Title

Machine learning model to predict oncologic outcomes for drugs in randomized clinical trials

Journal

International Journal of Cancer, 147(9)

ISSN

0020-7136

Authors

Schperberg, Alexander V
Boichard, Amélie
Tsigelny, Igor F
et al.

Publication Date

2020-11-01

DOI

10.1002/ijc.33240

Peer reviewed

Machine learning model to predict oncologic outcomes for drugs in randomized clinical trials

Authors:

Alexander V. Schperberg[1,3]; Amélie Boichard[2]; Igor F. Tsigelny[1,4,5]; Stéphane B. Richard[1,6]; Razelle Kurzrock[2*]

Affiliation:

[1]CureMatch, Inc., San Diego, 92121, CA, USA
[2]Center for Personalized Cancer Therapy and Division of Hematology and Oncology, University of California San Diego Moores Cancer Center, La Jolla, 92093, CA, USA
[3]University of California Los Angeles, Department of Mechanical and Aerospace Engineering, Los Angeles, 90095, CA, USA
[4]San Diego Supercomputer Center, University of California San Diego, La Jolla, 92093, CA, USA
[5]Department of Neurosciences, University of California San Diego, La Jolla, 92093, CA, USA
[6]Oncodesign, Inc., New York, 10004, NY, USA

***Corresponding Author:** Razelle Kurzrock, MD. Distinguished Professor of Medicine. Director, Center for Personalized Cancer Therapy. UCSD School of Medicine. 3855 Health Sciences Drive, MC #0658, La Jolla, California 92093-0658. Phone: (858) 246-1102. Fax: (858) 246-1915. Email: rkurzrock@ucsd.edu. Twitter: @Dr_R_Kurzrock.

Keywords: Machine learning, outcome prediction, molecular profiles, drug-related biomarkers, clinical trials

Abbreviations: DPYD: dihydropyrimidine dehydrogenase; JAK: janus kinase; KRAS: proto-oncogene; M1: cancer in distant sites; ML: machine learning; N1-3: number and location of lymph node that contain the cancer; NGS: next-generation sequencing; OS: overall survival; PDAC: pancreatic ductal adenocarcinoma; PD-L1: programmed death-ligand; PDS: probability of drug sensitivity; PFS: progression free survival; TCGA: the cancer genome atlas; TMB: tumor mutation burden; TYK: tyrosine kinase; TYMP: thymidine phosphorylase; TYMS: thymidylate synthetase; VEGF: vascular endothelial growth factor;

Article Category: Research Article

Novelty and Impact: We aim to predict oncologic outcomes of different drug combinations administered in clinical trials treating advanced cancer patients. This is achieved by a machine learning (ML) model trained on a dataset including clinical trial, drug-related biomarker, and molecular profile information. We validate our model by showing a significant correlation between oncologic outcomes and our model's predictive parameters in randomized trials. Our model may be useful to optimize new drug-development strategies.

ABSTRACT

Predicting oncologic outcome is challenging due to the diversity of cancer histologies and the complex network of underlying biological factors. In this study, we determine whether machine learning can extract meaningful associations between oncologic outcome and clinical trial, drug-related biomarker, and molecular profile information. We analyzed therapeutic clinical trials corresponding to 1,102 oncologic outcomes from 104,758 cancer patients with advanced colorectal adenocarcinoma, pancreatic adenocarcinoma, melanoma, and non-small-cell lung cancer. For each intervention arm, a dataset with the following attributes was curated: line of treatment, number of cytotoxic chemotherapies, small-molecule inhibitors, or monoclonal antibody agents, drug class, molecular alteration status of the clinical arm's population, cancer type, probability of drug sensitivity (PDS) (integrating the status of genomic, transcriptomic, and proteomic biomarkers in the population of interest), and outcome. A total of 467 progression-free survival (PFS) and 369 overall survival (OS) data-points were used as training sets to build our machine learning (random forest) model. Cross-validation sets were used for PFS and OS, obtaining correlation coefficients (r) of 0.82 and 0.70 respectively (outcome versus model's parameters). A total of 156 PFS and 110 OS data-points were used as test sets. The Spearman correlation (r_s) between predicted and actual outcome was statistically significant (PFS: r_s=0.879, OS: r_s =0.878, P<0.0001). The better outcome arm was predicted in 81% (PFS: N=59/73, z=5.24, P<0.0001) and 71% (OS: N=37/52, z=2.91, P=0.004) of randomized trials. The success of our algorithm to predict clinical outcome may be exploitable as a model to optimize clinical trial design with pharmaceutical agents.

BACKGROUND

Selecting drug regimens that are not optimized to the biology of individual cancers can lead to lower survival rates and is economically costly. Thus, new tools that can support or validate the decision making of drug selection are needed.

With the advent of next-generation sequencing (NGS) techniques, new predictive biomarkers are being discovered and are facilitating the trend in the literature of integrating clinical and biologic data for outcome predictions.[1–4] Larger precision medicine trials are now examining the use of predictive biomarkers across cancer types and with different agents including targeted therapy, conventional cytotoxic chemotherapy, and immunotherapy.[5,6] However, predicting clinical outcome for cancer treatments still remains a challenging endeavor due to the diversity of genetic and environmental factors influencing tumor biology. Thus, investigating complex datasets requires sophisticated nonlinear algorithms found in machine learning (ML) software.[7] The use of ML in oncology is a recently emerging innovative technology, and previous applications have shown promising results.[8,9] For example, ML has been used to identify cancer pathways and phenocopying variants affecting patient response-rates,[10] to distinguish high-risk and low-risk patient groups using an integrative network analysis of modules containing signature gene variants,[11] and to aid in cancer diagnosis through pattern recognition.[12]

Despite their potential, ML models for oncology face several substantial challenges, such as classifying which attributes contribute significantly to clinical predictions,[13] lack of external testing and validation,[1] and insufficient predictions based only on ‘macro-scale’ attributes (e.g., histology type).[14]

Additionally, one of the most important and difficult problems in computational biology is the process of expressing biological sequences with a discrete model (or vector), while still retaining sequence-order information or pattern characteristics. A comprehensive review[15] demonstrates that the majority of ML models can only incorporate vector-based data. However, finding solutions to this problem (which can avoid, for example, losing sequence-pattern information for proteins) led to the development of Chou's 'PseAAC' software.[16,17] In 2015, very powerful webservers such as 'Pse-in-One'[18] and its updated version 'Pse-in-One2.0'[19] were established and used to generate any desired feature vectors for protein/peptide and DNA/RNA sequences.[19]

We are inspired by the work done to extract relevant features, the use of webservers, and also by Chou's 5-step rule[20] as fully introduced in our method section (the rule has been widely employed in a variety of applications in driving proteome/genome analysis and drug development, as seen in recent papers[20–22]). In our work, we will follow the 5-step rule to formulate an algorithm that requires first obtaining important features (e.g., genomic, proteomic, transcriptomic, and clinical outcome information) and then applying ML to extract pattern information between our chosen attributes and clinical outcome. Importantly, our algorithm can be expanded and used as part of an open access webserver that provides useful clinical predictors for oncologists and research scientists.

To this end, in this study, we bridge databases that lack molecular profiles but contain clinical information (e.g., ClinicalTrials.gov) and databases that lack a variety of drugs tested and their outcome results but contain molecular data (e.g., The Cancer Genome Atlas or TCGA) through a variable termed as the probability of drug sensitivity (PDS). The PDS provides a numerical metric evaluating how well drug regimens and their biomarkers match to the

molecular signature of the disease. We validate PDS as a clinical predictor by correlating PDS of drug regimens to their outcome results using clinical trial data. We also assess the impact of other clinical predictors such as treatment line, molecular alteration status, and tumor mutation burden by adjusting the PDS calculation for each parameter individually. Finally, we aggregate our curated data into several attributes (including PDS) and apply a random forest model (ML algorithm) on this dataset to predict outcomes of randomized clinical trials.

METHODS

Chou's 5-steps rule

The methods section is organized by following Chou's 5-steps rule,[20] stated as follows: (1) build a benchmark dataset to train and test the predictor; (2) represent the dataset through an effective formulation that reflects the intrinsic correlation between samples and the target to be predicted; (3) introduce a powerful algorithm to make predictions; (4) perform statistical analysis to evaluate the prediction accuracy; (5) develop a user-friendly webserver for the predictor. Steps (1) - (4) are described below, while step (5) is left for future work.

Data collection process

Clinical trial selection

We initially analyzed the colorectal adenocarcinoma dataset and further validated our model on pancreatic adenocarcinoma, melanoma, and non-small cell lung cancer (including adenocarcinoma, squamous cell carcinoma, or mixed). From the U.S. National Institutes of Health (NIH) clinical trials database (https://clinicaltrials.gov/), we retrieved a total of 1,893 interventional trials (with results) using the following search terms: "colorectal adenocarcinoma", "colon cancer", "rectal cancer", "pancreatic adenocarcinoma", "pancreatic cancer", "melanoma", "skin cancer", and "non-small cell lung cancer" (colorectal adenocarcinoma retrieval date: 1/1/2018; all other cancers: 3/15/2018).

To control for study differences between clinical arms, several trials were filtered out from the initial dataset for the following reasons: less than 15 patients were treated, eligibility criteria permitted patients with different cancer types or uncommon histology (e.g., large-cell carcinomas), patients were not diagnosed with advanced cancer (stages III to IV), trials included non-pharmacologically-mediated and non-conventional interventions (i.e., radiation therapy, dietary procedures, chimeric antigen receptor-T-cells-based immunotherapy), or did not report

median progression-free survival (PFS) or overall survival (OS) values (since only median PFS or time-to-progression and OS results were considered in this study). Using our selection criteria, a total of 351 clinical trials remained (94 – colorectal adenocarcinoma, 51 – melanoma, 37 – pancreatic adenocarcinoma, 169 – lung cancer) and 1,102 PFS or OS outcome results (PFS: 212 – colorectal adenocarcinoma, 296 – lung cancer, 67 – melanoma, 48 – pancreatic cancer; OS: 143 – colorectal adenocarcinoma, 246 – lung cancer, 41 – melanoma, 49 – pancreatic adenocarcinoma) were derived from 104,758 patients (**Table 1**).

Drug-related biomarker collection

In total, 115 single agents were used alone or in combination in our list of clinical arms (we found and validated 208 unique biomarkers from the literature – see Supporting Information **Table S1**). In this study, a biomarker was a molecular aberration (includes genomic, transcriptomic, or proteomic data) altering either: (1) a direct target of the drug (e.g., growth factor receptors and response to tyrosine kinase inhibitors); (2) an indirect target of the drug (e.g., specific ligand/agonist of a receptor and response to drugs inhibiting the receptor itself); or (3) the metabolic mechanism of the drug (e.g., DNA-damage related markers and response to DNA-damaging agents).

Molecular profile collection

Molecular profiles were collected from the cancer genome atlas database (TCGA)[23] using the following cohorts: colorectal adenocarcinoma (277 tumors), pancreatic adenocarcinoma (129 tumors), skin cutaneous melanoma (169 tumors), lung adenocarcinoma (186 tumors), and lung squamous cell carcinoma (178 tumors). Only locally invasive tumors (N1, N2 or N3) or metastatic tumors (M1) were considered (to enforce greater parity between the TCGA and clinical trial populations). These cohorts included genomic, transcriptomic, and proteomic data.

Organizing our dataset to correlate input data with target for prediction

To formulate an algorithm that correlates inputs with a desired outcome, our dataset is organized into rows that represent drug regimens retrieved from different clinical intervention arms. For each intervention arm, the following attributes (or columns) were included: line of treatment, number of *cytotoxic chemotherapy* drugs (e.g., 5-fluorouracil), number of *small-molecule inhibitors* (e.g., ruxolitinib), number of *monoclonal antibody agents* (e.g., nivolumab), one column per *drug class* (e.g., VEGFR-inhibitors, platinums, taxanes), *molecular alteration status* of the clinical arm's population (e.g., 'EGFR-positive', 'KRAS-wild type' or 'random' if status was not provided), *probability of drug sensitivity* (see below), and clinical arm *outcome* (normalized from 0 – 1 per disease and for each outcome type: PFS or OS. Values were normalized to calculate correlation across diseases). See **Fig. 1** for a pictorial depiction of how our dataset was labeled and organized.

Algorithm used for prediction

In this study, we apply a supervised learning approach using the random forest classifier to correlate our chosen attributes to a clinical arm outcome result. This classifier was chosen as it prevents overfitting of the data, has proven effective in other studies,[24–28] and provided strong correlation coefficients. For example, using the PFS dataset, we compared the random forest classifier with other classifiers using the correlation coefficient (r) (outcome vs model's parameters): [random forest, r = 0.82; linear regression, r = 0.75; multilayer perceptron or feedforward neural networks, r = 0.71]. The random forest classifier was then used to predict the *outcome* attribute in the test sets (OS and PFS datasets were analyzed separately) (**Fig. 1**).

Application and summary of the random forest classifier

Using our curated dataset formatted for machine learning, we sequestered our data into two sets, 75% for training (to build our random forest classifier), and 25% for testing ('unseen'

data which is used to validate our classifier and provide our final results). On our training set (75% of our data), we build our random forest classifier using a 10-fold cross validation method. Specifically, our random forest classifier is calculated by averaging the performance of 10 random forest classifiers which are individually produced from 10 equal sized sets consisting of their own individual training and testing sets (i.e., each set consisting of 90% data for training, and 10% for testing). Further details of the 10-fold cross validation method can be found by the work by Rafaeilzadeh et al.[29]

We now provide a more detailed description of the random forest method which we use in a supervised learning approach. Step 1: we select random samples from a dataset of interest (in our case, the training set), where each sample can be chosen more than once to create a new 'bootstrapped' dataset (i.e., bootstrapping method[30]). Step 2: we create a decision tree (see decision tree algorithm for details[31]) on the bootstrapped dataset but only use a random subset of attributes (i.e., columns) at each decision step according to the following relationship: $int(log_2(\#\ attributes) + 1)$.[32] Step 3: we repeat steps 1 and 2 at a user-specified amount to create a variety of different decision trees (forming the 'random forest'). Step 4: the random forest classifier is applied by providing the input attributes and having each decision tree predict the outcome. The predicted outcome of each tree is then aggregated through the 'bagging' process[30] to provide a single (aggregated) outcome prediction. Step 5: due to random sampling from step 1, some sample of the data does not appear in the bootstrapped dataset (this 'lost' sample is called the out-of-bag sample, or OOB[33]). The OOB is used to evaluate the OOB error,[33] which is the proportion of OOB samples classified correctly by the random forest classifier. Step 6: The random forest classifier is further optimized by changing the number of columns considered at each decision tree (created in step 2) to try and reduce the OOB error. In

this study, we used the machine learning software WEKA[34] to automatically complete steps 1 – 6 and build our random forest classifier (only default settings were used with a 10-fold cross validation running on a laptop with an Intel Core i7-8850H CPU, and NVIDIA Quadro P3200 GPU). Note, that the mathematical details of the random forest classifier used by the WEKA software can be found in Leo Breiman's work.[35]

Probability of drug sensitivity

We hypothesized that one of the main attributes (considered in our random forest classifier) in predicting clinical outcome was predicated on the percentage of tumors (within the specified cancer type) carrying alterations in at least one of the biomarkers defined for the set of drugs within the regimen. This percentage of tumors carrying biomarker alterations was estimated by analyzing molecular profiles from the TCGA cohorts.

To this end, we formulate the 'probability of drug sensitivity' or PDS (equation (A)) to provide a numerical metric evaluating how well a drug regimen matches to the disease:

(A) $$PDS = (X + Y) * \frac{1}{2}$$

where X = *overall probability of sensitivity*; Y = *probability of sensitivity in absence of resistance*

The *overall probability of sensitivity* (X) to the drug regimen was estimated by the percentage of tumors presenting at least one biomarker of sensitivity to one of the drugs in the regimen (thus, a tumor may include an alteration known to increase drug sensitivity and an alteration known to increase drug resistance). The *probability of sensitivity in the absence of resistance* (Y) to the drug regimen was estimated by the percentage of tumors presenting at least one biomarker of sensitivity and no biomarker of resistance to one of the drugs in the regimen

(tumors only contain biomarkers that increase drug sensitivity). Biomarkers that increase drug sensitivity from deep-deletion, mRNA underexpression or inactivating mutations where labeled with a '-'. All other biomarkers were considered to increase drug sensitivity from gene amplification, mRNA overexpression, or activating mutations (Supporting Information **Table S1**).

To exemplify how we calculated the probability of drug sensitivity (PDS) (equation (A)), the following drug regimen was considered in the pancreatic adenocarcinoma cohort consisting of 129 patients: capecitabine (cytotoxic chemotherapy agent) + ruxolitinib (targeted agent). For this regimen the predictive biomarkers were *MBD4*, *TYMP*, *TYMS* -, *DPYD* - (for capecitabine)[36–39] and *JAK1, JAK2, TYK2* (for ruxolitinib).[40] Entering this gene set (which includes all selected biomarkers for all drugs in the regimen) into cBioPortal,[41] we found that 20 patients (20/129 or 15.5% = *overall probability of sensitivity*) carried at least one marker of sensitivity to the regimen. Out of these 20 patients, 5 patients carried a marker of sensitivity and a marker of resistance to the regimen; thus, 15 patients (15/129 or 11.6% = *probability of sensitivity in the absence of resistance*) carried only markers of sensitivity and no markers of resistance. To acknowledge the uncertainty of these 5 patients being sensitive or resistant to the drug regimen, the PDS is an average between 15.5% and 11.6% (=13.6%). Thus, we hypothesized that approximately 13.6% of pancreatic adenocarcinoma patients may be considered sensitive to a therapeutic regimen including capecitabine in combination with ruxolitinib (see Supporting Information **Fig. S1** for a visual depiction of this example).

Adjusting the PDS for confounding factors

In some instances, the PDS needs to be adjusted to compensate for factors that may impact clinical trial outcome. For example, drug regimens given in different lines of treatment

(with more advanced lines having attenuated response rates), treating population of patients with known molecular alteration status (EGFR-inhibitor to an EGFR-positive population), or using check-point inhibitors in cancer types with high tumor mutation burden (which is associated with increased response rates).[42]

Adjusting PDS for treatment line

We used equation (B) to adjust the PDS for drug regimens used as a second or more line of treatment:

$$\text{(B)} \qquad PDS\ Adjusted = PDS * \frac{Average\ 2nd\ or\ more\ line\ PFS/OS}{Average\ 1st\ line\ PFS/OS}$$

Average 1st line PFS/OS is the average PFS/OS of all 1st line PFS/OS outcomes in the specified cancer type, and the *Average 2nd line or more PFS/OS* is the average PFS/OS of all 2nd or more line PFS/OS outcomes in the specified cancer type.

Adjusting PDS for populations with known molecular alteration status

Some clinical trials present outcomes for drug regimens used in a population where the molecular alteration status is known. For example, erlotinib (inhibitor of EGFR) given to a population of patients with known EGFR-positive expression. In this case, the PDS must be calculated for a population of TCGA tumors that carry EGFR-positive aberrations (creating resemblance between the TCGA and clinical trial populations). Thus, for this example, the *overall probability of sensitivity* (see equation (A)) will be equal to 100%.

Adjusting PDS for tumor mutation burden

High tumor mutation burden (TMB) has been associated as a biomarker of response to check-point inhibitor immunotherapies (e.g., PD-L1 drugs such as nivolumab or prembrolizumab).[42] To incorporate this into the PDS score, we first retrieved 9,167 tumors (containing multiple cancer types) from the TCGA database using the GDAC Firehose website

(https://gdac.broadinstitute.org/ - standardized data run release 2016_01_28). All samples were published and available without restriction of use on May 1st, 2017. Because each tumor presented a total number of mutations, we calculated the 90^{th} percentile value across all tumors. This value was equal to 373 mutations per genome and was used as the threshold in determining what constituted as a 'high tumor mutation burden' in patients (i.e., patients with a TMB above this threshold were considered sensitive to any drug regimen including PD-L1 immunotherapies).

Statistical analysis

Statistical significance between first line and second or more line outcomes was evaluated by Welch's t-test of unequal variance.

The efficacy of PDS as a clinical predictor was verified by the nonparametric Spearman's correlation (r_s) between PDS of drug regimens and median PFS/OS outcomes of their corresponding treatment arms using the GraphPad Prism 6 software (version 6) – this was done for each cancer and outcome type separately. We demonstrated the difference in correlation when adjusting PDS for cofounding factors (treatment line, alteration status, TMB) versus not adjusting the PDS at all.

The nonparametric Spearman correlation was chosen as it evaluates the monotonic (or nonmonotonic) relationship between observed and predicted outcome (or in the case for the PDS evaluation, PDS vs outcome). Importantly, the correlation can be used for linear and nonlinear relationships and is less sensitive to strong outliers present in the data.

The correlation coefficients (r) of our machine learning algorithm, which determines the percentage of the variance described by the algorithm, was provided by the WEKA software for all training sets. We also calculated a nonparametric Spearman correlation (r_s) between predicted

outcomes using our machine learning model (random forest) versus actual outcomes for all test sets.

In addition to the Spearman correlation, we also use the root-mean-square error (RMSE) to measure the difference between values of the predicted model to the actual values observed (e.g., PDS vs outcome, or prediction by the random forest algorithm vs outcome). The RMSE was calculated on observed and predicted outcomes which were first normalized from 0 – 1. The reason for normalization is because the RMSE does not provide any information on whether the relationship between observation and prediction is monotonic or nonmonotonic. Instead, the RMSE is useful for comparing against other RMSE values by providing a relative measure of overall error of the prediction (since every cancer type has a different relationship between how outcome increases/decreases with the prediction, it is useful to normalize to make cross-cancer type evaluations more meaningful). Note, that RMSE may be inflated as RMSE is sensitive to outliers, which do occur in our data.

Lastly, we evaluated 73 PFS and 52 OS randomized trials (collected from a subset of our PFS and OS datasets which otherwise include both randomized and non-randomized trials) to examine whether we can predict which treatment arm will show a better PFS or OS outcome. A binomial probability test was used to evaluate how well our prediction method compared to flipping a coin of equal odds; this was assessed by calculating a z-score and corresponding P-value.

In all statistical P-value calculations, a P value $\leq .05$ was considered statistically significant.

RESULTS

Summary of our clinical trial collection

A total of 351 clinical trials corresponding to 1,102 PFS or OS outcome results were collected from ClinicalTrials.gov and used for analysis (**Table 1**). In the clinical trial data collection (351 clinical trials, 623 PFS outcomes, 479 OS outcomes, 104,758 patients), we included both randomized and non-randomized trials. The mean outcome and 95% confidence interval was calculated for each treatment line (first or second or more), cancer, and outcome type (PFS or OS): *1st line colorectal adenocarcinoma* [mean PFS = 8.5 (95% CI | 8.0, 9.1) months, mean OS = 19.6 (95% CI | 18.1, 21.1) months]; *2nd or more line colorectal adenocarcinoma* [mean PFS = 4.7 (95% CI | 4.2, 5.1) months, mean OS = 11.3 (95% CI | 10.2, 12.4) months]; *1st line pancreatic adenocarcinoma* [mean PFS = 3.5 (95% CI | 3.2, 3.9) months, mean OS = 6.7 (95% CI | 6.2, 7.2) months]; *2nd or more line pancreatic adenocarcinoma* [mean PFS = 2.1 (95% CI | 1.6, 2.7) months, mean OS = 4.8 (95% CI | 3.7, 5.9) months]; *1st line melanoma* [mean PFS = 4.9 (95% CI | 4.1, 5.7) months, mean OS = 13.5 (95% CI | 11.8, 15.1) months]; *2nd or more line melanoma* [mean PFS = 3.7 (95% CI | 2.8, 4.8) months, mean OS = 9.7 (95% CI | 8.4, 10.9) months]; *1st line lung cancer* [mean PFS = 6.2 (95% CI | 5.7, 6.6) months, mean OS = 13.2 (95% CI | 12.3, 14.1) months]; *2nd or more line lung cancer* [mean PFS = 3.7 (95% CI | 3.4, 4.1) months, mean OS = 10.1 (95% CI | 9.4, 10.7) months. The means between 1st line and 2nd or more line outcomes were statistically significant in all datasets analyzed ($P \leq .05$) (**Fig. 2**).

When we correlated PDS with outcome (**Fig. 3**), we considered clinical trials that were randomized and non-randomized (all outcomes were used). The training set (used for machine learning) consisted of 467 PFS outcomes and 369 OS outcomes, derived from clinical trials that

included both randomized and non-randomized trials (~75% of all outcomes). However, the test set consisted of only randomized trials and corresponded to 156 PFS and 110 OS outcomes (~25% of all outcomes) (**Figs. 4 and 5**).

Evaluating the probability of drug sensitivity as a variable of clinical prediction

A total of 115 single agents used alone or in combination (corresponding to 208 unique predictive biomarkers) were reviewed from the literature (see Supporting Information **Table S1**). A probability of drug sensitivity (PDS) (equation (A)) was calculated for each drug regimen administered in clinical arms. For each cancer and outcome type (PFS or OS) analyzed, we correlate the PDS of drug regimens to their median outcome results. Because the clinical arm's outcome may be affected by variables such as treatment line, molecular alteration status of the arm's population, or tumor mutation burden, we also adjusted the PDS for these parameters (see Methods).

The Spearman correlations (r_s) and RMSE values between adjusting PDS for the above parameters versus not adjusting PDS with median outcome were the following: *colorectal adenocarcinoma PFS dataset* [PDS adjusted: $r_s = 0.782$ ($P<0.0001$), RMSE=0.174; PDS not adjusted: $r_s = 0.647$ ($P<0.0001$), RMSE=0.203], *colorectal adenocarcinoma OS dataset* [PDS adjusted: $r_s = 0.860$ ($P<0.0001$), RMSE=0.134; PDS not adjusted: $r_s = 0.660$ ($P<0.0001$), RMSE=0.164], *melanoma PFS dataset* [PDS adjusted: $r_s = 0.711$ ($P<0.0001$), RMSE=0.177; PDS not adjusted: $r_s = 0.560$ ($P<0.0001$), RMSE=0.226], *melanoma OS dataset* [PDS adjusted: $r_s = 0.483$ ($P<0.004$), RMSE=0.235; PDS not adjusted: $r_s = 0.355$ ($P=0.042$), RMSE=0.262], *pancreatic adenocarcinoma PFS dataset* [PDS adjusted: $r_s = 0.650$ ($P<0.0001$), RMSE=0.197; PDS not adjusted: $r_s = 0.591$ ($P=0.002$), RMSE=0.212], *pancreatic adenocarcinoma OS dataset* [PDS adjusted: $r_s = 0.456$ ($P<0.007$), RMSE=0.218; PDS not adjusted: $r_s = 0.391$ ($P=0.022$),

RMSE=0.220], *lung cancer PFS dataset* [PDS adjusted: r_s = 0.615 (P<0.0001), RMSE=0.134; PDS not adjusted: r_s = 0.186 (P=0.0295), RMSE=0.201], *lung cancer OS dataset* [PDS adjusted: r_s = 0.414 (P<0.0001), RMSE=0.158; PDS not adjusted: r_s = 0.019 (P=0.824), RMSE=0.195] (**Fig. 3**).

Adjusting the probability of drug sensitivity (PDS) to treatment line, molecular alteration status, and tumor mutation burden individually

To quantify how each variable of interest (treatment line, molecular alteration status, and tumor mutation burden) individually affected the PDS calculation, the shift in Spearman correlation and RMSE between PDS adjusted versus not adjusted with median outcome was shown for several datasets: *colorectal adenocarcinoma OS dataset* [PDS adjusted only for treatment line: r_s = 0.761 (P<.0001), RMSE=0.154; PDS not adjusted for treatment line: r_s = 0.625 (P<.0001), RMSE=0.164], *lung cancer PFS dataset* [PDS adjusted only for molecular alteration status: r_s = 0.427 (P<.0001), RMSE=0.161; PDS not adjusted for molecular alteration status: r_s = 0.134 (P=.021), RMSE=0.201], *melanoma PFS dataset* [PDS adjusted only for tumor mutation burden: r_s = 0.719 (P<.0001), RMSE=0.211; PDS not adjusted for tumor mutation burden: r_s = 0.685 (P<.0001), RMSE=0.222] (Supporting Information **Fig. S2**).

Predicting clinical outcome using machine learning

The correlation coefficients (r) of our machine learning algorithm for the PFS and OS training sets were equal to 0.82 and 0.70 respectively. The Spearman correlations (r_s) and RMSE values between predicted outcome (using random forest) and actual outcome (results from clinical arms across diseases) for all regimens used in the PFS and OS datasets were the following: the PFS dataset [r_s = 0.879 (P<.0001), RMSE=0.101] and the OS dataset [r_s = 0.878 (P<.0001), RMSE=0.114] (**Fig. 4**).

We also used our random forest algorithm to predict the better outcome arm in 73 PFS and 52 OS randomized clinical trials (i.e., PFS and OS test sets) by comparing our predicted outcomes to the actual outcomes observed. The better outcome arm was predicted in 81% (59/73) of randomized trials in the PFS dataset, while the better outcome arm was predicted in 71% (37/52) of randomized trials in the OS dataset. Statistical significance of these results was evaluated by the binomial probability test: PFS dataset [$z = 5.24$ ($P<.0001$)] and the OS dataset [$z = 2.91$ ($P=.004$)] (**Fig. 5**).

DISCUSSION

Predicting clinical outcome can facilitate optimized drug selections in clinical trial design. In this study, we approached our prediction model using a combination of clinical trial, drug-related biomarker, and molecular profile information. We sequestered our data based on several attributes that were individually tested for their significance and then applied using a random forest algorithm to make predictions.

The first step in creating our prediction model was validating and collecting molecular biomarkers for several drugs. For targeted agents, most biomarkers include direct and/or indirect molecular targets (e.g., erlotinib and its target EGFR).[43] The literature reveals that biomarkers for cytotoxic chemotherapy agents are described by alterations that either upregulate transporters, increase phosphorylation, or enhance the metabolic rate of the agent.[44–46] Using the concept of biomarkers as the foundation, we hypothesized that a variable termed as the "Probability of Drug Sensitivity" or PDS (equation (A)) can be used as one of the main attributes to predict clinical outcome. Principally, PDS provides a numerical metric evaluating how well drug regimens and their biomarkers match to the molecular signature of the disease.

Overall, we show that PDS correlates with clinical outcome and is validated by data that includes numerous drug regimens across cancer types (**Fig. 3**). For example, in a randomized clinical trial (NCT00844649[47]) treating pancreatic ductal adenocarcinoma (PDAC), a combination of gemcitabine and albumin-bound paclitaxel demonstrated a longer PFS compared to the PFS of the standard-of-care gemcitabine alone (PFS = 5.5 versus 3.7 months respectively). The PDS for a drug regimen including gemcitabine and albumin-bound paclitaxel was equal to 34% compared to 21% for gemcitabine alone ($P<.001$)[48] – correctly predicting the better outcome result and further validated by other studies demonstrating the regimen's efficacy.[49–51]

Adjustment to the PDS calculation must be made in certain cases, especially for clinical trials that pre-select patients with specific molecular alterations. For example, a clinical trial aimed at lung cancer (NCT01609543) used erlotinib on EGFR mutation-bearing patients resulting in a PFS of 12.8 months, while another lung cancer trial (NCT01836133) used erlotinib on a random population (not pre-selected for alterations) and demonstrated a PFS of only 2.7 months. If the PDS was calculated on a population of TCGA tumors without considering the molecular alteration status of the clinical trial, it would equal 30%. In contrast, calculating the PDS on a population of TCGA tumors that carry EGFR alterations (creating parity between the clinical trial and TCGA populations) increases the PDS to 86% (which accurately reflects the improved outcome observed between both trials). By adjusting PDS for the molecular alteration status for all regimens and showing the increase in correlation with outcome, we validate molecular alteration status as a strong variable of clinical prediction (see Supporting Information **Fig. S2**).

Another variable tested for outcome prediction was the treatment line. We show that regimens administered as a first line of treatment had significantly improved outcome over regimens administered after the first line of treatment. This result may be unsurprising, as the majority of 2nd or more line of treatments are administered to patients whose initial treatments have failed and it is well known that PFS decreases with progressive lines of treatments.[52] By adjusting our PDS calculation (equation (B); Methods) for regimens administered after the first line, we significantly improved our correlation between PDS and outcome (see Supporting Information **Fig. S2**).

In this study, our prediction model is formulated using machine learning (ML), which has been shown to increase accuracy of predicting cancer mortality by 15-25% in multiple

studies.[14,53,54] Thus, we incorporated ML on a dataset consisting of the following attributes: treatment line, number of cytotoxic, small-molecule, and monoclonal antibody agents, cancer type, drug class, molecular alteration status of the clinical arm's population, PDS and outcome (PFS or OS). Overall, our ML algorithm excelled in predicting clinical outcomes across different cancer types, as the Spearman correlation (r_s) between predicted and actual outcome was 0.878 (P<.0001) and 0.879 (P<.0001) for the PFS and OS test sets respectively (**Fig. 4**). Additionally, our ML algorithm was able to predict the better outcome arm in the majority of randomized trials (PFS: 81% or 59/73 trials, P<.0001; OS: 71% or 37/52 trials, P=.004) (**Fig. 5**). This result is expected when analyzing the variables considered by the ML algorithm. For example, in this study, we verified that PDS alone is a significant variable for clinical prediction. However, our ML algorithm not only considers PDS but also additional important variables such as drug class. This is significant, as indicating the drug class for each drug in a regimen, our ML model considers potential drug-drug interactions.[55–58]

Limitations

Although PDS is one of the significant variables in our ML model, other factors may need to be considered to further improve PDS as a predictor of outcome. For one, the list of biomarkers found in the literature is likely to be incomplete, as the molecular mechanisms of each drug are not always fully understood. Additionally, PDS assumes that all biomarkers are equally 'predictive' of outcome even if that is not always the case. Another variable affecting clinical outcome may be drug dosage; for many types of chemotherapy, dose intensity may be an important correlate of better outcome, although this may not always be the case for some targeted agents.[59,60]

Another limitation is that statistics performed on a population-based dataset is difficult to approximate as differences in age, sex, ethnicity, or histology diagnosis can all affect the efficacy of drug treatments.[61] For example, many lung cancer trials included patients with mixed histology (with lung adenocarcinoma and squamous patients having very different survival rates).[62] Note that because the treated population from the TCGA database (used to derive our PDS score) and the clinical trial database (used to provide the clinical outcome) are different, can result in a less accurate prediction of outcome. Additionally, exclusion/inclusion criteria of clinical trials do not always follow a standard methodology, making comparisons between different trials and classification of trials challenging (e.g., number of patients receiving first line versus more lines of treatments is not always clear). Lastly, our ML approach was based on publicly available databases, thereby limiting the scope of this study as access to greater amounts of data may influence the prediction ability of ML models. Furthermore, in this context, a limitation of the work is that bulk data was used (because that type of data is publicly available whereas patient-level data is generally not available), and one cannot easily transfer from population-level to patient-level applications.

Conclusion

We illustrate the need to integrate multiple disease features including biomarker-based variables (i.e., PDS) in evaluating the clinical benefits of drug selection (i.e., precision medicine). We also show that an ML algorithm that considers a combination of variables such as treatment line, molecular alteration status, drug class, and PDS (among others) can impact clinical predictions significantly. These methods may be useful to predict success of randomized clinical trials and optimize drug development strategies.

In future work, we find that providing a publicly available webserver that can display new findings to be manipulated by users according to their need (e.g., list of drugs and their molecular biomarkers, or a score of predicted clinical outcome based on a user-specified drug regiment) to be valuable. As shown in recent publications,[63,64] delivering publicly accessible webservers can significantly enhance the impact of new findings or approaches, especially in regard to medicinal chemistry and data analysis.[65]

CONFLICT OF INTEREST

Dr. Kurzrock receives research funding from Genentech, Merck Serono, Pfizer, Boehringer Ingelheim, TopAlliance, Takeda, Incyte, Debiopharm, Medimmune, Sequenom, Foundation Medicine, Konica Minolta, Grifols, Omniseq, and Guardant, as well as consultant and/or speaker fees and/or advisory board for X-Biotech, Loxo, Neomed, Pfizer, Actuate Therapeutics, and Roche, has an equity interest in IDbyDNA and CureMatch Inc and serves on the Board of CureMatch and CureMetrix. Mr. Schperberg is an employee at CureMatch. Dr. Tsigelny is a stockholder and CSO of CureMatch. Dr. Richard is a shareholder and worked at CureMatch as CEO and COO. Dr. Boichard has no conflict of interest.

ACKNOWLEDGEMENTS

Funding: This work was supported in part by the Joan and Irwin Jacobs Fund philanthropic fund; and by National Cancer Institute at the National Institutes of Health [grant P30 CA023100] and by CureMatch Inc.

Authors' contributions: Conception and design: AS, AB, RK. Acquisition of data: AS. Analysis and interpretation of data: AS, RK. Development of methodology: AS, AB, RK. Writing, review, and revision of the manuscript: AS, AB, RK. Editing and approval of final version: AS, AB, IT, SR, RK.

DATA ACCESSIBILITY

All data used are publicly available (e.g., TCGA, clinicaltrials.gov), and any datasets used for algorithmic calculations may also be obtained by reasonable request from the corresponding author (note, that in the supplementary information we include our curated list of biomarkers utilized in the prediction score).

REFERENCES

1. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J*. 2014;13:8-17. doi:10.1016/j.csbj.2014.11.005

2. Linardou H, Pentheroudakis G, Varthalitis I, et al. Predictive biomarkers to chemotherapy in patients with advanced melanoma receiving the combination of cisplatin--vinblastine--temozolomide (PVT) as first-line treatment: a study of the Hellenic Cooperative Oncology Group (HECOG). *Anticancer Res*. 2015;35(2):1105-1113.

3. Ren Y, Yao Y, Ma Q, Zhong D. EGFR gene-mutation status correlated with therapeutic decision making in lung adenocarcinoma. *OncoTargets Ther*. 2015;8:3017-3020. doi:10.2147/OTT.S87146

4. Zhao H-Y, Huang H, Hu Z-H, et al. Evaluations of biomarkers associated with sensitivity to 5-fluorouracil and taxanes for recurrent/advanced breast cancer patients treated with capecitabine-based first-line chemotherapy. *Anticancer Drugs*. 2012;23(5):534-542. doi:10.1097/CAD.0b013e32834f7ef4

5. Sicklick JK, Kato S, Okamura R, et al. Molecular profiling of cancer patients enables personalized combination therapy: the I-PREDICT study. *Nat Med*. 2019;25(5):744-750. doi:10.1038/s41591-019-0407-5

6. Rodon J, Soria J-C, Berger R, et al. Genomic and transcriptomic profiling expands precision cancer medicine: the WINTHER trial. *Nat Med*. 2019;25(5):751-758. doi:10.1038/s41591-019-0424-4

7. Adashek JJ, Subbiah IM, Subbiah V. Artificial Intelligence Systems Assisting Oncologists? Resist and Desist or Enlist and Coexist. *The Oncologist*. Published online July 23, 2019. doi:10.1634/theoncologist.2019-0267

8. Taninaga J, Nishiyama Y, Fujibayashi K, et al. Prediction of future gastric cancer risk using a machine learning algorithm and comprehensive medical check-up data: A case-control study. *Sci Rep*. 2019;9(1):1-9. doi:10.1038/s41598-019-48769-y

9. Smith CC, Chai S, Washington AR, et al. Machine-Learning Prediction of Tumor Antigen Immunogenicity in the Selection of Therapeutic Epitopes. *Cancer Immunol Res*. Published online September 12, 2019. doi:10.1158/2326-6066.CIR-19-0155

10. Way GP, Sanchez-Vega F, La K, et al. Machine Learning Detects Pan-cancer Ras Pathway Activation in The Cancer Genome Atlas. *Cell Rep*. 2018;23(1):172-180.e3. doi:10.1016/j.celrep.2018.03.046

11. Li J, Lenferink AEG, Deng Y, et al. Identification of high-quality cancer prognostic markers and metastasis network modules. *Nat Commun*. 2010;1:34. doi:10.1038/ncomms1033

12. Mobadersany P, Yousefi S, Amgad M, et al. Predicting cancer outcomes from histology and genomics using convolutional networks. *Proc Natl Acad Sci U S A*. 2018;115(13):E2970-E2979. doi:10.1073/pnas.1717139115

13. Yousefi S, Amrollahi F, Amgad M, et al. Predicting clinical outcomes from large scale cancer genomic profiles with deep survival models. *Sci Rep*. 2017;7(1):11707. doi:10.1038/s41598-017-11817-6

14. Cruz JA, Wishart DS. Applications of Machine Learning in Cancer Prediction and Prognosis. *Cancer Inform*. 2007;2:59-77.

15. Chou K-C. Impacts of bioinformatics to medicinal chemistry. *Med Chem Shariqah United Arab Emir*. 2015;11(3):218-234. doi:10.2174/1573406411666141229162834

16. Chou KC. Prediction of protein cellular attributes using pseudo-amino acid composition. *Proteins*. 2001;43(3):246-255. doi:10.1002/prot.1035

17. Chou K-C. Using amphiphilic pseudo amino acid composition to predict enzyme subfamily classes. *Bioinforma Oxf Engl*. 2005;21(1):10-19. doi:10.1093/bioinformatics/bth466

18. Liu B, Liu F, Wang X, Chen J, Fang L, Chou K-C. Pse-in-One: a web server for generating various modes of pseudo components of DNA, RNA, and protein sequences. *Nucleic Acids Res*. 2015;43(Web Server issue):W65-W71. doi:10.1093/nar/gkv458

19. Liu B, Wu H, Chou K-C. Pse-in-One 2.0: An Improved Package of Web Servers for Generating Various Modes of Pseudo Components of DNA, RNA, and Protein Sequences. *Nat Sci*. 2017;9(4):67-91. doi:10.4236/ns.2017.94007

20. Chou K-C. Proposing 5-Steps Rule Is a Notable Milestone for Studying Molecular Biology. *Nat Sci*. 2020;12(3):74-79. doi:10.4236/ns.2020.123011

21. Use Chou's 5-Steps Rule to Predict Remote Homology Proteins by Merging Grey Incidence Analysis and Domain Similarity Analysis. Accessed May 27, 2020. https://www.scirp.org/journal/paperinformation.aspx?paperid=99082

22. Chou K-C. The Development of Gordon Life Science Institute: Its Driving Force and Accomplishments. *Nat Sci*. 2020;12(4):720-726. doi:10.4236/ns.2020.124018

23. The Cancer Genome Atlas Program - National Cancer Institute. Published June 13, 2018. Accessed May 26, 2020. https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga

24. Alam Z, Rahman S, Rahman S. A Random Forest based predictor for medical data classification using feature ranking. *Inform Med Unlocked*. 2019;15:100180. doi:10.1016/j.imu.2019.100180

25. Basu S, Kumbier K, Brown JB, Yu B. Iterative random forests to discover predictive and stable high-order interactions. *Proc Natl Acad Sci*. 2018;115(8):1943-1948. doi:10.1073/pnas.1711236115

26. Khalilia M, Chakraborty S, Popescu M. Predicting disease risks from highly imbalanced data using random forest. *BMC Med Inform Decis Mak*. 2011;11(1):51. doi:10.1186/1472-6947-11-51

27. Rigatti SJ. Random Forest. *J Insur Med*. 2017;47(1):31-39. doi:10.17849/insm-47-01-31-39.1

28. Rahman R, Dhruba SR, Ghosh S, Pal R. Functional random forest with applications in dose-response predictions. *Sci Rep*. 2019;9(1):1-14. doi:10.1038/s41598-018-38231-w

29. Refaeilzadeh P, Tang L, Liu H. Cross-Validation. In: LIU L, ÖZSU MT, eds. *Encyclopedia of Database Systems*. Springer US; 2009:532-538. doi:10.1007/978-0-387-39940-9_565

30. Breiman L. Bagging predictors. *Mach Learn*. 1996;24(2):123-140. doi:10.1007/BF00058655

31. Decision Tree Algorithm - an overview | ScienceDirect Topics. Accessed May 26, 2020. https://www.sciencedirect.com/topics/computer-science/decision-tree-algorithm

32. RandomForest. Accessed May 26, 2020. https://weka.sourceforge.io/doc.dev/weka/classifiers/trees/RandomForest.html

33. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning*. Vol 103. Springer New York; 2013. doi:10.1007/978-1-4614-7138-7

34. Frank E, Hall M, Witten I. *Data Mining: Practical Machine Learning Tools and Techniques*. Fourth. Morgan Kaufmann; 2016.

35. Breiman L. Random Forests. *Mach Learn*. 2001;45(1):5-32. doi:10.1023/A:1010933404324

36. Wisniewska-Jarosinska M, Sliwinski T, Kasznicki J, et al. Cytotoxicity and genotoxicity of capecitabine in head and neck cancer and normal cells. *Mol Biol Rep*. 2011;38(6):3679-3688. doi:10.1007/s11033-010-0482-7

37. Marangoni E, Laurent C, Coussy F, et al. Capecitabine Efficacy Is Correlated with TYMP and RB1 Expression in PDX Established from Triple-Negative Breast Cancers. *Clin Cancer Res Off J Am Assoc Cancer Res*. Published online February 20, 2018. doi:10.1158/1078-0432.CCR-17-3490

38. Joerger M, Huitema ADR, Boot H, et al. Germline TYMS genotype is highly predictive in patients with metastatic gastrointestinal malignancies receiving capecitabine-based chemotherapy. *Cancer Chemother Pharmacol*. 2015;75(4):763-772. doi:10.1007/s00280-015-2698-7

39. Henricks LM, Siemerink EJM, Rosing H, et al. Capecitabine-based treatment of a patient with a novel DPYD genotype and complete dihydropyrimidine dehydrogenase deficiency. *Int J Cancer*. 2018;142(2):424-430. doi:10.1002/ijc.31065

40. Works MG, Yin F, Yin CC, et al. Inhibition of TYK2 and JAK1 Ameliorates Imiquimod-Induced Psoriasis-like Dermatitis by Inhibiting IL-22 and the IL-23/IL-17 axis. *J Immunol Baltim Md 1950*. 2014;193(7):3278-3287. doi:10.4049/jimmunol.1400205

41. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. 2012;2(5):401-404. doi:10.1158/2159-8290.CD-12-0095

42. Goodman AM, Kato S, Bazhenova L, et al. Tumor Mutational Burden as an Independent Predictor of Response to Immunotherapy in Diverse Cancers. *Mol Cancer Ther*. Published online January 1, 2017:molcanther.0386.2017. doi:10.1158/1535-7163.MCT-17-0386

43. Targeted Cancer Therapies. National Cancer Institute. Accessed October 4, 2017. https://www.cancer.gov/about-cancer/treatment/types/targeted-therapies/targeted-therapies-fact-sheet

44. Voutsadakis IA. Molecular predictors of gemcitabine response in pancreatic cancer. *World J Gastrointest Oncol*. 2011;3(11):153-164. doi:10.4251/wjgo.v3.i11.153

45. Gautam A, Li Z-R, Bepler G. RRM1-induced metastasis suppression through PTEN-regulated pathways. *Oncogene*. 2003;22(14):2135-2142. doi:10.1038/sj.onc.1206232

46. Plunkett W, Huang P, Searcy CE, Gandhi V. Gemcitabine: preclinical pharmacology and mechanisms of action. *Semin Oncol*. 1996;23(5 Suppl 10):3-15.

47. Von Hoff DD, Ervin T, Arena FP, et al. Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. *N Engl J Med*. 2013;369(18):1691-1703. doi:10.1056/NEJMoa1304369

48. Vogel A, Römmler-Zehrer J, Li JS, McGovern D, Romano A, Stahl M. Efficacy and safety profile of nab-paclitaxel plus gemcitabine in patients with metastatic pancreatic cancer treated to disease progression: a subanalysis from a phase 3 trial (MPACT). *BMC Cancer*. 2016;16. doi:10.1186/s12885-016-2798-8

49. Blomstrand H, Scheibling U, Bratthäll C, Green H, Elander NO. Real world evidence on gemcitabine and nab-paclitaxel combination chemotherapy in advanced pancreatic cancer. *BMC Cancer*. 2019;19(1):40. doi:10.1186/s12885-018-5244-2

50. Vogel A, Pelzer U, Salah-Eddin A-B, Köster W. First-line nab-paclitaxel and gemcitabine in patients with metastatic pancreatic cancer from routine clinical practice. *In Vivo*. 2014;28(6):1135-1140.

51. Portal A, Pernot S, Tougeron D, et al. Nab-paclitaxel plus gemcitabine for metastatic pancreatic adenocarcinoma after Folfirinox failure: an AGEO prospective multicentre cohort. *Br J Cancer*. 2015;113(7):989-995. doi:10.1038/bjc.2015.328

52. Park IH, Lee KS, Ro J. Effects of Second and Subsequent Lines of Chemotherapy for Metastatic Breast Cancer. *Clin Breast Cancer*. 2015;15(1):e55-e62. doi:10.1016/j.clbc.2014.09.001

53. Colozza M, Cardoso F, Sotiriou C, Larsimont D, Piccart MJ. Bringing Molecular Prognosis and Prediction to the Clinic. *Clin Breast Cancer*. 2005;6(1):61-76. doi:10.3816/CBC.2005.n.010

54. Bottaci L, Drew PJ, Hartley JE, et al. Artificial neural networks applied to outcome prediction for colorectal cancer patients in separate institutions. *Lancet Lond Engl*. 1997;350(9076):469-472. doi:10.1016/S0140-6736(96)11196-X

55. Kanzawa F, Saijo N. In vitro interaction between gemcitabine and other anticancer drugs using a novel three-dimensional model. *Semin Oncol*. 1997;24(2 Suppl 7):S7-8-S7-16.

56. Giovannetti E, Danesi R, Mey V, Nannizzi S, Pasqualetti G, Del Tacca M. *In Vitro Studies on Gemcitabine Combinations with Other Antiblastics*. Vol 17 Suppl 5.; 2006. doi:10.1093/annonc/mdj943

57. Nightingale G, Pizzi LT, Barlow A, et al. The prevalence of major drug-drug interactions in older adults with cancer and the role of clinical decision support software. *J Geriatr Oncol*. Published online March 3, 2018. doi:10.1016/j.jgo.2018.02.001

58. Sorf A, Hofman J, Kučera R, Staud F, Ceckova M. Ribociclib shows potential for pharmacokinetic drug-drug interactions being a substrate of ABCB1 and potent inhibitor of ABCB1, ABCG2 and CYP450 isoforms in vitro. *Biochem Pharmacol*. 2018;154:10-17. doi:10.1016/j.bcp.2018.04.013

59. Jain RK, Lee JJ, Hong D, et al. Phase I Oncology Studies: Evidence That in the Era of Targeted Therapies, Patients on Lower Doses Do Not Fare Worse. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2010;16(4):1289. doi:10.1158/1078-0432.CCR-09-2684

60. Foote M. The Importance of Planned Dose of Chemotherapy on Time: Do We Need to Change Our Clinical Practice? *The Oncologist*. 1998;3(5):365-368.

61. Yasuda SU, Zhang L, Huang S-M. The role of ethnicity in variability in response to drugs: focus on clinical pharmacology studies. *Clin Pharmacol Ther*. 2008;84(3):417-423. doi:10.1038/clpt.2008.141

62. Kawase A, Yoshida J, Ishii G, et al. Differences Between Squamous Cell Carcinoma and Adenocarcinoma of the Lung: Are Adenocarcinoma and Squamous Cell Carcinoma Prognostically Equal? *Jpn J Clin Oncol*. 2012;42(3):189-195. doi:10.1093/jjco/hyr188

63. Chen W, Tang H, Ye J, Lin H, Chou K-C. iRNA-PseU: Identifying RNA pseudouridine sites. *Mol Ther Nucleic Acids*. 2016;5:e332. doi:10.1038/mtna.2016.37

64. Liu B, Fang L, Long R, Lan X, Chou K-C. iEnhancer-2L: a two-layer predictor for identifying enhancers and their strength by pseudo k-tuple nucleotide composition. *Bioinforma Oxf Engl*. 2016;32(3):362-369. doi:10.1093/bioinformatics/btv604

65. Chou K-C. An Unprecedented Revolution in Medicinal Chemistry Driven by the Progress of Biological Science. *Curr Top Med Chem*. 2017;17(21):2337-2358. doi:10.2174/1568026617666170414145508

66. Chou KC. Low-frequency collective motion in biomacromolecules and its biological functions. *Biophys Chem*. 1988;30(1):3-48. doi:10.1016/0301-4622(88)85002-6

FIGURE LEGENDS

Figure 1. Graphical representation of our overall design and methods of this study

a) We collected data from three sources: molecular profiles from TCGA, clinical trials from ClinicalTrials.gov, and drug-biomarker information from literature.

b) This data was entered into a table with the following attributes per intervention arm: treatment line (first or second or more), number of cytotoxic chemotherapies, number of small-molecule inhibitors, number of monoclonal antibodies, 1 column per drug class, disease, molecular alteration status of the clinical arm's population, PDS (equation (A)), and outcome (PFS or OS, normalized from 0 - 1). Machine learning software was used to build our random forest decision algorithm using ~75% of our data (training set). Our algorithm was used to predict the outcome for the rest of the data (test set).

This figure, along with Fig. S1, provide a graphical approach to our methods, which is useful for analyzing complication relationships as emphasized by a comprehensive review.[66]

Abbreviations: OS = overall survival; PDS = probability of drug sensitivity; PFS = progression-free survival; TCGA = the cancer genome atlas

Figure 2. Representation of PFS or OS outcomes among different cancer types and treatment lines

A total of 351 clinical trials with 1,102 PFS or OS outcomes were analyzed. The median PFS or OS outcomes for all clinical arms are shown by grey dots, the mean is represented by a black line, and the 95% confidence interval is represented by red bars (clinical trials were grouped based on treatment line and disease considered, and a mean outcome – PFS or OS – was calculated for each group). The outcome results for the PFS dataset (623 outcomes) are shown in graph **a)**, while the outcome results for the OS dataset (479 outcomes) are shown in graph **b)**. Overall, we demonstrate that drug regimens administered as a first line of treatment had better clinical outcome compared to regimens administered after the first line of treatment (confirmed by Welch's t-test).

Abbreviations: ALL = data from colorectal adenocarcinoma, pancreatic adenocarcinoma, melanoma and NSCLC combined; NSCLC = non-small cell lung cancer (adenocarcinoma, squamous cell carcinoma, or mixed); OS = overall survival; PFS = progression-free survival

Figure 3. Correlation between probability of sensitivity (adjusted and not adjusted) and median outcome (PFS or OS)

PDS was verified as an attribute of clinical prediction by calculating the Spearman correlation (r_s) between PDS of drug regimens and median PFS or OS outcomes from clinical arms administering these regimens (randomized and non-randomized trials). The median PFS/OS was also averaged across clinical trials administering the same regimen. This was done for each

cancer and outcome type separately: **a)** colorectal adenocarcinoma; **b)** melanoma; **c)** pancreatic adenocarcinoma; and **d)** NSCLC (adenocarcinoma, squamous cell carcinoma, or mixed). The blue line represents the PDS when not adjusted to treatment line, molecular alteration status, and TMB (see methods), and the red line represents the PDS when adjusted for these parameters. The green markers represent the median outcome (PFS or OS) in months for each regimen administered in the clinical arms. For every cancer and outcome type, the correlation between PDS and outcome increased when the PDS was adjusted for treatment line, molecular alteration status, and TMB. All datasets represented statistically significant correlations (P<.05) except for the lung OS dataset (graph **d)**) where PDS was not adjusted (r_s=.019, P=.824, RMSE=0.195). The colorectal OS dataset (graph **a)**) where PDS was adjusted presented the highest correlation (r_s=0.860, P<.0001, RMSE=0.134).

Abbreviations: NSCLC = non-small cell lung cancer (adenocarcinoma, squamous cell carcinoma, or mixed); OS = overall survival; PDS = probability of drug sensitivity; PFS = progression-free survival; TMB = tumor mutation burden

Figure 4. Correlation between predicted outcome using machine learning and actual outcome using the PFS and OS test sets

We derived a machine learning algorithm (random forest) using 75% of our data composed of several attributes including treatment line (first or second or more), disease, number of cytotoxic chemotherapies, number of small-molecule inhibitors, number of monoclonal antibodies, drug class, molecular alteration status of the clinical arm's treated population, PDS (equation (A) in the **Methods**), and outcome (PFS or OS, normalized from 0 - 1). The rest of the data (25%) was used as our test set (consisting of only randomized trials) where our algorithm predicted the outcome. This analysis was done for the PFS and OS datasets separately and both show similar and highly statistically significant results: **a)** The PFS test set was composed of 156 outcomes, and the Spearman correlation (r_s) between predicted and actual outcome was equal to 0.879 (P<.0001); **b)** The OS test set was composed of 110 outcomes, and the Spearman correlation (r_s) between predicted and actual outcome was equal to 0.878 (P<.0001); The r_s value was also provided for each cancer type separately and all datasets presented statistically significant results (P<.0001).

Abbreviations: OS = overall survival; PDS = probability of drug sensitivity; PFS = progression-free survival

Figure 5. Predicting the better outcome arm in randomized trials

a) Our PFS and OS test sets (see **Figure 4**) included 73 and 52 randomized clinical trials, respectively. We used our machine learning algorithm to predict the better outcome arm by comparing our predicted outcome to the actual outcome observed. For the PFS and OS test sets, we correctly predicted the better outcome arm in 81% (59/73) and 71% (37/52) of randomized trials respectively. To test if these predictions are better than flipping a coin, we used the binomial probability test (z-score) and obtained statistically significant results for both sets (PFS: P<.0001; OS: P=.004).

b) An example of predicting the better clinical arm is provided for a randomized clinical trial (NCT00326599) treating lung cancer (adenocarcinoma or squamous cell carcinoma). In this trial, the experimental arm (using AZD2171 (cediranib), carboplatin, and gemcitabine) demonstrated a better PFS outcome compared to the comparator arm (using carboplatin and gemcitabine). Correspondingly, our prediction score was higher for the experimental arm than the comparator arm.

Abbreviations: OS = overall survival; PFS = progression-free survival

Table 1. Summary of clinical trial and molecular data

Clinical trial collection summary*	**Total**
Number of clinical trials	351
Number of patients	104,758
Number of PFS outcomes	623
Number of OS outcomes	479
Molecular profile, biomarker, and drug collection summary**	Total
Number of molecular profiles	939
Number of unique drugs considered	115
Number of unique biomarkers considered	208

*A total of 351 clinical trials (colorectal adenocarcinoma: N = 94, melanoma: N = 51, pancreatic adenocarcinoma: N=37, lung cancer: N=169) were retrieved from ClinicalTrials.gov that also met our selection criteria (see methods)

**Using cBioPortal, we downloaded the TCGA cohorts per disease and collected 939 molecular profiles with genomic, transcriptomic, and proteomic information (colorectal adenocarcinoma: N=277, melanoma: N=169, pancreatic adenocarcinoma: N=129, lung adenocarcinoma: N=186, and lung squamous cell carcinoma: 178). All molecular profiles where sequenced from tumors with lymph nodes at the N1, N2, N3 locations or diagnosed with M1 cancer. In our study, 115 unique drugs corresponding to 208 biomarkers where collected and validated in literature and FDA documentation

Abbreviations: OS = overall survival; PFS = progression-free survival

Step 1 - Collect data from the following sources

1) Molecular profiles

2) Clinical trials with outcomes

3) Biomarker info from literature

Step 2 - Create a database for machine learning

Drug regimen	Treatment line	Disease	# of small-molecule inhibitors	# of cytotoxic chemotherapies	# of monoclonal antibodies	Drug-Class 1: VEGFR-inh	Drug-Class 2: EGFR-inh	Drug-Class 3: DNA-dmg	Molecular alteration status	PDS	Normalized Outcome (PFS or OS)
Bevacizumab, Gemcitabine	First	Pancreatic adenocarcinoma	0	1	1	YES	NO	YES	VEGFR-positive	32%	0.5
Erlotinib, Bevacizumab	Second or more	Colorectal adenocarcinoma	1	0	1	YES	YES	NO	random	40%	0.8
Erlotinib, Gemcitabine	First	Pancreatic adenocarcinoma	1	1	0	NO	YES	YES	random	36%	?

Machine learning (Random forest)

Training set: 75% of data

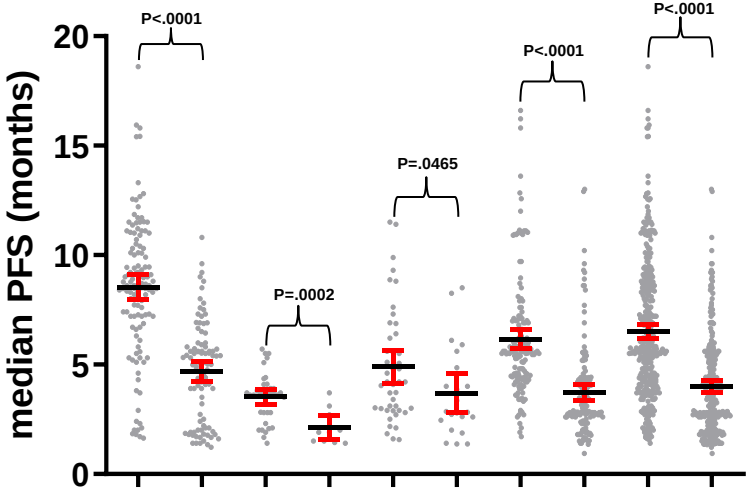
Test set: 25% of data

X = overall probability of sensitivity
Y = probability of sensitivity in absence of resistance

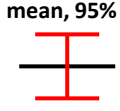
PDS = (X + Y)/2

a)

median PFS (months) between treatment lines

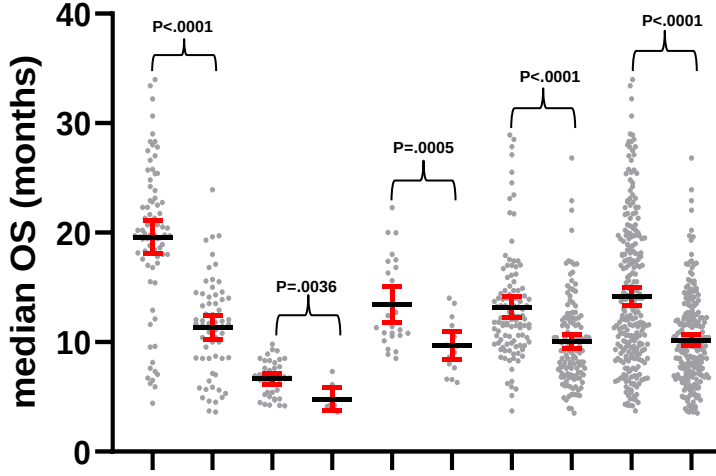


CRC		PANC		MEL		NSCLC		ALL	
1st	>=2nd	1st	>=2nd	1st	>=2nd	1st	>=2nd	1st	>=2nd
120	92	38	10	45	22	154	142	357	266



b)

median OS (months) between treatment lines



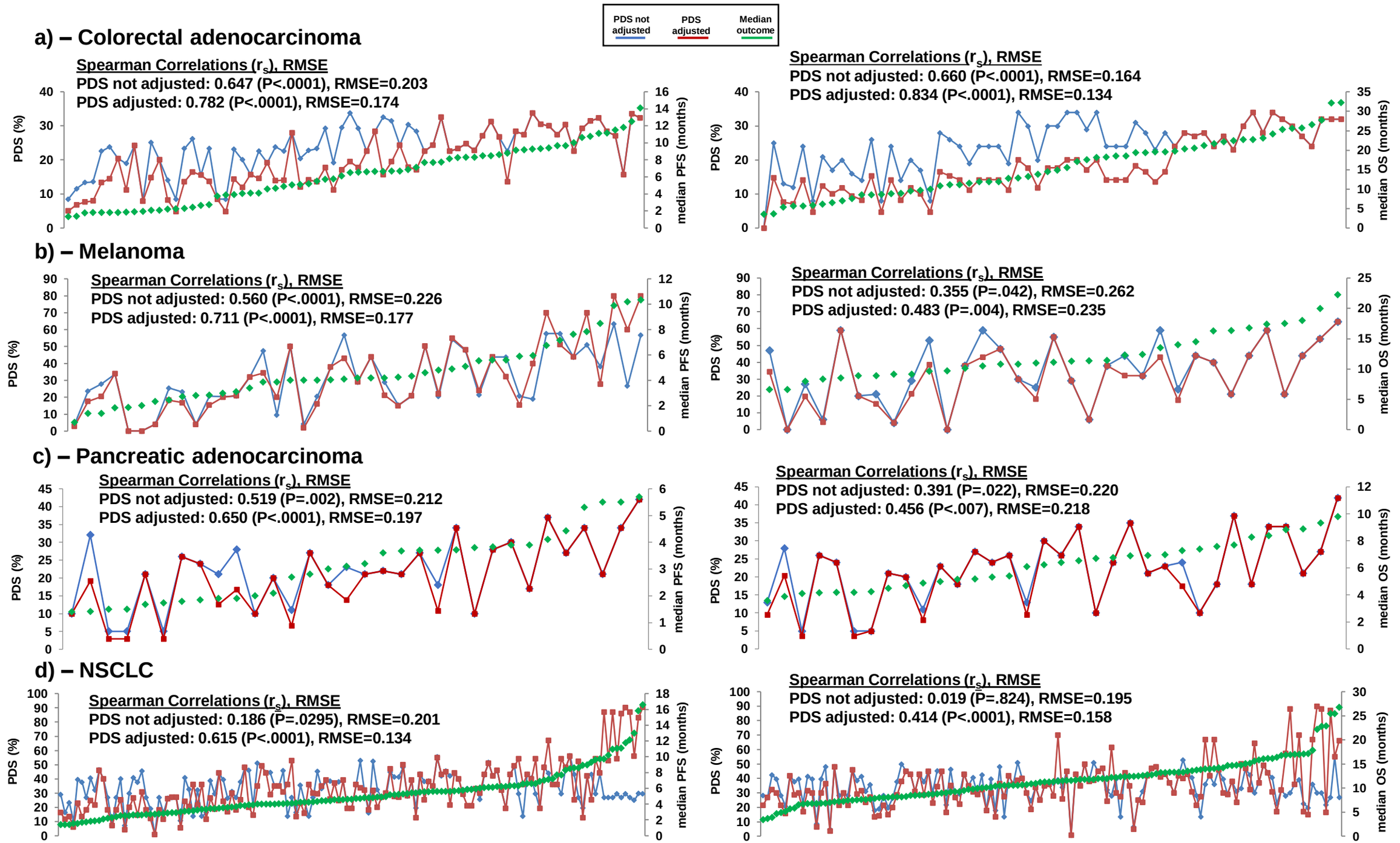
CRC		PANC		MEL		NSCLC		ALL	
1st	>=2nd	1st	>=2nd	1st	>=2nd	1st	>=2nd	1st	>=2nd
80	63	41	8	25	16	108	138	254	225

treatment line

# of outcomes

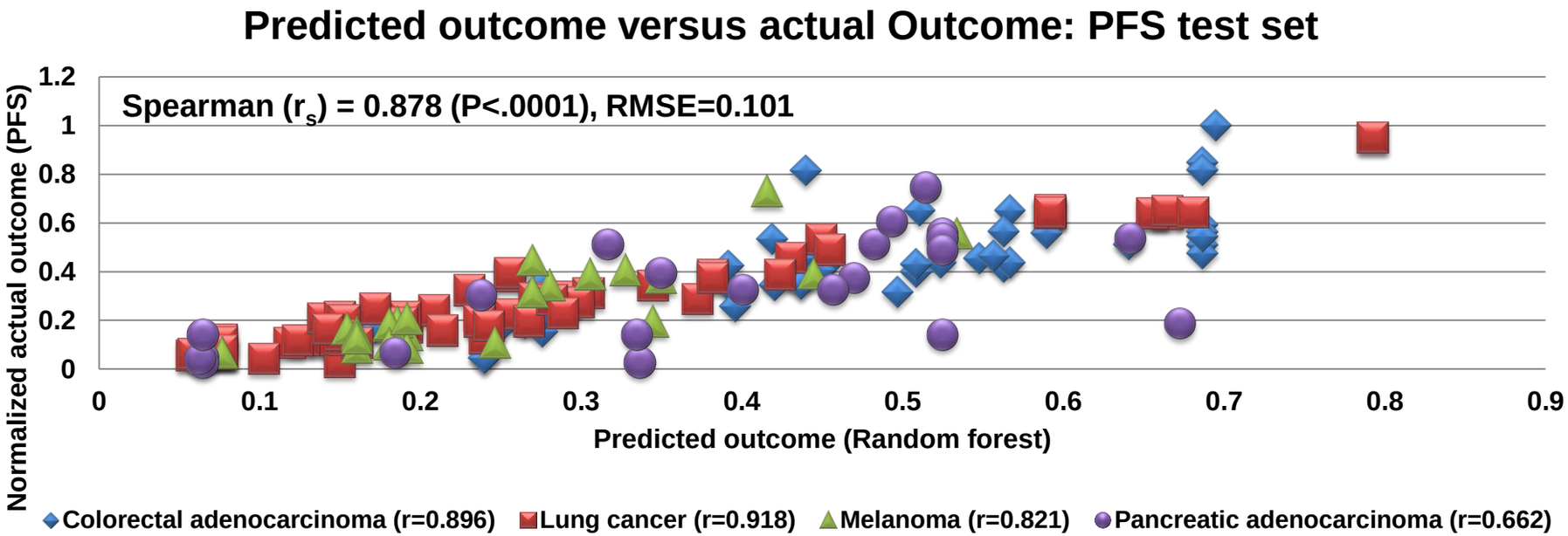
Correlation between PDS and median PFS

Correlation between PDS and median OS

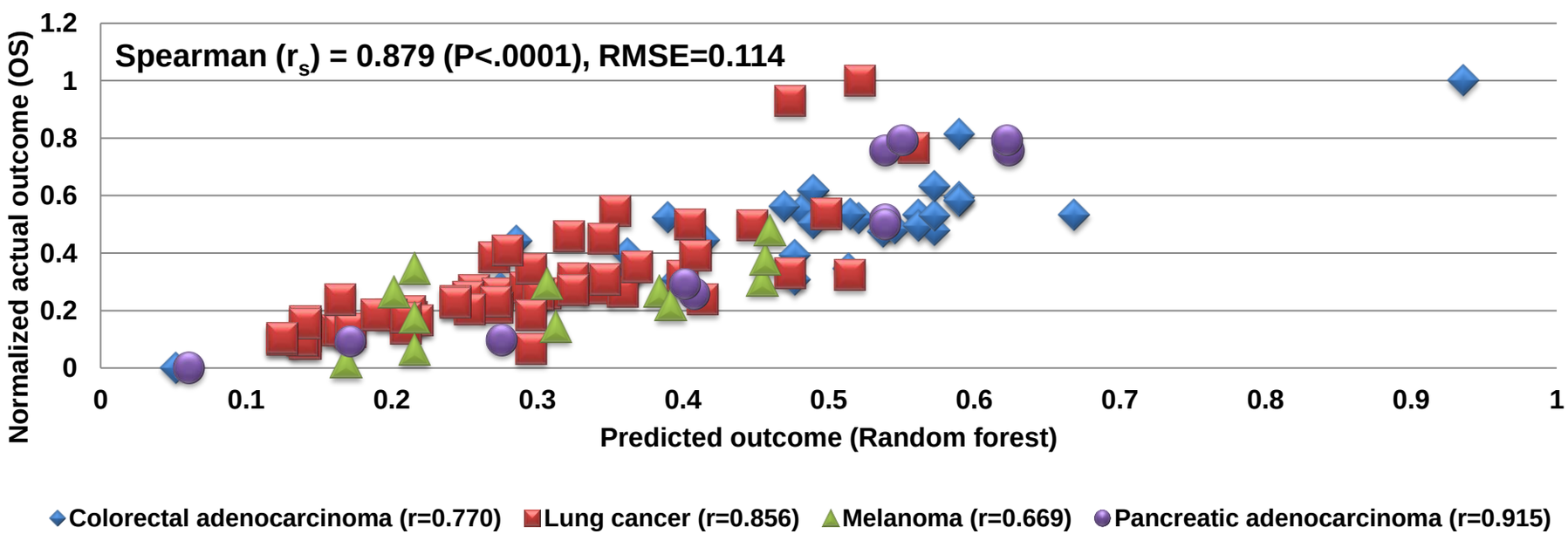


X – axis represents different drug regimens

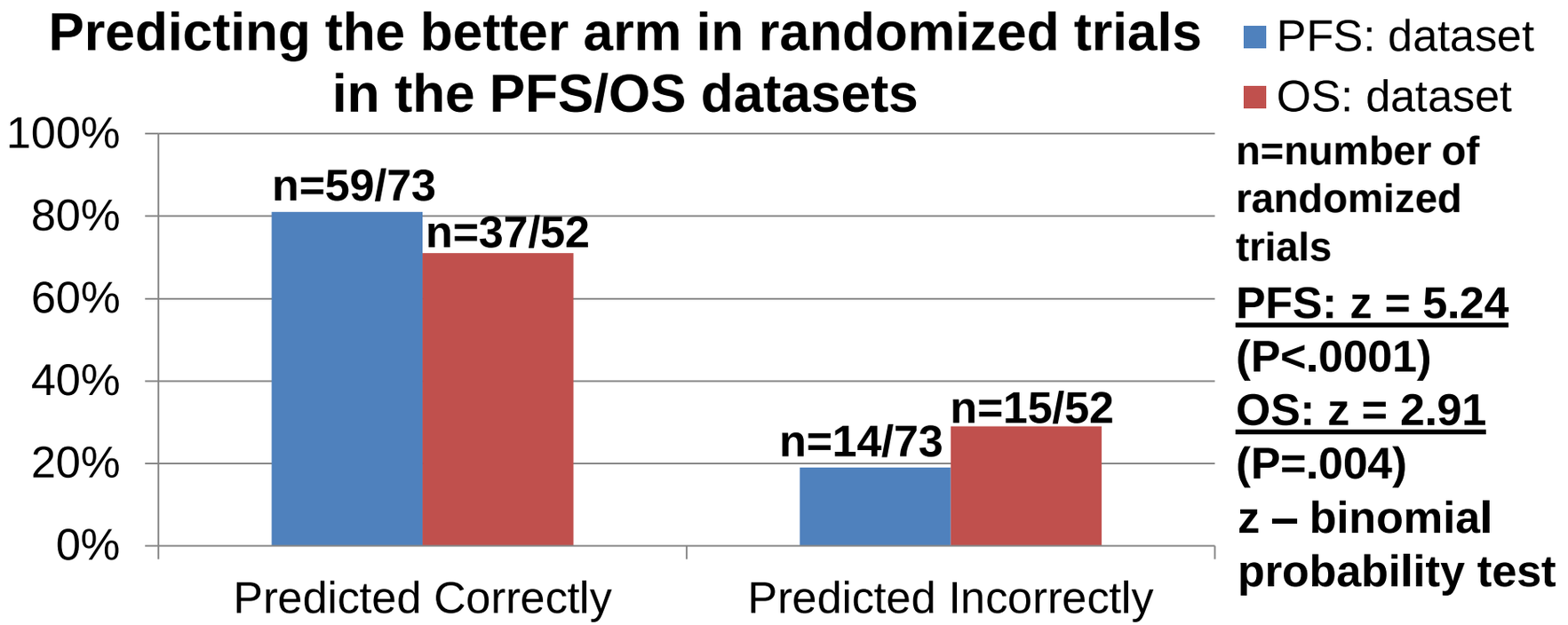
a)


Predicted outcome versus actual Outcome: PFS test set
Spearman (r_s) = 0.878 (P<.0001), RMSE=0.101
Normalized actual outcome (PFS)
1.2
1
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
Predicted outcome (Random forest)
Colorectal adenocarcinoma (r=0.896)
Lung cancer (r=0.918)
Melanoma (r=0.821)
Pancreatic adenocarcinoma (r=0.662)

b)


Predicted outcome versus actual Outcome: OS test set
Spearman (r_s) = 0.879 (P<.0001), RMSE=0.114
Normalized actual outcome (OS)
1.2
1
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Predicted outcome (Random forest)
Colorectal adenocarcinoma (r=0.770)
Lung cancer (r=0.856)
Melanoma (r=0.669)
Pancreatic adenocarcinoma (r=0.915)

a)



b)

NCT ID	Disease	Drugs Used	PFS (months)	Prediction Score	Predicted Correctly?
NCT00326599 (randomized clinical trial)	Lung cancer	Active Comparator Arm: Gemcitabine, Carboplatin	4.5	36%	YES
		Experimental Arm: Cediranib, Gemcitabine, Carboplatin	6.3	40%	

SUPPLEMENTARY FIGURE LEGENDS

Figure S1: Example of calculating the PDS for capecitabine + ruxolitinib administered to a pancreatic adenocarcinoma population

PDS is a percentage reflecting the status of genomic, transcriptomic, and proteomic biomarkers in the population of interest. Here, we show an example of calculating the PDS for a drug combination including capecitabine and ruxolitinib administered to a pancreatic adenocarcinoma population. The individual steps are the following: step 1) describe the biomarkers for capecitabine and ruxolitinib; steps 2-3) download molecular profiles from the pancreatic adenocarcinoma TCGA cohort (129 patients) and count the number of patients without biomarker alterations or biomarker alterations resistant to capecitabine and ruxolitinib (109 patients), the number of patients with biomarker alterations only sensitive to this regimen (15 patients), and the number of patients with biomarker alterations both sensitive and resistant to this regimen (5 patients); step 4) the PDS is calculated using equation (A).

Abbreviations: OS = overall survival; PDS = probability of drug sensitivity; PFS = progression-free survival; TCGA = the cancer genome atlas

Figure S2: Adjusting PDS for line of treatment, molecular alteration status, and tumor mutation burden individually

Three key variables were hypothesized to impact clinical outcome: line of treatment, molecular alteration status, and tumor mutation burden (TMB). We demonstrate the significance of these variables by their effect on the PDS calculation. These graphs show correlations between outcome (PFS or OS) and PDS (adjusted versus not adjusted). The blue diamonds represent data points for which PDS was not adjusted for the variable of interest. The green triangles represent regimens where PDS may be adjusted, and the red squares represent the results when PDS was adjusted. The Spearman correlation (r_s) increased when adjusting PDS versus not adjusting PDS in every case.

a) Equation (B) (see Methods) was used to adjust PDS for drug regimens administered in different lines of treatment. The colorectal adenocarcinoma (OS) dataset exemplifies the significance of this variable as this dataset contained the largest difference in outcome between first and second or more line of treatments.

b) For clinical arms where the molecular alteration status is known (e.g., EGFR inhibitor to EGFR-positive patients), the PDS needs to be adjusted. This was done by calculating the PDS from a cohort of molecular profiles that also carry the alteration described by the clinical arm. To demonstrate the significance of this variable, we use the lung cancer PFS dataset as this dataset included the largest number of clinical arms (66) where the molecular alteration status was known.

c) As patients with high TMB are sensitive to anti-PD-L1/PD-1 checkpoint immunotherapies, PDS needs to be adjusted to adequately reflect these cases. The melanoma PFS dataset is shown as patients with this cancer type typically present high TMB.

Abbreviations: OS = overall survival; PDS = probability of drug sensitivity; PFS = progression-free survival; TCGA = The Cancer Genome Atlas; TMB = tumor mutation burden

Table S1: List of drugs used and their biomarkers

Drugs	Drug-related Biomarkers*
dasatinib	ABL1, ABL2, KIT, SRC
nab-paclitaxel	SPARC, ERBB2, TUBB3 -, TLE3
mk2206	AKT1, AKT2, AKT3
linsitinib	IGF1R, IGF1, IGF2, INSRR
everolimus	FKBP1A, MTOR, RPTOR, TSC1, TSC2
capecitabine	BIRC5, CA9, TYMP, TYMS -, DPYD -
oxaliplatin	ATM -, BRCA1 -, BRCA2 -, ERCC1 -, ERCC2 -, ERCC6 -, FANCC -, PALB2 -, BRIP1 -, BAP1 -
cetuximab	EGFR, TGFA, HBEGF, AREG, BTC, EPGN, KRAS -
irinotecan	TOP1, APTX -, BRCA1 -
pembrolizumab	CD274, PDCD1LG2, PDCD1
azacitidine	DNMT1, DNMT3A, DNMT3B, TET1, TET2, TRDMT1
bevacizumab	FLT1, KDR, PGF, VEGFA, VEGFB, VEGFC, FIGF, CA9
panitumumab	EGFR, TGFA, HBEGF, AREG, BTC, EPGN, KRAS -
ganitumab	IGF1R, IGF1, IGF2, INSRR
ruxolitinib	JAK1, JAK2, JAK3, PTPRD, STAT3, FER, TYK2
regorafenib	BRAF, FLT1, FLT4, KDR, KIT, PDGFRB, RAF1, RET
aflibercept	FLT1, KDR, PGF, VEGFA, VEGFB, VEGFC, FIGF, VHL
cediranib	KDR, KIT, FLT4, FLT1, PDGFRB, FGFR1, PDGFRA
tivozanib	FLT1, FLT4, EPHB2, FLT3, PDGFRA, PDGFRB
vemurafenib	SRMS, TNK2, BRAF, RAF1, MAP4K5
nintedanib	FGFR2, FLT1, FLT3, FLT4, KDR, LCK
axitinib	FLT1, FLT4, KDR, KIT, PDGFRA, PDGFRB, FLT3
olaparib	ATM, ATR, BAP1, BARD1, BRCA1, BRCA2, BRIP1, CDK12, EMSY, ERCC1, FANCA, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCL, PALB2, PARP1, PARP2
saracatinib	SRC, LCK, YES1, EGFR, LYN, FYN, FGR, BLK, KRAS -
linifanib	FLT1, CSF1R, KDR, FLT3, KIT, PDGFRB, VEGFB, VEGFA
afatinib	EGFR, TGFA, HBEGF, AREG, BTC, EPGN, KRAS -
figitumumab	IGF1R, IGF1, IGF2, INSRR
sunitinib	CSF1R, FLT1, KIT, PDGFRB, CSF1R
conatumumab	TNFRSF10B
vorinostat	HDAC1, HDAC10, HDAC11, HDAC2, HDAC3, HDAC4, HDAC5, HDAC6, HDAC7, HDAC8, HDAC9
celecoxib	FAT1, SLIT2, PTGS2, COX20
pemetrexed	DHFR, TYMP, TYMS -
floxuridine	BIRC5, CA9, TYMP, TYMS -, DPYD -
gemcitabine	SLC29A1, SLC29A2, DCK, FKBP5, SLC28A1, SLC28A3, RRM1 -, RRM2 -
necitumumab	EGFR, TGFA, HBEGF, AREG, BTC, EPGN, KRAS -
ramucirumab	FLT1, KDR, PGF, VEGFA, VEGFB, VEGFC, FIGF, VHL
sorafenib	BRAF, FLT1, FLT4, RAF1
erlotinib	EGFR, TGFA, HBEGF, AREG, BTC, EPGN, KRAS -
dalotuzumab	IGF1R, IGF1, IGF2, INSRR
pertuzumab	AREG, EGFR, ERBB2, ERBB3, ERBB4, MAPK1, KRAS -
talimogene	ERBB2
ipilimumab	CTLA4, CD80, CD86, CD28
dabrafenib	BRAF, LYMK1, NEK1, RAF1, SIK1
temozolomide	MGMT
dacarbazine	MGMT
trametinib	MAP2K1, MAP2K2, PTPN11, RAF1
nilotinib	ABL1, ABL2, BCR, DDR1, KIT
selumetinib	MAP2K1, MAP2K2, PTPN11, RAF1
genasense	BCL2
carboplatin	ATM -, BRCA1 -, BRCA2 -, ERCC1 -, ERCC2 -, ERCC6 -, FANCC -, PALB2 -, BRIP1 -, BAP1 -
vinblastine	?
cyclophosphamide	?
sylatron	?
ril-21	?
pimasertib	MAP2K1, MAP2K2, PTPN11, RAF1
dinaciclib	CDK2, CDK5, CDK1, CDK9
gsk2132231a	?
nivolumab	CD274, PDCD1LG2, PDCD1
cobimetinib	MAP2K1, MAP2K2, PTPN11, RAF1
intetumumab	ITGA1, ITGAV, ITGA3, ITGA6
lenvatinib	FGFR2, FGFR4, FLT1, FLT4, KDR, PDGFRB, RET
temsirolimus	FKBP1A, MTOR, RPTOR, TSC1, TSC2
evofosfamide	?
etanercept	TNF
imexon	NR2E3, RRM1, RRM2
vismodegib	SMO, PTCH1, PTCH2, SHH, DISP1, GLI1
ro4929097	HES1
sta-9090	HSP90AA1
alvocidib	CDK1, CDK2, CDK4, CDK6, CDK9
docetaxel	SPARC, ERBB2, TUBB3 -
anti-CD40	CD40, CD40LG, TRAF1, TRAF2, TRAF5, TRAF6
apricoxib	COX20, APC, CTNNB1
ixabepilone	TUBB3
cixutumumab	IGF1R, IGF1, IGF2, INSRR
ispinesib	KSP
tanespimycin	HSP90AA1
imatinib	HMGCLL1
tivantinib	MET, HGF
atezolizumab	CD274, PDCD1LG2, PDCD1
bexarotene	RXRA
dexamethasone	IL6, IL8, CCL2
vinorelbine	ERBB2, EGFR, ALK
vandetanib	EGFR, KDR, LCK, LYN, PTK6, RET, TEK
gefitinib	EGFR, TGFA, HBEGF, AREG, BTC, EPGN, KRAS -
topotecan	TOP1, APTX -, BRCA1 -
fosbretabulin	?
zibotentan	ESR1, ERAL1, EDNRA, EDN1, EDN2
dacomitinib	EGFR, TGFA, HBEGF, AREG, BTC, EPGN, KRAS -
volasertib	PLK1, PDK1, MYC
pazopanib	CSF1R, FLT1, FLT4, ITK, KDR, LCK
olaratumab	PDGFC, PDGFD, PDGFRA
cbp501	?
seribantumab	AREG, EGFR, ERBB3, KRAS -
icotinib	EGFR, TGFA, HBEGF, AREG, BTC, EPGN, KRAS -
eribulin	TLE3 -

crizotinib	**ALK, HGF, MET, MST1R, ROS1**
rilotumumab	**MET, HGF**
buparlisib	**PIK3CA, PIK3CB, PIK3CG**
cabazitaxel	**SPARC, ERBB2, TUBB3 -**
veliparib	**ATM, ATR, BAP1, BARD1, BRCA1, BRCA2, BRIP1, CDK12, EMSY, ERCC1, FANCA, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCL, PALB2, PARP1, PARP2**
carbozantinib	**AXL, FLT1, FLT3, FLT4, KDR, KIT, MET, NTRK2, RET, TEK**
ganetespib	**HSP90AA1**
osimertinib	**EGFR, TGFA, HBEGF, AREG, BTC, EPGN, KRAS -**
neratinib	**EGFR, TGFA, HBEGF, AREG, BTC, EPGN, KRAS -**
ro5424802	**ALK, HGF, MET, MST1R, ROS1**
pg-csf	**?**
abp 215	**FLT1, KDR, PGF, VEGFA, VEGFB, VEGFC, FIGF, CA9**
brigatinib	**ALK, HGF, MET, MST1R, ROS1**
alectinib	**ALK, HGF, MET, MST1R, ROS1**
rabusertib	**CHK1, CHK2**
trastuzumab	**EGFR, TGFA, HBEGF, AREG, BTC, EPGN, KRAS -**
patritumab	**AREG, EGFR, ERBB3, KRAS -**
luminespib	**HSP90AA1**
ceritinib	**ALK, HGF, MET, MST1R, ROS1**
ganetespib	**HSP90AA1**

Overall, 115 unique drugs were used as either single-agents or in combination in our datasets and corresponded to 208 biomarkers that were individually validated in literature. Biomarkers were defined as an aberration altering either: 1) a direct target of the drug; 2) an indirect target of the drug; or 3) the metabolic mechanism of the drug. Some biomarkers were denoted by '-', which indicates that a 'negative' alteration (under expression or deep-deletion mutation) increased drug sensitivity. All other biomarkers increased drug sensitivity from 'positive' alterations (over expression or amplification mutations).

***Note, that while we obtained this list of biomarkers to the best of our knowledge using the literature, this list may be incomplete (we used a '?' for drugs were the drug-related biomarkers could not be determined)**

Figure S1. Example of calculating the PDS for capecitabine + ruxolitinib administered to a pancreatic adenocarcinoma population

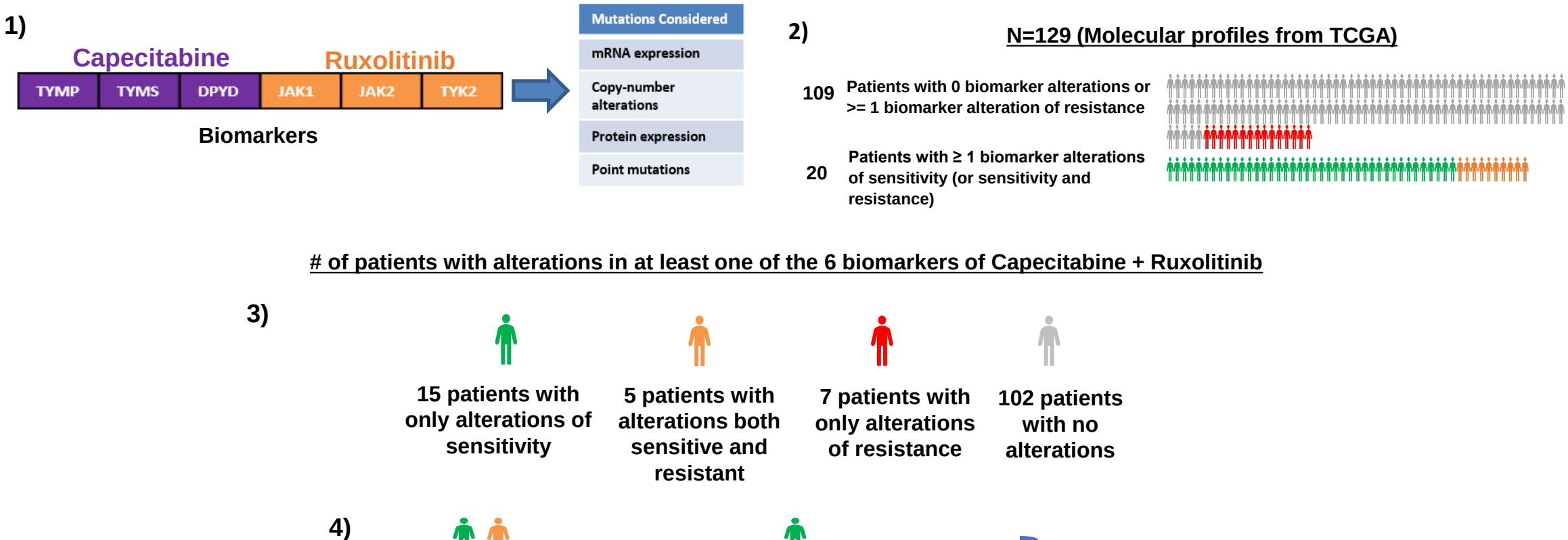
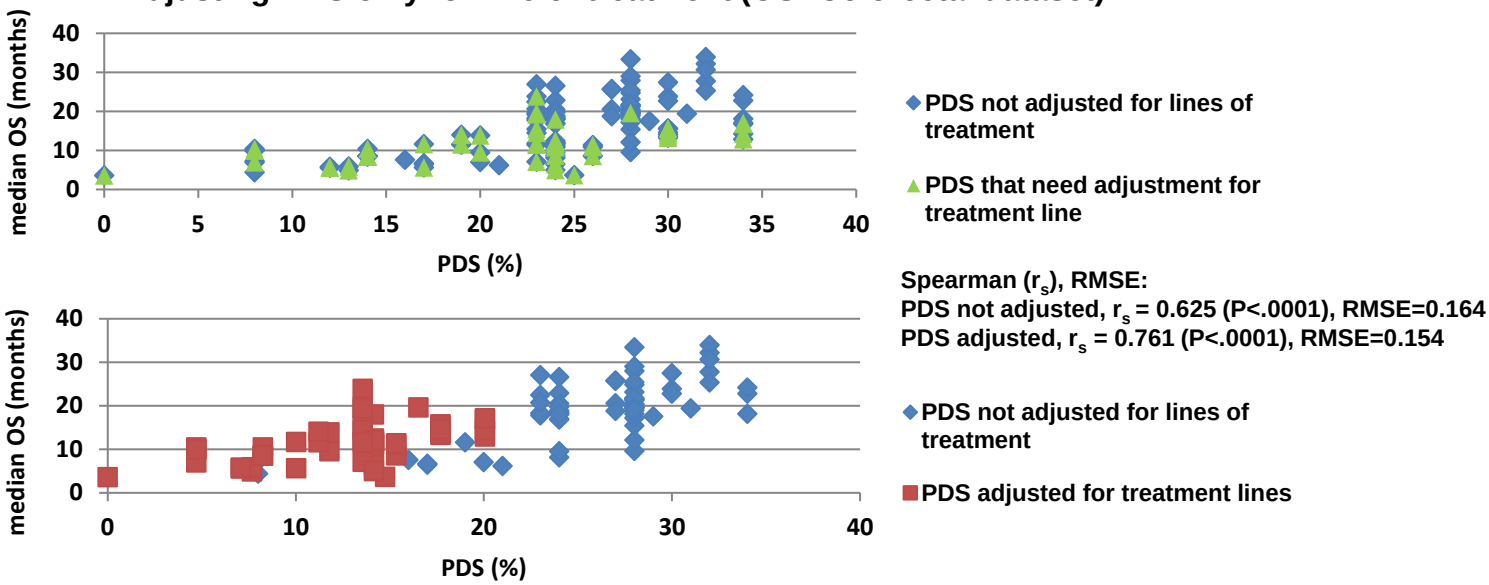
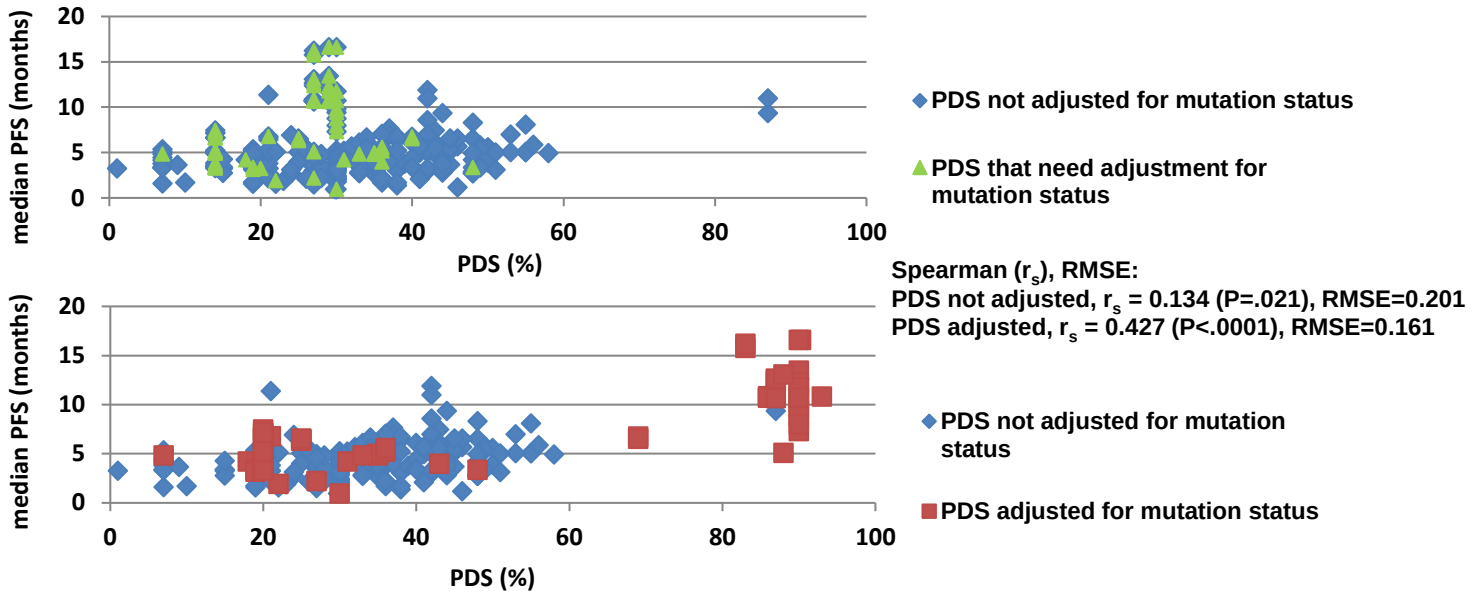


Figure S2. Adjusting PDS for line of treatment, molecular alteration status, and tumor mutation burden individually

a)



b)



c)