nature medicine

Article

https://doi.org/10.1038/s41591-023-02226-6

A longitudinal circulating tumor DNA-based model associated with survival in metastatic non-small-cell lung cancer

Received: 29 June 2022

Accepted: 23 January 2023

Published online: 16 March 2023

Check for updates

Zoe June F. Assaf[1] ✉, Wei Zou[1], Alexander D. Fine[2], Mark A. Socinski[3], Amanda Young[2], Doron Lipson[2], Jonathan F. Freidin[2], Mark Kennedy[2], Eliana Polisecki[2], Makoto Nishio[4], David Fabrizio[2], Geoffrey R. Oxnard[2], Craig Cummings[1], Anja Rode[5], Martin Reck[6], Namrata S. Patil[1], Mark Lee[1], David S. Shames[1] ✉ & Katja Schulze[1] ✉

One of the great challenges in therapeutic oncology is determining who might achieve survival benefits from a particular therapy. Studies on longitudinal circulating tumor DNA (ctDNA) dynamics for the prediction of survival have generally been small or nonrandomized. We assessed ctDNA across 5 time points in 466 non-small-cell lung cancer (NSCLC) patients from the randomized phase 3 IMpower150 study comparing chemotherapy-immune checkpoint inhibitor (chemo-ICI) combinations and used machine learning to jointly model multiple ctDNA metrics to predict overall survival (OS). ctDNA assessments through cycle 3 day 1 of treatment enabled risk stratification of patients with stable disease (hazard ratio (HR) = 3.2 (2.0–5.3), $P < 0.001$; median 7.1 versus 22.3 months for high- versus low-intermediate risk) and with partial response (HR = 3.3 (1.7–6.4), $P < 0.001$; median 8.8 versus 28.6 months). The model also identified high-risk patients in an external validation cohort from the randomized phase 3 OAK study of ICI versus chemo in NSCLC (OS HR = 3.73 (1.83–7.60), $P = 0.00012$). Simulations of clinical trial scenarios employing our ctDNA model suggested that early ctDNA testing outperforms early radiographic imaging for predicting trial outcomes. Overall, measuring ctDNA dynamics during treatment can improve patient risk stratification and may allow early differentiation between competing therapies during clinical trials.

One of the great challenges in therapeutic oncology is determining who might achieve survival benefits from a particular therapy. Cytotoxic agents, such as platinum-based alkylating agents or small-molecule inhibitors of receptor tyrosine kinases, can lead to observable reductions in overall tumor burden as measured by computerized tomography (CT) or magnetic resonance imaging. Imaging-based evaluation of the therapeutic effects of oncology drugs during the course of treatment informs on the response of a patient's tumor to the drug or drug combination, prognosis of the patient and aids in physician decision making. In addition, imaging-based modalities have been developed as surrogates of

[1]Genentech Inc., South San Francisco, CA, USA. [2]Foundation Medicine Inc., Cambridge, MA, USA. [3]AdventHealth Cancer Institute, Orlando, FL, USA. [4]The Cancer Institute Hospital, Japanese Foundation for Cancer Research, Tokyo, Japan. [5]F. Hoffman-La Roche AG, Basel, Switzerland. [6]LungenClinic Grosshansdorf, Airway Research Center North, German Center for Lung Research, Grosshansdorf, Germany. ✉e-mail: assaf.zoe-june@gene.com; shames.david@gene.com; schulze.katja.ks2@gene.com

overall survival (OS) and are widely used endpoints in oncology drug trials[1,2]. However, for certain types of drugs including cancer immunotherapies, progression-free survival (PFS) or overall response rate do not always correlate with OS[3,4]. Because of this lack of correlation between surrogate measures of drug efficacy and OS, oncology drug trials often depend on OS as a primary endpoint[5,6]. This means that trials can take many years to complete. Therefore, there is an important need to evaluate immunotherapy drug efficacy early in the course of therapy using alternative methods that are better associated with OS.

Circulating tumor DNA (ctDNA) testing has the potential to transform patient management by providing real-time assessments of patient prognoses and response to treatment using a simple blood draw[7]. ctDNA is a subset of the total cell-free DNA circulating in the bloodstream, thought to be shed by necrotic or apoptotic cells[8]. It can be profiled using next-generation sequencing as well as other methods and can be differentiated from background cell-free DNA by the presence of somatic tumor mutations[8].

In the surgically resectable cancer setting, a positive ctDNA test after surgery has shown to be a poor-prognostic factor[9–14] and changes in ctDNA correlate with treatment response[15–17]. In the metastatic setting, treatment response and survival times have been associated with changes in ctDNA levels during systemic treatment with chemotherapy[18,19], targeted therapies[20,21], immune checkpoint inhibitors (ICIs)[19,22] and combination chemo-ICI[23]. The relatively higher ctDNA levels in patients with metastatic cancer compared to early-stage disease[24] suggest this setting would be well suited for developing ctDNA as an early endpoint for new drug or combination evaluation, or to inform risk-based treatment decisions[25,26].

The clinical implementation of ctDNA dynamics as a surrogate of survival has thus far been limited by small sample sizes, study designs without randomization or a lack of clarity on which ctDNA collection time points and summary metrics are optimal for predicting survival outcomes. To address these challenges, we performed high-sensitivity longitudinal ctDNA testing of 311 genes including correction for clonal hematopoiesis of indeterminate potential (CHIP) variants in 466 patients across 5 time points (1,954 samples total) in the phase 3 IMpower150 trial (NCT02366143).

The IMpower150 study was a randomized, open-label study that evaluated the safety and efficacy of anti-PD-L1 atezolizumab in combination with carboplatin + paclitaxel with or without bevacizumab compared with treatment with carboplatin + paclitaxel + bevacizumab in chemotherapy-naive participants with Stage IV nonsquamous non-small-cell lung cancer (NSCLC). The IMpower150 study met its primary endpoints of PFS (PFS hazard ratio (HR) = 0.62; 95% confidence interval (CI), 0.52–0.74; $P < 0.001$) and of OS (OS HR = 0.78; 95% CI, 0.64–0.96; $P = 0.02$), which led to the approval of atezolizumab + carboplatin + paclitaxel + bevacizumab in 1 L NSCLC[27]. Clinical data used in this work are based on the final OS analysis for the study (OS HR = 0.80; 95% CI, 0.67–0.95; data cutoff September 13, 2019)[28]. Atezolizumab is also an approved treatment in the early lung cancer setting[29] as well as for other tumor types[30–34].

After performing longitudinal ctDNA testing in IMpower150, we (1) examined the utility of individual ctDNA metrics to risk stratify patients including those with stable disease (SD) or partial responses (PR), (2) leveraged a machine learning approach in a training/testing framework to jointly model multiple ctDNA metrics to predict landmark survival, and (3) performed simulations to investigate whether our ctDNA model could outperform early radiographic imaging to detect differences between treatment arms in early clinical trial scenarios.

Results

Experimental plan and assay development

Of the 1,202 patients enrolled in IMpower150, baseline plasma samples from 1,062 patients were evaluated using a prototype version of the FoundationOne Liquid CDx assay by Foundation Medicine Inc. (FMI), which sequenced >1.25 Mb of genomic content covering 394 genes[35] (Fig. 1a). Sequence data were processed by a cell-free DNA computational pipeline that corrected errors via the use of fragment barcodes as previously described[35]. After the algorithmic removal of common germline and CHIP mutations, putative tumor-derived somatic alterations were identified at this baseline time point (Methods).

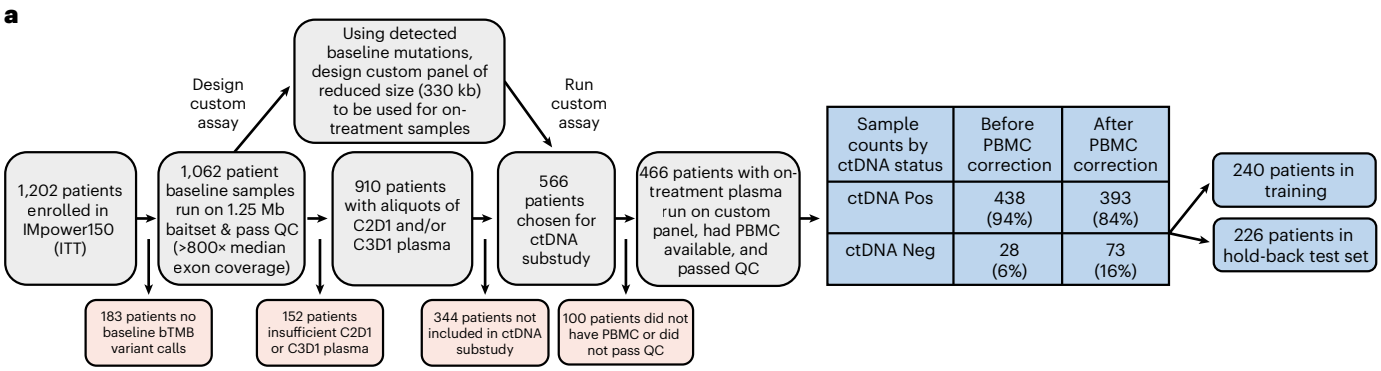
A subset of patients ($n = 466$) was chosen to be evaluated for on-treatment time points where we required the patient to have plasma available for C2D1 and/or C3D1, as well as peripheral blood mononuclear cells (PBMCs) available if putative tumor-derived variants were detected in baseline plasma ($n = 438$; Fig. 1a). Although common germline and CHIP variants were removed algorithmically in the original samples (see Methods), matched normal PBMCs were still required to remove less common germline and CHIP variants including those in canonical driver genes[36]. We expected a survivorship bias in the ctDNA-evaluable population due to our requirement for patients to have samples available after randomization, and while no strong PFS bias was found for ctDNA evaluable versus nonevaluable (HR = 0.92 (0.82–1.05)), we did detect a survivorship bias for OS (HR = 0.86 (0.75–0.99); Extended Data Fig. 1a). However, baseline characteristics were similar between the intention-to-treat (ITT) and the ctDNA evaluable population, including baseline Eastern Cooperative Oncology Group (ECOG), age, sex, race, region, among others (Supplementary Table 1).

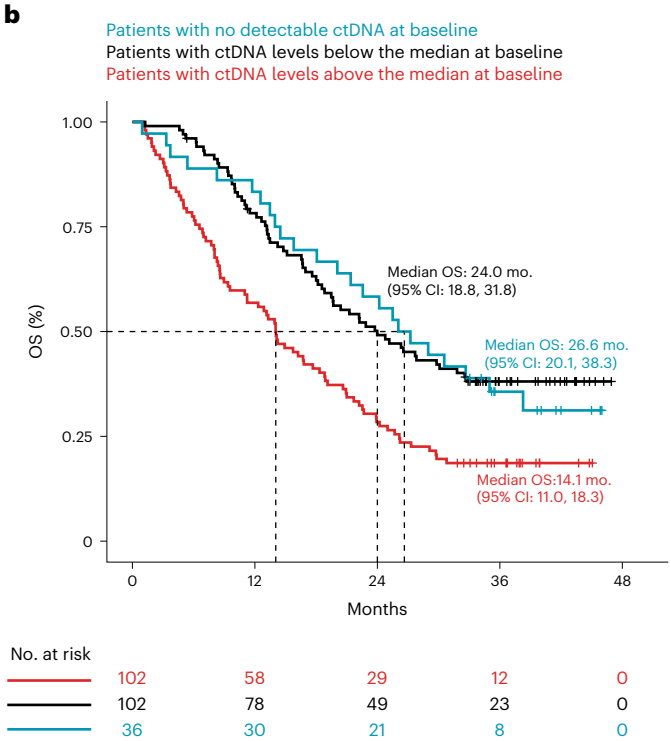
On-treatment time points were assessed using a custom fixed panel assay to track changes in ctDNA in response to therapy. A custom assay was used to allow a higher depth of coverage for a similar cost. The final hybridization capture panel reduced the total panel size to 330 kb while capturing mutations in 311 genes present in ~94% of the IMpower150 patients' baseline samples (Fig. 1a; Methods). Proprietary software developed by FMI was used to estimate on-treatment variant allele frequency (VAF) for every mutation detected at baseline (Methods). The final assay was experimentally validated to be highly concordant with the baseline assay and to have high sensitivities down to ~0.1% VAF (Extended Data Fig. 1b; Methods). The matched normal PBMCs were also run on this custom panel at high sequencing coverage (average mean target ~5,400× consensus deduplicated), and variants detected in both plasma and PBMCs were considered germline or CHIP mutations[37].

There were 282 (64%) patients who had plasma variants that were also detected in PBMCs, including 45 patients who switched from ctDNA positive (at least one mutation detected) to ctDNA negative (zero mutations detected) after this PBMC correction (Fig. 1a). The number of PBMC-derived variants detected among these 282 patients ranged from 1 to 7, with mean 1.8 and median of 1 variant. The PBMC-derived germline/CHIP mutations had allele frequencies in plasma that overlapped with somatic tumor mutations in plasma (range 0.175–69% for PBMC-derived mutations and 0.14–82% for somatic tumor mutations, medians 1.3 and 2.2%, respectively; Extended Data Fig. 1c). Common CHIP genes were excluded from our custom panel (*TET2, DNMT3A, CBL, PPM1D, CHEK2, JAK2, ASXL1* and *SF3B1*). Among the 311 genes included in our panel, the PBMC-derived mutations were most prevalent in *TP53* (5.3% of patients), followed by *MLL3* (4.1%), *FAT1* (3.6%) and *ATM* (3.0%; Extended Data Fig. 1d). All PBMC-derived variants were subtracted from the final plasma mutation dataset.

ctDNA was detected in 393 patients (84%) at the baseline time point, of which 348 (89%) had pathogenic alterations detected including in the genes *TP53* (52%), *KRAS* (23%), *STK11* (13%) and *EGFR* (10%; Extended Data Fig. 1e). For downstream analyses, the ctDNA-evaluable population was split into a training ($n = 240$) and test set ($n = 226$; Fig. 1a), which were similar in survival outcomes (Extended Data Fig. 1f), baseline clinical features and ctDNA status (Supplementary Table 2). We noted that the number of patients for each treatment arm in training and test set was well balanced; atezolizumab + bevacizumab + carboplatin + paclitaxel (ABCP) 35% ($n = 84$) and 32.3% ($n = 73$) in training

https://doi.org/10.1038/s41591-023-02226-6





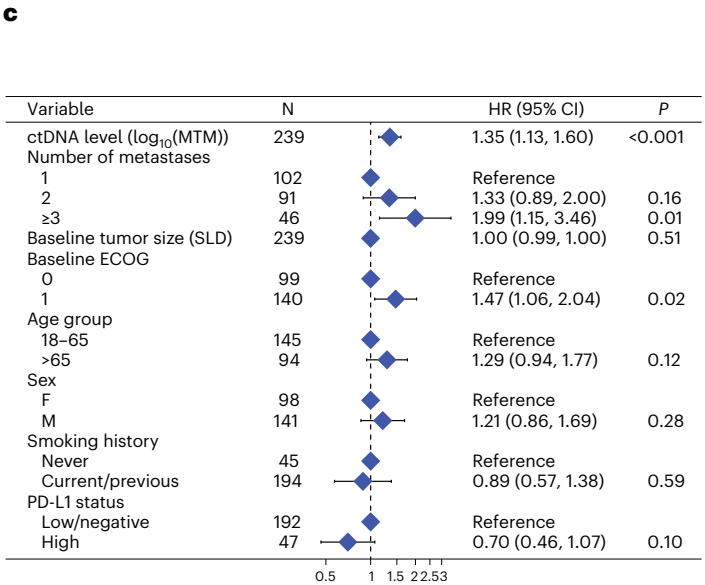


Fig. 1 | Design of ctDNA substudy and prognostic value of baseline ctDNA in training set. a, Consort diagram showing how the final 466 patients in the ctDNA evaluable population were identified and showing the prevalence of ctDNA positivity at the baseline time point before and after PBMC correction. **b**, Kaplan–Meier analysis showing the prognostic value of baseline ctDNA for OS in the training set of patients ($n = 240$), where blue curve indicates ctDNA negative patients (zero mutations detected), red curve indicates patients with ctDNA levels greater than or equal to the median (≥64 MTM) and black curve indicates patients with ctDNA levels less than the median. **c**, Multivariable Cox regression including baseline clinical features confirms that the ctDNA level is an independently poor prognostic factor for OS ($n = 239$ patients with nonmissing data available for all baseline clinical features). Two-sided Wald test P values are reported, and points and error bars indicate HR and 95% confidence interval, respectively. The exact P value for the first row '$P < 0.001$' is 0.000672. MTM, mean tumor molecules.

and test set, respectively, atezolizumab + carboplatin + paclitaxel (ACP) 31.2% (75) and 34.5% (78), and bevacizumab + carboplatin + paclitaxel (BCP) 33.8% (81) and 33.2% (75) (Supplementary Table 2). Final sample counts can be found in the supplement (Extended Data Fig. 1g). Note that all initial exploratory analyses and model building shown in Figs. 1–3 were performed in the training set of data, after which model validation was performed in the hold-back test set shown in Figs. 4 and 5.

ctDNA levels are prognostic in training data

The baseline prevalence of ctDNA positivity (at least one mutation detected) was 85% in the training split of data (204/240), and prevalence decreased to 79.3% at C2D1, 77.0% at C3D1, 77.3% at C4D1 and 76.4% at C8D1 (Extended Data Fig. 2a). Baseline ctDNA was assessed for its prognostic value and association with baseline clinical features. Patients with any detectable ctDNA ($n = 204$) trended toward worse OS compared to the 36 ctDNA negative patients (HR = 1.33 (0.87–2.06), log-rank $P = 0.19$; median OS 18.8 versus 26.6 months from baseline, respectively). Among baseline ctDNA-positive patients, the median ctDNA level was 64 MTM (mean tumor molecules per ml plasma), which corresponded to a median of 1.4% mean allele frequency (AF). Note that MTM and mean AF were highly correlated (Extended Data Fig. 2b). Patients who were positive for ctDNA could be risk stratified using the median ctDNA level, where patients with ctDNA above the median had shorter OS compared to patients with ctDNA levels below the median (Fig. 1b; HR = 1.9 (1.36–2.64), log-rank $P < 0.001$; median OS 14.1 versus 24.0 months from baseline). Higher ctDNA levels were found to be associated with poor prognostic features including age below 65 years (Wilcoxon $P = 0.0035$), positive history of smoking (Wilcoxon $P = 0.0044$), baseline tumor size by the sum of longest diameters (SLD) above the median (Wilcoxon $P < 0.001$) and higher number of metastatic sites

https://doi.org/10.1038/s41591-023-02226-6

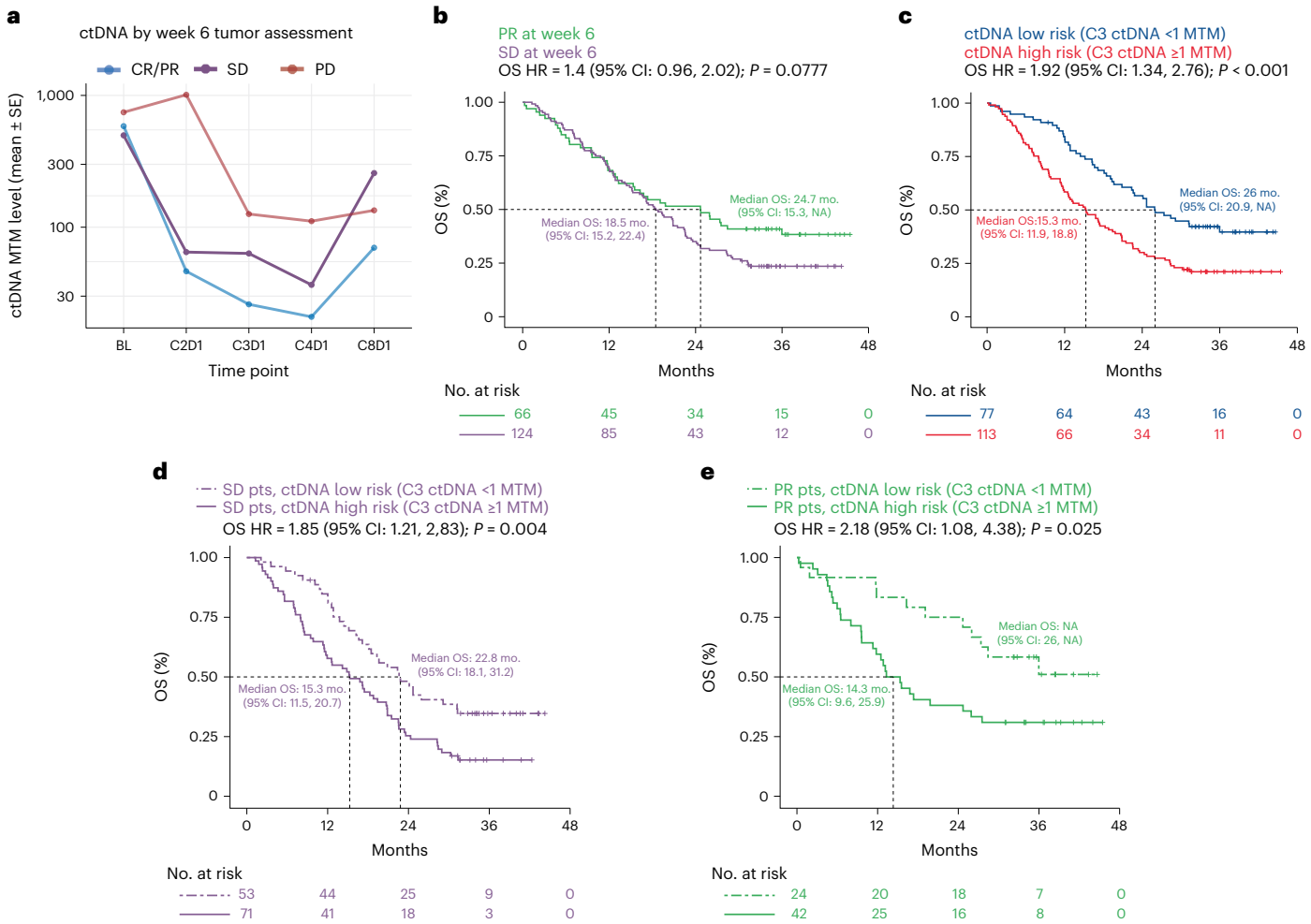


Fig. 2 | On treatment ctDNA dynamics associate with clinical outcomes in the training dataset. a, On-treatment ctDNA levels as measured by MTM (per milliliter plasma) across longitudinal time points for patients with week 6 radiographic assessments of treatment response of PD (red), SD (purple) and CR/PR (blue). **b**, KM curves showing OS for patients with SD (purple) versus PR (green) as determined at the week 6 radiographic assessment of treatment response. A univariable Cox proportional-hazards model was used to estimate HR and log-rank test to report *P* value. **c**, KM curves showing OS for patients with C3D1 ctDNA levels below the LOD of the assay (<1 MTM, ctDNA low risk, blue) versus near or above the LOD (≥1 MTM, ctDNA high risk, red). A univariable Cox proportional-hazards model was used to estimate HR and log-rank test to report *P* value. The exact *P* value for '*P* < 0.001' is 0.00029871. **d**,**e**, KM curves showing OS for patients with SD (**d**) and PR (**e**) at week 6 who are further risk stratified by ctDNA levels at C3D1. A univariable Cox proportional-hazards model was used to estimate HR and log-rank test to report *P* value. MTM, mean tumor molecules.

(Kruskal–Wallis $P < 0.001$; Extended Data Fig. 2c). However, the baseline ctDNA level was confirmed to be an independently poor prognostic factor for OS in a multivariable Cox regression model including baseline clinical features (HR = 1.35 (1.13–1.60), log-rank $P < 0.001$; Fig. 1c). Results were similar for PFS (Extended Data Fig. 2d,e).

Treatment initiation correlated with reductions in ctDNA levels, generally decreasing with each subsequent on-treatment time point through C4D1, which is the last time point in the series in which chemotherapy was given in combination with atezolizumab, bevacizumab or both (Extended Data Fig. 2f). Treatment responses as assessed by RECIST criteria at week 6 were associated with longitudinal ctDNA dynamics such that patients with CR or PR had lower ctDNA levels for all on-treatment ctDNA time points compared to patients with week 6 SD or progressive disease (PD; Fig. 2a). For example, baseline ctDNA-positive patients with a week 6 radiographic treatment response assessment of CR/PR ($n = 67$) tended to have greater reductions in ctDNA levels at C3D1 (mean −70% reduction in MTM level for CR/PR) compared to the 111 patients with week 6 SD (mean −39% reduction for SD) and the 16 patients with week 6 PD (mean +54% increase for PD; Kruskal–Wallis rank sum test $P = 0.079$). Radiographic tumor assessments were performed at baseline and every 6 weeks in the study, which is most contemporaneous with ctDNA collections at the BL and C3D1 (week 6) time points (Extended Data Fig. 2f). Levels of ctDNA were also correlated with tumor size (SLD) at BL and C3D1 (week 6; Pearson $R = 0.37$, $P < 0.001$ and Pearson $R = 0.16$, $P = 0.042$, respectively; Extended Data Fig. 2g,h). The percent change in ctDNA level from BL to C3D1 was correlated with the percent change in SLD from baseline to week 6 (Pearson $R = 0.24$, $P = 0.002$; Extended Data Fig. 2i).

Risk stratification using early radiographic tumor assessments alone showed numerically shorter OS for patients with week 6 SD compared to those with week 6 PR (median OS 18.5 versus 24.7 months; HR = 1.4 (0.96–2.02), log-rank $P = 0.078$; Fig. 2b). However, ctDNA data generated at a similar time point (C3D1) showed that patients who had ctDNA levels near or above the limit of detection (LOD) of the assay (≥1 MTM) had shorter OS compared to patients who maintained or reduced ctDNA to below the LOD (HR = 1.92 (1.34–2.76), log-rank $P < 0.001$; median 15.3 versus 26.0 months from C3D1 for patients with <1 and ≥1 MTM at C3D1, respectively; Fig. 2c).

We found that combining C3D1 ctDNA risk (≥1 versus <1 MTM) with week 6 treatment response by RECIST improved risk stratification further

https://doi.org/10.1038/s41591-023-02226-6

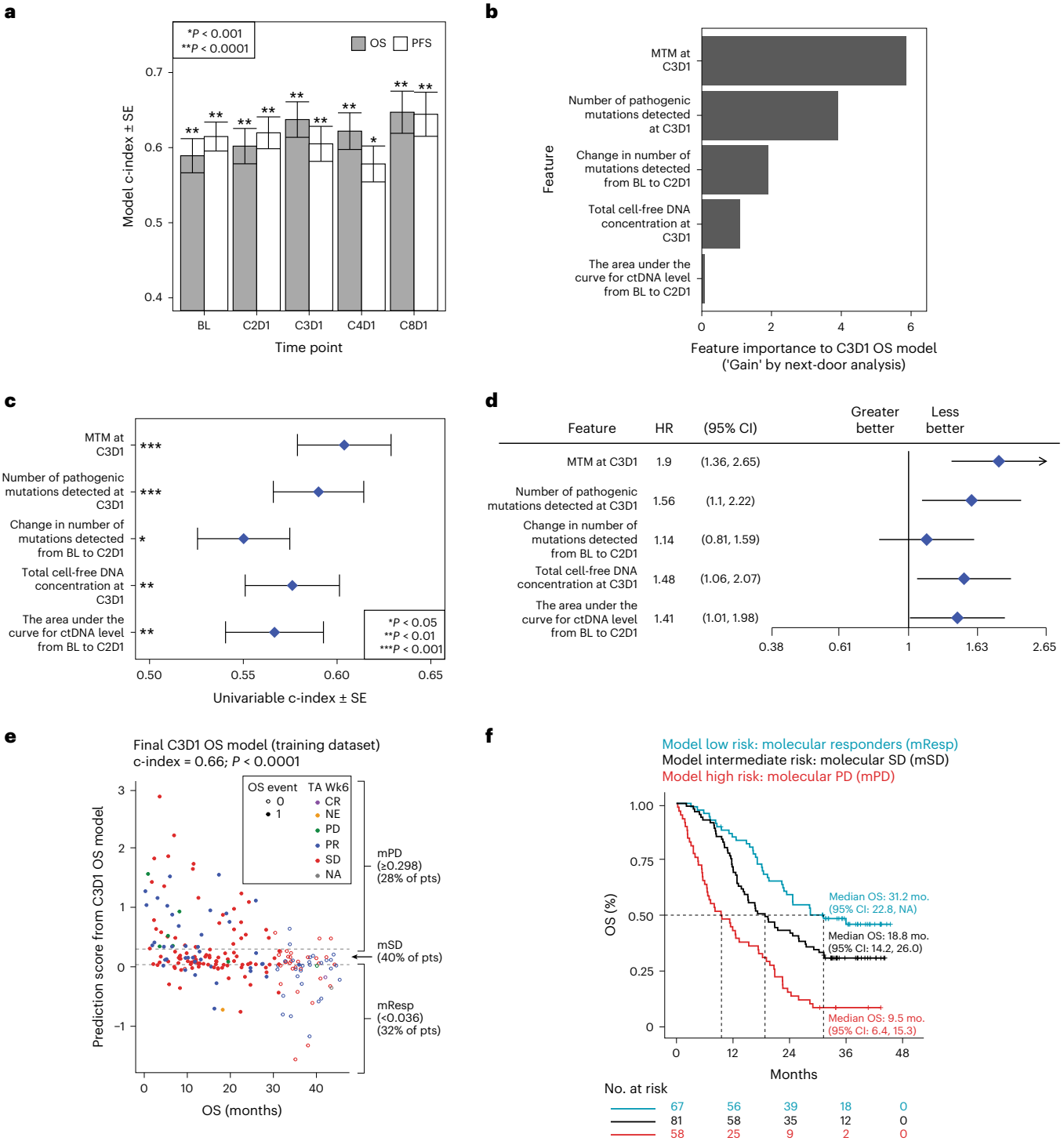


Fig. 3 | Building a machine learning model in the training dataset. a, Model performance for each survival outcome (PFS, OS) and plasma collection (BL thru C8D1) estimated by rank concordance (c-index) calculated from leave-one-out-cross-validation (LOOCV) to fit an elastic net model with ctDNA features. Bar height indicates c-index estimate, error bars indicate ± the standard error of the c-index, and two-sided *P* values are shown comparing each model's c-index to random classifier. Each model is built using patients in the training subset at-risk for the relevant landmarked survival endpoint, where the numbers from left to right are: 240, 240, 237, 237, 206, 202, 201, 196, 146 and 136. The exact *P* values from left to right are 6.69×10^{-5}, 9.50×10^{-10}, 1.06×10^{-5}, 7.97×10^{-9}, 2.87×10^{-9}, 4.16×10^{-6}, 3.35×10^{-7}, 0.000797098, 6.54×10^{-8}, 3.18×10^{-7}. **b**, Gain metric by next-door analysis for the five top features identified during LOOCV for the C3D1 OS ctDNA model. **c**, Univariable c-index showing the strength of association between OS from C3D1 ($n = 206$ patients) and each of the five top features for the C3D1 OS ctDNA model. Error bars indicate ± standard error of the c-index. Exact values from top to bottom for the two-sided *P* values comparing c-index to a random classifier are 2.23×10^{-5}, 1.35×10^{-4}, 0.0366, 0.0021 and 0.0093. **d**, Forest plot showing the HR for OS from C3D1 ($n = 206$ patients) estimated by univariable Cox proportional-hazards model, using the median value for the feature split, for the five top features for C3D1 OS ctDNA model. Higher feature values (above median) were generally associated with worse OS (HR above 1). Points and error bars indicate HR and 95% CI, respectively. **e**, Scatterplot showing final C3D1 OS ctDNA model predictions (*y* axis) versus OS time (*x* axis) in the training data, with dotted lines indicating the thresholds chosen in training set for mPD (≥0.298 prediction score), molecular response (mResp < 0.036) and molecular stable disease (mSD for (0.036, 0.298)). The exact value for the two-sided *P* value comparing the final C3D1 OS model's c-index to a random classifier *P* value indicated by '$P < 0.0001$' is 1.318316×10^{-12}. **f**, KM curve showing that the final C3D1 OS ctDNA model can risk stratify patients in the training data.

https://doi.org/10.1038/s41591-023-02226-6

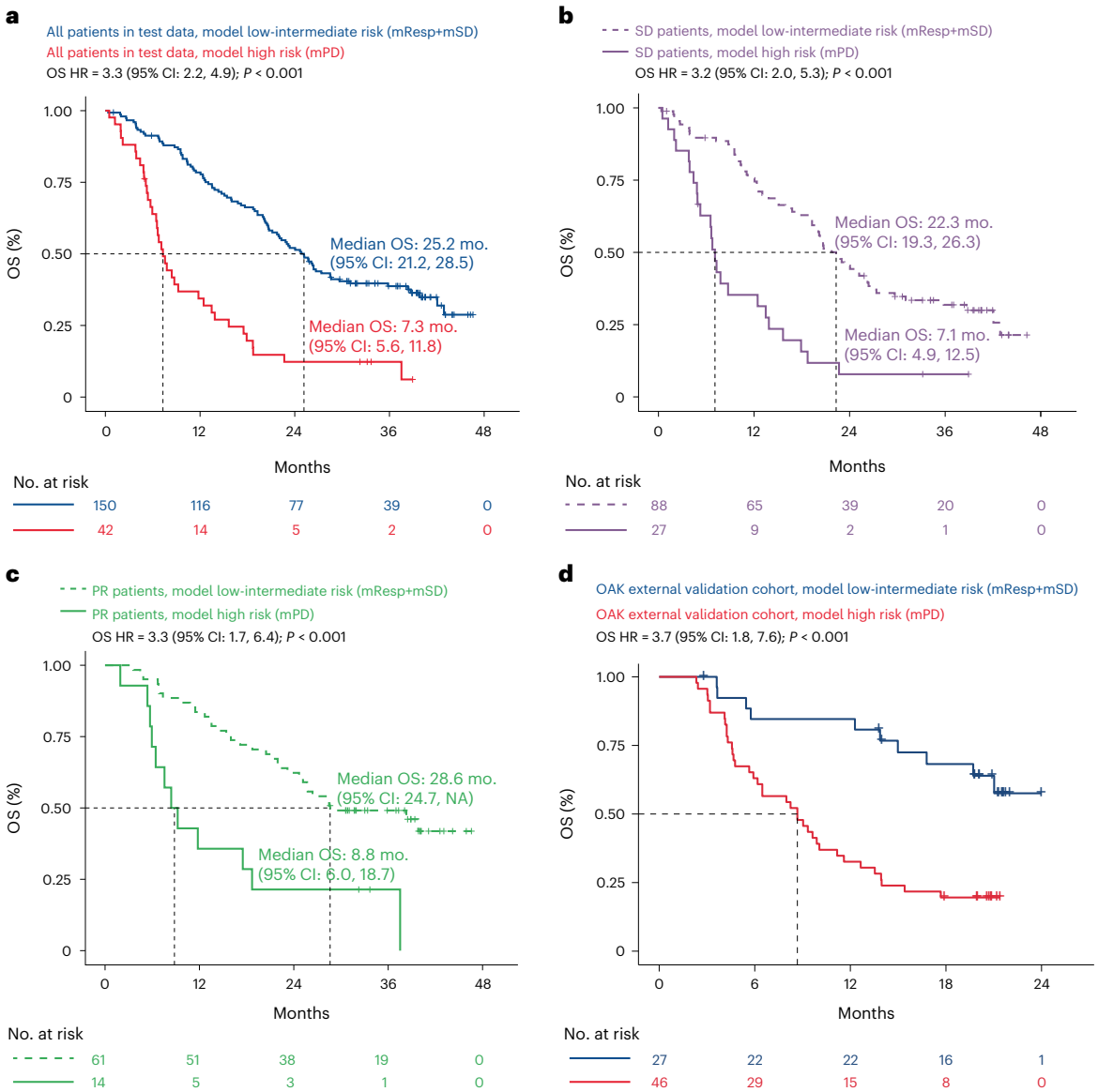


Fig. 4 | Machine learning model performs well for risk stratification in the hold-back test dataset and in the OAK external validation cohort. a, KM curve showing that the final C3D1 OS ctDNA model can be used for risk stratification in the hold-back test data, where patients with mPD (red) have worse OS compared to patients with a molecular response or molecular stable disease (mResp + mSD, blue). A univariable Cox proportional-hazards model was used to estimate HR and log-rank test to report *P* value. The exact *P* value indicated by '$P < 0.001$' is 3.7228×10^{-10}. **b**, KM curve showing that patients with radiographic treatment response of SD at the week 6 tumor assessment can be risk stratified using the final C3D1 OS ctDNA model in the hold-back test data, identifying SD/ctDNA high-risk patients (mPD, solid curve) and SD/ctDNA low-intermediate risk patients (mSD + mResp, dashed curve). A univariable Cox proportional-hazards model was used to estimate HR and log-rank test to report *P* value. The exact *P* value indicated by '$P < 0.001$' is 8.8076×10^{-7}. **c**, KM curve showing that patients with radiographic treatment response of PR at the week 6 tumor assessment can be risk stratified using the final C3D1 OS ctDNA model in the hold-back test data, identifying PR/ctDNA high-risk patients (mPD, solid curve) and PR/ctDNA low-intermediate risk patients (mSD + mResp, dashed curve). A univariable Cox proportional-hazards model was used to estimate HR and log-rank test to report *P* value. The exact *P* value indicated by '$P < 0.001$' is 0.0003018. **d**, KM curve showing that the C3D1 OS model applied to the external validation cohort of 73 patients from the OAK clinical trial can provide predictions that identify high-risk patients in this 2nd line mNSCLC setting that used a distinct ctDNA assay technology. A univariable Cox proportional-hazards model was used to estimate HR and log-rank test to report *P* value. The exact *P* value indicated by '$P < 0.001$' is 0.000119.

such that SD patients could be split into SD/ctDNA high-risk versus SD/ctDNA low-risk (Fig. 2d; OS HR = 1.85 (1.21–2.83), log-rank $P = 0.004$; median 15.3 versus 22.8 months from C3D1, respectively). Week 6 PR patients could also be split into PR/ctDNA high-risk versus PR/ctDNA low-risk (Fig. 2e; OS HR = 2.18 (1.08–4.38), log-rank $P = 0.025$; median 14.3 months from C3D1 versus median not reached, respectively). The week 6 PR patients who were high risk by ctDNA had numerically shorter duration of treatment response compared to patients who were low risk by ctDNA (median 5.6 versus 11.5 months duration of treatment response; duration of treatment response (DoR) HR = 0.59 (0.33–1.06), log-rank $P = 0.075$; Extended Data Fig. 3a). Results were similar when analyses were repeated using PFS (Extended Data Fig. 3b) or alternative ctDNA thresholds or metrics (Extended Data Fig. 3c,d). Visualizing the SLD dynamics parallel to ctDNA dynamics revealed that while ctDNA changes mirror SLD changes, patients with different response categories by radiographic imaging could still have very similar ctDNA patterns (Extended Data Fig. 2e). Overall, these results suggest that imaging-based risk stratification at early on-treatment time points may be improved upon by combining

https://doi.org/10.1038/s41591-023-02226-6

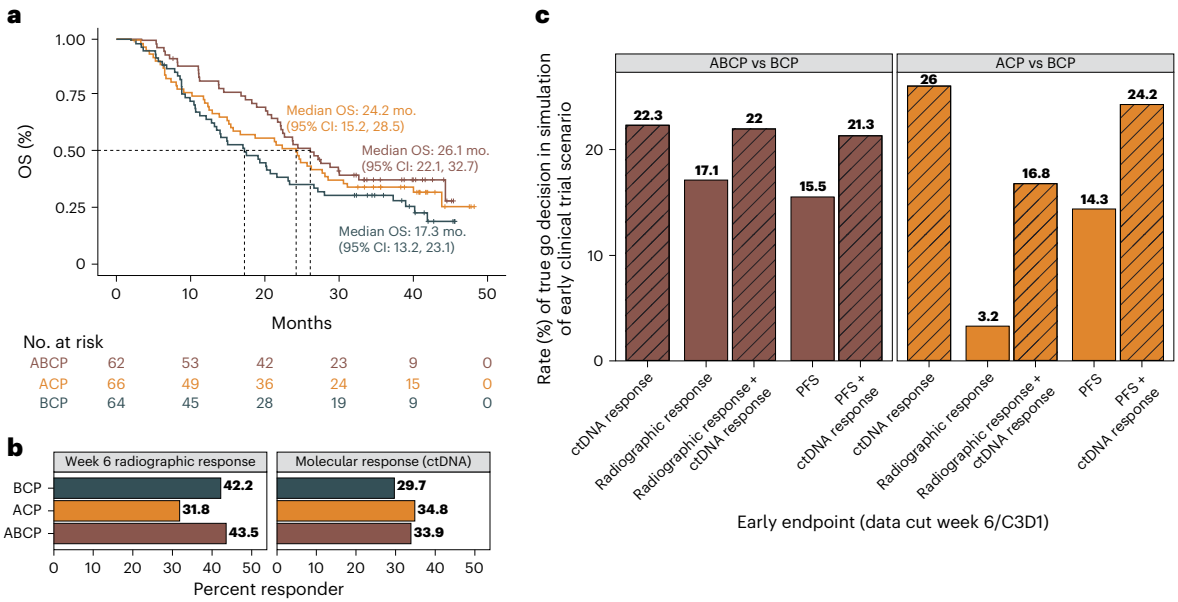


Fig. 5 | Machine learning model may be useful for detecting differences between treatment arms in early phase 2 clinical trial scenarios. a, KM curve showing OS in the test dataset for the three arms in the IMpower150 trial including ABCP (brown) versus ACP (orange) versus the control arm of BCP (black, control arm). **b**, Bar plot showing the rate of radiographic response at the week 6 tumor assessment for each treatment arm (left panel, CR/PR by RECIST criteria), and the rate of ctDNA molecular response for each treatment arm (right panel, mResp by C3D1 OS ctDNA model). **c**, Bar plot showing results from simulations of early phase 2 clinical trial scenario utilizing test data, where an early endpoint based on ctDNA (mResp by C3D1 OS model) is compared to early radiographic endpoints (week 6 RECIST response, week 6 PFS). Bar height corresponds to the proportion of simulations in which the active arm had higher rates of treatment response compared to control arm ('true go rate') for each early endpoint (*x* axis), where the left panel shows simulations comparing active ABCP arm to control BCP arm (left panel, brown colors), and right panel shows simulations comparing active ACP arm to control BCP arm (right panel, orange colors). *X* axis corresponds to which early endpoint is used in the simulation, comparing ctDNA criteria alone (mResp by C3D1 OS model), radiographic response alone (CR/PR by RECIST), PFS alone, or ctDNA added to radiographic response or PFS response.

radiographic imaging with ctDNA measurements, potentially informing risk-based treatment decisions made by clinicians.

Machine learning model predicts survival using ctDNA metrics

Multiple metrics can be used to describe ctDNA dynamics, for example, the ctDNA level can be measured using the AF, the MTM per milliliter plasma, the total cell-free DNA concentration, the number of mutations detected or the number of pathogenic mutations detected, among others[38]. We derived 19 metrics of the ctDNA level (measured for every time point) and 59 metrics of ctDNA change relative to baseline (measured for each on-treatment time point) (Supplementary Table 3; Methods). The performance of different metrics to individually predict landmark survival can be summarized using the rank concordance (Harrell's c-index), and we find that performance varies across metric types, time points and survival outcomes in the training data ($n = 240$; Extended Data Fig. 4a). We next jointly modeled these features in a machine learning framework using an elastic net approach[39] to predict landmark PFS and landmark OS from each plasma collection time point (see Methods). This permits many features related to ctDNA levels and changes to be included in the model initially; however, nested cross-validations during the training reduce the number of features to an optimal subset that minimizes prediction error.

Models were trained for each visit time point, where all measurements collected from baseline up until that particular visit were included as features, for example for C3D1 models, features from BL, C2D1 and C3D1 were included. Note that treatment arms were pooled to build a model useful as an early endpoint regardless of the treatment regime being used (similar concept to RECIST criteria). Pooling was also appealing because building a multivariable predictive model requires relatively large sample sizes in the training and testing splits. Results showed that model performance was generally best (higher c-index) for OS at the C3D1/C8D1 time points (Fig. 3a and Supplementary Table 4). We decided to focus on the OS model for the C3D1 time point given that C8D1 samples were available for only 60% of patients (which itself is a good prognostic/favored by immortal bias) and because OS is the most relevant metric for oncology therapeutics.

To investigate how a ctDNA-focused model compared to known prognostic factors, we compared C3D1 OS models trained using either clinical features alone or combined with ctDNA features. For clinical features, we included the baseline factors shown in Fig. 1c as well as measures of tumor size from early radiographic assessments at baseline and week 6 (see Methods). We find that when comparing C3D1 OS runs, the combined ctDNA + clinical feature set performs substantially better compared to the clinical features alone ($P = 0.0125$; Extended Data Fig. 4b and Supplementary Table 4).

To identify top features for the C3D1 OS ctDNA model, we chose ctDNA features that were chosen during training in >50% of cross-validations and which had a positive gain metric by next-door analysis[40]. This reduced the number of features in the model to just five metrics, which included metrics related to MTM, the number of detected mutations and the total cell-free DNA concentration (Fig. 3b, Supplementary Table 5). Individually, these metrics showed univariable c-indices for OS between 0.55 and 0.60 (Fig. 3c, Extended Data Fig. 4c and Supplementary Table 5), and the median split of each metric separately provided OS HRs between 1.14 and 1.90 (Fig. 3d).

The final C3D1 OS ctDNA model was fit using the five top ctDNA features in the entire training dataset with a final c-index of 0.66 ($P < 0.001$ for comparison to random classifier; Fig. 3e and Supplementary Table 6). Numeric thresholds for binning C3D1 OS model predictions were chosen in the training data (see Methods) to bin patients into three groups including those at high risk (molecular progressive disease (mPD), 28% of training set), those at low risk (molecular response

(mResp), 32% of training set) and those with intermediate risk (molecular stable disease (mSD), 40% of training set; Fig. 3e, Extended Data Fig. 4d,e). Kaplan–Meier analysis confirmed the prognostic value for OS of these three bins in the training data (median OS from C3D1 of 9.5, 18.8 and 31.2 months for mPD, mSD and mResp, respectively; Fig. 3f).

We tested the final C3D1 OS ctDNA model at a c-index of 0.67 ($P < 0.001$, Extended Data Fig. 5a) in the hold-back test set with evaluable C3D1 plasma ($n = 192$ patients, Fig. 1b) and found that the prespecified thresholds chosen in the training set also provided good separation between risk groups in the test data (Extended Data Fig. 5b). The ctDNA model high-risk group showed shorter OS compared to patients with low-intermediate risk (OS HR = 3.28 (2.2–4.9), log-rank $P < 0.001$; median 7.3 months versus 25.2 months from C3D1 for mPD and mResp + mSD, respectively; Fig. 4a). Model predictions in the test set also performed well in patients with a week 6 tumor assessment of SD (c-index in SD = 0.68; $P < 0.001$), and the prespecified threshold could risk stratify SD patients into high risk versus low-intermediate risk (OS HR in SD = 3.23 (1.98–5.29)], log-rank $P < 0.001$; median 7.1 versus 22.3 from C3D1 for mPD and mResp + mSD, respectively) (Fig. 4b). Similarly, ctDNA model predictions in the test set performed well in patients with PR (c-index in PR = 0.64; $P = 0.002$) and was able to identify a high-risk PR subgroup (OS HR in PR = 3.26 (1.66–6.42)), log-rank $P < 0.001$; median 8.8 versus 28.6 from C3D1 for mPD and mResp + mSD, respectively; Fig. 4c).

We further validated our C3D1 OS ctDNA model in an external patient cohort ($n = 73$), which is a subset of patients from the OAK clinical trial (NCT02008227) who had ctDNA assessed using the Avenio ctDNA assay at BL, C2D1 and C3D1[19]. In addition to a different ctDNA assay, the external cohort also included metastatic NSCLC patients from the 2nd line setting (versus 1st line in IMpower150) and different treatment regimes (monotherapy of atezolizumab or docetaxel versus chemo-ICI combinations in IMpower150). Applying our C3D1 OS predictor developed in the IMpower150 data to the OAK data (Methods), we find it validates well in this external cohort with a c-index of 0.69 ($P < 0.0001$; Extended Data Fig. 5c), and the prespecified cutoffs were also able to identify high risk (mPD) versus low-intermediate risk (mResp + mSD) patients in the OAK clinical trial (OS HR = 3.73 (1.83–7.60), log-rank $P = 0.00012$; Fig. 4d and Extended Data Fig. 5d). These findings support the potential clinical utility of our ctDNA model across multiple treatments and settings in NSCLC.

ctDNA model utility as an early endpoint in drug development

For a new early endpoint to be useful for clinical drug development in metastatic lung cancer, it should sensitively detect differences in OS between treatment arms at an early time point. Additionally, it should improve upon more traditional early radiographic-based endpoints like RECIST response or PFS. In this IMpower150 study, patients in the control arm received BCP and patients in the experimental arms received either ABCP or ACP (Fig. 5a, Extended Data Fig. 6a, test data shown). We note that the US approval was based on comparing ABCP arm to BCP arm[27,41], although the ABCP and ACP arms had similar OS results[28]. To explore whether our ctDNA C3D1 OS model could provide early signals of treatment efficacy, we compared rates of ctDNA mResp between the different treatment arms and contrasted with RECIST response and PFS assessed near week 6.

The rate of ctDNA mResp was 33.9% in ABCP and 34.8% in ACP versus 29.7% in BCP control arm (Fig. 5b). Whereas in week 6 the RECIST response rates were numerically higher for the active ABCP versus control BCP arms (43.5% in ABCP and 42.2% in BCP), the active ACP arm had a lower rate of radiographic response compared to control at this early week 6 time point (31.8% in ACP).

To quantify the utility of ctDNA as an early endpoint capable of informing drug development decision-making in an early phase 2 clinical trial scenario, we resampled IMpower150 to $n = 30$ patients per arm in $n = 2{,}000$ simulations using our test data. We then measured the proportion of simulations in which the active arm had higher rates of treatment response compared to control arm ('true go rate') using different early endpoints assessed near week 6 including ctDNA response, RECIST response, PFS, or a combination (see Methods section on operation characteristics simulations[42]). The results of this analysis showed that ctDNA response by itself had higher true go rates than either week 6 RECIST or PFS (Fig. 5c), suggesting that ctDNA response may be more useful than early radiographic assessments for detecting signals of drug efficacy. Additionally, combining the ctDNA response metric with RECIST improved the true go rate compared to RECIST by itself, as did combining ctDNA response metric with PFS (Fig. 5c). For completeness, we repeated the simulations in the training data and results were similar (Extended Data Fig. 6b), and we also modified simulations to use a ramp-up enrollment approach (see Methods) which showed the ctDNA response endpoint to have less utility (Extended Data Fig. 6c). Overall, these findings suggest that on-treatment ctDNA measurements may have utility as an early endpoint to support early decision-making in clinical trial scenarios.

Discussion

To our knowledge, this is the first study to systematically evaluate the utility of longitudinal ctDNA dynamics across a large, randomized phase 3 clinical trial. We show that ctDNA metrics collected across longitudinal time points can be used to risk stratify patients and predict survival in patients with metastatic nonsquamous NSCLC treated with chemo-immunotherapy combinations. In addition to the high prognostic value of baseline ctDNA levels for PFS and OS, on-treatment ctDNA changes are correlated with treatment response and can be combined with radiographic imaging to provide finer risk stratification of patients who achieve PR or SD. Notably, in an external cohort of patients our model predicted high-risk patients despite differences in treatment setting and ctDNA technology, further supporting the potential clinical utility of ctDNA for predicting OS for immunotherapy and immunotherapy combinations in multiple NSCLC settings.

There are diverse approaches used in the literature to summarize ctDNA levels and integrate ctDNA features for association with clinical outcomes, as well as the open question regarding which on-treatment time points may be optimal for longitudinal ctDNA analyses[19,38,43,44]. We leveraged a machine learning approach to address these questions by jointly modeling multiple ctDNA metrics to predict landmark OS and PFS. We found C3D1 to be a preferred time point for model performance, and due to occur relatively early in the treatment timeline before some patients have disease progression. Additionally, we found top ctDNA features to include metrics related to the number of detected variants, the MTM per milliliter plasma and the total cell-free DNA extracted from plasma.

We also explored the utility of ctDNA as an early endpoint for detecting differences between treatment and control arms in randomized clinical trial settings. Using our final C3D1 OS model, we compared ctDNA response to radiographic endpoints assessed at week 6 including PFS and RECIST response. We found that ctDNA response outperforms RECIST and PFS, particularly when evaluating drugs that are not cytotoxic agents like immunotherapies. mResp by ctDNA could be very useful in early clinical development decision-making for these types of drugs.

In terms of study limitations, the number of patients with samples available decreased with each consecutive time point due to coming off the study at disease progression, potentially limiting the statistical power of our approach at later time points. We note that an early endpoint will be of limited utility for patients that progress so rapidly because their clinical outcomes are also apparent early in the study timeline. Another caveat of our work is that while we assessed the association of ctDNA with clinical benefit for chemo-ICI combinations as well as monotherapies, additional data should be generated in other treatment contexts (for example, targeted therapies) to delineate treatment-dependent ctDNA

dynamics. Additionally, we focus on the association of ctDNA with OS which is an endpoint that can be confounded by poststudy treatment, and it is unknown how this may have affected our results. Lastly, this study used data from two separate ctDNA assays, both of which were panel-based approaches with high sensitivity down to ~0.1% ctDNA fraction and high specificity via PBMC correction. However, there are other assays currently available on the market and it is unclear whether assay choice may affect the correlation between ctDNA metrics and clinical outcomes. While we included a comparison of ctDNA features with baseline clinical factors, an interesting extension of our work would be to incorporate other on-treatment circulating biomarkers and assess whether they are superior to ctDNA for predicting outcomes.

Overall, we have mapped ctDNA dynamics in a large, randomized study with unprecedented resolution. We have shown that changes in ctDNA, as modeled in a machine learning framework and validated in both a hold-back test set and an external cohort, can improve patient risk stratification, as well as sensitively detect differences between treatment arms at early time points in clinical trial settings. ctDNA shows promise as an early endpoint for decision-making during drug development and, with prospective validation, potentially as a risk stratification tool to inform treatment decisions.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-023-02226-6.

References

1. Eisenhauer, E. A. et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).
2. Fojo, A. T. & Noonan, A. Why RECIST works and why it should stay—counterpoint. *Cancer Res.* **72**, 5151–5157 (2012).
3. Chiou, V. L. & Burotto, M. Pseudoprogression and immune-related response in solid tumors. *J. Clin. Oncol.* **33**, 3541–3543 (2015).
4. Tazdait, M. et al. Patterns of responses in metastatic NSCLC during PD-1 or PDL-1 inhibitor therapy: comparison of RECIST 1.1, irRECIST and iRECIST criteria. *Eur. J. Cancer* **88**, 38–47 (2018).
5. Petrelli, F. et al. Surrogate endpoints in immunotherapy trials for solid tumors. *Ann. Transl. Med.* **7**, 154–154 (2019).
6. Nie, R.-C. et al. Evaluation of objective response, disease control and progression-free survival as surrogate end-points for overall survival in anti-programmed death-1 and anti-programmed death ligand 1 trials. *Eur. J. Cancer* **106**, 1–11 (2019).
7. Corcoran, R. B. & Chabner, B. A. Application of cell-free DNA analysis to cancer treatment. *N. Engl. J. Med.* **379**, 1754–1765 (2018).
8. Wan, J. C. M. et al. Liquid biopsies come of age: towards implementation of circulating tumour DNA. *Nat. Rev. Cancer* **17**, 223–238 (2017).
9. Tie, J. et al. Circulating tumor DNA analyses as markers of recurrence risk and benefit of adjuvant therapy for stage III colon cancer. *JAMA Oncol.* **5**, 1710 (2019).
10. Garcia-Murillas, I. et al. Assessment of molecular relapse detection in early-stage breast cancer. *JAMA Oncol.* **5**, 1473 (2019).
11. Abbosh, C. et al. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. *Nature* **545**, 446–451 (2017).
12. Lau, E. et al. Detection of ctDNA in plasma of patients with clinically localised prostate cancer is associated with rapid disease progression. *Genome Med.* **12**, 72 (2020).
13. Cullinane, C. et al. Association of circulating tumor DNA with disease-free survival in breast cancer. *JAMA Netw. Open* **3**, e2026921 (2020).
14. Kuang, P.-P. et al. Circulating tumor DNA analyses as a potential marker of recurrence and effectiveness of adjuvant chemotherapy for resected non-small-cell lung cancer. *Front. Oncol.* **10**, 595650 (2021).
15. Powles, T. et al. ctDNA guiding adjuvant immunotherapy in urothelial carcinoma. *Nature* **595**, 432–437 (2021).
16. Christensen, E. et al. Early detection of metastatic relapse and monitoring of therapeutic efficacy by ultra-deep sequencing of plasma cell-free DNA in patients with urothelial bladder carcinoma. *J. Clin. Oncol.* **37**, 1547–1557 (2019).
17. Magbanua, M. J. M. et al. Circulating tumor DNA in neoadjuvant-treated breast cancer reflects response and survival. *Ann. Oncol.* **32**, 229–239 (2021).
18. Osumi, H., Shinozaki, E., Yamaguchi, K. & Zembutsu, H. Early change in circulating tumor DNA as a potential predictor of response to chemotherapy in patients with metastatic colorectal cancer. *Sci. Rep.* **9**, 17358 (2019).
19. Zou, W. et al. ctDNA predicts overall survival in patients with NSCLC treated with PD-L1 blockade or with chemotherapy. *JCO Precis. Oncol.* 5, 827–838.
20. Buder, A., Hochmair, M. J., Setinek, U., Pirker, R. & Filipits, M. EGFR mutation tracking predicts survival in advanced EGFR-mutated non-small cell lung cancer patients treated with osimertinib. *Transl. Lung Cancer Res.* **9**, 239–245 (2020).
21. Cheng, M. L. et al. Plasma ctDNA response is an early marker of treatment effect in advanced NSCLC. *JCO Precis. Oncol.* **16**, 393–402 (2021).
22. Bratman, S. V. et al. Personalized circulating tumor DNA analysis as a predictive biomarker in solid tumor patients treated with pembrolizumab. *Nat. Cancer* **1**, 873–881 (2020).
23. Ricciuti, B. et al. Early plasma circulating tumor DNA (ctDNA) changes predict response to first-line pembrolizumab-based therapy in non-small-cell lung cancer (NSCLC). *J. Immunother. Cancer* **9**, e001504 (2021).
24. Bettegowda, C. et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci. Transl. Med.* **6**, 224ra24 (2014).
25. Ferrara, R. et al. Do immune checkpoint inhibitors need new studies methodology? *J. Thorac. Dis.* **10**, S1564–S1580 (2018).
26. Chen, T.-T. Statistical issues and challenges in immuno-oncology. *J. Immunother. Cancer* **1**, 18 (2013).
27. Socinski, M. A. et al. Atezolizumab for first-line treatment of metastatic nonsquamous NSCLC. *N. Engl. J. Med.* **378**, 2288–2301 (2018).
28. Socinski, M. A. et al. IMpower150 final overall survival analyses for atezolizumab plus bevacizumab and chemotherapy in first-line metastatic nonsquamous NSCLC. *J. Thorac. Oncol.* **16**, 1909–1924 (2021).
29. Felip, E. et al. Adjuvant atezolizumab after adjuvant chemotherapy in resected stage IB–IIIA non-small-cell lung cancer (IMpower010): a randomised, multicentre, open-label, phase 3 trial. *Lancet* **398**, 1344–1357 (2021).
30. Herbst, R. S. et al. Atezolizumab for first-line treatment of PD-L1-selected patients with NSCLC. *N. Engl. J. Med.* **383**, 1328–1339 (2020).
31. Balar, A. V. et al. Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. *Lancet* **389**, 67–76 (2017).
32. Horn, L. et al. First-line atezolizumab plus chemotherapy in extensive-stage small-cell lung cancer. *N. Engl. J. Med.* **379**, 2220–2229 (2018).

33. Gutzmer, R. et al. Atezolizumab, vemurafenib, and cobimetinib as first-line treatment for unresectable advanced BRAF^{V600} mutation-positive melanoma (IMspire150): primary analysis of the randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* **395**, 1835–1844 (2020).
34. Finn, R. S. et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. *N. Engl. J. Med.* **382**, 1894–1905 (2020).
35. Gandara, D. R. et al. Blood-based tumor mutational burden as a predictor of clinical benefit in non-small-cell lung cancer patients treated with atezolizumab. *Nat. Med.* **24**, 1441–1448 (2018).
36. Yaung, S. J. et al. Clonal hematopoiesis in late-stage non-small-cell lung cancer and its impact on targeted panel next-generation sequencing. *JCO Precis. Oncol.* **4**, 1271–1279 (2020).
37. Razavi, P. et al. High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. *Nat. Med.* **25**, 1928–1937 (2019).
38. Friends of Cancer Research. Assessing the use of ctDNA as an early endpoint in early-stage disease. https://friendsofcancerresearch.org/wp-content/uploads/Assessing_Use_of_ctDNA_Early_Endpoint_Early-Stage_Disease-1.pdf (2021).
39. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R. Statist. Soc. Ser. B* **67**, 301–320 (2005).
40. Guan, L. & Tibshirani, R. Post model-fitting exploration via a 'next-door' analysis. *Can. J. Statist.* **48**, 447–470 (2020).
41. FDA. FDA approves atezolizumab with chemotherapy and bevacizumab for first-line treatment of metastatic non-squamous NSCLC. https://www.fda.gov/drugs/fda-approves-atezolizumab-chemotherapy-and-bevacizumab-first-line-treatment-metastatic-non-squamous (2018).
42. Fridlyand, J., Kaiser, L. D. & Fyfe, G. Analysis of tumor burden versus progression-free survival for phase II decision making. *Contemp. Clin. Trials* **32**, 446–452 (2011).
43. Nabet, B. Y. et al. Noninvasive early identification of therapeutic benefit from immune checkpoint inhibition. *Cell* **183**, 363–376 (2020).
44. Sanz-Garcia, E., Zhao, E., Bratman, S. V. & Siu, L. L. Monitoring and adapting cancer treatment using circulating tumor DNA kinetics: current research, opportunities, and challenges. *Sci. Adv.* **8**, eabi8618 (2022).

Methods

IMpower150 trial design, participants and endpoints

Details on the IMpower150 study plan and results have been published elsewhere[27], and the study Protocol and Statistical Analysis Plan can be found on clinicaltirals.gov (NCT02366143). The protocol for IMpower150 was approved by ethics committees at each site, and the study followed the International Conference on Harmonisation Good Clinical Practice guidelines and accorded with the principles of the Declaration of Helsinki. The list of 161 ethics committees can be found in the reporting summary. All patients provided written informed consent and were not compensated for participation.

IMpower150 was a phase 3 study that randomly assigned patients in a 1:1:1 ratio to receive every 3 weeks either ACP group, or ABCP group or BCP group. Patients underwent radiographic (CT scan) tumor assessments until the occurrence of disease progression (according to RECIST v1.1 criteria) or until the loss of clinical benefit among patients who continued to receive atezolizumab after the initial disease progression. These assessments were performed at screening and every 6 weeks from cycle 1 day 1 for the first 48 weeks, and every 9 weeks thereafter. The primary endpoints were PFS (as assessed by investigators according to RECIST criteria) and OS. The final analysis including distribution of poststudy treatment usage has been reported previously[28], although we note there is a lot of missingness in the poststudy treatment data due to patients coming off-study upon deviation from protocol-specified anticancer therapy.

Patient inclusion criteria were the following: stage IV or recurrent metastatic nonsquamous NSCLC without previously receiving chemotherapy, a ECOG performance-status score of 0 or 1 at baseline, available tumor tissue for testing and eligibility to receive bevacizumab. Patients who had received previous adjuvant or neoadjuvant chemotherapy were eligible if the last treatment was at least 6 months before randomization. Any PD-L1 immunohistochemistry status was eligible, and tumor PD-L1 expression (on tumor cells or tumor-infiltrating immune cells) was assessed in archival or freshly collected tissue (or both) with a PD-L1 immunohistochemistry assay (Ventana Medical Systems; clone SP142; N/A; predilute ready to use antibody product at 36 µg/5 ml). Patients having *EGFR* or *ALK* genomic alterations were included if they previously had treatment with 1+ approved tyrosine kinase inhibitor but had disease progression or unacceptable side effects. Exclusion criteria were untreated metastases of the central nervous system, autoimmune disease, or receiving previous immunotherapy or anti-CTLA-4 therapy within 6 weeks before randomization, or receiving systemic immunosuppressive medications within 2 weeks before randomization.

Sample collection and processing, assay development and splitting into training/testing sets

The PBMCs used for analysis were isolated from one 8.5 ml of whole blood collected in an acid citrate dextrose tube at a specialty vendor and from an 8 ml cell preparation tube containing sodium citrate. The plasma used for analysis was separated from two times 6 ml of whole blood collected in K2 EDTA vacutainers and was processed within 30 min after blood collection.

Baseline plasma samples from 1,062 patients were retrospectively analyzed using the assay method described previously[35].

On-treatment samples from 566 patients (C2D1, C3D1, C4D1 or C8D1) were evaluated with a custom 330 kb assay targeting 311 genes. The hybrid capture panel for this assay was designed by pooling the alterations found for all the samples in the baseline assay, filtering for known germline variants based on ExAC database (http://exac.broadinstitute.org/), known CHIP genes *TET2*, *DNMT3A*, *CBL*, *PPM1D*, *CHEK2*, *JAK2*, *ASXL1*, *SF3B1*, noncoding variants and repetitive regions, and <100× coverage. The resulting pool of alterations was clustered based on proximity within the genome, and clusters with four or more alterations plus smaller clusters that represented samples with less than three alterations were chosen for hybrid capture bait designs. The genomic regions of the clusters were compared to the baits in the baseline bait set, the corresponding baits were selected as the custom assay bait set.

The sequencing libraries were prepared using the same plasma extraction, library construction and hybrid capture-based methodology as FoundationACT with consistent analytical performance (that is, sensitivity, specificity) and has been previously described[35,45]. Briefly, between 1 ml and 5 ml of frozen plasma from each patient was sent to FMI. Once received, cfDNA was extracted, and isolated cfDNA was quantified using the 4,200 TapeStation (Agilent Technologies). A minimum of 20 ng of extracted cfDNA was required for a sample to undergo sequencing.

In this study, the assay LOD and lower limit of quantitation (LOQ) were determined to be 0.1% and 0.5%, respectively, as follows: 63 on-treatment samples with residual plasma were rerun through the sequencing library prep, construction, hybrid capture and sequencing pipeline. Based on the baseline time point, there was a total of 485 variants expected to be present in these samples. The AF of these 485 variants was then assessed in the replicate on-treatment samples (Extended Data Fig. 1b, right panel). The LOD of the assay, commonly defined as the lowest concentration of an analyte in a sample that can be consistently detected with a stated probability (typically near 90%), was then determined. It was found that the 32 variants with allele frequencies near 0.1% were detected with 85% probability across replicates. The LOQ of the assay, commonly defined as the lowest standard concentration that can be quantified with a % coefficient of variation (CV) value below a certain threshold (typically ~20–30%), was then determined. It was found that the 28 variants with allele frequencies near 0.5% had % CV of 18%. Therefore, the LOD and LOQ have been reported to be near ~0.1% and ~0.5%, respectively.

Matched whole blood or PBMC samples were sequenced to subtract germline and CHIP mutations. A total of 300 µl of sample was extracted using KingFisher platform (Thermo Fisher Scientific). The genomic DNA was sheared by ultrasonication to generate approximately 200-bp fragments (Covaris). Postshearing, 200 ng was used with the same protocol as the cfDNA samples mentioned above.

The training and test sets were initially chosen based on the sequencing batch for the set of 566 patients chosen for the ctDNA substudy (see Fig. 1a), where we put sequencing batch 1 in the training set and then added in patients from later batches to reach the target 50%/50% split. The sequencing lab decided which samples to include in batch 1 without any knowledge about the baseline characteristics, treatment or clinical outcomes of the patients. We then checked for imbalances, and it was found that RACE was not well distributed due to all Asian patients appearing in batch 1, and so we moved half of the Asian patients to the test set and replaced these spots in the training data with a random set of patients. As the analysis progressed (in the training subset of data), we decided to add in PBMC correction due to concern over germline/CHIP variants contaminating the ctDNA dataset, which reduced the number of patients to those with PBMC available for correction, giving a final *n* of 466 patients and a final split of 240/226 patients for train/test. The final training/test sets were well balanced in clinical features and survival outcomes as can be seen in Supplementary Table 2 and Extended Data Fig. 1f.

Variant calling

Sequence data for the baseline plasma samples were processed by a cfDNA computational pipeline that corrected errors via the use of fragment barcodes previously described[35]. Short variants called by the bTMB assay were evaluated.

Sequence data processing and variant calling for on-treatment plasma and matched whole blood or PBMC samples were performed similarly to methods previously described[46]. In brief, reads were demultiplexed and fragment barcodes used to reduce errors, deduplicated and merged into consensus reads representing all information in the set of reads for each fragment. The consensus reads were then aligned to the reference genome.

For sensitively estimating on-treatment VAF, aligned consensus reads were postprocessed using proprietary software also developed by FMI. Each baseline variant was left- and right-justified to determine the locus of all possible overlapping reads. Those reads were then re-aligned to both the reference genome and the reference genome as modified by the presence of the variant. Striped Smith–Waterman alignment was used to score each of those two alignments, classifying each consensus read as either

- Supporting the presence of the variant (NRv)
- Supporting the absence of the variant (NRr)
- Has no discriminating value (equivocal) (NRe)

Equivocal, duplicate and low mapping quality reads were ignored. VAF was estimated as NRv/(NRv + NRr).

To further refine the VAF, we assigned a classification by comparing it with the distribution of VAFs using the same method on a set of presumed normal samples that is samples whose baseline genomic profile had no nearby variant calls. The VAF classifications are:

- Positive—at least 2 variant supporting reads and VAF > maximum VAF found in presumed normal samples
- Negative—VAF < 0.95 quantiles of presumed normal VAFs
- Equivocal—neither of the above conditions is met.

Variants were called and classified in matched PBMCs using the same methodology described for ctDNA variant calling. Alterations that were classified as positive in PBMCs were presumed to be germline/CHIP and excluded from analyses.

Statistics

To remove immortality bias[47] when assessing the correlation between on-treatment measurements and PFS/OS, patients with events before the collection date were excluded from the analyses and PFS and OS time were recalculated from the on-treatment sample collection date. When utilizing both week 6 tumor response assessments and C3D1 ctDNA metrics, PFS/OS time was recalculated from the C3D1 date.

Landmark PFS and OS were compared between groups using a univariable Cox proportional-hazards model to estimate HR and log-rank test to report *P* values. Cox models used the 'exact' method for handling tied event times. *P* values reported in forest plots for multivariable Cox regression are using two-sided Wald test. The strength of the association between event time and a continuous predictor was measured by Harrell's c-index, which indicates the overall rank concordance between event time and the predictor. Standard errors for the c-index were computed by assuming asymptotic normality[48] and their *P* values test if an estimate is different from 0.5[49]. Two C indices were compared based on a U-statistic to test for whether one predictor is more concordant with the outcome than another (R package version 4.7-0. https://CRAN.R-project.org/package=Hmisc).

Descriptive statistics were used to summarize clinical characteristics and ctDNA metrics, including the mean, median and range for continuous variables and frequency and percentage for categorical variables. Association between ctDNA positivity and baseline prognostic factors was measured using a two-sided Wilcoxon Rank Sum test for numeric variables and Fisher's Exact test (two-sided) for categorical variables. The association between continuous ctDNA metrics and radiographic treatment response categories was measured using a two-sided Wilcoxon rank sum test or Kruskal–Wallis rank sum test. Correlations between two continuous metrics depicted in scatterplots reported Pearson's correlation.

Unless otherwise noted, all analyses combined patients across the three study arms and reported *P* values were two-sided and unadjusted for multiplicity or covariates. All statistical analyses were performed in R version 3.6 (https://www.R-project.org/). *P* values were reported for descriptive purposes and were unadjusted for multiple hypothesis testing.

ctDNA feature derivation for predictive modeling in IMpower150 training/test and OAK validation data

Every ctDNA mutation had an associated AF reported by the assay at each time point, and in this study, the assay limit of quantitation (LOQ) and LOD were determined to be 0.5% and 0.1%, respectively (Extended Data Fig. 1b). Reported mutations with AF below the LOQ were censored to LOQ/2, and reported mutations with AF below the LOD were censored to LOQ/4.

The ctDNA analysis plan for the machine learning model was finalized before the development of the model. ctDNA levels (AF, MTM, AUC, etc.) were quantified using 23 different metrics measured for each time point (BL, C2D1, C3D1, C4D1 and C8D1), and the change in ctDNA relative to baseline was quantified using 55 different metrics for each on-treatment time point (C2D1, C3D1, C4D1 and C8D1).

Additional feature processing before running the ML model included handling missingness and interquartile range (IQR) normalization. If a patient in the ctDNA evaluable population has a record of a blood sample collection for a given visit with an associated date for that sample collection, then the patient is included in the landmark analysis for that visit. However, if the ctDNA data are missing despite the record of a blood sample collection (for example due to failing ctDNA assay QC), then ctDNA features were imputed using the population median of the feature for that visit. We considered this imputation for patients with ctDNA collected but QC-failing samples to be important because models included ctDNA features from multiple time points (that is C3D1 OS run included BL and C2D1 features as well) and so theoretically a patient with a QC-failing sample for C3D1 still may have C2 and/or BL ctDNA data informative for predicting survival time from C3D1. Final sample counts for each visit time point can be found in Extended Data Fig. 1g. Individual features were scaled by the IQR of that feature before running the machine learning model.

A complete list of ctDNA features can be found in Supplementary Table 3, along with the rank concordance (c-index) of each metric with landmark OS and PFS for each visit.

To test the validity of our ctDNA C3D1 OS model in an external cohort, we leveraged the availability of ctDNA data for $n = 73$ patients from the OAK clinical trial (NCT02008227). A continuous predictor was derived from the 5 ctDNA metrics measured by the Avenio panel and their coefficients which were used in the final C3D1 OS model. Note that for the feature 'Number of pathogenic mutations detected at C3D1', we considered Avenio mutations to be pathogenic when they were both nonsilent and present in COSMIC database (https://cancer.sanger.ac.uk/cosmic). OAK ctDNA features were processed as detailed for IMpower150 above, including censoring of small values, imputation for missing data and IQR normalization. Please note the censoring of small values occurred before feature derivation and used the same approach as described above for IMpower150 in which the LOD and LOQ were considered to be 0.1% and 0.5%, respectively. We applied the same thresholds identified in IMpower150 training data to identify high-risk mPD patients, low-risk mResp patients and intermediate-risk mSD patients.

Training the ML model and choosing thresholds for mPD, mSD and mResp

At a visit, all ctDNA measurements collected from baseline up to the particular visit, among patients who are still at risk for PFS/OS, were used to associate with the rebaselined endpoint. All modeling was repeated with leave-one-out-cross-validation (LOOCV). Linear combinations of individual features were associated with landmark PFS and OS in an elastic network (R package glmnet v3.0-2) with an equal weight of lasso penalty and ridge penalty at each visit (alpha = 0.5). The optimal lasso penalty (lambda) was chosen for each LOOCV where nested cross-validation was repeated 10 times and the average of the lambda that minimizes the prediction error was used. Feature interaction was addressed in survival random forest modeling (R package

randomForestSRC v 2.9.3) with default parameters, but no improvements in model performance were detected. Feature importance was assessed as the number of cross-validations a feature was retained, and by the 'Gain' metric as assessed by the average worsening statistic from the next-door analysis[40] across LOOCV. Model performance, measured by c-index, was estimated after pooling all LOOCV predictions together to reconstruct the original training dataset. The final time point and endpoint were chosen to be C3D1 and OS due to numerically superior performance by c-index during LOOCV.

For model runs utilizing baseline clinical factors, the features included the following: ECOG score (0 or 1), age (continuous metric), number of metastatic sites (continuous metric), sex (M/F), history of tobacco use (y/n), PD-L1 high status (y/n) and SLD (of target lesions from radiographic assessment, a continuous metric). For the C3D1 model runs, we also included the week 6 radiographic tumor assessment data available, including week 6 SLD (continuous metric), difference in SLD between baseline and week 6 (continuous metric), and percent change in SLD between baseline and week 6 (continuous metric).

The final top features included in the ctDNA C3D1 OS model were features that were chosen in at least 50% of CV and with positive gain metric (5 features total) and can be found in Supplementary Table 5. The final C3D1 OS model was fit in the entire training set using these five features and coefficients can be found in Supplementary Table 6.

The threshold for the high-risk (mPD) group was chosen by visualizing different splits of the C3D1 OS model predictions for patients with week 6 SD and PR separately, choosing the optimal split within each and then taking the mean (Extended Data Fig. 4d) in the training dataset which corresponded to a numeric value of 0.298. The threshold for low-risk ctDNA responders (mResp) was chosen by finding the 32% percentile of the prediction scores, which corresponds to the proportion of patients who achieved durable (3 years) OS (Extended Data Fig. 4e). This 32% quantile of C3D1 OS model predictions corresponded to a numeric value of 0.036 in the training dataset (Extended Data Fig. 4e).

Simulation of operation characteristics

To assess the utility of the ctDNA model in early clinical decision-making, with and without the radiographic endpoints, we performed operational characteristics analyses in simulated randomized phase 2 studies[42].

Two routine endpoints used in early clinical developments are PFS and tumor response as assessed by the investigator according to RECIST criteria version 1.1. PFS is the time from randomization until tumor progression or death. We note that in phase 3 IMpower150 study patients have ~39.8 months of median follow-up; thus, these two endpoints are mature. In contrast, for this simulation, we are interested in early ctDNA and early PFS or tumor response signals observed within the first ~6 weeks of treatment initiation and whether these early endpoints can predict the outcome of the clinical trial (in the case of IMpower150, superior OS of treatment ABCP versus control BCP, or of treatment ACP versus control BCP).

The operation characteristics were assessed as follows: after running the final C3D1 OS model in the test dataset to obtain ctDNA model predictions, patients with predictions below the predefined threshold for mResp were identified. To characterize true go rates, we sampled 30 random patients (n = 2,000 simulations) from active arms (ABCP/ACP) arm and control (BCP) arm with replacement, which mimics those developmental settings where a Go decision is favorable. To characterize False Go Rates, we sampled two sets of 30 patients from the control arm, with one set as the standard of care treatment and the other as the new treatment, which mimics those development settings where a No-Go decision is favorable. In each simulated study, we compared the number of mResp patients between the treatment and control arms (Fisher's exact test), as well as the distribution of PFS times (log-rank test), or number of radiographic response by RECIST (Fisher's exact test). All P values were one-sided. When combining a ctDNA criterion with a RECIST criterion, the smaller one of the two P values is used for Go/No-Go decisions. For a single metric, the 15% (the desired False Go Rate) percentile of P values for an arm BCP versus BCP comparison was found (n = 2,000 simulations) and was considered the cutoff value for a Go decision for the arm ABCP/ACP versus BCP comparison. If the P value for ABCP-versus-BCP comparison is less than the cutoff value, then a Go decision is made, and the true go rate is the proportion of 2,000 P values less than the cutoff value. Under instantaneous enrollment scenario, all patients have the same length of follow. Under ramp-up enrollment, the clinical cutoff date is 14 days (expected ctDNA assay turnaround time) plus the C3D1 day of the last enrolled patient in a cohort of random samples. For simulation purposes, when selecting 30 patients, the actual enrollment of the chosen 30 must be within 12 months of each other.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All clinical and ctDNA data for IMpower150 are deposited to the European Genome-Phenome Archive under accession number EGAS00001006703. Qualified researchers may request access to individual patient-level data through the clinical study data request platform (https://vivli.org/). Further details on Roche's criteria for eligible studies are available at https://vivli.org/members/ourmembers. For further details on Roche's Global Policy on the Sharing of Clinical Information and how to request access to related clinical study documents, see https://www.roche.com/research_and_development/who_we_are_how_we_work/clinical_trials/our_commitment_to_data_sharing.htm.

Code availability

The documented code for the R statistical computing environment for analyses related to ctDNA in IMpower150 is deposited to the European Genome-Phenome Archive under accession number EGAS00001006703.

References

45. Clark, T. A. et al. Analytical validation of a hybrid capture-based next-generation sequencing clinical assay for genomic profiling of cell-free circulating tumor DNA. *J. Mol. Diagn.* **20**, 686–702 (2018).
46. Woodhouse, R. et al. Clinical and analytical validation of FoundationOne Liquid CDx, a novel 324-Gene cfDNA-based comprehensive genomic profiling assay for cancers of solid tumor origin. *PLoS ONE* **15**, e0237802 (2020).
47. Beyersmann, J., Gastmeier, P., Wolkewitz, M. & Schumacher, M. An easy mathematical proof showed that time-dependent bias inevitably leads to biased effect estimation. *J. Clin. Epidemiol.* **61**, 1216–1221 (2008).
48. Penciana, M. J. & D'Agostino, R. B. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat. Med.* **23**, 2109–2123 (2004).
49. Schröder, M. S., Culhane, A. C., Quackenbush, J. & Haibe-Kains, B. survcomp: an R/bioconductor package for performance assessment and comparison of survival models. *Bioinformatics* **27**, 3206–3208 (2011).

Acknowledgements

The study is sponsored by F. Hoffmann-La Roche Ltd. and Genentech, Inc. We thank the patients and their families who participated in the trial and the clinical site investigators. We would also like to

acknowledge several individuals for medical writing assistance for this manuscript including J. Fioretti of Health Interactions, H. Debnath of Anshin Biosolutions and A. Bruce of A.K. Bruce Design. This study was funded by F. Hoffmann-La Roche Ltd and Genentech, Inc., a member of the Roche Group.

Author contributions

D.S., K.S., M.L., W.Z., Z.J.A., A.Y., D.L., C.C., D.F., M.S., and M.R. contributed to project concept, design, and/or methodological approach. Z.J.A., D.S., K.S., W.Z., A.Y., A.F., J.F.F., M.K., D.L., G.O., D.F., N.P., and M.S. contributed to data analysis and interpretation. M.N., J.F.F., A.R., M.S., M.R., and E.P. contributed to provision of study materials or patients. A.F., J.F.F., and E.P. contributed to collection and assembly of data. Z.J.A., D.S., M.L., A.Y., M.N., A.F., J.F.F., C.C., G.O., W.Z., D.F., K.S., N.P., M.S., M.R., and E.P. contributed to manuscript writing, review and revision. Z.J.A., D.S., M.L., A.Y., M.N., A.F., J.F.F., D.L., C.C., G.O., W.Z., D.F., K.S., N.P., M.S., M.R., and M.K. contributed to the final approval of the manuscript.

Competing interests

Z.J.A., D.S., K.S., C.C., Z.J.A., A.R., M.L., N.P., and W.Z. disclose current or recent employment with Roche. Z.J.A., D.S., K.S., G.O., C.C., M.L., D.F., N.P., W.Z., A.F., D.L., A.R., A.Y., and J.F. disclose stock or other ownership interests with Roche. G.O., A.Y., A.F., D.F., D.L., M.K., E.P., and J.F. disclose current or recent employment with FMI. M.R. discloses funding from Genentech, Pfizer, Spectrum, Takeda, Daiichi Sankyo, AstraZeneca, and Speaker Bureau for Genentech, AstraZeneca, Guardant, Jazz, Janssen and GI Therapeutics. M.S. discloses funding from Genentech, Pfizer, Spectrum, Takeda, Novartis, Beigene, AstraZeneca, and Daiichi Sankyo and discloses speaking fees from Genentech, Lilly, Blueprint, Guardant, BMS, Jazz, GI Therapeutics, Janssen and Amgen. M.L. discloses board member and shareholder in Foresight Diagnostics, scientific advisory board member and shareholder in Delfi Diagnostics and Prognomiq, Inc. The other authors declare no competing interests.

Additional information

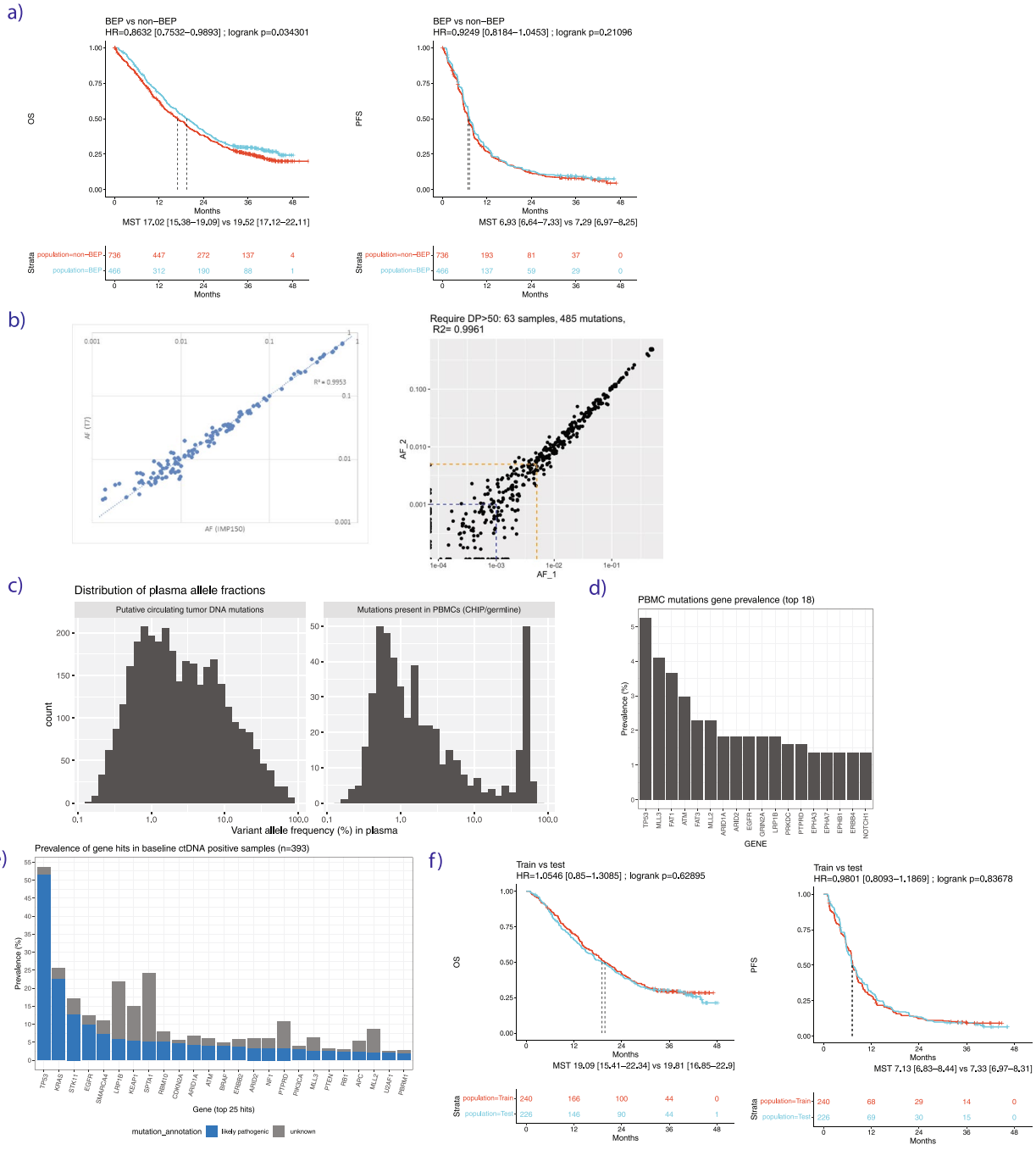
Extended data is available for this paper at https://doi.org/10.1038/s41591-023-02226-6.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41591-023-02226-6.

Correspondence and requests for materials should be addressed to Zoe June F. Assaf, David S. Shames or Katja Schulze.

Peer review information *Nature Medicine* thanks Tero Aittokallio, Keunchil Park, and Dominic Rothwell for their contribution to the peer review of this work. Primary Handling Editor: Ulrike Harjes, in collaboration with the *Nature Medicine* team.

https://doi.org/10.1038/s41591-023-02226-6



g)

TIME POINT for Landmark Analysis	BL	C2D1	C3D1	C4D1	C8D1
Plasma collected (466 patients)	466	458	398	386	289
Training subset (240 patients)	240	237	206	201	146
ctDNA data pass QC	240	227	200	194	140
Testing subset (226 patients)	226	221	192	185	143
ctDNA data pass QC	226	219	188	182	138

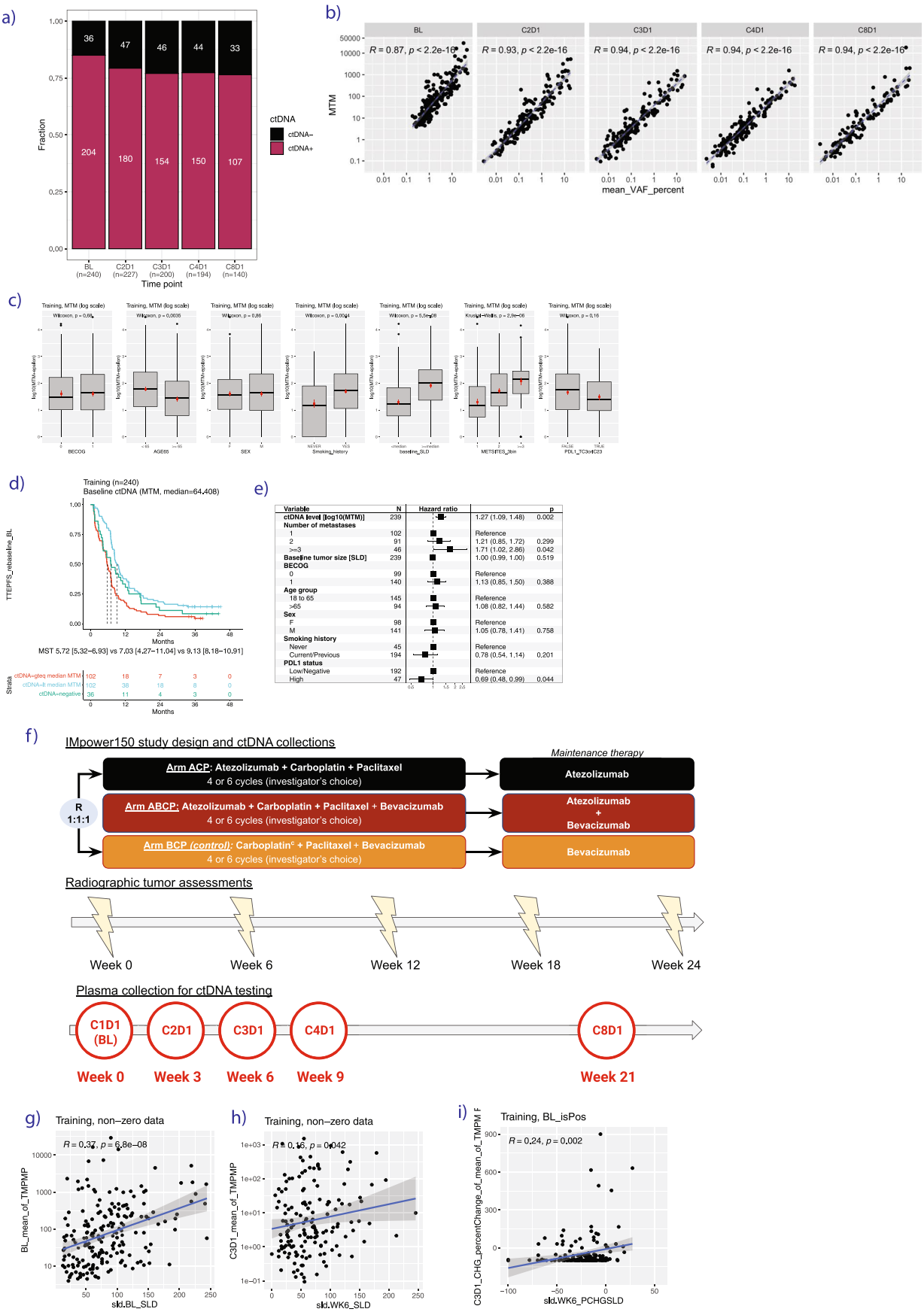
C3D1/Week 6 Tumor Assessment	CR	PR	SD	PD	N/A
C3D1 Plasma collected (398 patients)	1	144	242	9	2
Training subset (206 patients)	1	69	127	7	2
ctDNA data pass QC (n=200)	1	66	124	7	2
Test subset (192 patients)	0	75	115	2	0
ctDNA data pass QC (n=188)	0	75	111	2	0

Extended Data Fig. 1 | See next page for caption.

https://doi.org/10.1038/s41591-023-02226-6

Extended Data Fig. 1 | (a) KM curves showing OS (left) and PFS (right) for IMpower150 patients in the ctDNA biomarker evaluable population (BEP, blue) versus the ctDNA non-biomarker-evaluable-population (non-BEP, red). (b) Quality control experiments to show (left panel) high concordance of 330kb custom assay ('IMP150') compared to larger 1.25Mb assay ('T7'), and to show (right panel) high reproducibility and sensitivity of 63 samples run in replicate on the 330kb custom assay where the LOD of the assay is found to be near 0.1% (where 85% of mutations near this frequency are detected reproducibly, blue dashed line) and the LOQ of the assay is near 0.5% (where the % CV of mutations near this frequency is 18%, orange dashed line). **(c)** Histogram of variant allele frequencies (%) for mutations identified using the custom 330kb panel, showing mutations present in plasma cell-free DNA and absent from PBMCs (left), and for mutations identified in plasma cell-free DNA and present in PBMCs (right). **(d)** Bar plot showing the genes in which PBMC-derived mutations (CHIP/germline) were most prevalent (*y* axis, percent of patients). PBMC-derived mutations are defined as those which were identified in both cell-free DNA and PBMCs for genes included in the custom 330kb panel. **(e)** Bar plot showing the genes in which tumor-derived mutations were most prevalent (*y* axis, percent of patients). Mutations that are known or likely pathogenic alterations (blue) are delineated from those which are variants of unknown significance (gray). Tumor-derived mutations are defined as those detected in cell-free DNA and absent from PBMCs for genes included in the custom 330kb panel. **(f)** KM curves showing OS (left) and PFS (right) for patients in the ctDNA biomarker evaluable population in the training split (blue) versus the test split of data (red). **(g)** Table showing the number of plasma samples collected at each time point, including a breakdown of the number in the training and test subsets which passed ctDNA assay QC, which were used for model training and testing. The bottom table shows number of patients who had C3D1 plasma samples which passed ctDNA assay QC and also treatment response assessments available for week 6 tumor assessment.

https://doi.org/10.1038/s41591-023-02226-6



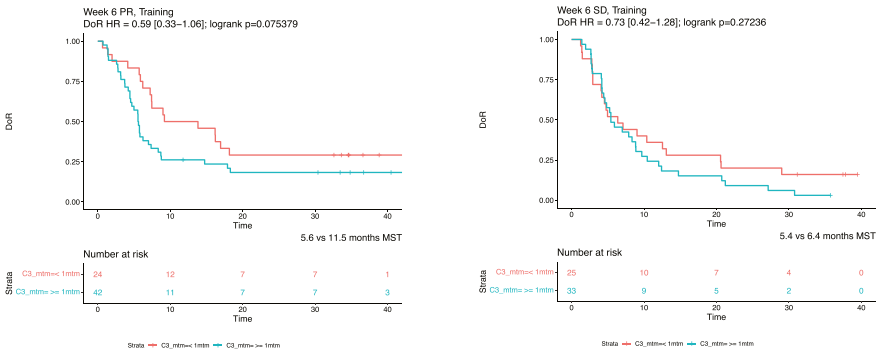
Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41591-023-02226-6

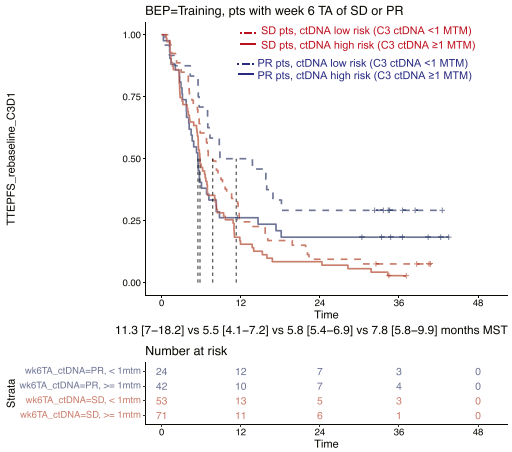
Extended Data Fig. 2 | (a) Bar plot of number of ctDNA positive and negative samples at each time point in training. (b) Scatterplots showing correlation between mean variant allele frequencies versus mean tumor molecules per ml plasma in training. The baseline timepoint, in addition to having higher patient ctDNA tumor fractions due to occurring prior to treatment initiation, was sequenced with a 1.25Mb assay with reportable VAF range down to ~0.5%. On-treatment time points were sequenced with a 330kb assay with a reportable range down to ~0.01%, and restricted to only mutations detected at baseline. Pearson's correlation coefficient is reported and its *P* value based on Pearson's product moment correlation. **(c)** Boxplots showing association between baseline clinical features (6 panels, one for each feature) and ctDNA levels as measured by MTM (*y* axis) in training, where *P* values reported using a two-sided Wilcoxon rank sum test. The box plots depict the median at the middle line, with the lower and upper hinges at the first and third quartiles, respectively, the whiskers showing the minima to maxima no greater than 1.5× the interquartile range, and the remaining outlying data points plotted individually. Additionally, the mean and standard error are overlayed as red points. Sample sizes for the box plots from left to right are n = 99, 140; 133, 107; 98, 142; 45, 195; 120, 120; 102, 92, 46; 177, 63. **(d)** KM curve showing the prognostic value of baseline ctDNA MTM levels for PFS in training data. **(e)** Multivariable cox regression for PFS in training data. Two-sided Wald test *P* values are reported, and points and error bars indicate HR and 95% confidence interval, respectively. **(f)** Study schema showing when radiographic and plasma collections were performed in the treatment course. (g-h) Scatterplots showing association between radiographic assessment of tumor size by SLD measurement (*x* axis) versus ctDNA levels measure by MTM (*y* axis) for **(g)** baseline time point, **(h)** C3D1 time point, **(i)** change from BL to C3D1. Plots restrict to patients with ctDNA detected at baseline. Error band indicates 95% confidence interval. Pearson's correlation coefficient is reported and its *P* value based on Pearson's product moment correlation.

https://doi.org/10.1038/s41591-023-02226-6

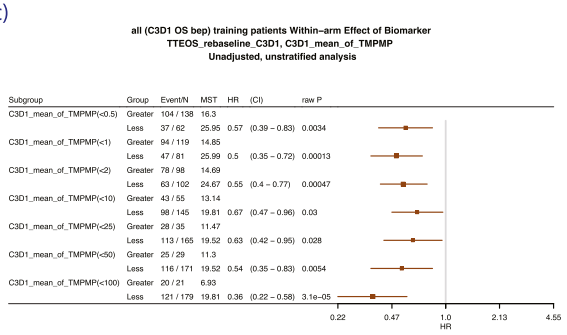
a)



b)



c)

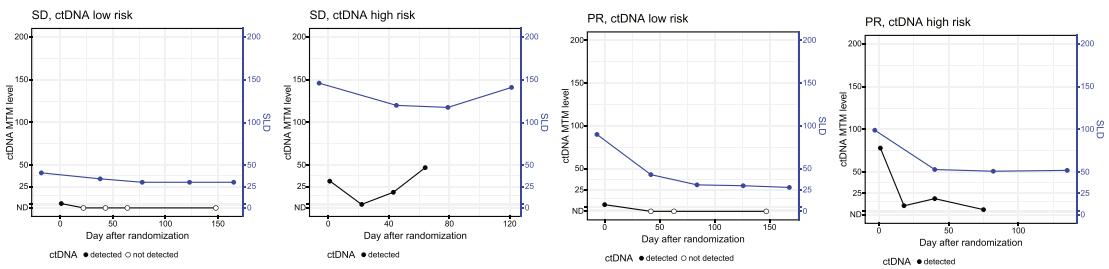


d)

Progostic value of different ctDNA metrics (landmarked C3D1)

RespGrp	Biomarker (bm)	BEP	N	outcome	Grp1_ID	Evnts/Pts	MST [95CI]	Grp2_ID	Evnts/Pts	MST [95CI]	Hazard Ratio [95CI]
PR	mean_of_AF	all	66	OS	bm=gt median	21/33	15.5 [11.8–NR]	bm=leq median	19/33	27.4 [16.8–NR]	0.75 [0.4–1.4]
PR	mean_of_TMPMP	all	66	OS	bm=gt median	22/33	15.3 [9.6–27.5]	bm=leq median	18/33	28.4 [19.1–NR]	0.62 [0.33–1.16]
PR	logRatio_of_mean_of_AF	BL_ctDNApos	60	OS	bm=gt median	17/30	20.1 [11.8–NR]	bm=leq median	20/30	19.4 [11.8–NR]	1.14 [0.6–2.18]
PR	logRatio_of_mean_of_TMPMP	BL_ctDNApos	60	OS	bm=gt median	17/30	22.2 [12.5–NR]	bm=leq median	20/30	18.2 [11.7–NR]	1.22 [0.64–2.33]
PR	mean_of_AF	all	66	PFS	bm=gt median	25/33	5.5 [4.2–8.3]	bm=leq median	26/33	7.2 [5.5–13.8]	0.91 [0.53–1.59]
PR	mean_of_TMPMP	all	66	PFS	bm=gt median	26/33	5.5 [4.1–8.3]	bm=leq median	25/33	7.3 [5.7–15.8]	0.76 [0.44–1.32]
PR	logRatio_of_mean_of_AF	BL_ctDNApos	60	PFS	bm=gt median	23/30	5.9 [4.9–14.6]	bm=leq median	24/30	6.2 [4.5–13.8]	1.06 [0.6–1.88]
PR	logRatio_of_mean_of_TMPMP	BL_ctDNApos	60	PFS	bm=gt median	23/30	6.5 [4.9–14.6]	bm=leq median	24/30	5.6 [4.5–8.8]	1.13 [0.63–2]
SD	mean_of_AF	all	124	OS	bm=gt median	51/62	15.3 [11.5–20.9]	bm=leq median	43/62	21.6 [17.1–24.7]	0.62 [0.41–0.93]
SD	mean_of_TMPMP	all	124	OS	bm=gt median	53/62	15.2 [11.3–20.7]	bm=leq median	41/62	22.6 [17.2–29.1]	0.54 [0.36–0.81]
SD	logRatio_of_mean_of_AF	BL_ctDNApos	102	OS	bm=gt median	42/51	16.8 [9.5–20.9]	bm=leq median	36/51	18.8 [15.2–23.6]	0.66 [0.42–1.02]
SD	logRatio_of_mean_of_TMPMP	BL_ctDNApos	102	OS	bm=gt median	41/51	17.1 [9.9–21.5]	bm=leq median	37/51	18.1 [14.8–22.6]	0.8 [0.51–1.24]
SD	mean_of_AF	all	124	PFS	bm=gt median	60/62	5.7 [4.9–6.7]	bm=leq median	58/62	8.2 [5.8–9.7]	0.71 [0.49–1.02]
SD	mean_of_TMPMP	all	124	PFS	bm=gt median	60/62	5.8 [5.4–6.9]	bm=leq median	58/62	7 [5.8–9.6]	0.71 [0.49–1.03]
SD	logRatio_of_mean_of_AF	BL_ctDNApos	102	PFS	bm=gt median	49/51	5.6 [4.3–6.9]	bm=leq median	48/51	7.8 [5.8–9.9]	0.7 [0.47–1.04]
SD	logRatio_of_mean_of_TMPMP	BL_ctDNApos	102	PFS	bm=gt median	49/51	5.7 [5.3–6.9]	bm=leq median	48/51	6.9 [5.8–9.7]	0.79 [0.53–1.18]

e)

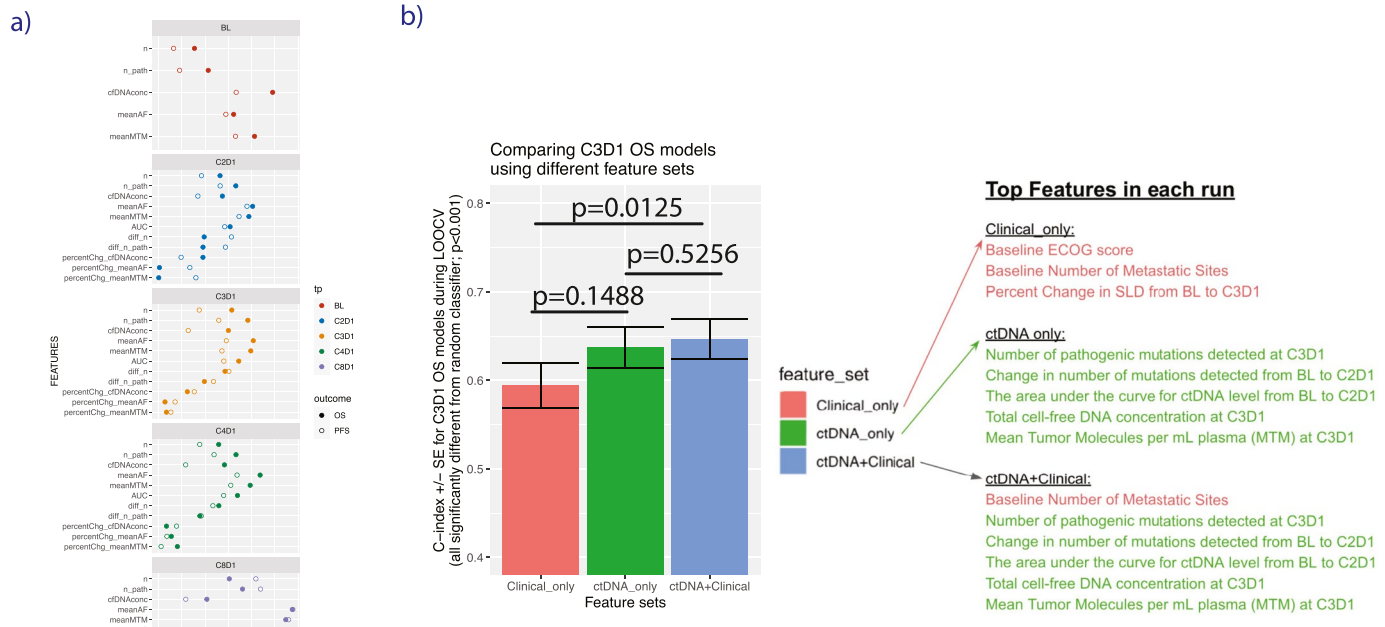


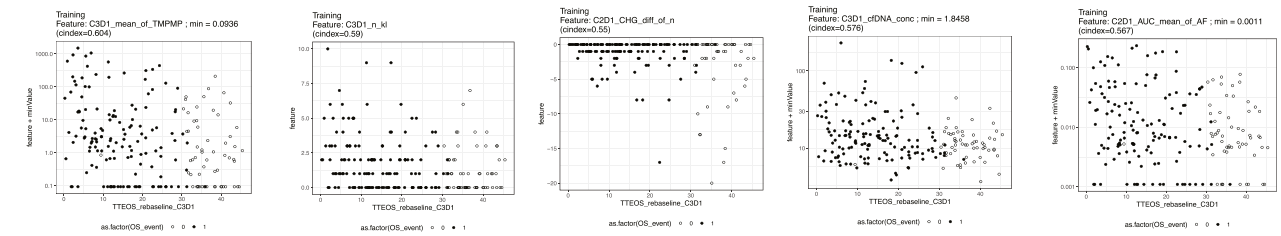
Extended Data Fig. 3 | See next page for caption.

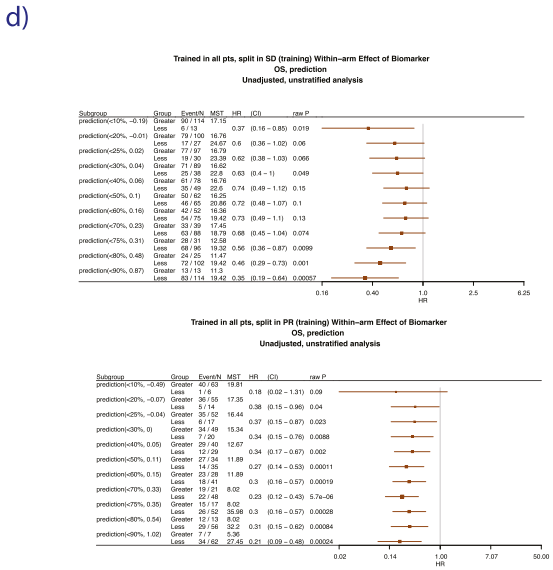
https://doi.org/10.1038/s41591-023-02226-6

Extended Data Fig. 3 | (a) KM analysis for duration of treatment response (DoR) in patients with PR (left) or SD (right) at week 6 tumor assessment who are risk stratified using ctDNA levels above or below 1 MTM, in training. **(b)** KM analysis for PFS in patients with SD or PR at week 6 tumor assessment who are risk stratified using ctDNA levels above or below 1 MTM, in training. **(c)** Forest plot showing prognostic value of other thresholds of MTM splits at C3D1 timepoint for risk stratification for OS in entire training dataset. Note that here MTM is labeled mean_of_TMPMP (for mean tumor molecules per ml plasma). HRs are comparing patients with MTM level below ('Less') versus above ('Greater') each threshold for splitting C3D1 MTM, where the number of patients can be found in the third column ('N'). MST indicates median survival time. Points and error bars indicate HR and 95% confidence interval, respectively. Univariable Cox proportional-hazards model was used to estimate HR and logrank test to report *P* values. **(d)** Forest plot showing prognostic value of other ctDNA metrics for OS and PFS in PR and SD patients in training. Note that here MTM is labeled mean_of_TMPMP (for mean tumor molecules per ml plasma). HRs are comparing patients with feature values ≤ versus > than the median value for that feature. RespGrp column indicates whether the subset for the risk stratification analysis is the PR or SD patients. BEP column indicates 'biomarker evaluable population', meaning the subset of patients included in the analysis, which is either 'all' patients (for features summarizing ctDNA levels), or for patients who are ctDNA positive at the baseline time point ('BL_ctDNApos', for features summarizing ctDNA change). Outcome column indicates if the HR is for OS or PFS. MST indicates median survival time. Points and error bars indicate HR and 95% confidence interval, respectively. **(e)** Four example patient time courses showing longitudinal ctDNA MTM level and tumor size by SLD for 4 example patients.

https://doi.org/10.1038/s41591-023-02226-6







e)

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
6	176	30	0.854	0.0246	0.808	0.904
12	139	36	0.679	0.0326	0.618	0.746
18	108	31	0.528	0.0349	0.464	0.601
24	83	25	0.405	0.0343	0.344	0.479
30	66	17	0.322	0.0327	0.264	0.393
36	32	5	0.293	0.0324	0.236	0.364
42	12	0	0.293	0.0324	0.236	0.364

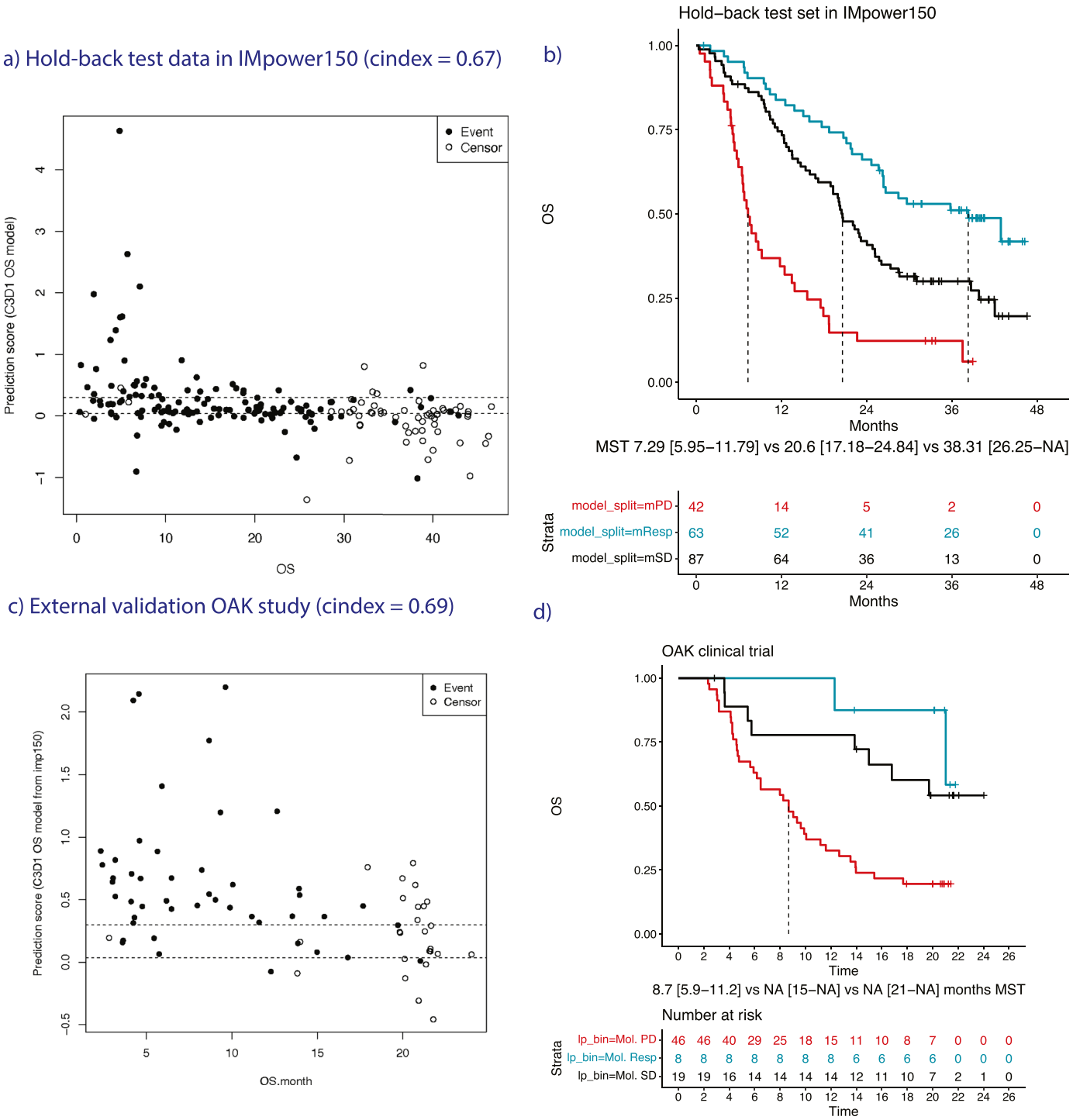
	OS_bin	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
OS < 30mo		-0.7565173	0.04020225	0.15322778	0.35864832	0.5386004	2.8987659
OS >= 30mo		-1.5566374	-0.19240854	0.03650846	-0.06797818	0.1555975	0.6394948

Extended Data Fig. 4 | See next page for caption.

https://doi.org/10.1038/s41591-023-02226-6

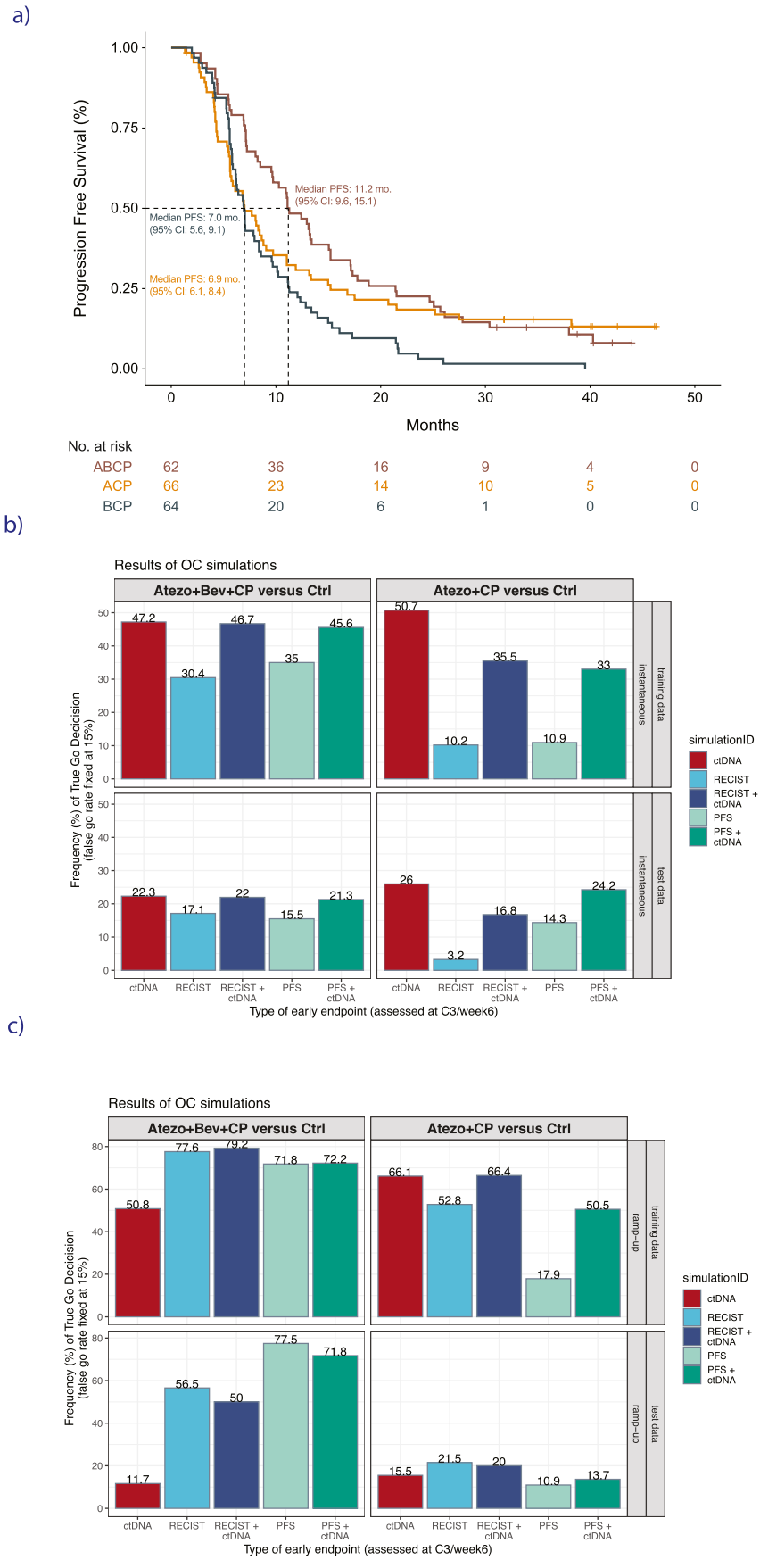
Extended Data Fig. 4 | (a) Scatterplots showing the univariable rank concordance (c-index, *x* axis) for each individual ctDNA feature for landmarked OS and PFS estimated at each time point (panels), in training. 'n' indicates number of detected variants', 'n_path' indicates number of detected known/likely pathogenic variants, 'percChg' and 'diff' indicate percent change and difference in ctDNA level from baseline. **(b)** Comparison of models trained using either clinical features alone (red), ctDNA features alone (green), or ctDNA+Clinical features (blue), with annotation as to which metrics were top features in each run. The bar height is rank concordance (c-index) calculated from leave-one-out-cross-validation (LOOCV) to fit an elastic net model, error bars are standard error of the c-index, *P* values are two-sided and based on a U-statistic to compare two predictors. Models were built using $n = 206$ patients in the training subset at-risk for an OS event at C3D1. **(c)** Scatterplots for the 5 top features from C3D1 OS model, showing the association between each feature value with landmark OS. **(d)** Forest plots showing prognostic value of C3D1 OS ctDNA model predictions in training data for patients with SD (top forest plot) and Partial Response (bottom forest plot), where the number of patients can be found in the third column ('N'), 'MST' indicates median survival time. Points and error bars indicate HR and 95% confidence interval, respectively. Univariable Cox proportional-hazards model was used to estimate HR and logrank test to report *P* values. Note that the threshold chosen for categorizing a patient as having molecular progressive disease (mPD) was done by taking the mean of the optimal split in SD patients (75th percentile, top forest plot) and the optimal split in PR patients (70th percentile, bottom forest plot). **(e)** Choosing a threshold in training data of C3D1 OS ctDNA model predictions for categorizing a patient as having mResp was done by identifying the patients in training data who achieved durable OS of ≥ 30 months (which was 32.2% of the population, see top table), and then taking the median prediction score of this population (which was 0.036, see bottom table) in training data.

https://doi.org/10.1038/s41591-023-02226-6



Extended Data Fig. 5 | (a) Scatterplot showing the final C3D1 OS ctDNA model predictions (*y* axis) versus OS time (*x* axis) in the hold-back test data for IMpower150 (c-index, 0.67). Dotted lines show thresholds for mPD (≥ 0.298 prediction score), mResp (< 0.036 prediction score), and mSD (for [0.036, 0.298) prediction scores), which were thresholds chosen in the training set of data. **(b)** KM curves for OS in hold-back test set showing the final subgroups identified using the C3D1 OS model prediction thresholds chosen in training data. Subgroups include mPD (red line), mResp (blue line), and mSD (black line), all confirmed to have prognostic value in this test data. **(c)** Scatterplot showing the final C3D1 OS ctDNA model predictions (*y* axis) versus OS time (*x* axis) in the external validation OAK cohort of 73 patients (c-index, 0.69). Dotted lines show thresholds for mPD (≥ 0.298 prediction score), mResp (< 0.036 prediction score), and mSD (for [0.036, 0.298) prediction scores), which were thresholds chosen in the training set of data. **(d)** KM curves for OS in external validation OAK cohort of 73 patients showing the final subgroups identified using the C3D1 OS model prediction thresholds chosen in training data. Subgroups include mPD (red line), mResp (blue line), and mSD (black line), all confirmed to have prognostic value in this external validation data.

https://doi.org/10.1038/s41591-023-02226-6



Extended Data Fig. 6 | See next page for caption.

https://doi.org/10.1038/s41591-023-02226-6

Extended Data Fig. 6 | (a) KM curve showing PFS in the test dataset for the three arms in the IMpower150 trial including ABCP (brown), ACP (orange), and control arm BCP (black, control arm). (**b**–**c**) Complete results of operation characteristics simulations showing the rate of true 'Go' decisions in **(b)** instantaneous enrollment scenario (every patient has their clinical data cut at their respective C3D1 time point), versus **(c)** ramp-up enrollment scenario (use all clinical data available for patient after last patient enrolls, so some patients could have additional radiographic data available after the week 6 time point). Training data shown in top rows, test data shown in bottom rows. The early endpoint is either ctDNA criteria alone (red bar), RECIST criteria alone (light blue) or RECIST criteria combined with ctDNA (dark blue), PFS alone (light green bar) or PFS combined with ctDNA criteria (dark green bar).

nature portfolio

Corresponding author(s): Zoe June Assaf
David Shames
Katja Schulze

Last updated by author(s): 01/11/2023

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	No software used for data collection
Data analysis	All statistical analyses were performed in R (https://www.R-project.org/), version 3.6. R libraries included Hmisc (version 4.7), glmnet (v3.0-2), randomForestSRC (v 2.9.3). Code is deposited in a public repository (EGAS00001006703) and the link is in the Code Availability statement of the manuscript.

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

Public datasets that were used during data processing included ExAC database (http://exac.broadinstitute.org/) and IDs from COSMIC database (https://cancer.sanger.ac.uk/cosmic).

The data and code required to reproduce results are deposited to the European Genome-Phenome Archive under accession number EGAS00001006703, and can be made available upon request. Qualified researchers may request access to individual patient-level data through the clinical study data request platform (https://vivli.org/). Further details on Roche's criteria for eligible studies are available at https://vivli.org/members/ourmembers. For further details on Roche's Global Policy on the Sharing of Clinical Information and how to request access to related clinical study documents, see https://www.roche.com/research_and_development/who_we_are_how_we_work/clinical_trials/our_commitment_to_data_sharing.htm.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Gender was not collected in this study, and accordingly we do not use this term in the paper. Please refer to Extended Data Tables 1 and 2 for covariate-relevant population characteristics of the human research participants in the IMpower150 ctDNA study, which includes Sex.
Population characteristics	Patient metadata are included in the manuscript, including tables of population characteristics for both the full IMpower150 population and this retrospective exploratory ctDNA substudy reported in this manuscript (Supp Table ST1). Briefly, IMpower150 patients had stage IV or recurrent metastatic nonsquamous NSCLC for which they had not previously received chemotherapy, a baseline Eastern Cooperative Oncology Group (ECOG) performance-status score of 0 or 1, and tumor tissue available for biomarker testing and if they were eligible to receive bevacizumab; patients with any PD-L1 immunohistochemistry status were eligible. Please refer to Extended Data Tables 1 and 2 for covariate-relevant population characteristics of the human research participants in the IMpower150 ctDNA study, including baseline ECOG score, Age, Sex, Tobocco use, Race, Region, Number of metastatic sites, and PDL1 status.
Recruitment	In the IMpower150 study patients were eligible for inclusion in this retrospective exploratory ctDNA substudy reported in this manuscript if they had plasma samples available for ctDNA testing at both baseline and an early on-treatment timepoint (either cycle 2 day 1 or cycle 3 day 1), as well as PBMCs available in order to perform germline subtraction. This reduced the number of patients in this ctDNA substudy to 466. We expected a survivorship bias in this ctDNA-evaluable population due to our requirement for patients to have samples available after randomization, and while no strong PFS bias was found for ctDNA evaluable versus non-evaluable (HR=0.92 [0.82-1.05]), we did detect a survivorship bias for OS (HR=0.86 [0.75 – 0.99]) (Extended Data Figure ED1a). However, baseline characteristics were similar between the full IMpower150 and ctDNA evaluable population, including baseline ECOG, age, sex, race, region, among others (Supplementary Table ST1). We do not expect this OS survivorship to strongly impact our results, because this exploratory ctDNA substudy is concerned with building ctDNA-based models to predict overall survival, and typically patients who do not have evaluable on-treatment samples available are those who unfortunately had very rapid disease progression, and therefore for which a ctDNA model to predict survival is of limited utility.
Ethics oversight	The study was conducted in accordance with the International Conference on Harmonisation Good Clinical Practice guidelines and with the principles of the Declaration of Helsinki. All patients provided written informed consent, and the protocol was approved by independent ethics committees at each participating site. The study protocol is in the appendix of the primary clinical manuscript for IMpower150 (Socinski et al, NEJM 2018) and can also be found at https://clinicaltrials.gov/ct2/show/NCT02366143. The are 161 institutes/organizations whose ethical committees approved the protocol, including: [1] "Copernicus Group Independent Review Board" [2] "Western Institutional Review Board" [3] "Melbourne Health Human Research Ethics Committee" [4] "Tasmania Health and Medical Human Research Ethics Committee" [5] "Asociacion Benefica Prisma" [6] "Comité Institucional de ética en Investigación Instituto Nacional de Enfermedades Neoplasicas" [7] "Singhealth Centralised Institutional Review Board" [8] "Institutional Review Board Kaohsiung Medical University Chung-Ho Memorial Hospital" [9] "Institutional Review Board Taipei Veterans General Hospital" [10] "Institutional Review Board" [11] "Institutional Review Board of the Chi Mei Medical Center" [12] "The Institutional Review Board of Taichung Veterans General Hospital" [13] "Institutional Review Board of Tri-Service General Hospital" [14] "Cheng-Hsin General Hospital Institutional Review Board" [15] "Consejo de Evaluación Ética de Investigación en Salud - CoEIS" [16] "Ministerio de Salud - Provincia de Río Negro" [17] "Comité Independiente de Etica en investigación clínica \"Dr. Carlos A. Barclay" [18] "Comité de ética en investigación Fundación Oncosalud" [19] "Comisión Conjunta de Investigación en Salud (CCIS)" [20] "Comité de Ética Centro de Oncología e Investigación Buenos Aires" [21] "Comité de Ética Independiente - Fundación Sanatorio" [22] "Institutional Review Board of Chang Gung Medical Foundation" [23] "Ethikkommission für das Bundesland Salzburg" [24] "The Ethics Committee for Clinical Trials of Medicinal Products" [25] "Chesapeake Institutional Review Board" [26] "St. Luke's Hospital & Health Network IRB" [27] "St Charles Medical Center" [28] "Saint Luke's Hospital Institutional Review Board" [29] "Kantonale Ethikkommission Bern (KEK)" [30] "Comite de Etica Cientifico del Servicio de Salud Metropolitano Norte" [31] "Comitato Etico Area Vasta Nord Ovest presso Azienda Ospedaliero Universitaria Pisana di Pisa"

[32] "Comitato Etico Cardarelli-Santobono"
[33] "Comitato Etico San Martino IST 2"
[34] "Comitato Etico Lazio 1"
[35] "COMITATO ETICO DELL´UNIVERSITA´ CAMPUS BIO-MEDICO DI ROMA"
[36] "Medical Research Ethics Committees United"
[37] "CPP Sud-Méditerranée 2"
[38] "Sir Charles Gairdner Hospital HREC"
[39] "Bellberry Human Research Ethics Committee"
[40] "Cabrini Human Research Ethics Committee"
[41] "Comitato Etico Catania 1 presso A.O. Universitaria Policlinico Vittorio Emanuele di Catania"
[42] "Comissão de Ética para a Investigação Clínica - CEIC"
[43] "Ethikkommission an der Universität Regensburg"
[44] "Ethikkommission an der Universität Regensburg"
[45] "Lakeridge Health REB"
[46] "St. Joseph Mercy Health System Institutional Review Board #2 - Oncology Central IRB"
[47] "Mercy Saint Vincent Medical Center Institutional Review Board"
[48] "US Oncology Inc. Institutional Review Board"
[49] "Western Institutional Review Board (WIRB)"
[50] "CEIC de la Corporacion Sanitaria del Parc Tauli"
[51] "CEIC de Cantabria"
[52] "CEIC de Andalucia (CCEIBA)"
[53] "CEIC Hospital Universitario La Paz"
[54] "CEIC Fundación Jiménez Díaz"
[55] "CEIC Hospital Clínico Universitario de Valencia"
[56] "CEIC Hospital General Universitario Gregorio Marañon"
[57] "CEIC Hospital Clinico San Carlos"
[58] "CEIC Grupo Hospital de Madrid"
[59] "CEIC Parc de Salut Mar"
[60] "CEIC Islas Baleares (CEIC-IB)"
[61] "CEIC de Galicia (CAEI)"
[62] "CEIC Hospital Clinic de Barcelona"
[63] "Kaiser Permanente Southern California Institutional Review Board"
[64] "Research Ethics Committee of National Taiwan University Hospital"
[65] "Institutional Review Board of Chung Shan Medical University Hospital"
[66] "CHU de Liège - Comité d'Ethique"
[67] "CEQ of MI Kryvyi Rih Oncology Dispensary of Dnipropetrovsk Regional Council"
[68] "CEQ of Treatment and Prevention Institution Volyn Regional Oncology Dispensary"
[69] "CEQ of MI Dnipropetrovsk City Multifield Clinical Hospital #4 of Dnipropetrovsk Regional Council"
[70] "CEQ of Poltava Regional Clinical Oncology Dispensary of Poltava Regional Council"
[71] "Commission on Ethics Questions of Uzhgorod Central City Clinical Hospital"
[72] "Commission of Ethics Questions on the basis of the Chernivtsi Regional Clinical Oncology Dispensary"
[73] "CEQ of MI of Zaporizhzhia Regional Council Zaporizhzhia Regional Clinical Oncology Dispensary"
[74] "CEQ of Transcarpathian Regional Clinical Oncology Dispensary"
[75] "CEQ of SI Institute of Medical Radiology n.a. S.P. Hryhoriev of NAMS of Ukraine"
[76] "CEQ of Municipal Noncommercial Institution Regional Center of Oncology"
[77] "CEQ of Regional Municipal Institution Sumy Regional Clinical Oncology Dispensary"
[78] "Comitato Etico Azienda Ospedaliera Universitaria Maggiore della Carità"
[79] "Comitato Etico Dell Universita Cattolica del Sacro Cuore Policlinico Universitario Agostino Gemelli"
[80] "Comitê de Ética em Pesquisa em Seres Humanos da Faculdade de Medicina de São José do Rio Preto"
[81] "Comitê de Ética em Pesquisa em Seres Humanos do Hospital Socor"
[82] "Comitê de Ética em Pesquisa em Seres Humanos da Universidade de Ribeirão Preto (UNAERP)"
[83] "Comitê de Ética em Pesquisa em Seres Humanos da Irmandade da Santa Casa de Londrina"
[84] "Comitê de Ética em Pesquisa em Seres Humanos da Liga Norte Riograndense Contra o Câncer"
[85] "Comitê de Ética em Pesquisa Fundação Pio XII Hospital de Câncer de Barretos"
[86] "Comité de Ética en Investigación de la Facultad de Medicina y Hospital Universitario"
[87] "Missouri Baptist Medical Center Institutional Review Board"
[88] "University of California Irvine Institutional Review Board"
[89] "University of California"
[90] "Frederick Memorial Hospital Institutional Review Board"
[91] "Mercy Medical Center IRB"
[92] "University of Chicago Hospitals Institutional Review Board"
[93] "Mount Sinai Medical Center IRB"
[94] "Comité Provincial de Bioética - Ministerio de Salud de la Provincia de Santa Fé"
[95] "Comite de Etica Investigacion de la Clinica Bajío"
[96] "Comite de Etica en Investigacion de Mexico Centre for Clinical Research SA de CV"
[97] "Comitê de Ética em Pesquisa da Universidade de Caxias do Sul"
[98] "Comitê de Ética em Pesquisa do Hospital de Clínicas de Porto Alegre"
[99] "Comite de Etica em Pesquisa da Universidade Federal de Sao Paulo - Hospital Sao Paulo"
[100] "Comitê de Ética em Pesquisa - Hospital Mãe de Deus"
[101] "Comitê de Ética em Pesquisa da Fundação Antônio Prudente – Hospital do Câncer A. C. Camargo"
[102] "Ethics Committee for Multi-Centre Trials"
[103] "Ethics Committee at Clinical Oncology Dispensary"
[104] "Ethics Committee at City Clinical oncologic dispensary"
[105] "Ethics Committee at Russian Oncology Research Center n.a. N.N.Blokhin"
[106] "Ethics Committee at Moscow City Oncology Hospital #62 of Moscow Healthcare Department"
[107] "Ethics Committee at Volzhskiy regional clinical oncology dispensary #3"

[108] "CEQ of Ivano-Frankivsk Regional Oncology Dispensary"
[109] "CEIC Hospital Universitario Insular Materno-Infantil de Las Palmas"
[110] "CEIC Hospital Universitari Vall d'Hebron"
[111] "CEIC Hospital Universitario Ramon y Cajal"
[112] "CEIC Hospital Universitario 12 de Octubre"
[113] "Eticka komisia Presovskeho samospravneho kraja"
[114] "Eticka komisia Univerzitna nemocnica Bratislava"
[115] "Lithuanian Bioethics Committee"
[116] "CEIC Hospital Universitari de Bellvitge"
[117] "Eticka komisia pri Narodnom onkologickom ustave"
[118] "WIRB Copernicus Group"
[119] "Kaiser Permanente of Colorado Institutional Review Board"
[120] "Ethics Committee at Russian Medical Military Academy n.a. S.M.Kirov"
[121] "Comite de Etica en Investigacion del Instituto Regional de Enfermedades Neoplasicas"
[122] "National Cheng Kung University Hospital Human Experiment and Ethic Committee"
[123] "Houston Methodist Research Institute IRB"
[124] "Ingalls Memorial Hospital IRB"
[125] "Rush University Medical Center Institutional Review Board"
[126] "Mercy Health Springfield Communities Institutional Review Board"
[127] "Mayo Clinic Institutional Review Board"
[128] "Salus IRB"
[129] "Mackay Memorial Hospital Institutional Review Board"
[130] "Commission on Ethics Questions of Vinnytsya Regional Clinical Oncology Dispensary"
[131] "Comite Etico Cientifico Clinica Santa Maria"
[132] "BRANY IRB"
[133] "Yale University Human Research Protection Program"
[134] "Maimonides Med Ctr Institutional Review Board"
[135] "Scripps Health Institutional Review Board"
[136] "University of Texas Health Science Center San Antonio Institutional Review Board"
[137] "Park Nicollet Institute Institutional Review Board"
[138] "Eticka komisia NsP Sv. Jakuba"
[139] "Ethikkommission der Bayerischen Landesärztekammer"
[140] "CEQ of Kyiv City Clinical Oncological Center"
[141] "The Institutional Review Board of China Medical University Hospital"
[142] "Ethics Committee at Private Medical Institution \"Evromedservis\""
[143] "Comitato di Bioetica dell'AUSL 1 di Sassari"
[144] "Eticka komisia Onkologicky ustav sv. Alzbety"
[145] "Ethics Committee at Railway Clinical Hospital JSC RZhD"
[146] "Sault Area Hospital Research Ethics Board"
[147] "Toranomon Hospital and Toranomon Hospital Kajigaya IRB"
[148] "Kitasato University Sagamihara IRB"
[149] "Niigata Cancer Center Hospital IRB"
[150] "Kyoto University Hospital IRB"
[151] "Osaka City University Hospital IRB"
[152] "National Hosptial Organization Toneyama National Hospital IRB"
[153] "Wakayama Medical University IRB"
[154] "Kurume University IRB"
[155] "National Hospital Organization Kyushu Cancer Center"
[156] "Kanagawa Cardiovascular and Respiratory Center IRB"
[157] "National Hospital Organization Kyushu Medical Center IRB"
[158] "National Hospital Organization Shikoku Cancer Center IRB"
[159] "Miyagi Cancer Center IRB"
[160] "Kyorin University Hospital IRB"
[161] "Center Hospital of the National Center for Global Health and Medicine IRB"

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	Details on the IMpower150 study plan are published elsewhere. No sample size calculations were done for the ctDNA subsidy of IMpower150. The sample size was determined by the number of patients in IMpower150 (n=1201) which had baseline samples run on a 1.25 Mb ctDNA assay (n=1062), and those were subset to those patients with aliquots of C2D1 or C3D1 plasma available (n=910). 566 patients were then

	randomly chosen for the ctDNA subsidy, only 466 of which had PBMC available for CHIP/germline correction who were included in the final ctDNA analysis population.
Data exclusions	Patients were excluded if they did not have baseline plasma and C2D1/C3D1 and PBMC available. Please refer to Figure 1 for sample flow diagram.
Replication	Biomarkers measured in patient plasma samples were not replicated. Replication was not possible due to the limited patient samples available. The one experiment that had replication was during assay development in which 63 replicates were run to measure the concordance between the baseline assay and on-treatment assay. This experiment showed high concordance between the replicates (R2=0.9961; Supplementary Figure 1B right). No other experiments were replicated.
Randomization	Details on the original IMpower150 study plan including randomization has been published elsewhere and can also be found at https://clinicaltrials.gov/ct2/show/NCT02366143. For this exploratory retrospective ctDNA substudy reported in this manuscript, the 466 ctDNA-evaluable patients were split into a training cohort and a testing cohort for model development. The training and test sets were initially chosen based on sequencing batch for the set of 566 patients chosen for the ctDNA substudy (see Figure 1a), where we put sequencing batch1 in training set and then added in patients from later batches to reach the target 50%/50% split. The sequencing lab decided which samples to include in batch 1 without any knowledge about the baseline characteristics, treatment, or clinical outcomes of the patients. We then checked for imbalances and it was found that RACE was not well distributed due to all Asian patients appearing in batch1, and so we moved half of the Asian patients to the test set and replaced these spots in the training data with a random set of patients. As the analysis progressed (in the training subset of data) we decided to add in PBMC correction due to concern over germline/CHIP variants contaminating the ctDNA dataset, which reduced the number of patients to those with PBMC available for correction, giving a final n of 466 patients and a final split of 240/226 patients for train/test. The final training/test sets were well balanced in clinical features and survival outcomes as can be seen in Supplementary Table ST2 and Extended Data Figure ED1f.
Blinding	Details about blinding during arm allocation for the IMpower150 study itself can be found in the IMpower150 primary clinical manuscript or at https://clinicaltrials.gov/ct2/show/NCT02366143 . For this retrospective exploratory ctDNA substudy, our model development analysis was not initiated until after samples were allocated to either the training or testing subgroup. The allocation of training/testing was not performed blinded, as the goal was to ensure that the training and testing subgroups were similar in baseline characteristics. Sample collection was performed prior to this retrospective analysis, and so investigators were blinded during sample collection. Data was generated by a diagnostic company separate from the analysts who performed the model development, and so data generation was also performed by blinded individuals.

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

Methods

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

Antibodies

Antibodies used	PD-L1 expression on tumor cells or tumor-infiltrating immune cells was analyzed in archival or freshly collected tumor tissue (or both) with the use of a PD-L1 immunohistochemistry assay (Ventana Medical Systems; clone SP142; catalog number N/A; pre-dilute ready to use antibody product at 36ug/5mL)
Validation	Ventana PD-L1 (SP142) assay validation can be found at the following link: https://www.accessdata.fda.gov/cdrh_docs/pdf16/p160002c.pdf

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	NCT02366143
Study protocol	The protocol was published with the primary clinical manuscript and can be found here: https://www.nejm.org/doi/suppl/10.1056/NEJMoa1716948/suppl_file/nejmoa1716948_protocol.pdf

Data collection

Details of clinical data collection can be found in primary clinical manuscript for IMpower150. Briefly, IMpower150 data was collected between March 2015 and December 2019, patients from 26 countries or regions were enrolled at 240 sites. The list of 240 sites is the following:
United States, Arizona
Ironwood Cancer & Research Centers
Chandler, Arizona, United States, 85224
Arizona Oncology Associates
Flagstaff, Arizona, United States, 86001
United States, California
Southern CA Permanente Med Grp
Bellflower, California, United States
Marin Cancer Care Inc
Greenbrae, California, United States, 94904
Scripps Health
La Jolla, California, United States, 92037
Chao Family Comprehensive Cancer Center UCI
Orange, California, United States, 92868
United States, Colorado
Rocky Mountain Cancer Center
Denver, Colorado, United States, 80218
Kaiser Permanente
Lonetree, Colorado, United States, 80124
United States, Connecticut
Danbury Hospital
Danbury, Connecticut, United States, 06810
Yale Cancer Center
New Haven, Connecticut, United States, 06520
United States, Florida
Holy Cross Hospital Inc
Fort Lauderdale, Florida, United States, 33308
Cancer Specialists of North Florida - Baptist South
Jacksonville, Florida, United States, 32258
Mount Sinai Medical Center
Miami Beach, Florida, United States, 33140
Hematology Oncology Associates of the Treasure Coast
Port Saint Lucie, Florida, United States, 34952
United States, Georgia
Piedmont Cancer Institute, PC
Atlanta, Georgia, United States, 30318
United States, Illinois
Rush University Medical Center
Chicago, Illinois, United States, 60612
Univ of Chicago
Chicago, Illinois, United States, 60637
Ingalls Memorial Hospital
Harvey, Illinois, United States, 60426
Oncology Specialists, S.C.
Park Ridge, Illinois, United States, 60068
United States, Iowa
Hematology-Oncology; Associates of the Quad Cities
Bettendorf, Iowa, United States, 52722
United States, Kentucky
Norton Cancer Institute
Louisville, Kentucky, United States, 40202
United States, Maine
New England Cancer Specialists
Scarborough, Maine, United States, 04074
United States, Maryland
Mercy Medical Center
Baltimore, Maryland, United States, 21202
Regional Cancer Care Associates
Bethesda, Maryland, United States, 20817
Maryland Oncology Hematology, P.A.
Columbia, Maryland, United States, 21044
United States, Michigan
St. Joseph Mercy Health System
Ann Arbor, Michigan, United States, 48106
United States, Minnesota
St. Luke's Regional Cancer Center
Duluth, Minnesota, United States, 55805
Park Nicolett - Frauenshuh Cancer Center
Saint Louis Park, Minnesota, United States, 55426
United States, Missouri
St. Luke's Cancer Institute
Kansas City, Missouri, United States, 64111

Missouri Baptist Medical Center
Saint Louis, Missouri, United States, 63131
United States, Montana
Billings Clinic
Billings, Montana, United States, 59102
Montana Cancer Specialists
Missoula, Montana, United States, 59802
United States, Nevada
Comprehensive Cancer Centers of Nevada
Henderson, Nevada, United States, 89014
United States, New Jersey
Summit Medical Group
Berkeley Heights, New Jersey, United States, 07922
Valley Hospital; Oncology Research
Paramus, New Jersey, United States, 07652
Regional Cancer Care Associates LLC
Sewell, New Jersey, United States, 08080
United States, New York
Montefiore Medical Center
Bronx, New York, United States, 10467
Maimonides Medical Center
Brooklyn, New York, United States, 11219
United States, North Carolina
First Health of the Carolinas
Pinehurst, North Carolina, United States, 28374
United States, Ohio
University of Cincinnati
Cincinnati, Ohio, United States, 45203-0542
Mercy St Anne Hospital
Toledo, Ohio, United States, 43623
United States, Oregon
Bend Memorial Clinic
Bend, Oregon, United States, 97701
St. Charles Medical Center Bend; Cancer Care Of The Cascades
Bend, Oregon, United States, 97701
Willamette Valley Cancer Insitute and Research Center
Springfield, Oregon, United States, 97477
United States, Pennsylvania
St. Luke's Cancer Care Associates
Bethlehem, Pennsylvania, United States, 18015
Allegheny Cancer Center
Pittsburgh, Pennsylvania, United States, 15212
Univ of Pittsburgh Medical Ctr
Pittsburgh, Pennsylvania, United States, 15232
United States, Tennessee
West Clinic
Germantown, Tennessee, United States, 38138
Tennessee Cancer Specialists
Knoxville, Tennessee, United States, 37920
United States, Texas
Houston Methodist Cancer Center
Houston, Texas, United States, 77030
Longview Cancer Center
Longview, Texas, United States, 75601
University of Texas Health Science Center at San Antonio
San Antonio, Texas, United States, 78229
United States, Virginia
Virginia Cancer Specialists, PC
Fairfax, Virginia, United States, 22031
Virginia Oncology Associates
Norfolk, Virginia, United States, 23502
Virginia Cancer Institute
Richmond, Virginia, United States, 23226
Blue Ridge Cancer Care
Roanoke, Virginia, United States, 24014
United States, Washington
MultiCare Regional Cancer Center - Auburn
Auburn, Washington, United States, 98002-4117
Providence Regional Cancer Partnership
Everett, Washington, United States, 98201
Virginia Mason Medical Center
Seattle, Washington, United States, 98101
Medical Oncology Associates
Spokane, Washington, United States, 99208
United States, West Virginia
West Virginia University; Mary Babb Randolph Can Ctr

Morgantown, West Virginia, United States, 26506
Argentina
Centro Medico Austral
Buenos Aires, Argentina, 1019
Fundación CENIT para la Investigación en Neurociencias
Buenos Aires, Argentina, C1125ABD
Sanatorio Allende
Cordoba, Argentina, X5000JHQ
Centro Oncologico Riojano Integral (CORI)
La Rioja, Argentina, F5300COE
Hospital Provincial del Centenario
Rosario, Argentina, 2000
Fundacion Koriza
Santa Rosa, Argentina, 6300
Centro de Investigacion; Clinica - Clinica Viedma S.A.
Viedma, Argentina, R8500ACE
Australia, New South Wales
Chris O'Brien Lifehouse
Camperdown, New South Wales, Australia, 2050
Concord Repatriation General Hospital
Concord, New South Wales, Australia, 2139
Nepean Cancer Care Centre
Sydney, New South Wales, Australia, 2747
Australia, Queensland
Prince Charles Hospital; Department of Medical Oncology
Chermside, Queensland, Australia, 4032
Townsville Hospital
Townsville, Queensland, Australia, 4810
Princess Alexandra Hospital
Woolloongabba, Queensland, Australia, 4102
Australia, South Australia
Royal Adelaide Hospital
Adelaide, South Australia, Australia, 5000
Adelaide Cancer Centre
Kurralta Park, South Australia, Australia, 5037
Australia, Tasmania
Royal Hobart Hospital
Hobart, Tasmania, Australia, 7000
Launceston General Hospital
Launceston, Tasmania, Australia, 7250
Australia, Victoria
Frankston Hospital
Frankston, Victoria, Australia, 3199
Austin Health
Heidelberg, Victoria, Australia, 3084
Cabrini Hospital Malvern
Malvern, Victoria, Australia, 3144
The Alfred Hospital
Prahan, Victoria, Australia, 3181
Sunshine Hospital
St Albans, Victoria, Australia, 3021
Australia, Western Australia
Sir Charles Gairdner Hospital
Nedlands, Western Australia, Australia, 6009
Austria
Paracelsus Medizinische Privatuniversität
Salzburg, Austria, 5020
Klinikum Wels-Grieskirchen
Wels, Austria, 4600
Belgium
CHU de Liège
Liège, Belgium, 4000
Clinique Ste-Elisabeth
Namur, Belgium, 5000
Brazil
CETUS Hospital Dia Oncologia
Uberaba, MG, Brazil, 38082-049
Instituto Do Cancer Delondrina_X; Unidade De Pesquisa Clinica
Londrina, PR, Brazil, 86 015 520
Liga Norte Riograndense Contra O Câncer
Natal, RN, Brazil, 59040150
IPCEM; Instituto de Pesquisa de Estudos Multicêntricos
Caxias do Sul, RS, Brazil, 95070-560
Hospital Mae de Deus
Porto Alegre, RS, Brazil, 90470-340
Hospital de Cancer de Barretos

Barretos, SP, Brazil, 14784-400
Instituto Ribeirãopretano de Combate Ao Câncer; Centro Especializado De Oncologia
Ribeirão Preto, SP, Brazil, 14015-130
Hospital de Base de Sao Jose do Rio Preto
Sao Jose do Rio Preto, SP, Brazil, 15090-000
Hospital A. C. Camargo; Oncologia
Sao Paulo, SP, Brazil, 01509-010
Bulgaria
Multiprofile Hospital for Active Treatment Central Onco Hospital OOD
Plovdiv, Bulgaria, 4000
MHAT Serdika, EOOD
Sofia, Bulgaria, 1303
Canada, Ontario
Lakeridge Health Center
Oshawa, Ontario, Canada, L1J 2J2
Chile
Clinica Santa Maria
Santiago, Chile, 0
Health & Care SPA
Santiago, Chile, 7500006
Sociedad de Investigaciones Medicas Ltda (SIM)
Temuco, Chile, 4810469
France
Institut Bergonié Centre Régional de Lutte Contre Le Cancer de Bordeaux Et Sud Ouest
Bordeaux, France, 33076
CHU de Grenoble
Grenoble, France, 38043
Centre Jean Bernard Clinique Victor Hugo
Le Mans, France, 72015
Hôpital Saint Joseph
Marseille, France, 13008
Hopital Nord AP-HM
Marseille, France, 13015
Hôpital Européen Georges Pompidou
Paris, France, 75908
CHU de Bordeaux
Pessac, France, 33600
Service de Pneumologie Centre Hospitalier Régional La Réunion Site Felix Guyon
Saint Denis Cedex, France, 97405
CH de Saint Quentin
Saint Quentin, France, 2100
Centre Hospitalier Intercommunal Toulon - La Seyne sur Mer
Toulon, France, 83000
Hôpital d'Instruction des Armées de Sainte Anne; Service de Pneumologie
Toulon, France, 83000
Hôpital Larrey;Université Paul Sabatier
Toulouse, France, 31059
Germany
Zentralklinikum Augsburg
Augsburg, Germany, 86156
Helios Klinikum Emil von Behring GmbH
Berlin, Germany, 14165
Ev.Krankenhaus Bielefeld gGmbH; Klinik für Innere Medizin und Geriatrie
Bielefeld, Germany, 33611
Augusta Kranken-Anstalt gGmbH
Bochum, Germany, 44791
Universitätsklinikum "Carl Gustav Carus" der Technischen Universität Dresden
Dresden, Germany, 01307
St. Elisabethen Krankenhaus
Frankfurt am Main, Germany, 60487
LungenClinic Großhansdorf GmbH
Großhansdorf, Germany, 22927
Krankenhaus Martha-Maria; Halle-Dolau gGmbH
Halle, Germany, 06120
Asklepios Klinik Harburg
Hamburg, Germany, 21075
Lungenklinik Hemer
Hemer, Germany, 58675
Universität Des Saarlandes; Klinik für Innere Medizin V
Homburg, Germany, 66421
Kliniken der Stadt Koln gGmbH; Lungenklinik Onkologische Ambulanz
Koln, Germany, 51109
Klinik Loewenstein gGmbH; Onk & Pal
Loewenstein, Germany, 74245
Klinikum Bogenhausen
München, Germany, 81925

Krankenhaus Barmherzige Bruder Regensburg
Regensburg, Germany, 93049
Klinikum der Universität Regensburg
Regensburg, Germany, 93053
Stiftung Mathias-Spital Rheine
Rheine, Germany, 48431
Italy
AORN A Cardarelli
Napoli, Campania, Italy, 80131
Policlinico Universitario Campus Biomedico; Uoc Oncologia Medica
Roma, Lazio, Italy, 00128
Azienda Ospedaliera San Camillo Forlanini
Roma, Lazio, Italy, 00152
Università Cattolica Del S Cuore
Roma, Lazio, Italy, 00168
Istituto Nazionale per la Ricerca sul Cancro di Genova
Genova, Liguria, Italy, 16132
ASL 3 Genovese; DSM
Genova, Liguria, Italy, 16147
A.O.U. Maggiore della Carità
Novara, Piemonte, Italy, 28100
Azienda Unita Sanitaria Locale N1 Sassari; Unita Operativa Di Oncologia Medica
Sassari, Sardegna, Italy, 07100
Policlinico Vittorio Emanuele
Catania, Sicilia, Italy, 95123
Ospedale Versilia
Lido Di Camaiore, Toscana, Italy, 55043
Ospedale Civile - Livorno
Livorno, Toscana, Italy, 57124
Japan
National Hospital Organization Shikoku Cancer Center
Ehime, Japan, 791-0280
National Hospital Organization Kyushu Medical Center
Fukuoka, Japan, 810-8563
NHO Kyushu Cancer Center
Fukuoka, Japan, 811-1395
Kurume University Hospital
Fukuoka, Japan, 830-0011
Kanagawa Cardiovascular and Respiratory Center
Kanagawa, Japan, 236-0051
Kitasato University Hospital
Kanagawa, Japan, 252-0375
Kyoto University Hospital
Kyoto, Japan, 606-8507
Miyagi Cancer Center
Miyagi, Japan, 981-1293
Niigata Cancer Center Hospital
Niigata, Japan, 951-8566
Osaka City University Hospital
Osaka, Japan, 545-8586
National Hospital Organization Osaka Toneyama Medical Center
Osaka, Japan, 560-8552
Toranomon Hospital
Tokyo, Japan, 105-8470
Center Hospital of the National Center for Global Health and Medicine
Tokyo, Japan, 162-0052
Kyorin University Hospital
Tokyo, Japan, 181-8611
Wakayama Medical University Hospital
Wakayama, Japan, 641-8510
Latvia
Riga East Clinical University Hospital Latvian Oncology Centre
Riga, Latvia, LV-1079
Pauls Stradins Clinical University Hospital
Rīga, Latvia, LV-1002
Lithuania
National Cancer Institute
Vilnius, Lithuania, 08660
Mexico
Centro Universitario Contra El Cancer
Monterrey, Mexico, 64020
Cancerologia de Queretaro
Queretaro, Mexico, 76090
Centro Hemato Oncologico Privado; Oncologia
Toluca, Mexico, 50080
Netherlands

Jeroen Bosch Ziekenhuis
'S Hertogenbosch, Netherlands, 5223 GZ
Amsterdam UMC Location VUMC
Amsterdam, Netherlands, 1081 HV
Amphia Ziekenhuis; Afdeling Longziekten
Breda, Netherlands, 4818 CK
Ziekenhuis Gelderse Vallei
EDE, Netherlands, 6716 RP
Tergooiziekenhuizen
Hilversum, Netherlands, 1201 DA
Spaarne Gasthuis; Spaarne Ziekenhuis
Hoofddorp, Netherlands, 2134 TM
Maastricht University Medical Center
Maastricht, Netherlands, 6229 HX
St. Antonius Ziekenhuis; R&D Long
Nieuwegein, Netherlands, 3435 CM
Erasmus MC; Afdeling Longziekten
Rotterdam, Netherlands, 3015 GD
Maasstad ziekenhuis
Rotterdam, Netherlands, 3079 DZ
Universitair Medisch Centrum Utrecht
Utrecht, Netherlands, 3584 CX
Gelre Ziekenhuizen, Zutphen
Zutphen, Netherlands, 7207 AE
Peru
Centro Medico Monte Carmelo
Arequipa, Peru, 04001
Centro Especializado de Enfermedades Neoplásicas
Arequipa, Peru
Instituto Nacional de Enfermedades Neoplasicas
Lima, Peru, Lima 34
Portugal
Centro Hospitalar E Universitário de Coimbra EPE
Coimbra, Portugal, 3000-602
Instituto Portugues Oncologia de Lisboa Francisco Gentil EPE
Lisboa, Portugal, 1099-023
Hospital Pulido Valente; Servico de Pneumologia
Lisboa, Portugal, 1796-001
Centro Hospitalar do Porto - Hospital de Santo António
Porto, Portugal, 4099-001
Instituto Portugues de Oncologia Do Porto Francisco Gentil Epe
Porto, Portugal, 4200-072
Hospital de Sao Joao; Servico de Pneumologia
Porto, Portugal, 4200
Russian Federation
Moscow City Oncology Hospital #62
Moscovskaya Oblast, Moskovskaja Oblast, Russian Federation, 143423
Russian Oncology Research Center n.a. N.N. Blokhin
Moscow, Russian Federation, 115478
Clinical Oncology Dispensary
Omsk, Russian Federation, 644013
Evromedservis LCC
Pushkin, Russian Federation, 196603
City Clinical Oncology Dispensary
Saint-Petersburg, Russian Federation, 197022
Singapore
National Cancer Centre; Medical Oncology
Singapore, Singapore, 169610
Slovakia
Univerzitna nemocnica Bratislava
Bratislava, Slovakia, 813 69
Narodny onkologicky ustav
Bratislava, Slovakia, 833 10
POKO Poprad s.r.o.
Poprad, Slovakia, 058 01
Spain
Instituto Catalan de Oncologia de Hospitalet (ICO); Servicio de Farmacia
L'Hospitalet de Llobregat, Barcelona, Spain, 08908
Corporacio Sanitaria Parc Tauli; Servicio de Oncologia
Sabadell, Barcelona, Spain, 8208
Hospital Universitario Marques de Valdecilla
Santander, Cantabria, Spain, 39008
Hospital Universitario Son Espases
Palma De Mallorca, Islas Baleares, Spain, 07014
Hospital Universitario Insular de Gran Canaria
Las Palmas de Gran Canaria, LAS Palmas, Spain, 35016

Complejo Hospitalario U. de Ourense
Ourense, Orense, Spain, 32005
Hospital Lluis Alcanyis De Xativa
Xativa, Valencia, Spain, 46800
Hospital del Mar
Barcelona, Spain, 08003
Hospital Univ Vall d'Hebron; Servicio de Oncologia
Barcelona, Spain, 08035
Hospital Clinic de Barcelona
Barcelona, Spain, 08036
Hospital Universitario Reina Sofia
Cordoba, Spain, 14004
Hospital Lucus Augusti; Servicio de Oncologia
Lugo, Spain, 27003
Hospital General Universitario Gregorio Marañon; Servicio de Oncologia
Madrid, Spain, 28007
Hospital Universitario La Paz
Madrid, Spain, 280146
Hospital Ramon y Cajal; Servicio de Oncologia
Madrid, Spain, 28034
Hospital Clinico San Carlos; Servicio de Oncologia
Madrid, Spain, 28040
Hospital Universitario Fundación Jimenez Díaz
Madrid, Spain, 28040
Hospital Universitario 12 de Octubre
Madrid, Spain, 28041
Hospital Universitario HM Sanchinarro-CIOCC
Madrid, Spain, 28050
Hospital Clinico Universitario de Valencia
Valencia, Spain, 46010
Switzerland
Kantonsspital Baselland
Bruderholz, Switzerland, 4101
Luzerner Kantonsspital Sursee
Luzern, Switzerland, 6000
Kantonsspital St. Gallen; Onkologie/Hämatologie
St. Gallen, Switzerland, 9007
Taiwan
Changhua Christian Hospital; Hematology-Oncology
Changhua, Taiwan, 500
Kaohsiung Medical University Hospital; Department of Urology
Kaohsiung City, Taiwan, 807
Chi Mei Medical Center Liou Ying Campus
Liuying Township, Taiwan, 736
Chang Gung Memorial Hospital Chiayi
Putzu, Taiwan, 613
National Cheng Kung Univ Hosp
Tainan, Taiwan, 00704
National Taiwan Uni Hospital
Taipei City, Taiwan, 10041
Cheng Hsin General Hospital
Taipei, Taiwan, 112
Tri-Service General Hospital
Taipei, Taiwan, 11490
Chang Gung Medical Foundation Linkou Branch
Taoyuan City, Taiwan, 333
Taichung Veterans General Hospital
Xitun Dist., Taiwan, 40705
Ukraine
ME Bukovinian Clinical Oncology Center
Chernivtsi, Chernihiv Governorate, Ukraine, 58013
Municipal Institution City Clinical Hospital #4 of Dnipro City Council - PPDS; Dept of Chemotherapy
Dnipropetrovsk, Katerynoslav Governorate, Ukraine, 49102
Uzhgorod Central City Clinical Hospital
Uzhhorod, Katerynoslav Governorate, Ukraine, 88000
MNPE Zaporizhzhia Regional Antitumor Center ZRC
Zaporizhzhia, Katerynoslav Governorate, Ukraine, 69040
Communal Non profit Enterprise Regional Center of Oncology; Oncosurgical dept of thoracic organs
Kharkiv, Kharkiv Governorate, Ukraine, 61070
MNPE Transcarpathian Antitumor Center of the Transcarpathian Regional Council; Chemotherapy Dept
Uzhhorod, Kherson Governorate, Ukraine, 88014
Municipal Institution SubCarpathian Clinical Oncological Centre; Surgical department#2
Ivano-Frankivsk, KIEV Governorate, Ukraine, 76018
Communal Nonprofit Enterprise Podilsky Regional Center Of Oncology OfTheVinnytsia Regional Council
Vinnytsia, KIEV Governorate, Ukraine, 21029
SI Institute of Medical Radiology n.a. S.P. Hryhoriev of NAMS of Ukraine

Kharkiv, Ukraine, 61024
ME Kryviy Rih Oncology Dispensary of Dnipropetrovs'k Regional Council; Chemotherapy Department
Kryvyi Rih, Ukraine, 50048
Kyiv City Clinical Oncological Center
Kyiv, Ukraine, 03115
Poltava Regional Clinical Oncology Dispensary of Poltava Regional Council; Thoracic department
Poltava, Ukraine, 36011
Regional Municipal Institution Sumy Regional Clinical Oncology Dispensary
Sumy, Ukraine, 40005

Outcomes

The two primary end points of IMpower150 were progression-free survival (as assessed by investigators according to RECIST criteria) both among patients in the intention-to-treat population who had a wild-type genotype (WT population; patients with EGFR or ALK genomic alterations were excluded) and among patients in the WT population who had high expression of an effector T-cell (Teff) gene signature in the tumor (Teff-high WT population) and overall survival in the WT population. The ctDNA substudy also used the Overall Survival and radiographic response as assessed by RECIST v1.1. Additional details on the IMpower150 study can be found in the primary clinical manuscript.