, (3) no treatment with another antitumour treatment between biopsy and treatment, and (4) available follow-up data on progression. Radiological response was assessed in each centre according to the Response Evaluation Criteria in Solid Tumours criteria (version 1.1).

All datasets included hepatocellular carcinomas of varying origins including hepatitis B virus (HBV), hepatitis C virus (HCV), and alcohol. Data regarding sex were retrieved from medical records. Data related to race and ethnicity were available for the TCGA-LIHC series, and were not available for the three validation series.

Experiments were done in accordance with the Declaration of Helsinki. All necessary written informed consents were obtained from the patients. The study and biocollection were approved by ethics committees of the CPP Ile de France III and IV (collection reference number: APHP220112 ID-RCB 2015-A01530-49), and the overall study was conducted between Oct 1, 2021, and Aug 20, 2023.

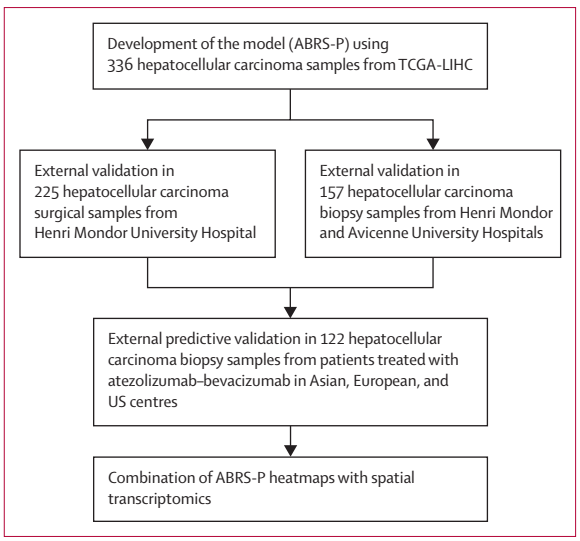

Figure 1: **Workflow of the study**
We developed the ABRS-P model with use of whole-slide scanned histology images and ABRS score as the labels (determined by RNA sequencing) from TCGA-LIHC public data. The model was validated in two external series of surgical samples and liver biopsies, respectively. The predictive value of ABRS-P was subsequently tested in a multicentre cohort of patients treated with atezolizumab–bevacizumab. We then matched ABRS-P prediction heatmaps with spatial transcriptomics to examine the in situ gene expression of predictive areas in four surgical resection samples. ABRS=atezolizumab–bevacizumab response signature. ABRS-P=ABRS-prediction. TCGA-LIHC=The Cancer Genome Atlas Liver Hepatocellular Carcinoma data collection.

RNA sequencing and spatial transcriptomics

Detailed protocols are provided in appendix 1 (p 2). For bulk RNA sequencing of the first two validation series, tumour areas of samples from the first two validation series were macrodissected and RNA was extracted. Libraries were prepared with use of the Lexogen QuantSeq-FWD 3'mRNASeq kit (version 1; Lexogen, Vienna, Austria). Sequencing was performed with an Illumina NovaSeq 6000 instrument (Illumina, San Diego, CA, USA).

For spatial transcriptomics, tumour sections were cut from FFPE blocks of four randomly selected tumours from the surgical resection validation set and placed within FFPE Visium Spatial Gene Expression Slides (10× Genomics, Pleasanton, CA, USA). Random selection was done with the ALEA function (also known as RAND) in Microsoft Excel 2016. Sequencing of libraries was performed on a NextSeq 2000 instrument (P3 flow cell, 100 cycles; Illumina, San Diego, CA, USA). The gene expression analysis pipeline is reported in appendix 1 (p 3).

Image preprocessing and deep learning networks

For all datasets, tissue regions were detected using the clustering-constrained attention multiple instance

(Prof H Rhee MD); **Department of Pathology, Severance Hospital, Yonsei University College of Medicine, Seoul, South Korea** (Prof Y N Park MD); **Graduate School of Medical Science, Brain Korea 21 Project, Yonsei University College of Medicine, Seoul, South Korea** (Prof Y N Park); **Liver Unit** (M Iñarrairaegui MD), **Department of Internal Medicine** (M Iñarrairaegui MD), **and Department of Pathology** (G Garcia-Porrero MD), **Clínica Universidad de Navarra, Pamplona, Spain; Instituto de Investigación Sanitaria de Navarra, Pamplona, Spain** (M Iñarrairaegui); **Centro de Investigación Sanitaria en Red de Enfermedades Hepáticas y Digestivas, Madrid, Spain** (M Iñarrairaegui); **Liver Unit and HPB Oncology Area, Clínica Universidad de Navarra and CIBEREHD, Pamplona, Spain** (J Argemi MD, Prof B Sangro MD); **Hepatology Program, Centro de Investigación Médica Aplicada, Universidad de Navarra, Pamplona, Spain** (J Argemi, Prof B Sangro); **Department of Surgery and Cancer, Imperial College London, Hammersmith Hospital, London, UK** (A D'Alessio MD, D J Pinato MD); **Department of Biomedical Sciences, Humanitas University, Milan, Italy** (A D'Alessio MD, L Rimassa MD, Prof L Di Tommaso MD); **Division of Gastroenterology and Hepatology, Department of Internal Medicine III, Medical University of Vienna, Vienna, Austria** (B Scheiner MD, M Pinter MD); **Division of Oncology, Department of Translational Medicine, University of Piemonte Orientale "A Avogadro", Novara, Italy** (D J Pinato); **Centre de Recherche sur l'Inflammation, INSERM 1149, Université Paris Cité, Paris, France** (Prof V Paradis MD, A Beaufrère MD); **Pathology Department, Beaujon Hospital, AP-HP Nord, Clichy, France** (Prof V Paradis, A Beaufrère); **Department of Gastroenterology, Hepatology and Endocrinology and Center for Personalized Medicine, Hannover Medical School, Hannover, Germany** (S Peter PhD, Prof A Vogel MD); **Medical Oncology and Hematology Unit** (L Rimassa)

and Pathology Unit (Prof L Di Tommaso), Humanitas Cancer Center IRCCS Humanitas Research Hospital, Milan, Italy; Laboratoire HIFIH, EA 3859, Université d'Angers, Angers, France (S Michalak MD); Angers University Hospital, Department of Pathology, Angers, France (S Michalak); Service d'Hépato-Gastroentérologie et Oncologie Digestive, Centre Hospitalier Universitaire d'Angers, Angers, France (Prof J Boursier MD); Laboratoire Hémodynamique, Interaction Fibrose et Invasivité Tumorales Hépatiques, University Paris Research, Structure Fédérative de Recherche, Interactions Cellulaires et Applications Thérapeutiques 4208, University of Angers, Angers, France (Prof J Boursier MD); Centre de Ressources Biologiques (BB-0033-00027) Hôpitaux Universitaires Paris-Seine-Saint-Denis, AP-HP, Bobigny, France (Prof M Ziol MD); Service d'Anatomie Pathologique, Hôpital Avicenne, Hôpitaux Universitaires Paris-Seine-Saint-Denis, AP-HP, Bobigny, France (Prof M Ziol MD)

Correspondence to: Prof Julien Calderaro, Department of Pathology, Henri Mondor-Albert Chenevier University Hospital, AP-HP, 94010 Créteil, France
julien.calderaro@aphp.fr

See Online for appendix 1

For the **TCGA-LIHC dataset** see https://www.cbioportal.org/study/summary?id=lihc_tcga

For the **deep learning code** see https://github.com/qinghezeng/ABRS-P

learning (CLAM) pipeline.[14] Polygonal regions of interest in tumour areas were annotated with QuPath (version 0.1.2)[15] by an expert pathologist (JC). The intersecting areas of annotated tumour areas and tissues detected by our pipeline were further exhaustively split into 256×256 pixel patches with no overlap, at a resolution of approximately 0·5 µm per pixel. Our deep learning network was based on the CLAM architecture that we modified to predict numerical continuous values for regression analysis.[14] We used an unsupervised contrastive learning transformer (CTransPath) to encode each patch into a 768-dimensional feature embedding.[16] The feature embedding of each patch was further reduced to 512 dimensions and an attention score was learned for each patch. A fully connected layer with the Softplus activation function was used to predict a non-negative continuous value for each patch on the basis of the 512-dimensional features. Appendix 1 provides further details on the preprocessing and deep learning stages (pp 4–5) and an overview of the workflow for the prediction of ABRS expression from whole-slide digitised histological images (p 10). The code used for deep learning is available online.

Network training, performance evaluation, and model interpretability

The model was trained by minimising the mean squared error loss between the predicted and true ABRS score (the arithmetic mean of the signature gene expression levels).[4] Training was performed with a ten-times Monte Carlo cross-validation strategy, with 60%, 20%, and 20%, respectively, of patients in the development (TCGA-LIHC) series randomly partitioned into training, validation, and test sets. Pearson's correlation coefficient (r), 95% CIs, and p values were used for performance evaluation. For the first two external validation sets, the predictions of the ten-times cross-validation were ensembled by average pooling to lessen the possibility of obtaining misleading attention weights.[17] For the development set, we assessed the performance of both the best model and the ensemble model. Performance of the ensemble model was also assessed after stratification according to common clinical features (eg, diagnosis age, sex, and disease stage) and pathological features (eg, tumour size, vascular invasion, and differentiation) when data were available, using the same methodology (Pearson's correlation). For the series of patients treated with atezolizumab–bevacizumab, the predictions of each cross-validation were classified as having high ABRS-P values (ABRS-P-high) or low ABRS-P values (ABRS-P-low) relative to the median split thresholds in the first biopsy series, and majority voting was used to determine the final classification (patient samples predicted as ABRS-P-high in five or more iterations were identified as ABRS-P-high). Classification was performed by QZ, who was masked to the endpoints of the study. Further details are provided in appendix 1 (p 6).

For model intepretability, patch predictions were rescaled to (0, 1) for each cross-validation, with 1 representing the highest prediction and 0 the lowest prediction, and then ensembled by average pooling for all ten iterations. The spatial coordinates of each patch were used to create a colour map (red for 1 and blue for 0) to illustrate final patch predictions.

Outcomes and statistical analysis

The primary endpoint for the cohort of patients treated with atezolizumab–bevacizumab was the association of ABRS-P with progression-free survival, defined as the interval between initiation of atezolizumab–bevacizumab treatment and the occurrence of disease progression or death. Patients who did not progress or die were censored on the date of their last evaluable tumour assessment. Overall survival, as a supplementary endpoint, was defined as the interval between initiation of systemic therapy and death or last follow-up. Patients who did not die were censored on the date of their last follow-up. Progression-free survival and overall survival of patients with ABRS-P-high samples versus those with ABRS-P-low samples were compared. Survival curves were represented with use of the Kaplan-Meier method and compared with log-rank statistics, using R (version 4.2.0) and packages survminer (version 0.4.9), ggplot2 (version 3.4.2), and survival (version 3.5-5). Univariable analysis of progression-free survival according to variables related to clinical presentation and pathology was performed with a Cox proportional-hazards regression model using R (version 4.2.0) and packages rms (version 6.4-1) and dplyr (version 1.1.2). All variables in the univariable analysis were included in a multivariable analysis to avoid bias. Pearson's correlation between ABRS-P and ABRS score was computed in Python (version 3.7.12) with packages pandas (version 1.3.5), NumPy (version 1.21.4), and SciPy (version 1.7.3). For the spatial transcriptomics, the Mann-Whitney U test was used to compare ABRS score of the 100 patches with the highest ABRS-P values and the 100 patches with the lowest ABRS-P values for each sample, using the R stats package (version 4.1.2). Differentially expressed genes across the transcriptome between the top 100 ABRS-P-high patches and top 100 ABRS-P-low patches in each sample were identified with use of the FindMarkers function in Seurat (version 4.1.1) with the test.use="MAST" parameter (MAST version 1.20.0). All tests were two-tailed and a p value of less than 0·05 was considered to indicate statistical significance.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

We first trained our model using the public TCGA-LIHC dataset, which includes patients treated with surgical

	Development series (N=336)				External validation series of resection samples (N=225)				External validation series of biopsy samples (N=157)				Series of patients treated with atezolizumab–bevacizumab (N=122)			
	Female (n=108)		Male (n=228)		Female (n=46)		Male (n=179)		Female (n=27)		Male (n=130)		Female (n=18)		Male (n=104)	
	Available data (n)	n (%)	Available data (n)	n (%)	Available data (n)	n (%)	Available data (n)	n (%)	Available data (n)	n (%)	Available data (n)	n (%)	Available data (n)	n (%)	Available data (n)	n (%)
Clinical features																
Diagnosis age >60 years	108	59 (55%)	227	109 (48%)	46	28 (61%)	179	105 (59%)	27	21 (78%)	130	104 (80%)	18	15 (83%)	104	71 (68%)
Risk factors for hepatocellular carcinoma*																
Hepatitis B virus infection	98	19 (19%)	217	82 (38%)	41	8 (20%)	171	61 (36%)	22	1 (5%)	124	18 (15%)	16	3 (19%)	89	12 (13%)
Hepatitis C virus infection	98	11 (11%)	217	37 (17%)	41	18 (44%)	171	37 (22%)	22	6 (27%)	124	25 (20%)	16	5 (31%)	89	20 (22%)
Alcohol consumption	98	13 (13%)	217	98 (45%)	41	4 (10%)	171	58 (34%)	22	4 (18%)	124	53 (43%)	16	3 (19%)	89	33 (37%)
Non-alcoholic steatohepatitis	98	7 (7%)	217	13 (6%)	41	2 (5%)	171	32 (19%)	22	5 (23%)	124	54 (44%)	16	5 (31%)	89	26 (29%)
No history of primary risk factor or undetermined cause	98	48 (49%)	217	27 (12%)	41	9 (22%)	171	17 (10%)	22	6 (27%)	124	4 (3%)	16	3 (19%)	89	9 (10%)
α-fetoprotein at diagnosis >20 ng/mL	84	49 (58%)	169	71 (42%)	41	22 (54%)	142	58 (41%)	20	6 (30%)	91	34 (37%)	18	10 (56%)	99	52 (53%)
Barcelona Clinic of Liver Cancer stage before intervention																
0 or A	NA	..	NA	..	46	32 (70%)	179	117 (65%)	26	18 (69%)	125	70 (56%)	15	0	94	2 (2%)
B or C	NA	..	NA	..	46	14 (30%)	179	62 (35%)	26	8 (31%)	125	55 (44%)	15	15 (100%)	94	92 (98%)
Treatment before biopsy (yes)	NA	..	NA	..	NA	..	NA	..	NA	..	NA	..	15	5 (33%)	87	26 (30%)
Tumour pathological features																
Largest nodule diameter >50 mm	NA	..	NA	..	46	18 (39%)	179	73 (41%)	22	8 (36%)	108	29 (27%)	13	9 (69%)	80	56 (70%)
Multinodularity (yes)	NA	..	NA	..	46	4 (9%)	179	32 (18%)	26	10 (38%)	115	51 (44%)	13	7 (54%)	82	54 (66%)
Satellite nodules (yes)	NA	..	NA	..	45	15 (33%)	179	60 (34%)	NA	..	NA	..	NA	..	NA	..
Tumour grade																
G1	106	14 (13%)	226	32 (14%)	NA	..	NA	..	NA	..	NA	..	NA	..	NA	..
G2	106	48 (45%)	226	110 (49%)	NA	..	NA	..	NA	..	NA	..	NA	..	NA	..
G3	106	40 (38%)	226	77 (34%)	NA	..	NA	..	NA	..	NA	..	NA	..	NA	..
G4	106	4 (4%)	226	7 (3%)	NA	..	NA	..	NA	..	NA	..	NA	..	NA	..
Vascular invasion†																
Macrovascular	95	6 (6%)	190	8 (4%)	46	4 (9%)	179	36 (20%)	NA	..	NA	..	5	1 (20%)	35	15 (43%)
Microvascular	95	24 (25%)	190	59 (31%)	46	20 (43%)	179	81 (45%)	NA	..	NA	..	NA	..	NA	..
No	95	65 (68%)	190	123 (65%)	46	26 (57%)	179	96 (54%)	NA	..	NA	..	NA	..	NA	..
Residual tumour																
R0	103	100 (97%)	208	197 (95%)	46	39 (85%)	179	154 (86%)	NA	..	NA	..	NA	..	NA	..
R1 or R2	103	3 (3%)	208	11 (5%)	46	7 (15%)	179	25 (14%)	NA	..	NA	..	NA	..	NA	..
WHO differentiation[19]																
Poor	NA	..	NA	..	46	9 (20%)	179	24 (13%)	26	3 (12%)	130	8 (6%)	18	6 (33%)	104	26 (25%)
Moderate	NA	..	NA	..	46	25 (54%)	179	126 (70%)	26	17 (65%)	130	99 (76%)	18	11 (61%)	104	60 (58%)
Well	NA	..	NA	..	46	12 (26%)	179	29 (16%)	26	6 (23%)	130	23 (18%)	18	1 (6%)	104	18 (17%)

(Table continues on next page)

	Development series (n=336)				External validation series of resection samples (n=225)				External validation series of biopsy samples (n=157)				Series of patients treated with atezolizumab–bevacizumab (n=122)			
	Female (n=108)		Male (n=228)		Female (n=46)		Male (n=179)		Female (n=27)		Male (n=130)		Female (n=18)		Male (n=104)	
	Available data (n)	n (%)	Available data (n)	n (%)	Available data (n)	n (%)	Available data (n)	n (%)	Available data (n)	n (%)	Available data (n)	n (%)	Available data (n)	n (%)	Available data (n)	n (%)
(Continued from previous page)																
Ishak fibrosis score[20]																
0–2	64	40 (62%)	134	58 (43%)	NA	..	NA	..	NA	..	NA	..	NA	..	NA	..
3 or 4	64	7 (11%)	134	19 (14%)	NA	..	NA	..	NA	..	NA	..	NA	..	NA	..
5 or 6	64	17 (27%)	134	57 (43%)	NA	..	NA	..	NA	..	NA	..	NA	..	NA	..
Cirrhosis	NA	..	NA	..	43	19 (44%)	174	63 (36%)	19	12 (63%)	90	58 (64%)	17	7 (41%)	101	50 (50%)
Extrahepatic metastasis	NA	..	NA	..	NA	..	NA	..	NA	..	NA	..	15	5 (33%)	97	41 (42%)

NA=not available. *Patients could have multiple risk factors. †Some patients had both macrovascular and microvascular invasion.

Table: **Clinical and pathological features of each series of patients investigated**

resection (figure 1).[18] A total of 336 patients were included; their main clinical, biological, and pathological characteristics are summarised in the table. In this development series, 228 (68%) patients were male and 108 (32%) were female, and median age at diagnosis was 61 years (IQR 51–69). The most frequent risk factor for hepatocellular carcinoma was alcohol consumption (111 [35%] of 315 with available data). Race distribution of the patients was as follows: White, 159 (49%) of 327 with available data; Asian, 152 (46%); Black or African American, 14 (4%); and American Indian or Alaska Native, two (1%). Fibrotic livers (Ishak score 5–6)[20] was observed in 37% patients (74 of 198 with available data).

Our deep learning architecture was adapted from a previously published pipeline that we modified to perform regression analysis, which we denoted the ABRS-P model, for ABRS prediction (appendix 1 p 10). Features were extracted from patches located in tumour areas using CTransPath. The features were fed, along with ABRS score as the labels, into our architecture. Using cross-validation, the best model reached a Pearson's coefficient (*r*) of 0·71 (95% CI 0·58–0·81; p<0·0001; figure 2A). The mean Pearson's coefficient for all cross-validations was 0·62 (SD 0·09; mean p<0·0001, SD<0·0001). The ensemble ABRS-P model (average prediction of all ten-times cross-validation models) had a Pearson's coefficient of 0·86 (0·83–0·89; p<0·0001) for the whole development set (figure 2B). The predictive value remained significant after stratification according to common clinical and pathological features (appendix 1 pp 7, 11). Interestingly, pathological review of image patches with high ABRS-P values (ie, predicted to have high ABRS values) showed an enrichment of immune cells (figure 2C, 2D).

We aimed to investigate how the ABRS-P model performed in two independent series of patients with hepatocellular carcinoma. Importantly, these external datasets differed from the development set in terms of staining protocols, histological sampling, gene expression profiling technologies, and slide encoding formats.

The first validation dataset included 225 patients treated by surgical resection in Henri Mondor University Hospital; their main clinical, biological, and pathological features are presented in the table. The cohort comprised 179 (80%) men and 46 (20%) women, and median age was 64 years (IQR 56–72). Barcelona Clinic of Liver Cancer (BCLC) stage[21] was 0 or A for 149 (66%) of 225 patients and B or C for 76 (34%) patients. The most frequent risk factors for hepatocellular carcinoma were HBV infection (69 [33%] of 212 with available data) and alcohol intake (62 [29%]). Among the 225 patients, morphological features of aggressiveness (poor differentiation, macrovascular invasion, microvascular invasion, and satellite nodules) were identified in 33 (15%), 40 (18%), 101 (45%), and 75 (33%) patients, respectively. As observed in the TCGA-LIHC series, the

majority of tumours developed in non-cirrhotic livers (table). Slides were processed using our pipeline, and ABRS-P values remained significantly correlated with ABRS score, with a Pearson coefficient of 0·60 (95% CI 0·51–0·68; p<0·0001; figure 3A, 3C).

As biopsy is the only type of sample that can be obtained from patients with advanced hepatocellular carcinoma, we next sought to validate the ABRS-P model in another external dataset that only included biopsies (n=157). The clinical and pathological features of this cohort are reported in the table. The cohort comprised 130 (83%) male patients and 27 (17%) female patients, and median age was 69 years (IQR 61–75). Briefly, BCLC stage was 0 or A for 88 (58%) of 151 patients with available data, and B or C for 63 (42%) patients. The most frequent cause of hepatocellular carcinoma was non-alcoholic steatohepatitis (59 [40%] of 146 with available data). Features extracted from tumour areas were analysed by our deep learning architecture and an ABRS score was computed for each case. We observed a small decrease in the performance of our ABRS-P model (Pearson's coefficient of 0·53, 95% CI 0·40–0·63) compared with the performance in the first external validation set, but the correlation was still significant (p<0·0001; figure 3B, 3D). In both validation series, the ABRS-P model generally retained its value after stratification according to common clinical and pathological features (appendix 1 pp 7, 12–13). Notable decreases in performance were observed for patients aged 60 years or younger and for patients with cirrhosis.

We subsequently evaluated if ABRS-P values were associated with progression-free survival in patients with hepatocellular carcinoma after atezolizumab–bevacizumab initiation. Although hepatocellular carcinoma biopsy samples are rare due to the existence of non-invasive diagnostic criteria, we were able to gather slides from 122 patients treated in 20 different centres (table). The cohort comprised 104 (85%) male patients and 18 (15%) female patients, and median age was 67 years (IQR 59–74). To classify samples as ABRS-P-high or ABRS-P-low, we set a fixed threshold using the median of the ABRS-P output values for each cross-validation iteration from the previous biopsy validation series, and aggregated the predictions by majority voting for all iterations.

Median follow-up for progression-free survival was 6 months (IQR 3–10). Among the 122 patients included, 74 (61%) had a tumour classified as ABRS-P-high and 48 (39%) had a tumour classified as ABRS-P-low after majority voting. Patients with ABRS-P-high tumours showed significantly longer median progression-free survival after treatment initiation than patients with ABRS-P-low tumours (12 months [95% CI 7–not reached] *vs* 7 months [4–9]; p=0·014; figure 4A). The predictive value remained significant in multivariable analysis (appendix 1 p 8). The model predictions were not associated with overall survival (figure 4B; overall survival median follow-up 9 months [IQR 6–13]).

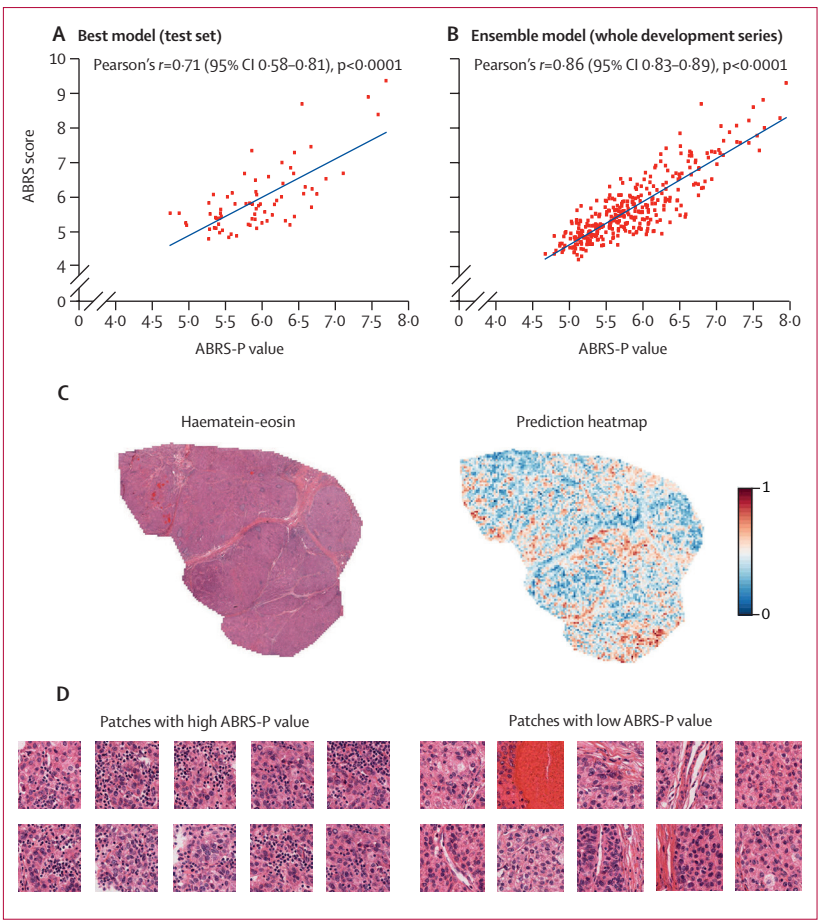

Figure 2: **Prediction of ABRS expression with deep learning**
(A) Correlation between ABRS scores and predicted values from the best ABRS-P model in the test set of patients from the development series. (B) Correlation, using the final ABRS-P model (average prediction from the ten-times cross-validation), between ABRS scores and predictions in the whole development series. (C) Haematein-eosin stained section and prediction heatmap of a case from the development series test set: in the heatmap, areas predicted as having high ABRS-P values are highlighted red and those with low ABRS-P values are highlighted blue (model predictions rescaled between 0 and 1). (D) Patches from the tumour in part C with high and low ABRS-P values: microscopic review showed an enrichment of immune cells in areas with high ABRS-P values. ABRS=atezolizumab–bevacizumab response signature. ABRS-P=ABRS-prediction.

To obtain insights on the gene expression profile associated with ABRS-P values, we performed spatial transcriptomics on four hepatocellular carcinoma samples from the surgical validation series, and developed a pipeline that matches prediction heatmaps with in situ expression profiles. We then compared, for each tumour, the top 100 patches with high ABRS-P values against the top 100 patches with low ABRS-P values. In three of the four cases investigated, we identified significantly higher ABRS scores in tumour areas with high ABRS-P values versus areas with low ABRS-P values (mean expression, sample D2: 0·273 [SD 0·121] *vs* 0·192 [0·088], p<0·0001; sample E4: 0·131 [0·112] *vs* 0·054 [0·085], p<0·0001; and sample E6: 0·228 [0·082] *vs* 0·182 [0·090], p<0·0001). The

p value for the fourth sample (E7) was 0·058 (mean expression, 0·027 [0·047] *vs* 0·016 [0·038]). Results for samples D2 and E4 are shown in figure 5, and results for samples E6 and E7 are provided in appendix 1 (p 14). We also analysed the correlation between ABRS score and ABRS-P predictions in these patches. Pearson coefficients were low (0·14–0·37), but p values were significant in three of four cases (samples D2, E4 and E6, all p<0·0001; sample E7, p=0·052; appendix 1 p 9).

We next analysed, for each of the four cases, the differences in gene expression across the whole transcriptome. We observed significant differences between ABRS-P-high and ABRS-P-low patches, with the full lists of differentially expressed genes in ABRS-P-high patches provided in appendix 2. For example, in sample D2, we detected upregulation of T-cell markers (*CD3E, CD3D*), hepatocytic markers (*ALB, FABP1, APOB*), genes involved in inflammation (*CXCL5, IL15-RA, IL32*), interferon signalling (*IFITM2, IFITM3, OAS2*), and antigen presentation (*HLA-A, HLA-DQA1*) in ABRS-P-high patches versus ABRS-P-low patches. In sample E4, there was upregulation of T-cell markers (*CD3E, CD3D, CD3G*), genes involved in antigen presentation (*HLA-A, HLA-E, HLA-DRA, B2M*), inflammation (*LTB, CXCR4*), chemoattraction (*CCL5*), and extracellular matrix production (*COL1A1, COL1A2, COL3A1, COL4A1, VIM*) in ABRS-P-high patches. In sample E6, we detected upregulation of genes related to hepatocytic differentiation (*ALB, APOE, APOC2, APOA2*) in ABRS-P-high patches. In sample E7, there was upregulation of genes involved in chemoattraction (*CXCL8, CXCL12*), extracellular matrix production (*TGFBI, COL4A1, COL5A1, COL6A1, VIM*), inflammation (*CRP, IL32*), and antigen presentation (*B2M, HLA-DRA,*

See Online for appendix 2

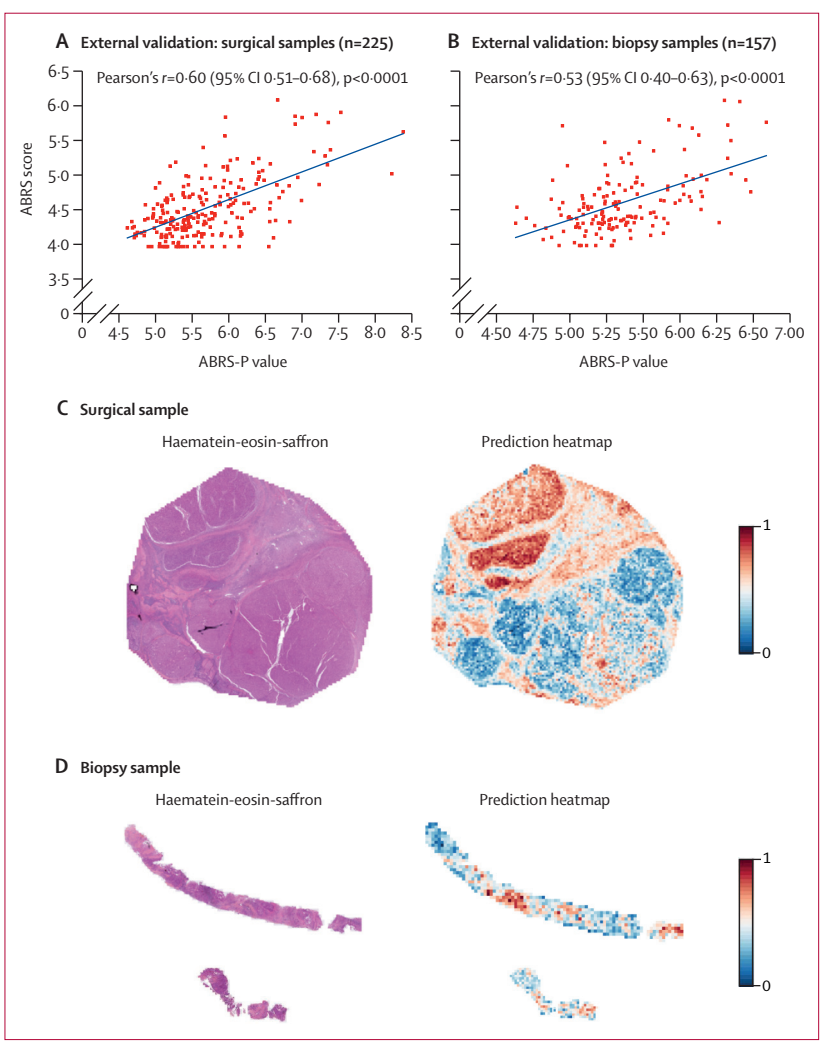

Figure 3: **External validation of the ABRS-P model**
Correlation between ABRS scores and predicted values from the ABRS-P model in patients with surgical resection (A) or liver biopsy (B). Example of a haematein-eosin-saffron-stained section and the corresponding prediction heatmap for a surgical sample (C) and a biopsy sample (D). In the heatmaps, areas predicted as having high ABRS-P values are highlighted in red and those with low ABRS-P values are highlighted blue (model predictions rescaled between 0 and 1). ABRS=atezolizumab–bevacizumab response signature. ABRS-P=ABRS-prediction.

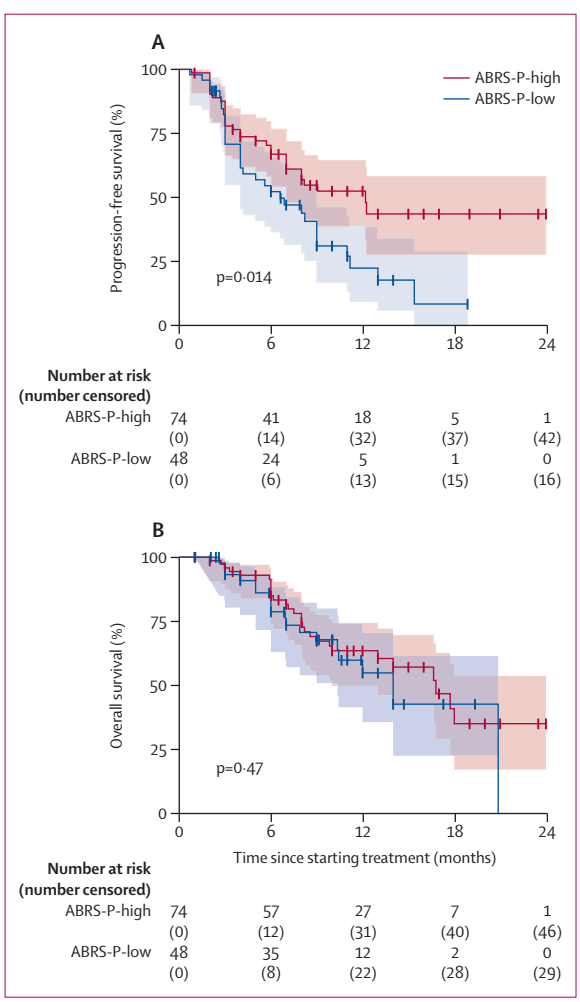

Figure 4: **Progression-free survival (A) and overall survival (B) of patients with ABRS-P-high versus ABRS-P-low tumours after atezolizumab–bevacizumab initiation**
Shaded regions are 95% CIs. ABRS=atezolizumab–bevacizumab response signature. ABRS-P=ABRS-prediction.

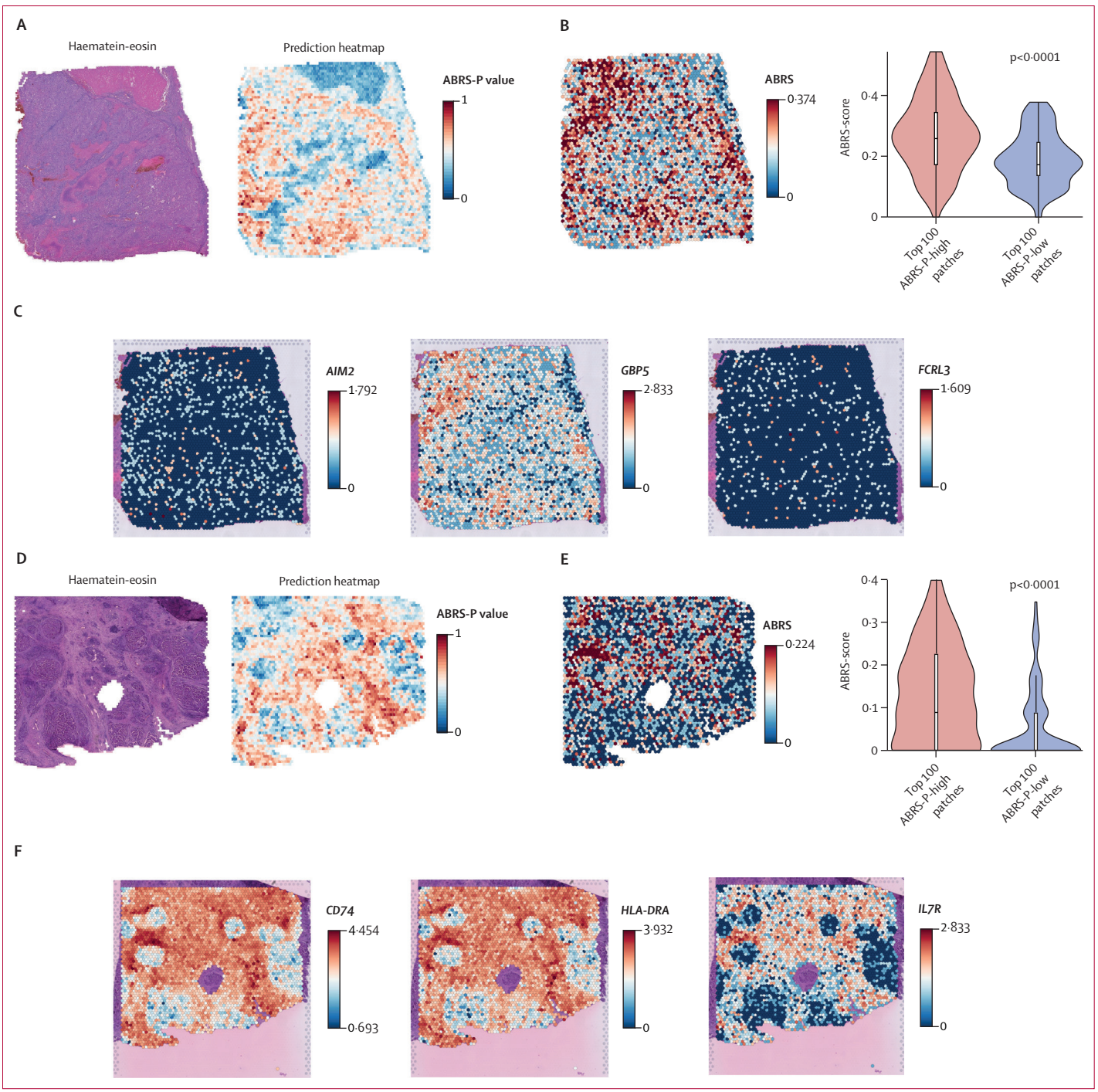

Figure 5: **Combination of ABRS-P heatmaps with spatial transcriptomics**
(A) Haematein-eosin-stained section and prediction heatmap of sample D2 (ABRS-P values were rescaled between 0 and 1). (B) Heatmap of ABRS scores assessed by spatial transcriptomics: the areas with high ABRS scores visually match the areas with high ABRS-P values. Higher ABRS scores were observed in the top 100 ABRS-P-high patches compared with the top 100 ABRS-P-low patches. (C) Expression heatmaps for three genes included in ABRS highlighting the different expression levels of ABRS genes. (D) Haematein-eosin-stained section and corresponding prediction heatmap for sample E4. Areas predicted to have high ABRS-P values mainly consisted of stroma and fibrous tissue. (E) Heatmap of ABRS values assessed by spatial transcriptomics: the match between areas with high ABRS scores and areas with high ABRS-P values was not visually obvious, however statistical analysis showed significantly higher ABRS score in patches predicted as being ABRS-P-high. (F) Upregulation of immune-related genes not included in ABRS, with some overlap with areas predicted as being ABRS-P-high. For all heatmaps, areas predicted as having high ABRS-P values are highlighted in red and those with low ABRS-P values are highlighted blue (predictions rescaled between 0 and 1). For all violin-box plots, the centre line is the median and box limits are the first and third quartiles. ABRS=atezolizumab–bevacizumab response signature. ABRS-P=ABRS-prediction.

HLA-DPA1, HLA-DPB1) in ABRS-P-high patches. Altogether, these findings suggested that the model, in three of the four cases investigated, identified areas characterised by a particular immune profile.

Discussion

Our study indicates that AI applied on digitised hepatocellular carcinoma histological slides can, by estimating the activation of ABRS, serve as a biomarker for progression-free survival in patients treated with atezolizumab–bevacizumab. Although hepatocellular carcinoma is a highly heterogeneous disease in terms of aetiology and molecular alterations, the former work by Zhu and colleagues did not identify any impact of risk factors or genetic defects on clinical outcomes of patients treated with atezolizumab–bevacizumab.[4] We therefore chose to focus on gene expression (ABRS). Several gene signatures have been tested to predict the outcomes of patients treated with systemic or targeted therapies, but most have not been implemented in clinical practice, probably because they are complex to set up and expensive.

To our knowledge, our study is one of the first to predict a continuous biomarker of gene signature expression from pathological data using weakly supervised learning, and the first to validate the predictions in situ using spatial transcriptomics. We also showed that the AI model, when applied on samples from treated patients, had predictive value.

Deep learning networks are particularly prone to overfit the data they are trained on and we believe that one of the strengths of our work is the broad external validation of our model. The robustness of our model was tested in patients from different centres, in validation series with different staining protocols, encoding formats, and expression profiling technologies compared with the development set, and on different types of histological specimens (resection or biopsy). Importantly, we used a fixed threshold for the series of treated patients, and this type of approach is crucial to demonstrate that such biomarkers could be implemented in clinical practice, where patients need to be individually classified. The results are encouraging and suggest that the potential technical biases are limited. Interestingly, the model generalised in patients with different risk factors such as alcohol consumption and HBV infection, which are known to induce molecularly distinct hepatocellular carcinoma subtypes.[22]

AI models are often criticised for their lack of transparency. Features extracted from images are indeed processed through deep complex layers of mathematical computing and it is currently impossible to precisely identify the characteristics used for prediction. Deep neural networks are therefore widely regarded as black boxes (ie, a system without any knowledge of its internal workings). We aimed to address this issue via the development of an original pipeline that combines AI heatmaps and spatial transcriptomics. In three of four cases, we were able to confirm that tumour areas associated with high ABRS-P values were overexpressing ABRS. Additionally, ABRS-P-high areas were over-expressing other immune effectors in three cases. Although this analysis does not make the model completely explainable, we believe that this type of approach is interesting and aids the identification of molecular features that affect prediction. This workflow is not meant to be performed in daily practice, but its use during the development of future AI-based models might help to build the confidence and trust needed to adopt AI-based pathological biomarkers.

Our study has some limitations. It is retrospective, with some missing or incomplete data, and prospective validation will be essential before the use of this type of predictive model can be considered in clinical practice. Investigation of model performance in populations not included in our study (eg, patients from Africa and a larger sample from Asia) will also be important. To completely exclude a prognostic value non-specific to atezolizumab–bevacizumab treatment, the model should also be tested in patients treated with other targeted therapies. As previously observed by Zhu and colleagues,[4] we found that ABRS-P values were not associated with overall survival in patients treated with atezolizumab–bevacizumab. The use of progression-free survival as a surrogate marker of overall survival in patients with advanced hepatocellular carcinoma has been challenged.[23–26] Several trials investigating different systemic therapies, such as REFLECT or COSMIC-312,[27,28] showed significantly longer median progression-free survival but did not find an improvement in median overall survival.

These discrepancies in treatment effects on survival outcomes might be explained by post-progression treatments affecting overall survival; or, in patients treated with immune checkpoint inhibitors, a transient increase in tumour size due to a massive intratumoural immune infiltration (pseudoprogression) being recorded as progression. Future studies should also establish if longer follow-up might reveal differences in overall survival. Finally, RNA molecules from FFPE are fragmented, and one limitation of in situ expression profiling is that genes with very low baseline expression might not be detected. Spatial transcriptomics also does not provide full explainability and further research in the field of explainable AI (known as XAI) will be important to better understand ABRS-P behaviour and detect its potential flaws.

In conclusion, pending prospective validation, we believe that the type of AI-based biomarker tested herein could enable rapid identification of patients who are likely to benefit from specific treatments. The combination of prediction heatmaps with in situ gene expression profiling can also provide insights on the molecular features of tumour areas with high predictive value. This

approach could be applied to other cancers or diseases and improve understanding of the biological mechanisms that drive outcomes or responses to treatments.

Contributors

JC conceptualised the study and acquired funding. QZ and SC curated data. QZ, CK, SC, PM, and MZ formally analysed data. All authors screened and collected patient data. QZ, CK, SC, and PM developed the methodology. QZ and JC were project administrators. All authors provided study resources. QZ, SC, NL, and CK developed the software. JC, SC, NL, CK, and MZ supervised the study. QZ, SC, and CK visualised the data. QZ, CK, NL, MZ, and JC wrote the original draft. All authors reviewed and edited the manuscript. QZ, SC, and JC accessed and verified the data. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Declaration of interests

LR reports grants from Agios, AstraZeneca, BeiGene, Eisai, Exelixis, Fibrogen, Incyte, Ipsen, Lilly, MSD, Nerviano Medical Sciences, Roche, and Zymeworks; consulting fees from AstraZeneca, Basilea, Bayer, Bristol Myers Squibb, Eisai, Exelixis, Genenta, Hengrui, Incyte, Ipsen, IQVIA, Jazz Pharmaceuticals, MSD, Nerviano Medical Sciences, Roche, Servier, Taiho Oncology, and Zymeworks; honoraria from AstraZeneca, Bayer, Bristol Myers Squibb, Eisai, Incyte, Ipsen, Merck Serono, Roche, and Servier; support for attending meetings from AstraZeneca; is Treasurer for the International Liver Cancer Association and member of the Executive Council of the International Liver Cancer Association; and is Co-chair of the European Organisation for Research and Treatment of Cancer task force on hepatobiliary and neuroendocrine tumours. SL reports grants from Institut National du Cancer, Fondation ARC, Ligue Contre le Cancer du Val de Marne, Fondation de l'Avenir, Ipsen, and Fondation Bristol Myers Squibb. MP reports consulting fees from Bayer, Bristol Myers Squibb, Eisai, Roche, AstraZeneca, Ipsen, Lilly, and MSD; honoraria from Bayer, Bristol Myers Squibb, Eisai, Roche, and MSD; and support for attending meetings from Bayer, Bristol Myers Squibb, Ipsen, and Roche. VP reports honoraria from Servier and participation in an advisory board for Bristol Myers Squibb. FP reports grants from Institut National du Cancer, Fondation ARC, Ligue Contre le Cancer du Val de Marne, Fondation de l'Avenir, Ipsen, and Fondation Bristol Myers Squibb. JB reports grants from Diafir, Echosens, Intercept, Inventiva, and Siemens; consulting fees from AstraZeneca, Echosens, Intercept, and Siemens; speaker fees from AbbVie, Gilead, Intercept, and Siemens; and participation in advisory boards for Bristol Myers Squibb, Intercept, Pfizer, MSD, and NovoNordisk. MM reports grants from Institut National du Cancer, Fondation ARC, Ligue Contre le Cancer du Val de Marne, Fondation de l'Avenir, Ipsen, and Fondation Bristol Myers Squibb. AV reports consulting fees from Astra Zeneca, Amgen, Beigene, Böhringer Mannheim, Bristol Myers Squibb, BTG, Daichi-Sankyo, Eisai, Incyte, Ipsen, MSD, Pierre Fabre, Roche, Servier, Sirtex, Tah, and Terumo; honoraria from AstraZeneca, Amgen, BeiGene, Böhringer Mannheim, BMS, BTG, Daichi-Sankyo, Eisai, GSK, Imaging Equipment (acquired by AAA), Incyte, Ipsen, Jiangsu Hengrui Medicines, MSD, Pierre Fabre, Roche, Servier, Sirtex, Tahio, and Terumo; and participation in advisory boards for AstraZeneca, Amgen, BeiGene, Böhringer Mannheim, BMS, BTG, Daichi-Sankyo, Eisai, Incyte, Ipsen, MSD, PierreFabre, Roche, Servier, Sirtex, Tahio, and Terumo. DJP reports grants from Associazone Per la Ricerca sul Cancro; consulting fees from AstraZeneca, MiNa Therapeutics, Avamune, Roche, Mursla, LiFt Biosciences, and Boeringher Ingelheim; honoraria from Roche, AstraZeneca, and Ipsen; and participation in advisory boards for AstraZeneca, Exact Sciences, and Roche. BM reports competitive grants from Instituto de Salud Carlos III (grant numbers PI18/00961 and PI21/00714), cofounded by the EU; grants from Laboratorios Viñas; honoraria from Bayer, Roche, Eisai, and AstraZeneca; support for attending meetings from Roche and Eisai; and participation in advisory boards for Bayer, Eisai, and Roche. JC reports grants from Institut National du Cancer, Fondation ARC, Ligue Contre le Cancer du Val de Marne, Fondation de l'Avenir, Ipsen, and Fondation Bristol Myers Squibb; honoraria from Servier, AstraZeneca, and Ipsen; participation in an advisory board for Bristol Myers Squibb; and owning stock in Clarapath. BrS reports grants from Instituto de Salud Carlos III and Bristol Myers Squibb; consulting fees from Adaptimmune, AstraZeneca, Bayer, Bristol Myers Squibb, Boston Scientific, Eisai, Incyte, Ipsen, MSD, Roche, Sanofi, Sirtex Medical, and Terumo; honoraria from AstraZeneca, Bayer, Bristol Myers Squibb, Eisai, Incyte, Ipsen, Roche, Sirtex Medical, and Terumo; advisory board participation for the GOING and ACTION trials; and support for attending meetings from AstraZeneca. BeS reports grants from Eisai and travel grants from AbbVie, Ipsen, Gilead, and AstraZeneca. MN reports honoraria from Astellas, Ipsen, and Roche. J-CN reports grants from Ipsen and Bayer. DSA reports grants from Ventana Medical Systems, Novo Nordisk, Hamni, Pfizer, Viking Therapeutics, and the US National Institutes of Health (grant numbers NIAA P50 and NIDDK U01); and is Secretary/Treasurer for the Hans Popper Hepatopathology Society. AD'A reports consulting fees, honoraria, and support to attend meetings from Roche. JA reports honoraria from Pfizer and Roche; and participation in a steering committee for Roche. MI reports grant funding from the Italian Ministry of Health; honoraria from Roche, AstraZeneca, Ipsen, Bayer, Gilead, Eisai, and MSD; support for attending meetings from AstraZeneca and Roche; and participation in advisory boards for AstraZeneca and Roche. All other authors declare no competing interests.

Data sharing

The codes used for gene expression analysis, deep learning, and spatial transcriptomics processing are available according to GitHub criteria at https://github.com/qinghezeng/ABRS-P, and the full commands used are provided at https://github.com/qinghezeng/ABRS-P/blob/main/docs/tutorial.md. Raw spatial sequencing data will be available 3 months after publication (by Feb 8, 2024) at the European Nucleotide Archive (accession number PRJEB62485, accession according to European Nucleotide Archive criteria). Data from proprietary cohorts cannot be made available due to privacy and ethical or legal issues.

Acknowledgments

We acknowledge the following institutions for supporting the study: Institut National du Cancer, Fondation ARC (TRANSCAN 2021 Joint Transnational Call for research proposals, project TANGERINE), Ligue Contre le Cancer du Val de Marne, Fondation de l'Avenir (project number AP-RM-20-011), Ipsen, and Fondation Bristol Myers Squibb Pour la Recherche en Immuno-Oncologie (project "Microenvironnement immunitaire du carcinome hépatocellulaire et intelligence artificielle"), and the China Scholarship Council (grant number 201908070052). The Institut Curie Genomics of Excellence Next Generation Sequencing Platform is supported by grants ANR-10-EQPX-03 (Equipex) and ANR-10-INBS-09-08 (France Génomique Consortium) from the Agence Nationale de la Recherche (Investissements d'Avenir programme), by the Aviesan Thematic Multi-Organization Institute for Cancer (Plan Cancer III) and by the Site de Recherche Intégrée sur le Cancer-Curie programme (SIRIC grants INCa-DGOS-465 and INCa-DGOS-Inserm_12554). DJP is supported by the Wellcome Trust Strategic Fund (grant number PS3416) and the Fondazione AIRC per la Ricerca sul Cancro (grant number AIRC MFAG 25697); and acknowledges grant support from the Cancer Treatment and Research Trust and the Foundation for Liver Research, and infrastructural support from the Imperial Experimental Cancer Medicine Centre and the National Institute for Health and Care Research Imperial Biomedical Research Centre. LDT is supported by the Associazione Italiana per la Ricerca sul Cancro (Individual Grant 2020; project reference 25087).

References

1 Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021; **71:** 209–49.

2 Sangro B, Sarobe P, Hervás-Stubbs S, Melero I. Advances in immunotherapy for hepatocellular carcinoma. *Nat Rev Gastroenterol Hepatol* 2021; **18:** 525–43.

3 Finn RS, Qin S, Ikeda M, et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. *N Engl J Med* 2020; **382:** 1894–905.

4 Zhu AX, Abbas AR, de Galarreta MR, et al. Molecular correlates of clinical response and resistance to atezolizumab in combination with bevacizumab in advanced hepatocellular carcinoma. *Nat Med* 2022; **28:** 1599–611.

5 Cardoso F, van't Veer LJ, Bogaerts J, et al. 70-gene signature as an aid to treatment decisions in early-stage breast cancer. *N Engl J Med* 2016; **375:** 717–29.

6 Zeng Q, Klein C, Caruso S, et al. Artificial intelligence predicts immune and inflammatory gene signatures directly from hepatocellular carcinoma histology. *J Hepatol* 2022; **77:** 116–27.

7 Calderaro J, Kather JN. Artificial intelligence-based pathology for gastrointestinal and hepatobiliary cancers. *Gut* 2021; **70:** 1183–93.

8 Kather JN, Calderaro J. Development of AI-based pathology biomarkers in gastrointestinal and liver cancer. *Nat Rev Gastroenterol Hepatol* 2020; **17:** 591–92.

9 Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol* 2019; **16:** 703–15.

10 Kather JN, Pearson AT, Halama N, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med* 2019; **25:** 1054–56.

11 Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* 2018; **24:** 1559–67.

12 Li B, Li F, Liu Z, et al. Deep learning with biopsy whole slide images for pretreatment prediction of pathological complete response to neoadjuvant chemotherapy in breast cancer: a multicenter study. *Breast* 2022; **66:** 183–90.

13 Li B, Yang L, Zhang H, et al. Outcome-supervised deep learning on pathologic whole slide images for survival prediction of immunotherapy in patients with non-small cell lung cancer. *Mod Pathol* 2023; **36:** 100208.

14 Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat Biomed Eng* 2021; **5:** 555–70.

15 Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: open source software for digital pathology image analysis. *Sci Rep* 2017; **7:** 16878.

16 Wang X, Yang S, Zhang J, et al. Transformer-based unsupervised contrastive learning for histopathological image classification. *Med Image Anal* 2022; **81:** 102559.

17 Haab J, Deutschmann N, Martínez MR. Is attention interpretation? A quantitative assessment on sets. *arXiv* 2022; published online July 26. https://doi.org/10.48550/arXiv.2207.13018 (preprint).

18 Cancer Genome Atlas Research Network. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. *Cell* 2017; **169:** 1327–41.e23.

19 WHO Classification of Tumours Editorial Board. WHO Classification of Tumours, 5th edn, vol 1. Digestive system tumours. https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/Digestive-System-Tumours-2019 (accessed Nov 2, 2023).

20 Ishak K, Baptista A, Bianchi L, Callea F, De Groote J, Gudat F. Histological grading and staging of chronic hepatitis. *J Hepatol* 1995; **22:** 696–99.

21 Reig M, Forner A, Rimola J, et al. BCLC strategy for prognosis prediction and treatment recommendation: the 2022 update. *J Hepatol* 2022; **76:** 681–93.

22 Calderaro J, Ziol M, Paradis V, Zucman-Rossi J. Molecular and histological correlations in liver cancer. *J Hepatol* 2019; **71:** 616–30.

23 Lim M, Muquith M, Miramontes B, et al. Surrogate and modified endpoints for immunotherapy in advanced hepatocellular carcinoma. *Hepatology* 2023; published online June 1. https://doi.org/10.1097/HEP.0000000000000494.

24 Cabibbo G, Bruix J. Radiological endpoints as surrogates for survival benefit in hepatocellular carcinoma trials: all that glitters is not gold. *J Hepatol* 2023; **78:** 8–11.

25 Kelley RK, Rimassa L, Cheng A-L, et al. Cabozantinib plus atezolizumab versus sorafenib for advanced hepatocellular carcinoma (COSMIC-312): a multicentre, open-label, randomised, phase 3 trial. *Lancet Oncol* 2022; **23:** 995–1008.

26 Kudo M, Finn RS, Qin S, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *Lancet* 2018; **391:** 1163–73.

27 Kudo M, Finn RS, Qin S, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *Lancet* 2018; **391:** 1163–73.

28 Kelley RK, Rimassa L, Cheng AL, et al. Cabozantinib plus atezolizumab versus sorafenib for advanced hepatocellular carcinoma (COSMIC-312): a multicentre, open-label, randomised, phase 3 trial. *Lancet Oncol* 2022; 23: 995–1008.