*Submitted to the Annals of Applied Statistics*
2005 Vol. 0, No. 0, 1–8
arXiv: arXiv:0000.0000

# PREDICTION OF CANCER CELL SENSITIVITY TO DRUGS*

BY JAMES NEWLING[†‡], SUPERVISED BY SACH MUKHERJEE[‡]

*Warwick University*[†] *and The Netherlands Cancer Institute* [‡]

We consider the problem of using high-dimensional genomic covariates to predict the drug response of cancer cell lines. The cell lines are grouped according to tissue type, which we show may be an important factor in determining the dependence of drug response on genomic covariates. We develop predictors using $l_1$-penalized linear regression models, and develop a novel variant which we call the *indicator lasso* which exploits inherent group structure. The superior performance of this new method in simulations over other methods which neglect group structure is illustrated. We finish by presenting the drug response prediction results.

## 1. Introduction.

1.1. *Biological Introduction.* The systematic use of genomic data in guiding therapies remains a challenging problem in oncology. However, there are several simple cases where drugs are directed to patients presenting specific genomic indicators. Consider the case of breast cancer, where currently the two main non-chemotherapeutical drugs in use are tamoxifen and trastuzumab. The decision of which of these two drugs to use rests mainly on the abundance of certain receptors present in cancerous cells. There are three receptors which are used as biomarkers in this way: ER (estrogen receptor), PR (progesterone receptor) and HER2/neu. Cancers exhibiting an overexpression of ER are commonly treated with tamoxifen, which binds to the estrogen receptors and blocks a vital pathway by which cancerous cells replicate. Cancers in which HER2/neu is over expressed are commonly targeted with trastuzumab, which binds to HER2 receptors and in turn deactivates another important pathway. A third and heterogeneous class of breast cancers, in which none of ER, PR and HER2/neu are amplified, is referred to as the triple negative breast cancer subgroup. There are no targeted treatments for triple negative breast cancers, and it currently has a relatively poor prognosis. Breast cancers associated with mutated BRCA1 and BRCA2 genes fall into this subgroup.

Recently, the classification of breast cancer by receptor abundance has been enhanced by the use of microarray data. Over the last decade, analyses of DNA microarray data have suggested differences between normal and cancerous cells in hundreds of genes. Several of these genes control the expression of proteins related to the ER and HER2/neu pathways, and classifications by the expression of these genes strongly overlap those provided by receptor abundance. Several companies have started to provide tests based on gene expression to determine which drugs are appropriate, and the probability of a cancer recurring after surgery. One example is the diagnostic test Oncotype

*submitted on 1 July 2013 as the final second year mini-project report for the Erasmus Masters in Complexity Science

DX provided by the company Genomic Health, which uses 21 genes in its forecasting. Genomic Health also provides diagnostics for several cancer types, and has to date analysed samples from 200,000 patients and several tissue types. The field of genomic analysis for predictive personalized medicine is rapidly growing.

In this project, we attempt to develop a novel method for drug efficacy prediction, presented in Section 2. The data we analyse is the Cancer Cell Line Encyclopedia[1](CCLE). The CCLE dataset consists of 947 human cancer cell lines each with the following features,

**mRNA expression** the amount of a particular mRNA molecule in a cell is a good proxy for the abundance of certain proteins in that cell. mRNA expression is also more easily and accurately measured than protein expression. We have for each cell line the mRNA expression of $\sim 18,000$ genes.

**chromosomal copy number** the number of times a gene appears on a chromosome is also correlated with the overall expression of certain proteins. Each cell line has the copy number of $\sim 350$ genes.

**mutation data** these are binary variables indicating if the gene is of a common variety or if the gene is a *mutant*. Each cell line has the mutation status of $\sim 100$ genes.

Along with these covariates, we have augmented the dataset with protein expression measurements of certain proteins which are available for some of the breast cell lines in the dataset.

**protein expression** we have the protein expression of 42 proteins in 15 breast cancer lines.

In addition to abundant genomic data, the CCLE data set contains the response of 479 of the 947 cell lines to 24 anticancer drugs. The responses of the cell lines are available at several doses. A commonly used proxy for overall drug efficacy is the IC 50, which is the concentration of drug necessary to cause 50% inhibition of biological activity. A low IC 50 is indicative of an effective treatment, all else being equal. We will refer to IC 50 generically as *the response*.

We wish to develop a drug response predictor based on genomic data. Such prediction is difficult due to the dimensions involved: there are far more possible explanatory covariates than there are cell lines. A resulting difficulty is that there are covariates which appear to be strongly correlated with response, but the correlations turns out to be artefacts of the large number of covariates. Another concern is that it is easy to overfit models, and standard approaches such as simple linear regression fail. In 1.2 we discuss techniques for avoiding overfitting, and methods for performing linear regression in high dimensions.

1.2. *Statistical Introduction.* Here we introduce statistical ideas used in later sections. In 1.2.1 we briefly discuss cross-validation, a technique used to guard against overfitting. In 1.2.2 we discuss penalized regression, in particular the lasso and elastic net. Then in 1.3, ideas used in multiple hypothesis testing are presented.

[1]Publicly available from http://www.broadinstitute.org/

1.2.1. *Cross-validation.* Cross-validation is a technique used to test a model's predictive power. The key idea is that model fitting and model testing are always done on mutually exclusive partitions of the data. The data is separated into several partitions and predictions in each partition are obtained from a model fitted using all the remaining partitions.

1.2.2. *The lasso.* The lasso [3] is a method in linear regression for parameter shrinkage. It can be defined as standard least squares minimization with a penalty on the $\ell_1$-norm of the coefficients. For a set of observations $(x_i, y_i) : i = 1 \dots N_{\texttt{obs}}$, where the $x_i$'s are vectors of covariates and the $y_i$'s are scalar responses, the lasso estimate is defined as

$$a_{\texttt{la}}, \beta_{\texttt{la}} = \arg\min_{a,\beta} \sum_{i=1}^{N_{\texttt{obs}}} \left(y_i - a - x_i^T \beta\right)^2 + \lambda \|\beta\|_1. \tag{1.1}$$

Here $a_{\texttt{la}}$ is the intercept and $\beta_{\texttt{la}}$ is the coefficient vector. When $\lambda = 0$, this reduces to standard least squares regression. As $\lambda$ increases, $\|\beta_{\texttt{la}}\|_1$ shrinks towards zero. A similar method for parameter shrinkage is ridge regression [3], where the estimate is defined as

$$a_{\texttt{ri}}, \beta_{\texttt{ri}} = \arg\min_{a,\beta} \sum_{i=1}^{N_{\texttt{obs}}} \left(y_i - a - x_i^T \beta\right)^2 + \lambda \|\beta\|_2^2. \tag{1.2}$$

The penalty terms in 1.1 and 1.2 shrink $\beta$ towards 0, but they do so in different ways. The lasso tends to reduce many coefficients to exactly zero, resulting is sparse coefficient vectors, while ridge regression tends to radically shrink larger coefficients while leaving smaller coefficients unchanged. From a Bayesian perspective both 1.1 and 1.2 are viewed as log posterior distributions, where the prior for the lasso is a Laplace distribution and the prior for ridge regression is a Normal distribution. From a frequentest perspective 1.1 and 1.2 are *penalized likelihoods*. In the presence of noise covariates[2], penalized likelihoods act to reduce mean square error (mse), defined as

$$\text{mse} = \text{E}\left[\left(y - a - x^T \beta\right)^2\right].$$

The mse can be estimated by cross-validation. We consider *adjusted mse*, which is defined as the mse divided by the variance of the response, and is a measure of the mse relative to the best constant predictor.

In addition to reducing prediction error, the sparsity[3] of the lasso solution is desirable for variable selection. Ridge regression on the other hand works particularly well when a set of predictors[4] are correlated, by effectively producing better predictions by averaging covariates [2]. The desirable properties of lasso and ridge regression are combined in the *elastic net*,

[2] by *noise covariates* we mean covariates which are independent of the response

[3] many coefficients of the lasso estimate are exactly zero

[4] by *predictors* we mean the non-noise covariates

$$a_{\text{en}}, \beta_{\text{en}} = \arg\min_{a,\beta} \sum_{i=1}^{N_{\text{obs}}} \left(y_i - a - x_i^T \beta\right)^2 + \alpha\lambda\|\beta\|_1 + (1-\alpha)\,\lambda\|\beta\|_2^2. \tag{1.3}$$

In 1.3, the terms $\alpha$ and $\lambda$ can be chosen to minimize mse in a layer of cross-validation, although in practise $\alpha$ is often set beforehand to reduce computation. We considered both elastic net and lasso estimation in this project, but found no significant differences and for brevity present only lasso results in this report.

1.3. *Correcting for Multiple Hypotheses.* It is common to test several hypotheses simultaneously when exploring high-dimensional genomic datasets. For example, one may wish to simultaneously test the independence of all gene expressions to response. In single hypothesis testing, the type I error rate is defined as the probability of rejecting a null hypothesis when it is true. In multiple hypothesis testing there are two commonly used equivalents of the type I error rate, the Familywise Error Rate (FWER) and the False Discovery Rate (FDR). The FWER is the probability that any one of the true null hypotheses is rejected, while the FDR is the expected proportion of rejected null hypotheses which are true, or zero if no hypotheses are rejected. Both of these reduce to the type I error rate when there is only one hypothesis. When all the null hypotheses are true, it is easy to see that FWER and FDR are equivalent. However, in general the FDR is a less stringent error rate to control for, and is widely considered to be more practical for genomic applications.

There are several methods available for increasing or *correcting* single hypothesis $p$-values to account for multiplicity, and the correct manner in which to interpret such corrected $p$-values depends on whether the correction is to control FWER or FDR. In Section 4 we are interested in considering the meta-null hypothesis: $H_0$*: all m null hypotheses are true.* We use the Benjamini-Hochberg FDR correction method [1], and will reject $H_0$ at level $\alpha$ if one of the post FDR-corrected $p$-values is less that $\alpha$. To verify the sense in this, note that it follows from the definition of FDR correction that under $H_0$ rejection happens with probability at most $\alpha$, and thus $\alpha$ is the level of the test.

One weakness of FWER and FDR corrections is their failure to take into account dependencies between null hypotheses. In general, when the hypotheses are strongly dependent, these corrections are unduly conservative. In many situations there is an elegant alternative to these corrections in the form of the Westfall-Young permutation approach, which asymptotically (in the number of hypotheses) provides the most powerful test [4]. The idea of the Westfall-Young permutation approach is best explained in the context of genomic correlations, as will be done in Section 4.

**2. Modelling the response.** The most important factor in determining which treatment a patient receives is currently the tissue type of the cancer. It would thus be reasonable to expect tissue type to be a key covariate in predicting drug efficacy. However, the fact that tissue type is an important factor in drug prescription does not necessarily imply that it is the most important factor in drug efficacy prediction, as until only recently it was the only factor available to medical practitioners. It may be the case that tissue type is a proxy of more powerful genomic predictors which

are only now becoming available. This is speculative, and for now we are still interested in giving tissue type an elevated status amongst covariates.

For each tissue type, indexed[5] by $k$ ($1 \leq k \leq K$), we model the response $(Y^k)$ as an offset linear combination of the $p$ covariates $(X^k)$, with a Gaussian noise term,

$$Y^k = a^k + \sum_{j=1}^{p} \beta_j^k X_j^k + \epsilon^k.$$

We expect there to be similarities between the coefficient vectors ($\beta^k$s) of the different types, based on a partial understanding of how the drugs work. Indeed, we verify that significant similarities between certain tissue types do exist in section 4. Depending on how similar the $\beta^k$s are, one may consider different model fitting approaches.

2.1. *The local lasso.* With the local lasso approach, estimates for each type are obtained independently,

$$a_{\texttt{lo}}^k, \beta_{\texttt{lo}}^k = \arg\min_{a,\beta} \ell\left(a, \beta | \underline{y}^k, \underline{x}^k\right) + \lambda_{\texttt{lo}}^k \|\beta\|_1. \tag{2.1}$$

where

$$\ell\left(a, \beta | \underline{y}^k, \underline{x}^k\right) = \sum_{i=1}^{N_k} \left(y_i^k - a^k - \beta^k \cdot x_i^k\right)^2.$$

The lasso penalty factor $\lambda_{\texttt{lo}}^k$ in eqn. 2.1 is chosen by cross-validation to minimize mse. The local lasso is the optimal[6] approach to take if the parameters in the models of the different types are independent. It can also be optimal when the number of cell lines is very large and the models are not independent, as will be discussed after the following method has been presented.

2.2. *The global lasso.* The global lasso approach ignores group differences completely and finds a single lasso estimate from all the data,

$$a_{\texttt{gl}}, \beta_{\texttt{gl}} = \arg\min_{a,\beta} \ell\left(a, \beta | \underline{y}, \underline{x}\right) + \lambda_{\texttt{gl}} \|\beta\|_1. \tag{2.2}$$

The global lasso is of course the optimal approach if there is no difference between the type coefficients. It is also the optimal approach in the case of mild correlations between $\beta^k$s when the group sizes are small.

The global versus local dichotomy can be viewed as a variance versus bias trade-off. The local fits are unbiased[7] with high variance, while the global fit is biased with low variance. The bias is independent of sample sizes, but the variance can be reduced arbitrarily with sufficiently large samples. The next two lasso based approaches attempt to find a balance between variance and bias.

[5] there are several indices to be introduced. To remember that $k$ is for type, $k$ sounds like $c$, and $c$ for class (type)

[6] *optimal* here means resulting in a lower mse than the alternative approaches

[7] actually the local fits are like all lasso estimates biased, but there is an additional bias incurred in the global fit which local fits do not incur

2.3. *The mixture lasso.* With the mixture lasso approach, coefficients are taken to be a linear combination of the optimal local and global coefficients. More specifically,

$$\beta_m \leftarrow \alpha_o \beta_{gl} + (1-\alpha_o)\,\beta_{lo}, \quad a_m \leftarrow \alpha_o a_{gl} + (1-\alpha_o)\,a_{lo},$$

where,

$$\alpha_o = \arg\min_{\alpha} \left\{ \text{mse}(\underline{\hat{y}}, \underline{y}) \right\}$$

and where $\underline{\hat{y}}$ is the vector of predictions using $a_m$ and $\beta_m$.

2.4. *The indicator lasso.* The indicator lasso approach uses both local and global penalties simultaneously. Let us denote by $(a^k_{\texttt{in}}, \beta^k_{\texttt{in}})$ the $k$'th type fit using the indicator method. We decompose these into global and local[8] components,

$$\beta^k_{\texttt{in}} = \beta_{\texttt{gl}} + \beta^k_{\texttt{lo}}$$
$$a^k_{\texttt{in}} = a_{\texttt{gl}} + a^k_{\texttt{lo}}.$$

The indicator fitted coefficients are chosen to be those that minimize the penalized likelihood,

$$\ell_{\texttt{in}}\,(a,\beta) = \ell\left(a,\beta|\underline{x},\underline{y}\right) + \lambda_{\texttt{in}}\|\beta_{\texttt{gl}}\|_1 + \sum_{k=1}^{K} \lambda^k_{\texttt{lo}}\|\beta^k_{\texttt{lo}}\|_1. \tag{2.3}$$

We consider the special case where in 2.3 the local penalty factors are all set to be equal, and it proves convenient to express this local penalty factor as a multiple of the global penalty factor,

$$\lambda^k_{\texttt{lo}} = \alpha_{\texttt{in}}\lambda_{\texttt{in}} \qquad \text{for } 1 \le k \le K.$$

Thus 2.3 reduces to

$$\ell_{\texttt{in}}\,(a,\beta) = \ell\left(a,\beta|\underline{x},\underline{y}\right) + \lambda_{\texttt{in}}\|\beta_{\texttt{gl}}\|_1 + \alpha_{\texttt{in}}\lambda_{\texttt{in}} \sum_{k=1}^{K} \|\beta^k_{\texttt{lo}}\|_1. \tag{2.4}$$

One can choose the penalty terms $\lambda_{\texttt{in}}$ and $\alpha_{\texttt{in}}$ by cross-validation. If one fixes $\alpha_{\texttt{in}}$ and chooses only $\lambda_{\texttt{in}}$ by cross-validation, then in the limiting cases of $\alpha_{\texttt{in}} \to 0$ and $\alpha_{\texttt{in}} \to \infty$, the indicator lasso is equivalent to the local and a global[9] lasso respectively, a proof of which is given in Appendix A. The penalty of the indicator lasso closely resembles the penalty of the fused lasso [7], where the $\ell_1$-norm of the coefficients as well as the $\ell_1$-norm of all differences are penalized,

$$P_{\texttt{fu}}\left(\beta^1 \cdots \beta^k\right) = \lambda_{\texttt{fu-1}} \sum_{k=1}^{K} \|\beta^k\|_1 + \lambda_{\texttt{fu-2}} \sum_{k_1=1}^{K} \sum_{k_2=k_1+1}^{K} \|\beta^{k_1} - \beta^{k_2}\|_1. \tag{2.5}$$

[8] $\beta_{\texttt{gl}}$ and $\beta^k_{\texttt{lo}}$ here are not to be confused with the global and local lasso fits which we described earlier

[9] when $\alpha_{\texttt{in}} \to \infty$ we do not recover the global lasso in the form of 2.2, as the class specific intercepts are left unpenalized

We do not consider the fused lasso in this report, but discuss its relationship to the indicator lasso in Appendix A. In the following section we compare the four methods described thus far through a simulation designed to mimic true genomic structures. When reference is made to the indicator lasso without specifying the value of $\alpha_{\texttt{in}}$, it can be assumed that $\alpha_{\texttt{in}} = 1$.

**3. Simulation.** The design of the simulation and a description of the free parameters is first given in 3.1. Simulation results are then given in 3.2.

3.1. *Simulation design.* Covariates (genes) are indexed by $j$ for $1 \leq j \leq p$, groups are indexed by $k$ for $1 \leq k \leq K$ and observations by $i$ for $1 \leq i \leq N_k$, where there are $N_k$ observations per group. The indices of covariates which have non-zero coefficients in all groups (the globally important genes) belong to the set $S_0$,

$$S_0 = \{j \,|\, \forall k : \beta_j^k \neq 0\}.$$

The indices of covariates which are non-zero only in group $k$ belong to the set $S_k$,

$$S_k = \{j \,|\, \beta_j^k \neq 0 \ \wedge \ \forall k' \neq k : \beta_j^{k'} = 0\}.$$

The total proportion of indices in $S_0$ is given by the free parameter $\phi_0$. We constrain the proportion of indices in $S_k$ to be constant across groups, giving one more free parameter $\phi_1$.

$$\phi_0 = \frac{|S_0|}{p}, \ \phi_1 = \frac{|S_k|}{p} \ (1 \leq k \leq K).$$

All coefficients are non-zero in either all, one or no groups, so we have the constraint

$$\phi_0 + K\phi_1 \leq 1.$$

We now discuss the actual generation of the $K$ sparse coefficient vectors. First we present the simpler setup for doing this and then the more complicated and realistic one.

**Simple coefficient generation.** For $j \in S_0$, we generate the coefficients of covariate $j$ for the $K$ types, that is $\beta_j = (\beta_j^1 \cdots \beta_j^k)$, as

$$\beta_j \sim N(0, \Sigma),$$

where $\Sigma_{k,k'} = \mathbf{1}_{k=k'} + \rho \mathbf{1}_{k \neq k'}$. When $\rho = 0$, the coefficients are independent $N(0,1)$ random variables and when $\rho = 1$, the coefficients are the same across all groups. For $j \in S_k$, we generate $\beta_j$ as,

$$\beta_j^k \sim N(0,1), \text{ and for } k' \neq k, \ \beta_j^{k'} = 0.$$

**More complicated coefficient generation.** We believe that it is more accurate to model coefficients as having non-zero modal values. This belief is incorporated into the model by generating coefficient values as

$$\mu_j^k = S_j^k \mu_0 + Z_j^k, \tag{3.1}$$

where the mean in 3.1 is composed of an offset of magnitude $\mu_0$ and sign $S_j^k$, and a perturbation term $Z_j^k \sim N(0,1)$. The sign $S_j^k$ is $+1$ or $-1$ with equal probability so as to have $\mathbf{E}(\mu_j^k) = 0$. For coefficients in $S_0$ we need to consider the correlations between coefficients. To ensure that we still have $\text{cor}(\mu_j^k, \mu_j^{k'}) = \rho$, we take $Z_j \sim N(0, \Sigma)$ where $\Sigma_{k,k'} = \mathbf{1}_{k=k'} + \rho \mathbf{1}_{k \neq k'}$, and additionally $\mathbf{P}\left(S_j^k S_j^{k'} = 1\right) = (1 - \sqrt{\rho})/2$. Once we have the $K$ coefficient vectors, we generate observations for $1 \leq k \leq K, 1 \leq i \leq N_k$ from

$$Y_i^k = \mu_i^k + \epsilon_i^k, \tag{3.2}$$

where $\mu_i^k = a_0^k + \beta^k \cdot X_i^k$ and $X_i^k$ is the vector of covariates of the $i$'th observation in the $k$'th group. For the simulation we draw $X_i^k \sim N(0,1)$, and then normalize the covariates such that for each group $k$,

$$\sum_{i=1}^{N_k} X_i^k = 0 \quad \text{and} \quad \frac{1}{N-1} \sum_{i=1}^{N} (X_i^k)^2 = 1. \tag{3.3}$$

In addition to considering generated covariates, we considered extracting covariates from the CCLE genomic data. This did not not result in significantly different simulation results and we do not discuss this further. The noise term $\epsilon_i^k$ in 3.2 is drawn as $\epsilon_i^k \sim N(0, \sigma_k^2)$. The relationship between the variance of the noise $\sigma_k^2$ and the estimated signal-to-noise ratio $\text{SNR}_k$ in type $k$ is,

$$\sigma_k^2 = \frac{\frac{1}{N_k - 1} \sum_{i=1}^{N} \left(\mu_i^k - \bar{\mu}^k\right)^2}{\text{SNR}_k}. \tag{3.4}$$

In this simulation we wish to avoid having different noise variances between groups, and thus define a global noise variance based on a global SNR,

$$\sigma^2 \leftarrow \frac{\frac{1}{N-K} \sum_{k=1}^{K} \sum_{i=1}^{N_k} \left(\mu_i^k - \bar{\mu}^k\right)^2}{\text{SNR}}, \tag{3.5}$$

where $N = \sum_{k=1}^{K} N_k$. Each group noise variance is then set to the global value, $\sigma_k^2 \leftarrow \sigma^2$. As the signal strength is dependent on randomly generated coefficient vectors, the constant noise variance 3.5 results in different signal-to-noise ratios in the different groups. Let us summarize what the simulation parameters which need to be set are,

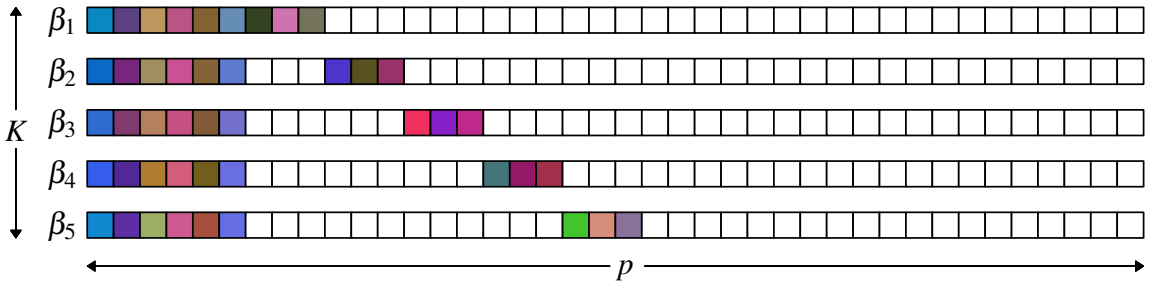


Fig 1: An example realization of $K = 5$ coefficient vectors. Similar colours depict similar real values, and white cells depict zeros. The first 6 covariate coefficients are non-zero in all types, and each type additionally has 3 unique non-zero covariate coefficients. With $p = 40$ covariates, we thus have $\phi_0 = 6/40$ and $\phi_1 = 3/40$. The within-column similarity of colours for the first six coefficients is a result of a high correlation: $\rho = 0.8$.

| | |
|---|---|
| $K$ | number of groups with distinct coefficients (number of tissue types) |
| $p$ | number of covariates (number of genes) |
| $\mu_0$ | offset of the coefficients (mean 'activity' of active genes) |
| $\phi_0$ | proportion of covariates whose coefficients are non-zero in all groups (globally active genes) |
| $\phi_1$ | proportion of covariates whose coefficients are non-zero in one group (locally active genes) |
| $SNR$ | signal-to-noise ratio |
| $N_k$ | number of observations in group $k$ (number of cell lines of type $k$ with drug response) |
| $\rho$ | for coefficients which are non-zero in all groups, the correlation between coefficients across types. |

The proportion of zero coefficients in each group is $1 - \phi_0 - \phi_1$. The simulation is outlined as Algorithm 1, and the parameters parameters are illustrated in Figure 1.

**Algorithm 1** Simulation Outline

Indices run as: $1 \leq j \leq p,\ \ 1 \leq k \leq K,\ \ 1 \leq i \leq N_k,\ \ 1 \leq i' \leq 10N_k$

1: Set all the free parameters
2: Generate coefficient realizations $\beta_j^k$
3: Generate training (TR) and test (TE) covariate arrays $X_{i,j}^k$ and $X_{i',j}^k$
4: Calculate the mean responses $\mu_i^k$ and then $\sigma^2$
5: Generate the TR responses $Y_i^k$ and the TE responses $Y_{i'}^k$
6: Fit the different models using TR covariates and responses
7: Test the predictive power of the fitted models on TE

3.2. *Simulation results.* In Figures 2, 3 and 4 comparisons are made between the local, indicator (with $\alpha_{\texttt{in}} = 1$) and global lassos over various combinations of the simulation parameters. In Figure 5 the performance with several $\alpha_{\texttt{in}}$ values are compared, and in Figure 6 the performance of the mixture lasso over several $\alpha_0$ values are illustrated alongside the local and global variants.

In all the Figures, the x-axis is $\phi_0/(\phi_0 + \phi_1)$. This is the proportion of active[10] covariates which are active in all types. Thus large values of $\phi_0/(\phi_0+\phi_1)$ correspond to high similarity across types, which is favourable for a global fit.

In almost all situations we observe the indicator method resulting in lower prediction errors than the global and local methods. The variables which are varied in Figure 2 are the signal-to-noise ratio and $\rho$, then in Figure 3 we show the effect of varying $\mu_0$ and $K$, and finally in Figure 4 the parameters $p$ and $\phi_0 + \phi_1$ are varied. The results of varying each of these parameters are discussed in the Figure captions.

Figure 5 illustrates how varying $\alpha_{\texttt{in}}$ affects the indicator fit. Low $\alpha_{\texttt{in}}$ values result in mses similar to those obtained with the local fit, while high $\alpha_{\texttt{in}}$ values result in an approach to global fit mses. This is in agreement with the results of Appendix A, where it is shown that $\alpha_{\texttt{in}} \to 0$ results in predictions from the indicator fit resembling those of the local lasso, and similarly those for $\alpha_{\texttt{in}} \to \infty$ resembling those of a global fit.

In Figure 6 the mse of the mixture method for different values of $\alpha_o$ is compared with local, indicator and global mses. We observe that the mixture method mse is not significantly better than either the global or local mses for any $\alpha_o$ values, whereas the indicator mse is significantly lower over a large range of $\phi_0/(\phi_0 + \phi_1)$ values.

**4. Correlation structures in the CCLE.** Before applying the different penalized linear regression techniques to the CCLE data, we perform preliminary analyses of correlations in the data. We do so by posing a series of progressively weaker null hypotheses which we then attempt to disprove. Ideally, the null hypotheses should all be rejected for the use of the indicator lasso to be appropriate.

We first introduce some notation which better enables us to state the null hypotheses. In the simulation we introduced indices $i, j$ and $k$ to index cell line, gene and type respectively. We continue with this indexing system, and introduce a fourth index $l$ for drugs. Accordingly we write $Y^k_{i,l}$ for the response to drug $l$ of the $i$'th cell line of type $k$ and $X^k_{i,j}$ for the expression of the $j$'th gene, in the $i$'th cell of type $k$. We will present the null hypotheses in the context of specific type-drug pairs. The first null hypothesis is trivial, and included for the sake of completeness,

($H_{0,1}$) Drug $l$ has no effect on cell lines of type $k$.

[10] *active* here means having a non-zero coefficient

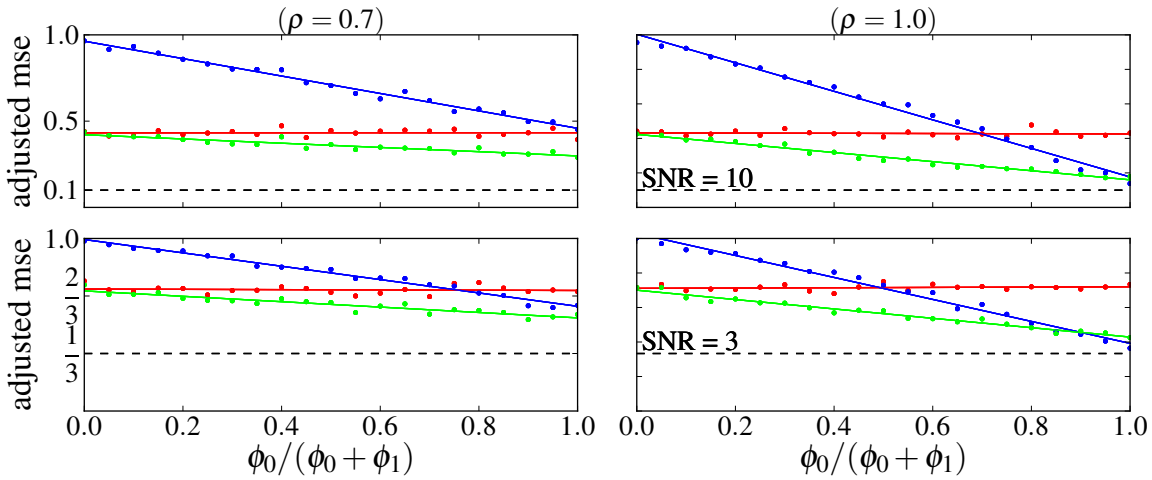


Fig 2: Comparison of the adjusted mse for local (red), global (blue) and indicator (green) lassos. On the x-axis of the four subfigures is $\phi_0/(\phi_0+\phi_1)$, which is the proportion of active sites which are globally active. In the left column $\rho = 0.7$ and in the right column $\rho = 1.0$. The top two Figures have signal-to-noise ratio 10, while the bottom two have signal-to-noise ratio 3. Other parameters which are kept constant in all four subfigures are: $K = 3, N_1 = N_2 = N_3 = 60, \phi_0 + \phi_1 = 0.1, p = 200, \mu_0 = 0$. Each point corresponds to an average over 20 simulation runs. The lines are least square fits, although we have no reason to expect a linear relationship. There are several unsurprising behaviours observed. (1) The mse of the local lasso fit is independent of both $\rho$ and $\phi_0/(\phi_0+\phi_1)$. (2) The global mse improves as $\rho$ and $\phi_0/(\phi_0+\phi_1)$ increase. (3) The relative mse is bound below by the SNR. The indicator method appears to be at least as good as the local and global lassos for all parameter settings. When $\rho = 1$ and $\phi_0/(\phi_0+\phi_1) = 1$ (far right), the group models are identical, and thus the global lasso should is the correct lasso fit to use. However even here the indicator lasso appears to perform at least as well as the global lasso.

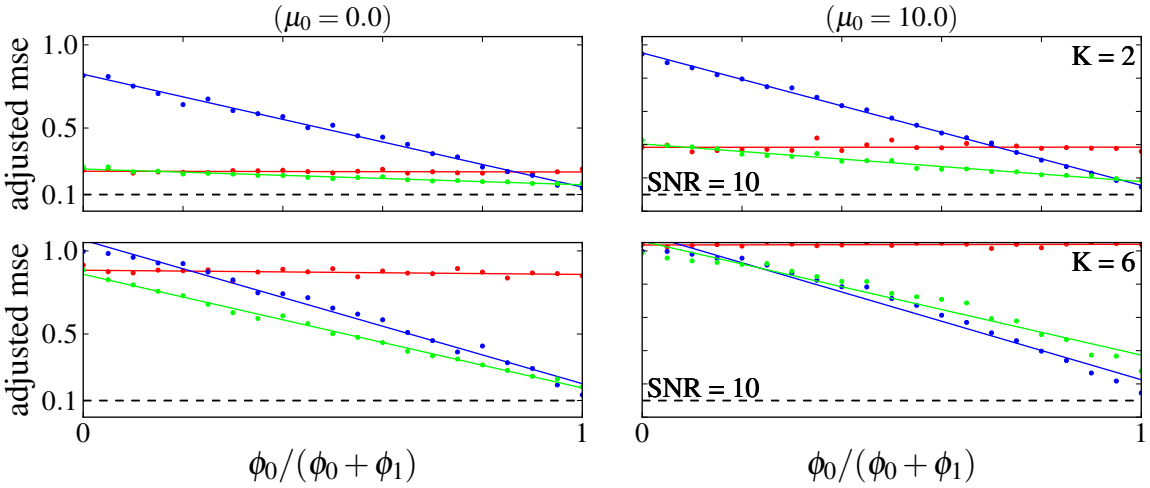


Fig 3: As in Figure 2, we have the local lasso (red), global lasso (blue) and indicator lasso (green). In the left column $\mu_0 = 0$ and in the right column $\mu_0 = 10$. In the top row $K = 2$, while in the bottom row $K = 6$. In all subfigures the total number of observations is $\sum_{k=1}^{K} N_k = 180$. Other parameters kept constant in all four are: $\phi_0 + \phi_1 = 0.1, p = 200, \text{SNR} = 0$ and $\rho = 1$. On comparing the left and right panels, we observe that larger $\mu_0$ values result in poorer relative predictions than small $\mu_0$ values. On considering the definition of $\mu_0$ this may seem strange as a larger $\mu_0$ implies a larger signal which, all else being equal, results in an easier reconstruction and better predictions. But the SNR is fixed, and not the noise variance, and so a larger signal coincides with a larger noise variance. Comparing the top and bottom panels we note that for a fixed number of observations, a larger number of groups results in a more challenging reconstruction. We also observe the global lasso outperforming the indicator lasso when $\mu_0 = 10, K = 6$ (bottom right).

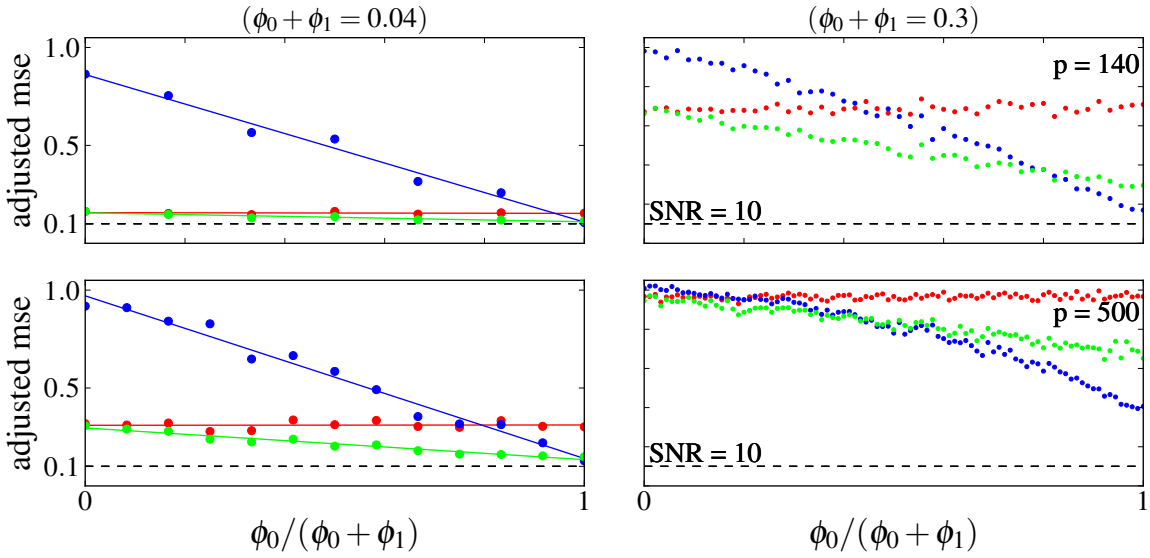


Fig 4: We have the local (red), global (blue) and indicator (green) lassos. In the left column the sparsity of the true model is $\phi_0 + \phi_1 = 0.04$ and in the right column it is $\phi_0 + \phi_1 = 0.3$. In the top panel, the total number of covariates is $p = 140$, while in the bottom panel it is $p = 500$. In all subfigures there are $K = 3$ groups of 60 observations. Other parameters kept constant in all four are: $\mu_0 = 0, \text{SNR} = 10$ and $\rho = 1$. On comparing the left and right columns, we observe lower adjusted mses in the sparser models, all other parameters equal. Similarly, comparing the top and bottom Figures we observe lower adjusted mses in the case of fewer covariates. Neither of these results is surprising. For the dense, high-dimensional case (bottom right) we observe the global lasso outperforming the indicator lasso with $\alpha_{\texttt{in}} = 1$, although in Figure 5 we see that with larger $\alpha_{\texttt{in}}$ values the indicator method is more competitive with the global lasso.

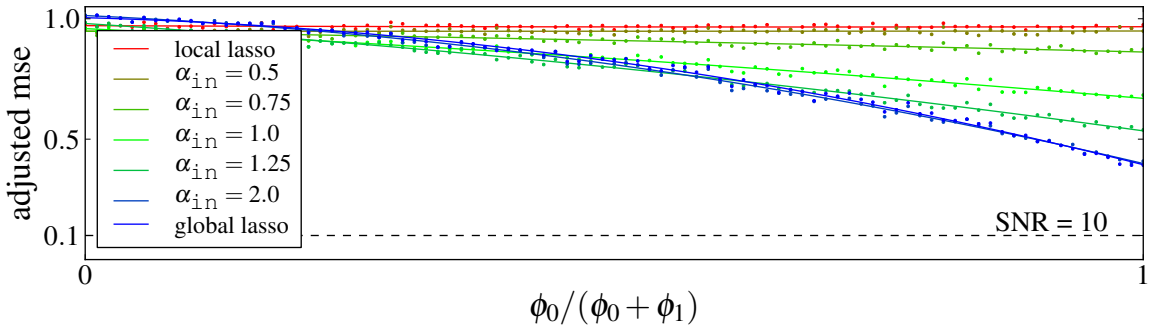


Fig 5: The same setup as for the bottom-right subfigure of Figure 4: $\phi_0 + \phi_1 = 0.3, p = 500, N_1 = N_2 = N_3 = 60, \mu_0 = 0, \text{SNR} = 10$ and $\rho = 1$. This Figure contains in addition to the local lasso (red), global lasso (blue) and the indicator lasso with $\alpha_{\mathtt{in}} = 1$, the indicator lasso with varying weight between global and local penalty, with $\alpha_{\mathtt{in}}$ running from 0.5 to 2.0. We see that $\alpha_{\mathtt{in}} = 2$ is quite close to the global fit.

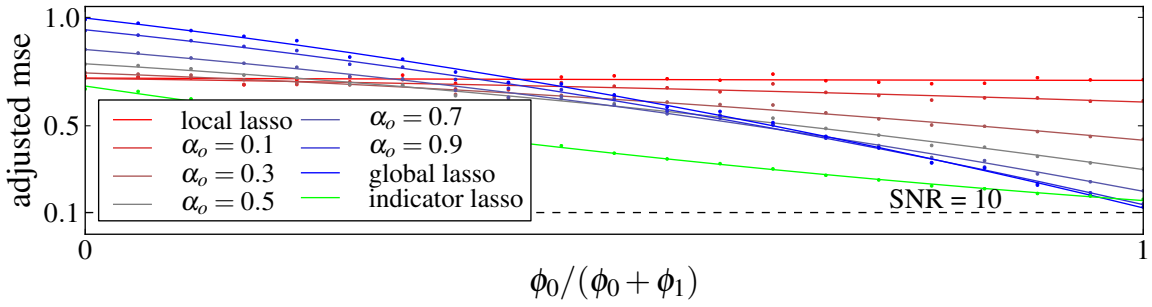


Fig 6: Comparison of adjusted mses for methods: indicator (green), local (red), global (blue), and mixture lasso with varying $\alpha_o$ values. Setup parameters $K = 6$ types, with $N_1 = \cdots = N_6 = 40, \rho = 1, \text{SNR} = 10, \phi_0 + \phi_1 = 0.1, p = 200$ and $\mu_0 = 0$.We observe that the mixture mse does only slightly better than simply interpolating between the global and local mses, but it is significantly poorer than the indicator lasso.

Rejection of $H_{0,1}$ follows from the fact that we have finite IC 50 responses. The second null hypothesis concerns correlations between genomic covariates and drug response,

($H_{0,2}$) For drug $l$ and type $k$, and for all genes $j$: cor $\left(X_j^k, Y_l^k\right) = 0$.

Hypothesis $H_{0,2}$ being true would paint a bleak picture for drug response prediction and personalized medicine in general. Fortunately for several type-drug pairs we have sufficiently small $p$-values, obtained via different methods, to be able to confidently reject $H_{0,2}$. We consider both Pearson and Spearman correlations. The Pearson correlation is defined as

$$\text{cor}_P(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}. \tag{4.1}$$

In finite samples this is estimated in the obvious way by replacing the covariance and variance terms with sample estimates. The Spearman correlation is defined for finite samples as the Pearson correlation of the ranks of $X$ and $Y$ in eqn. 4.1. We assess the validity of $H_{0,2}$ in three different ways. The first and most direct way is by estimating the correlation of all genes with drug response, and then calculating a global $p$-value associated with these correlations. We consider four variants of the first approach, by using Benjamini-Hochberg FDR and Westfall-Young multiple testing corrections with both Pearson and Spearman correlations to obtain global $p$-values. The second approach involves comparing cross-validated *background*[11] correlations with the cross-validated strongest correlation. The use of cross-validation makes multiple-correction unnecessary, but with it there is an associated loss of power which makes this approach weaker than the first approach. The third approach is the least direct, and involves comparing the sparsity of coefficients of bootstrapped lasso fits. All three approaches will be described in detail.

4.1. *$p$-values for multiple hypotheses and the Westfall-Young approach.* If $X$ and $Y$ in eqn. 4.1 are normally distributed, exact $p$-values for the Pearson correlation coefficient can be calculated from the Student's t-distribution. For the Spearman correlation, exact $p$-values can be obtained without any distributional assumptions. These $p$-values relate to single gene-response pairs. As already mentioned, these can be adjusted for by using the Benjamini-Hochberg FDR correction. The smallest post-correction $p$-value, which we denote by $p_B$, turns out to be

$$p_B = m \min_{1 \le j \le J} \frac{p^{(j)}}{j},$$

where $p^{(j)}$ is the $j$'th smallest pre-correction $p$-value, and $m$ is the number of genes. Under $H_{0,2}$,

$$\text{P}\left(p_B < \alpha\right) < \alpha. \tag{4.2}$$

[11] by background we mean genomic covariates are the covariates which are not correlated with drug response, by far the majority. We referred to these as noise covariates in the simulation

The degree to which inequality 4.2 is conservative depends on the dependence structure of the genes. Of course for 4.2 to hold at all requires the gene-response $p$-values to be reliable, and so a deviance from normality may render 4.2 invalid in the case where gene-response $p$-values are obtained from Pearson t-statistics. These concerns of lack of power and lack of robustness can be avoided if one adopts a Westfall-Young permutation approach. The idea of Westfall and Young was to compare the correlations obtained using the true response vector to those obtained using permutations of it. To test the null hypothesis $H_{0,2}$ for subtype $k$ and drug $l$, 1000 pseudo-responses $\tilde{Y}^k_{l,q}, 1 \leq q \leq 1000$ are generated by permuting $Y^k_l$. The highest correlation obtained with the real drug is then then compared to the 1000 highest correlations obtained with the permutations. The $p$-value is estimated as $n/1000$ where $n$ is the rank of the real correlation amongst the 1000 top permutation correlations. In general, one need not compare the top correlation with the top permutation correlations. One could compare a lower rank correlation, or use some combination of top ranks. The most powerful combination of ranks to use depends on the underlying data, and is beyond the scope of this project but is addressed in [5]. Note that with the Westfall-Young approach, there are no distributional assumptions made in calculating the global $p$-value.

4.2. *Corrected $p$-values to assess $H_{0,2}$.* We consider Benjamini-Hochberg corrected Spearman correlations (BS), Benjamini-Hochberg corrected Pearson correlations (BP), Westfall-Young corrected Spearman correlations (WS) and Westfall-Young corrected Pearson correlations (WP). The only one of these which assumes normality in calculating a global $p$-value is BP. In Figure 7 and in Figure 16 in Appendix D, the $p$-values using these four variants are illustrated for each type-drug pair. We see that for several type-drug pairs, there is a significant correlation. For the smaller groups, correlations need to be large to appear significant, and it is possible that several type-drug pairs with real correlations do not appear significant due to small sample size. The BP method predicts the largest number of significant correlations. In Figure 14 in Appendix D, the significance of each drug-tissue type is illustrated in $p$-value panels. We see a strong similarity in the $p$-values when using the four approaches. We have included two random responses, which we have called placebos, as a further check that $p$-values are reliable. The first placebo contains i.i.d draws from a Normal distribution, and the second is a permutation of lapatinib responses. Further comments can be found in the captions, and detailed tables of $p$-values and correlations using the described method can be found in Appendix E.

4.3. *Cross-validated $p$-values to assess $H_{0,2}$.* In this second approach a type of $N$-fold cross-validation is used to obtain estimates of how significant the largest correlation is for each drug-type pair. For each fold[12], we choose the most strongly correlated gene in the remaining $N-1$ folds, and calculate the correlation of this chosen gene in the original fold. The average of the absolute value of the $N$ correlations thus obtained is then compared to equivalent values calculated on 500 background genes. The optimal number of folds to use is not clear. In general applications of cross-validation, $N$ is taken as large as computationally feasible, thus allowing the training sets[13] to be as

[12] *fold* is the commonly used term for what in 1.2.1 we called a *partition*

[13] The $N-1$ folds used to fit the model for the remaining fold are collectively called a *training set*

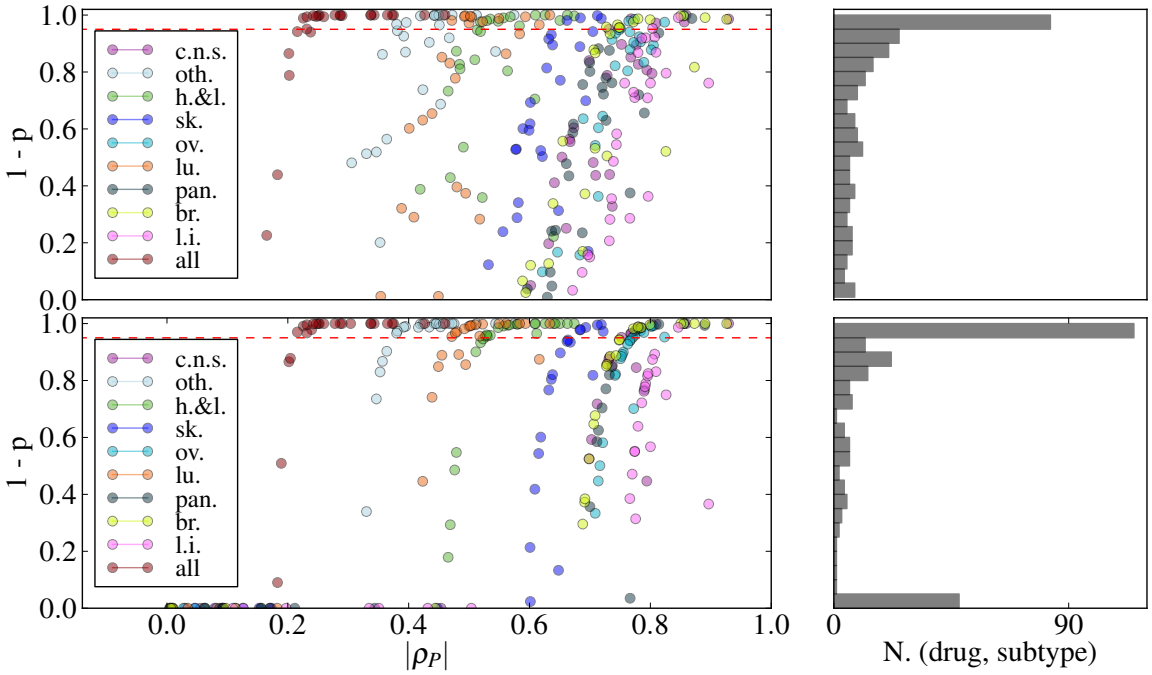


Fig 7: The $p$-values associated with type-drug specific null hypotheses $H_{0,2}$, that all genes are uncorrelated with response. In the top Figure, $p$-values are obtained using the Westfall-Young approach, while in the bottom Figure $p$-values are obtained from a Benjamini-Hochberg FDR correction. Each circle represents a type-drug pair, and is color coded by type. The dotted red line represents the 5% significance level. On the x-axis is the absolute value of the Pearson correlation coefficient. The apparent stripy clustering by tissue type is due to the difference in number of cell lines between the different types. For the smaller groups (large intestine, ovary, pancreas, breast) a larger correlation is required to be declared significant. Even correlations as large as 0.7 are insignificant in the smaller types. On the right are histograms of the $p$-values. The mild dispersion within the types is due to the fact that not all cell lines received certain drugs. We observe that the Benjamini-Hochberg correction obtains more significant genes, however, for reasons mentioned in the text, the Westfall-Young $p$-values are more reliable in this plot as the Pearson $p$-values rely on the normality assumption.

close as possible in size to the full data set. However, in the current application we wish to calculate correlations on a fraction $1/N$ of the data, and thus we have an incentive to keep $N$ small so as to reduce the variance in these correlations. We choose $N = 3$. It is worth mentioning that the gene selected in each fold may be different.

The results of using this cross-validation technique are illustrated in Figure 12 and in Figure 15 in Appendix D. In addition to calculating cross-validated top correlation, we have calculated cross-validated rank 2 to 5 correlations. We see that this cross-validation approach is less powerful than the multiple-correction approaches of section 4.2. Significant correlations are observed when all types are combined but few individual type-drug correlations appear significant with this approach. Further comments are presented in the Figure captions.

4.4. *Bootstrapped lasso counts to assess $H_{0,2}$.* The third approach we consider to assess $H_{0,2}$ is to compare lasso fits on true responses with lasso fits on permuted responses. The idea is that, if there were particular genomic covariates which obtain non-zero coefficients more frequently on bootstrapped[14] data than on bootstrapped permuted data, this would suggest that those covariates were truly correlated with response. The scheme is described in Algorithm 2. While we do not obtain a $p$-value using Algorithm 2, it could easily be extended so as to obtain one. For example, the permutation step on line 13 could be repeated 1000 times, and the distribution of the largest non-empty bin could be compared to the largest non-empty bin without permutation. In Figure 9, we compare the frequency distribution using bootstrap sampling between the true responses and permuted responses for the drugs Panobinostat and PD-0325901 for selected tissue types. In several cases we observe top covariates being selected more frequently when the true response is used as compared to the permuted response, suggesting that these covariates carry real predictive value.

**Algorithm 2** Boostrapped lasso for $H_{0,2}$

1: Choose the subtype $\tau$ and drug to be tested
2: Set the optimal $\lambda$ for the lasso Algorithm by cross-validation (glmnet in R)
3: **for** i =1:100 **do**
4: Take a bootstrapped sample of cell lines of type $\tau$
5: Fit the lasso penalized model and record which genes have non-zero coefficients
6: **end for**
7: **for** j =1:J **do** ▷ Where J is the number of genes
8: $c_j \leftarrow$ number of times gene j was non-zero in the 100 bootstraps
9: **end for**
10: **for** n = 1:100 **do**
11: Count how many times $c_j = n$ ▷ green bars in Figure 9
12: **end for**
13: Perform the above steps for the case where the response vector has been randomly permuted ▷ blue bars in Figure 9

[14] random sampling with replacement

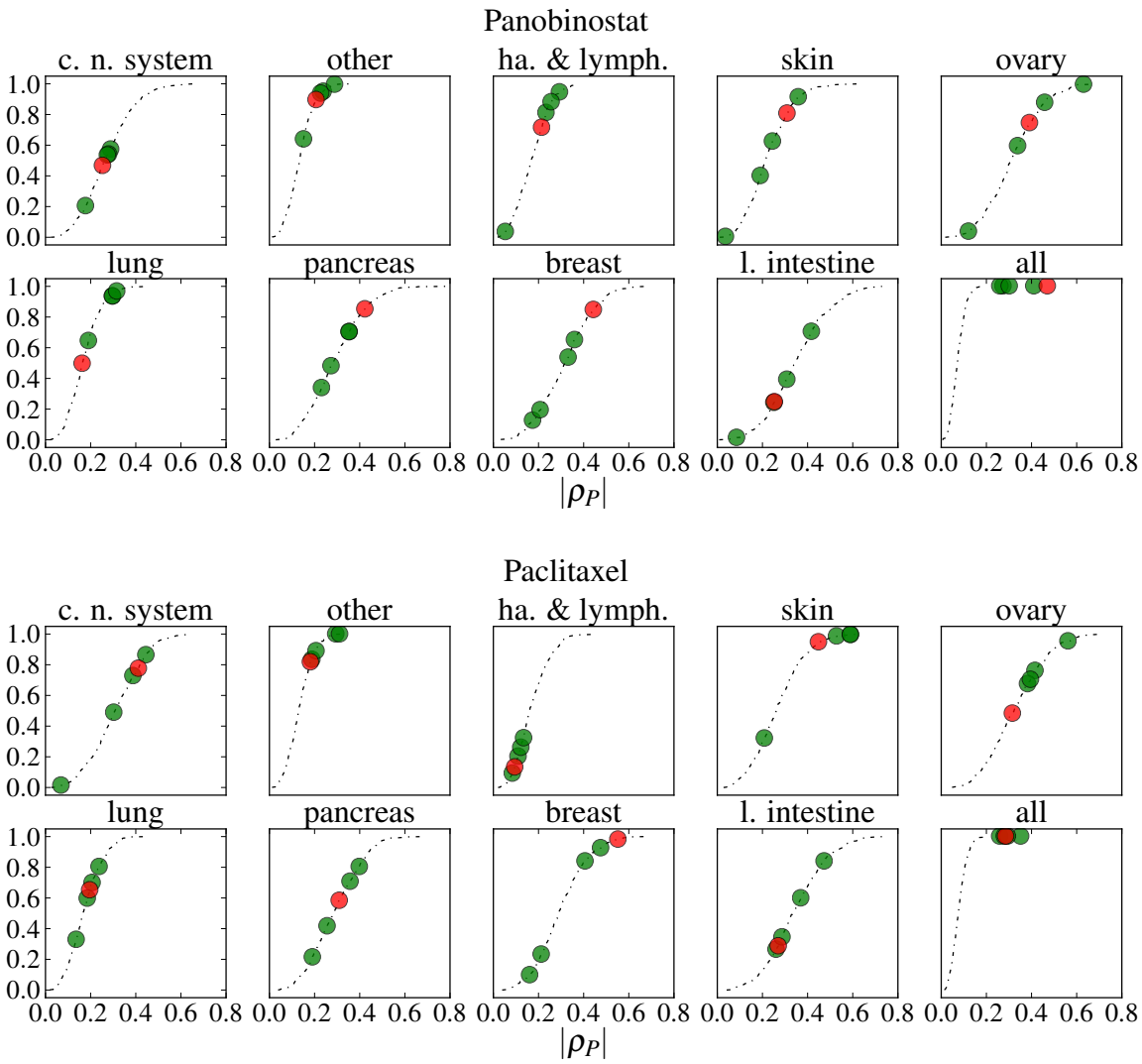


Fig 8: Cross-validated correlations and permutation $p$-values of the five largest correlations for different types, for the drugs Panobinostat and Paclitaxel. The largest correlation is coloured red. On the $x$-axis is the absolute value of the Pearson correlation, averaged over test sets. The dashed line is the cumulative distribution of background genes, as estimated by 500 randomly selected genes. For the drug Panobinostat, only when all types are considered simultaneously is there a significant $p$-value. This is in contrast to, for example, the $p$-values from the Westfall-Young permutation method as presented in Table 7. The drug Paclitaxel shows a significant correlation in type skin, and perhaps breast.

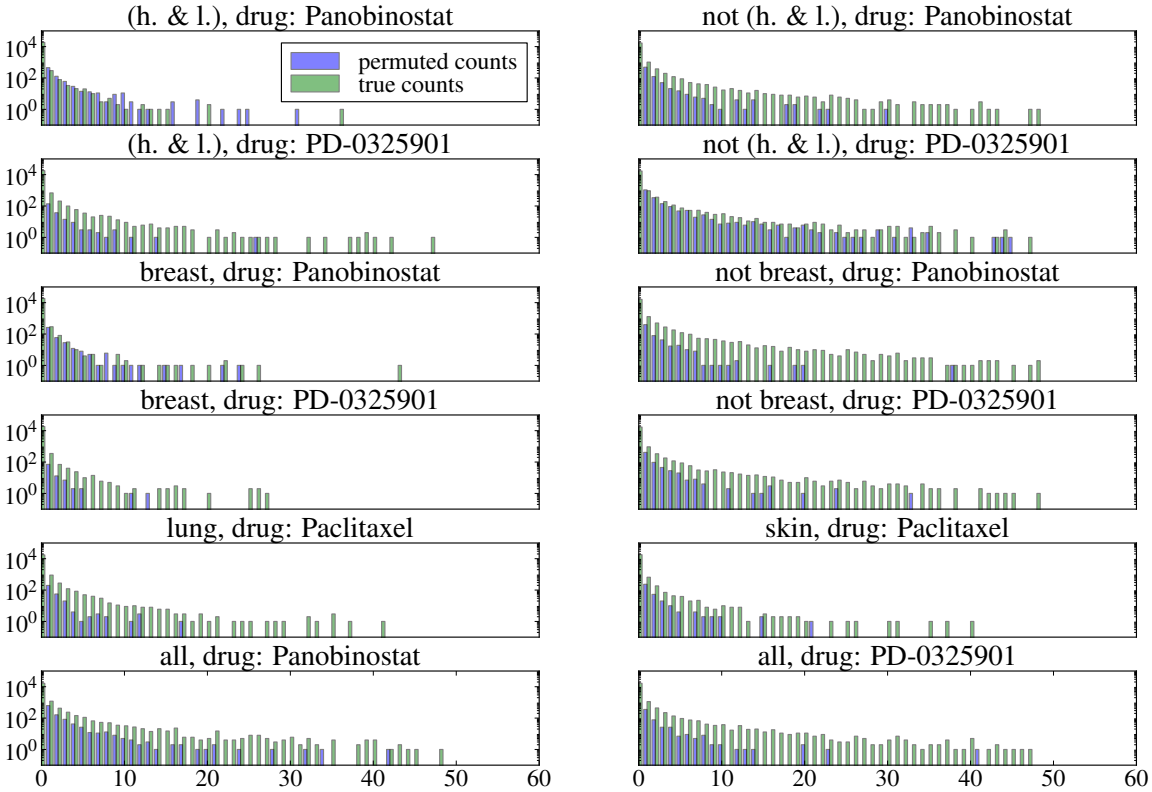


Fig 9: The frequency with which covariates have non-zero coefficient using the lasso, over 100 bootstrapped samples. The green bars are the counts using non-permuted responses. The blue bars use permuted responses. In certain cases there appear to be several covariates being selected more frequently using the true responses than would be expected if there were no real correlations (i.e. the green bars).

We have now considered the null hypothesis $H_{0,2}$ at length, and seen that it can be rejected for several type-drug pairs. The final two null hypotheses consider the similarity between different types. Hypothesis $H_{0,3}$ postulates a form of independence of the types, and $H_{0,4}$ that types are on the other hand indistinguishable. We assume that for type $k$ and drug $l$,

$$Y_{i,l}^k = \beta_l^k \cdot X_i^k + \epsilon_i. \tag{4.3}$$

In hypothesis $H_{0,3}$ we consider the independence of model coefficients of two types denoted $k$ and $k'$. To clarify what we mean by independence of model coefficients, we take a Bayesian view and consider model coefficient vectors as random variables. Let the sparsity of the coefficient vector $\beta_l^k$ be denoted by,

$$S_{i,l}^k = \mathbf{1}_{[\beta_{i,l}^k \neq 0]}.$$

Assume that the sparsity vector is a random variable with probability mass function,

$$\mathbf{P}\left(S_l^k = s_l^k\right) = \prod_{i=1}^{m} (p_l^k)^{s_{i,l}^k} (1 - p_l^k)^{1 - s_{i,l}^k}. \tag{4.4}$$

where $m$ is the number of genes, so that each coefficient is non-zero independently of other coefficients with probability $p_l^k$. Assume too that the size of non-zero coefficients are independent within and between types. Null hypothesis $H_{0,3}$ then states that the sparsity vectors are independent in the following sense:

$$\mathrm{P}\left(S_l^k = s_l^k, S_l^{k'} = s_l^{k'}\right) = \mathrm{P}\left(S_l^k = s_l^k\right) \mathrm{P}\left(S_l^{k'} = s_l^{k'}\right). \tag{$H_{0,3}$}$$

We have a strong prior belief that $H_{0,3}$ is false, based on a biological knowledge of how drugs function. Note that we do not observe the true vector $S_l^k$ but rather an estimate of it, denoted $\hat{S}_l^k$, obtained via the lasso. Roughly speaking, we expect to be able to disprove $H_{0,3}$ if $\hat{S}_l^k$ and $\hat{S}_l^{k'}$ have a large overlap. However, $H_{0,3}$ being true does not imply the validity of the analogous version for the sparsity estimate,

$$\mathrm{P}\left(\hat{S}_l^k = \hat{s}_l^k, \hat{S}_l^{k'} = \hat{s}_l^{k'}\right) = \mathrm{P}\left(\hat{S}_l^k = \hat{s}_l^k\right) \mathrm{P}\left(\hat{S}_l^{k'} = \hat{s}_l^{k'}\right). \tag{$H_{0,3a}$}$$

For $H_{0,3}$ to imply $H_{0,3a}$, and thus for a rejection of $H_{0,3a}$ to imply a rejection of $H_{0,3}$, we need to make further assumptions on the distribution of the covariates. A sufficient although probably unnecessarily strong assumption is that covariates within each model are independent[15].

We will not test hypothesis $H_{0,3a}$ directly, but rather one of its consequences. We will consider the vector of counts $c$ as introduced on line 8 of Algorithm 2, and denote by $c_l^k$ the vector $c$ for type $k$ and drug $l$. It is easy to see that $H_{0,3a}$ implies,

$$\mathrm{E}\left[\rho_P\left(c_l^k, c_l^{k'}\right)\right] = 0., \tag{$H_{0,3b}$}$$

A rejection of $H_{0,3b}$ will imply that types $k$ and $k'$ are not independent in the sense of $H_{0,3}$, and thus endorse the sharing of information between types $k$ and $k'$ while fitting. We choose to

[15]This is not a realistic assumption, and we need to develop weaker and more realistic assumptions to make our analysis more rigorous. The need for assumptions arises from the possibility that certain covariate structures may encourage certain coefficients to be selected more frequently than others, even given 4.4. However, one could argue that $H_{0,3a}$ is a more relevant null hypothesis and propose ignoring $H_{0,3}$ altogether, thus making connections between $H_{0,3}$ and $H_{0,3a}$ irrelevant.

TABLE 1
*The p-values of type independence (see $H_{0,3}$) for the drug Paclitaxel.*

| | central n.s. | skin | lung | ovary | l. intestine |
|---|---|---|---|---|---|
| central n.s. | - | 0.029 | 0.791 | 0.208 | 0.283 |
| skin | - | - | 0.031 | 0.573 | 0.014 |
| lung | - | - | - | 0.795 | 0.000 |
| ovary | - | - | - | - | 0.873 |
| l. intestine | - | - | - | - | - |

test $H_{0,3b}$ for all (type, drug) pairs which have a correlation significant at 1% in Table 5. We obtain $H_{0,3b}$ $p$-values through a permutation test, that is by comparing the correlation in $H_{0,3b}$ to equivalent correlations when $c_l^k$ is permuted. The results for the drug Paclitaxel are presented in Table 1, and the results for other drugs are presented in Appendix E. We notice that for several pairs we can reject $H_{0,3}$, and thus conclude that correlations between type coefficients do exist. It should be noted $H_{0,3b}$ is not the null hypothesis providing the most powerful test for dependence, as it ignores sign and magnitude of coefficients. We now turn our attention to the null hypothesis $H_{0,4}$,

$$(H_{0,4}) \qquad \beta_l^k = \beta_l^{k'}.$$

Disproving $H_{0,4}$ directly requires distinguishing between differences in estimates caused by noise in the data and any true differences that might exist. One way to go about disproving $H_{0,4}$ would be to compare the in-group variance of estimates obtained on mutually exclusive subsets of the data, with inter-group variances. If the inter-group variance is in some sense large compared to in-group variance, then $H_{0,4}$ can be rejected. However we have already seen that the lasso estimate within a group is strongly dependant on the data sample, and so this approach is not feasible with the small samples at our disposal.

It should be noted that the mean genomic covariates and responses are significantly different between types, as illustrated in Figure 10 and shown in our discussion on clustering in C. This is however not sufficient to reject $H_{0,4}$.

A second approach for disproving $H_{0,4}$ is to compare prediction mse by crossing training and test set types. More specifically, for types $k$ and $k'$ we can compare the four mses obtained by training and testing on one of $k$ and $k'$. The mses obtained by crossing all types for the drug Paclitaxel are illustrated in Table 2. The first observation is that the adjusted mses are overall quite poor, as will be discussed in the following section. What is important in the context of $H_{0,4}$ is that certain models fit on certain tissue types appear to perform badly when applied to other tissue types. In particular, haematopoietic and lymphoid tissue and the central nervous system appear to provide poor fits for the other types. Tables equivalent to Table 2 for other drugs can be found in Appendix E.

**5. Drug Response Prediction.** In this section we discuss the results obtained when using the local, global and indicator versions of the lasso. We do not discuss two additional aspects of

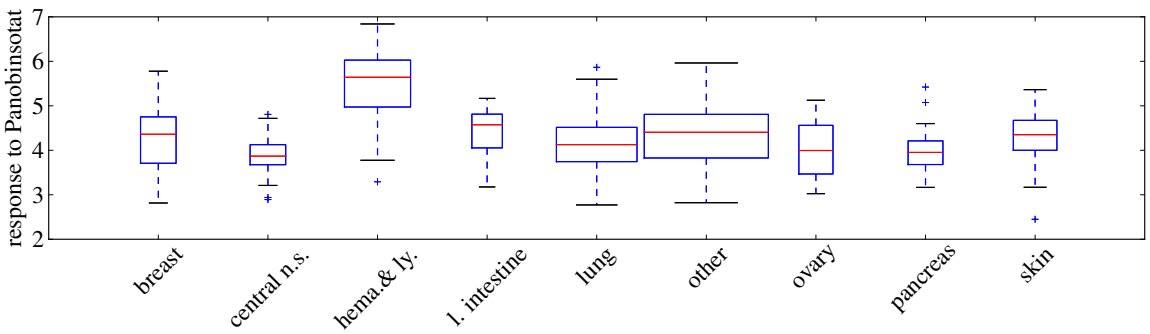


Fig 10: Boxplots of the response of cell lines to the drug Panobinostat, separated by tissue type. Box widths are proportional to number of cell lines. The difference between haematopoietic and lymphoid tissue, and the other tissue types is significant.

TABLE 2

*Using the drug Paclitaxel, the adjusted mse obtained with a lasso estimate trained on the row tissue type, and tested on the column tissue type. The diagonal terms are cross validated estimates, and off diagonal terms use all cell lines of the row type in fitting. Further comments are made in the text.*

| | central n.s. | other | hema.& ly. | skin | ovary | lung | pancreas | breast | l. intestine |
|---|---|---|---|---|---|---|---|---|---|
| central n.s. | 0.842 | 1.420 | 2.237 | 1.375 | 2.081 | 1.508 | 1.632 | 1.435 | 1.571 |
| other | 1.488 | 0.659 | 1.233 | 1.091 | 0.828 | 1.059 | 1.051 | 1.108 | 1.062 |
| hema.& ly. | 3.110 | 1.717 | 1.198 | 2.526 | 2.446 | 2.133 | 3.261 | 2.279 | 2.269 |
| skin | 1.714 | 1.103 | 1.349 | 0.921 | 0.719 | 1.026 | 1.210 | 0.899 | 1.112 |
| ovary | 1.414 | 1.817 | 2.157 | 1.261 | 0.809 | 1.524 | 2.356 | 1.584 | 1.660 |
| lung | 2.631 | 1.156 | 2.265 | 0.863 | 0.852 | 0.817 | 1.264 | 0.954 | 0.936 |
| pancreas | 1.379 | 1.093 | 2.437 | 1.000 | 1.011 | 1.001 | 1.151 | 1.004 | 1.012 |
| breast | 1.601 | 1.188 | 2.364 | 0.939 | 1.074 | 1.399 | 1.119 | 1.189 | 0.944 |
| l. intestine | 3.505 | 2.605 | 2.215 | 1.816 | 2.554 | 2.367 | 1.918 | 2.462 | 1.034 |

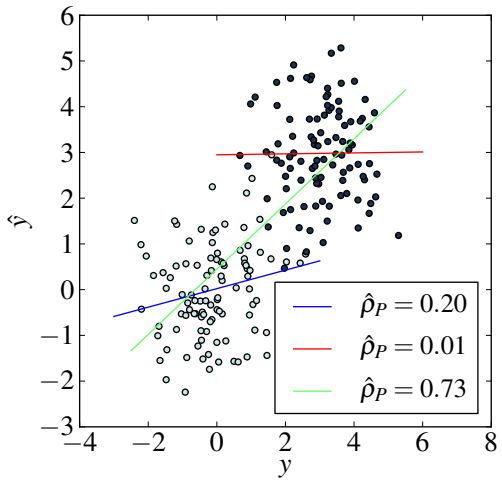


Fig 11: Two populations with (unknown) responses $y$ and predicted responses $\hat{y}$. Considering each population individually, we have Pearson correlations between predicted and true responses of 0.20 and 0.01. When the two populations are pooled, the Pearson correlation is 0.73. Of course the prediction errors are unchanged when the two populations are grouped, illustrating that the comparison of correlations between populations can be misleading.

creating predictive models, namely filtering and clustering, but refer the reader to Appendices B and C for more information about these complementary aspects.

5.1. *Evaluating prediction performance on the CCLE data.* In Section 3, we presented the quality of predictions in terms of an adjusted mse. There are other measures which may be considered, for example a Pearson or Spearman correlation between predicted and true response. In the one paper most closely resembling our work [6], the authors consider the Pearson correlation coefficient and the mse. The novel approach they use is to combine molecular features of the drugs with the genomic data and then predict response using the extended covariate. They claim that their approach results in substantially better predictions than when each drug is treated separately, and present a spectacular increase in Pearson correlation with their novel approach. However their prediction mse does not improve when drugs are pooled, suggesting to us that the observed increase in correlation is an artefact of pooling. Our postulate is illustrated in Figure 11. To avoid such a pitfall we focus on mse in only the breast cell lines. By more carefully considering the objective of drug prediction, which is to direct appropriate drugs to patients, it may be possible to derive a more meaningful performance measure, but this is beyond the scope of this project.

5.2. *Prediction results with the CCLE.* We do not succeed in making accurate drug response predictions using any of the methods presented in the report. However, we do observe that predictions on breast cell lines based on a global fit are significantly better than predictions based on a

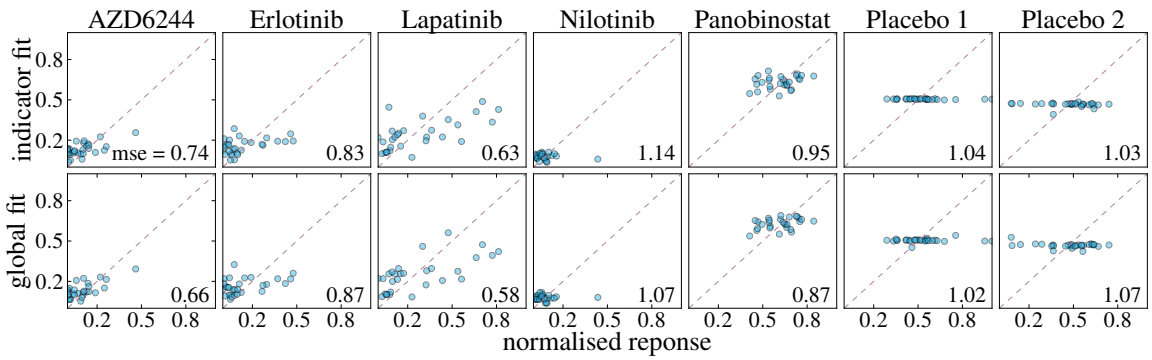


Fig 12: Scatter plots of true response (x-axis) versus predicted response (y-axis) of breast cell lines to several drugs, using the indicator lasso fit (above) and the global lasso fit (below). In the bottom right hand corners we give the adjusted mse associated with the points in the plots. The response of the two 'placebo' drugs are more or less constant, reflecting the ability of the cross-validated penalized regression fits to avoid overfitting. In all drugs other than Nilotinib we observe mild success at prediction. There does not appear to be a significant difference between the performance of the indicator and global methods.

local fit. We observed the poor predictive value of local fits in Table 2. In Figure 12 we present scatter plots of the response and predicted response to several drugs of breast cell lines, along with mses. We see there that the indicator method does not provide improved predictions over the global fit. We believe this is due to the low number of breast cell lines in the CCLE, which renders any bias reduction gained with the indicator method insignificant given the resulting increased variance.

**6. Discussion.** We have analysed correlations between genomic covariates and drug responses within and between different tissue types. We have then considered using the genomic covariates to predict the drug response of cancer cell lines, and have attempted to harness the similarity across tissue types to make improved predictions. The proposed method has been shown to perform well in simulation studies, but has failed to provide significant improvements when applied to the real drug prediction problem. We suspect that the failure of the use of tissue type via the indicator method to improve predictions is due to too small a number of cell lines available for model fitting. The extra variance incurred by dividing an already small number of cell lines into separate types outweighs any potential gain in bias reduction. We hypothesize that future datasets containing larger numbers of cell lines will prove move amenable to approaches making use of tissue type differences.

## APPENDIX A: INDICATOR LASSO PROOFS

The main result we prove present here is that in the limit $\alpha_{\texttt{in}} \to 0$ and $\alpha_{\texttt{in}} \to \infty$, the indicator lasso reduces to forms of local and global lassos. We also present a result pertaining to intermediate values of $\alpha_{\texttt{in}}$, and present a proof highlighting a major difference between the local indicator and fused lassos. Let us repeat what the expression is which we wish to minimize with the indicator lasso,

$$\ell_{\texttt{in}}\left(a,\beta\right) = \ell\left(a,\beta|\underline{x},\underline{y}\right) + \lambda_{\texttt{in}}\|\beta_{\texttt{gl}}\|_1 + \alpha_{\texttt{in}}\lambda_{\texttt{in}}\sum_{k=1}^{K}\|\beta_{\texttt{lo}}^k\|_1. \tag{A.1}$$

**A.1. In the limit $\alpha_{\texttt{in}} \to \infty$.** In A.1, as $\alpha_{\texttt{in}} \to \infty$ any perturbations from the global fit are penalized to such an extent as to make them not worthwhile, resulting in $\beta_{\texttt{lo}}^1 = \ldots \beta_{\texttt{lo}}^K = 0$. However we do not recover the global fit as presented in 2.2, as we still the constant terms $a_{\texttt{in}}^k, 1 \leq k \leq K$ which are unpenalized. Thus in the limit, the coefficients are the same between models but the intercept terms are free to differ.

**A.2. In the limits $\alpha_{\texttt{in}} \to 0$.** First we note that for a constant $\alpha_{\texttt{in}}$ the solution found with the indicator lasso as presented in A.1 for the optimal $\lambda_{\texttt{in}}$, is the same as that found when the penalty term is divided through by $\alpha_{\texttt{in}}$,

$$\ell_{\texttt{in}}\left(a,\beta\right) = \ell\left(a,\beta|\underline{x},\underline{y}\right) + \frac{\lambda_{\texttt{in}}}{\alpha_{\texttt{in}}}\|\beta_{\texttt{gl}}\|_1 + \lambda_{\texttt{in}}\sum_{k=1}^{K}\|\beta_{\texttt{lo}}^k\|_1, \tag{A.2}$$

albeit the optimal $\lambda_{\texttt{in}}$ in A.2 is a factor $\alpha_{\texttt{in}}$ larger than that in A.1. We now apply a similar argument to that in case $\alpha_{\texttt{in}} \to \infty$ to show that as As $\alpha_{\texttt{in}} \to 0$ the penalty on the global term increases to such an extent as to reduce $\beta_{\texttt{gl}}$ to 0. A more precise proof follows as a corollary of the following lemma.

**A.3. A lemma relating $\beta_{\texttt{gl}}$ and the $\beta_{\texttt{lo}}^k$s.** We present a result which puts a bound on $\beta_{\texttt{gl}}$ coefficients in terms of $\beta_{\texttt{lo}}^k$s coefficients. We first introduce some notation. For a particular covariate index $j$, let the sets of types with positive, zero and negative coefficients for a given $(\lambda,\alpha)$-indicator penalty be given as

$$\begin{aligned}
\mathcal{U}_j^+(\lambda,\alpha) &= \{k \,|\, \beta_{\texttt{in},j}^k > 0\},\\
\mathcal{U}_j^0(\lambda,\alpha) &= \{k \,|\, \beta_{\texttt{in},j}^k = 0\},\\
\mathcal{U}_j^-(\lambda,\alpha) &= \{k \,|\, \beta_{\texttt{in},j}^k < 0\}.
\end{aligned}$$

We define the excess of positive coefficients as

$$K_j(\lambda,\alpha) = |\mathcal{U}_j^+(\lambda,\alpha)| - |\mathcal{U}_j^0(\lambda,\alpha)| - |\mathcal{U}_j^-(\lambda,\alpha)|.$$

We define the $n$'th smallest coefficient in $\mathcal{U}_j^+$ to be $\beta_j^{+[n]}$. $K_j(\lambda,\alpha)$ may be negative, but we assume without loss of generality that $|\mathcal{U}_j^+(\lambda,\alpha)| \geq |\mathcal{U}_j^-(\lambda,\alpha)|$.

LEMMA 1. *If $K_j < \frac{1}{\alpha}$, then $\beta_{gl} = 0$. Otherwise for integer $n \leq \frac{K_j}{2}$, if $\dfrac{1}{K_j - 2n} < \alpha < \dfrac{1}{K_j - 2(n-1)}$, then $\beta_{gl} = \beta_j^{+[n]}$. Finally, for integer $n \leq \frac{K_j}{2}$ if $\alpha = K_j - 2n$, then $\beta_j^{+[n]} \leq \beta_{gl} \leq \beta_j^{+[n+1]}$.*

PROOF. *(sketch)* We assume that the minimum (in covariate $j$) is obtained at $\beta_{\mathtt{in},j}$, and then consider how best to divide $\beta_{\mathtt{in},j}$ between global and local components so as to minimize the penalty term. Consider the respective global and local components of the penalty,

$$\begin{aligned}\mathcal{P}_{\mathtt{in}}\left(\beta_{\mathtt{in},j}\right) &= \lambda_{\mathtt{in}}\|\beta_{\mathtt{gl}}\|_1 \qquad + \alpha_{\mathtt{in}}\lambda_{\mathtt{in}}\sum_{k=1}^{K}\|\beta_{\mathtt{lo}}^k\|_1 \\ &= \mathcal{P}_{\mathtt{in\text{-}gl}}\left(\beta_{\mathtt{gl}}\right) + \mathcal{P}_{\mathtt{in\text{-}lo}}\left(\beta_{\mathtt{lo}}\right).\end{aligned}$$

The global penalty is minimized when $\beta_{\mathtt{gl},j} = 0$ The local penalty $\mathcal{P}_{\mathtt{lo}}$ is minimized if $\beta_{\mathtt{gl}}$ is the median of $\beta_{\mathtt{in},j}$. In the case $K_j < 0$ these minima coincide at $\beta_{\mathtt{in},j} = 0$. In other cases, the value of $\alpha_{\mathtt{in}}$ dictates which is more important. When $\alpha_{\mathtt{in}}$ is large, the optimal $\beta_{\mathtt{gl}}$ is near the median of $\beta_{\mathtt{in},j}$. If $\alpha_{\mathtt{in}}$ is small, $\beta_{\mathtt{gl}}$ approaches zero. Indeed, if $\alpha_{\mathtt{in}} < 1/K_j$ then $\beta_{\mathtt{gl}}$. □

**A.4. Difference between indicator and fused.** We can rearrange the fused lasso to get,

$$\begin{aligned}P_{\mathtt{fu}}\left(\beta^1 \cdots \beta^k\right) &= \lambda_{\mathtt{fu\text{-}1}}\sum_{k=1}^{K}\|\beta^k\|_1 + \lambda_{\mathtt{fu\text{-}2}}\sum_{k_1=1}^{K}\sum_{k_2=k_1+1}^{K}\|\beta^{k_1} - \beta^{k_2}\|_1 \\ &= \lambda_{\mathtt{fu\text{-}1}}\sum_{k=1}^{K}\|\beta^k\|_1 + \lambda_{\mathtt{fu\text{-}2}}\sum_{k=1}^{K}\left(2k - N - 1\right)\beta^{[k]}\end{aligned}$$

where $\beta^{[k]}$ is the $k$'th largest coefficient. Thus we see that the coefficients which are further away from the median are penalized more. To make a clearer comparison with the indicator method, suppose that all the $\beta^k$'s are positive. We then have,

$$P_{\mathtt{fu}}\left(\beta^1 \cdots \beta^k\right) = \lambda_{\mathtt{fu\text{-}1}}K\|\mu_\beta\|_1 + \lambda_{\mathtt{fu\text{-}2}}\sum_{k=1}^{K}\left(2k - N - 1\right)\left(\beta^{[k]} - \mu_\beta\right)$$

where $\mu_\beta$ is the mean of the $\beta_k$'s. This is equivalent to a version of the Indicator lasso, where $\beta_0$ is constrained to be the mean, and the penalty applied to each perturbation term depends on its rank. Extremely ranked perturbations have a larger penalty. We saw in A.3 that $\beta_{\mathtt{gl}}$ is in general not the mean.

## APPENDIX B: FILTERING

In many ways the lasso can be viewed as a type of filter, but performing the lasso on all $\sim 20,000$ covariates is computationally burdensome[16]. By performing a fast pre-filter of covariates, exploring the different variants of the lasso fits becomes easier. A second potential advantage of filtering could be better predictions, although we have not observed this. We have attempted filtering for both genes which are most correlated within a specific type and across all types simultaneously, within a cross-validated framework much like that presented in 4.3.

## APPENDIX C: CLUSTERING

Tissue type does seem like the natural variable by which to group cell lines, but we have explored other possible groupings to see if better predictions can arise when using the indicator lasso. The alternative groupings were defined according to the two clusterings below.

**C.1. Clustering by covariates.** Using the $k$-medoid clustering technique, cell lines are clustered by the $\sim 20,000$ genomic covariates. Clusters are similar to the real tissue types, as illustrated in Figure 13.

**C.2. Clustering by response.** Here the cell lines are clustered according to their responses to the $\sim 50$ drugs. In an application setting, the response to a drug is of course not available. For this reason, we perform initial clustering on responses but then redefine clusters in terms of their mean genomic covariates. This is done in a cross-validated framework, so that the final cluster of a cell line does not depend (directly) on it's drug responses. Matching up cluster names from different folds can be done at the end to minimise inter-fold cluster mean variance.

## APPENDIX D: ADDITIONAL FIGURES

We present three additional Figures. Figure 14 compares $H_{0,2}$ $p$-values using four varieties of multiple correcting. Figure 15 presents the results of the cross-validated correlation test applied to the drugs Lapatinib and PD-0325901. Figure 16 presents the results of Benjamini-Hochberg FDR correction and Westfall-Young permutation for drug-type correlation significance.

## APPENDIX E: TABLES

We present additional tables. Tables with precise $p$-values correlations for each type-drug pair using both Spearman and Pearson correlations, and Benjamini-Hochberg FDR corrections and Westfall-Young permutation tests are first presented, followed by Tables of $p$-values for to the test of null hypothesis $H_{0,3b}$ for several drugs are presented.

[16]For a $n \times p$ covariate matrix, the operations required in $O(n\,p\min(n,p))$

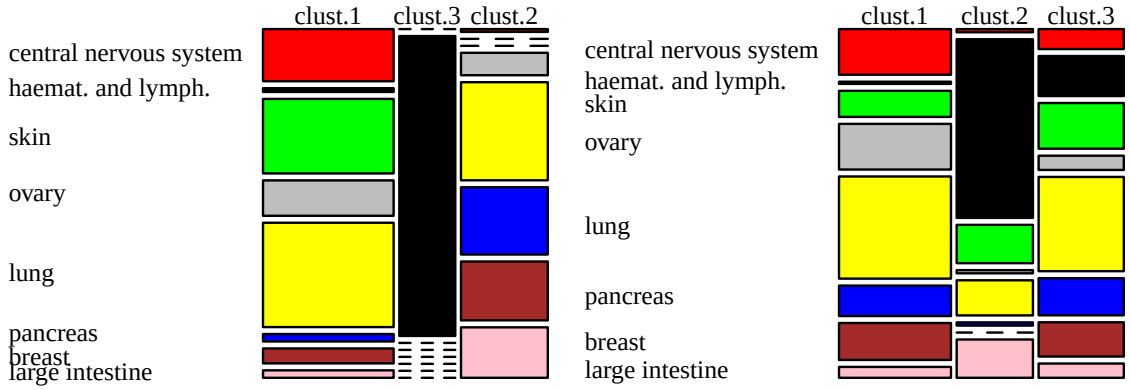


Fig 13: Results of clustering the $\sim 500$ cell lines. On the left, the $\sim 20,000$ covariates are used for clustering and on the right the $\sim 50$ responses are used. The clusters are strongly dependent on type ($p$-value 10e-28 on right using $\chi^2$ g.o.f test).

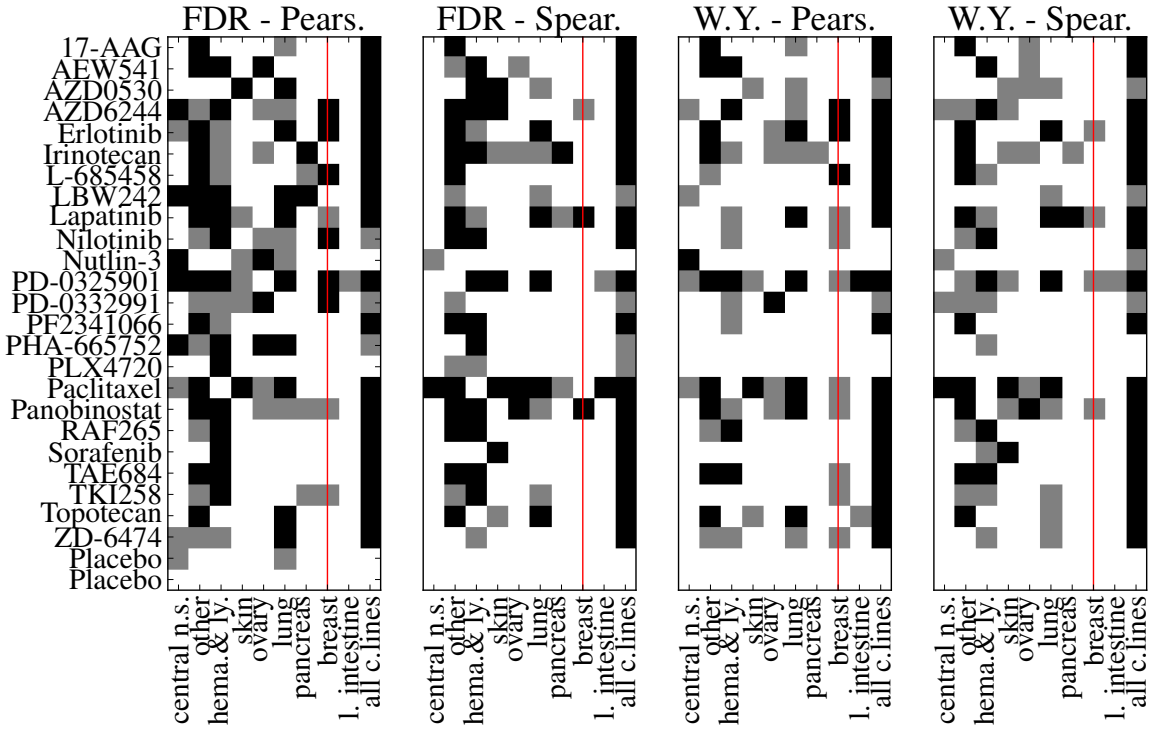


Fig 14: Hypothesis $H_{0,2}$ $p$-values using the four techniques described in 4.2. $p$-values which are less than 5% are shown in gray, and $p$-values of below 1% are shown in black. There is a strong similarity of $p$-values using the different approaches. The tissue types with more cell lines generally exhibit more significant correlations, and when all types are pooled (all c.lines, on the right of each inset) the largest number of significant drugs are reported. Reassuringly, the two random responses (placebos) are the only 'drugs' not to show any significant correlations.

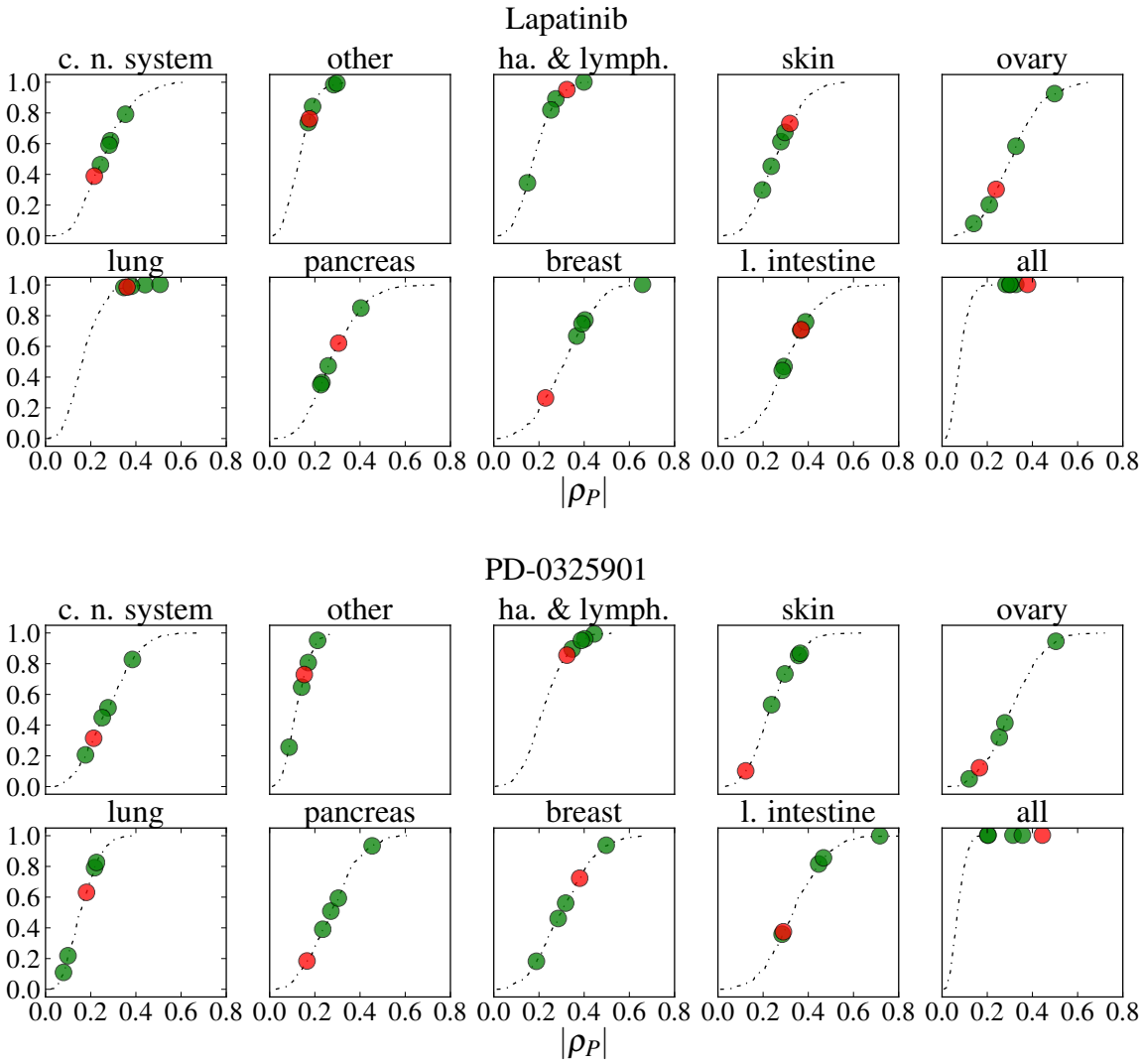


Fig 15: Permutation method applied to Lapatinib and PD-0325901. Few subtypes appear to have significant genomic correlates with drug response, but when all types are considered the correlations become significant.

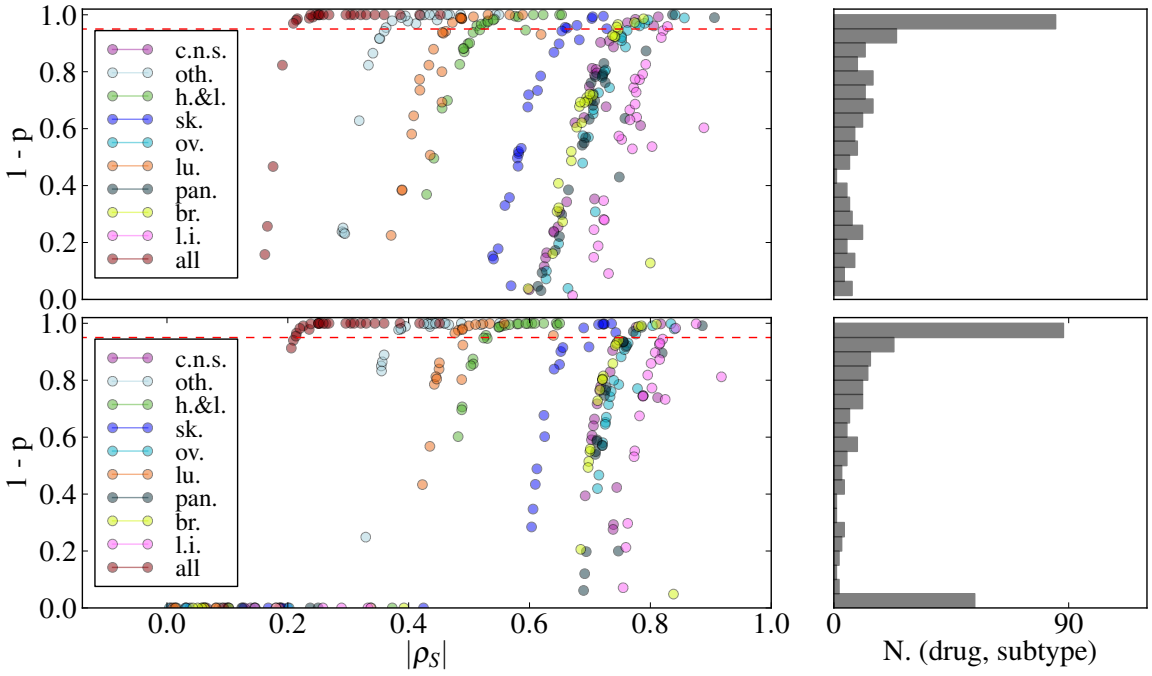


Fig 16: As in Figure 7, but using the Spearman correlation coefficient. Above: Westfall-Young and below: Benjamini-Hochberg FDR correction.

TABLE 3
*Benjamini-Hochberg FDR corrected p-values of largest Pearson correlation. Red values signify 1% significance.*

| | cent: | othe: | haem: | skin: | ovar: | lung: | panc: | brea: | larg: | all: |
|---|---|---|---|---|---|---|---|---|---|---|
| 17-AAG | 100 | 0.019 | 6.72 | 58.19 | 10.03 | 4.57 | 100 | 62.76 | 68.59 | 0 |
| AEW541 | 40.72 | 0.027 | 0.031 | 100 | 0.935 | 100 | 37.48 | 100 | 100 | 0.032 |
| AZD0530 | 100 | 66.05 | 5.41 | 0.333 | 17.68 | 0.207 | 100 | 11.48 | 100 | 0.895 |
| AZD6244 | 0.006 | 1 | 0 | 19.57 | 2.52 | 1 | 100 | 0.007 | 21.08 | 0 |
| Erlotinib | 4.8 | 0 | 4.47 | 5.76 | 12.36 | 0.001 | 18.02 | 0.01 | 18.42 | 0 |
| Irinotecan | 55.33 | 0.011 | 3.49 | 18.15 | 4.74 | 12.55 | 0.644 | 100 | 63.4 | 0 |
| L-685458 | 11.86 | 0.027 | 4.4 | 6.55 | 41.87 | 100 | 1.81 | 0 | 16.89 | 0.059 |
| LBW242 | 0 | 0 | 0.586 | 10.32 | 55.31 | 0.003 | 0.262 | 70.44 | 44.85 | 0 |
| Lapatinib | 100 | 0.003 | 0.864 | 1.37 | 100 | 0 | 22.93 | 1.8 | 44.99 | 0 |
| Nilotinib | 100 | 1.15 | 0.003 | 86.67 | 3.15 | 4.42 | 96.49 | 0 | 25.03 | 2.15 |
| Nutlin-3 | 0.005 | 13.28 | 9.93 | 2.24 | 0.782 | 2.85 | 100 | 14.33 | 23.8 | 13.42 |
| PD-0325901 | 0.202 | 0.158 | 0.001 | 2.14 | 5.96 | 0.655 | 100 | 0.373 | 1.33 | 0 |
| PD-0332991 | 100 | 1.24 | 4.12 | 2.74 | 0.024 | 14.43 | 23.24 | 0.005 | 43.26 | 3.29 |
| PF2341066 | 14.17 | 0.114 | 1.43 | 100 | 18.21 | 15.12 | 29.63 | 35.3 | 100 | 0.011 |
| PHA-665752 | 0.308 | 1.24 | 0.218 | 100 | 0.967 | 0.017 | 15.12 | 61.58 | 61.49 | 2.96 |
| PLX4720 | 14.69 | 9.74 | 0.004 | 100 | 100 | 55.42 | 100 | 100 | 17.55 | 12.19 |
| Paclitaxel | 3.48 | 0.001 | 82.11 | 0.483 | 4.17 | 0.438 | 64.32 | 14.95 | 36.1 | 0 |
| Panobinostat | 28.24 | 0.002 | 0.224 | 23.33 | 4.93 | 1.83 | 1.42 | 1.48 | 52.89 | 0 |
| RAF265 | 15.46 | 1.33 | 0.005 | 97.65 | 10.81 | 10.87 | 100 | 11.23 | 100 | 0.001 |
| Sorafenib | 100 | 26.48 | 0.033 | 17.94 | 49.92 | 25.89 | 100 | 100 | 27.84 | 0.064 |
| TAE684 | 100 | 0.019 | 0.01 | 39.89 | 6.06 | 100 | 100 | 5.8 | 22.2 | 0.432 |
| TKI258 | 100 | 2.6 | 0.499 | 78.65 | 66.65 | 11.12 | 2.5 | 1.19 | 10.73 | 0.115 |
| Topotecan | 12.56 | 0.007 | 45.27 | 6.24 | 11.53 | 0.751 | 100 | 47.55 | 12.63 | 0 |
| ZD-6474 | 3.56 | 3.35 | 3.99 | 100 | 29.89 | 0.033 | 100 | 5.66 | 100 | 0.001 |
| Placebo | 1.16 | 16.99 | 70.68 | 45.65 | 7.24 | 3.3 | 100 | 32.3 | 100 | 90.98 |
| Placebo | 47.49 | 100 | 51.44 | 100 | 100 | 100 | 41.51 | 100 | 100 | 49.1 |

TABLE 4
*Largest Pearson correlations, appearing in red if it is significant at 1% using Benjamini-Hochberg FDR correction.*

| | cent: | othe: | haem: | skin: | ovar: | lung: | panc: | brea: | larg: | all: |
|---|---|---|---|---|---|---|---|---|---|---|
| 17-AAG | 0.03 | 0.446 | -0.52 | -0.61 | -0.76 | 0.47 | 0.16 | -0.69 | -0.78 | 0.375 |
| AEW541 | -0.7 | 0.442 | 0.61 | 0.08 | -0.802 | -0 | 0.71 | -0.1 | -0.33 | 0.252 |
| AZD0530 | 0.09 | 0.33 | -0.52 | 0.719 | 0.74 | 0.522 | 0.09 | 0.73 | 0.43 | 0.226 |
| AZD6244 | 0.856 | 0.39 | 0.674 | 0.64 | 0.78 | 0.5 | -0.09 | 0.855 | 0.79 | 0.343 |
| Erlotinib | 0.75 | 0.512 | -0.53 | -0.66 | 0.75 | 0.593 | 0.73 | 0.85 | 0.8 | 0.338 |
| Irinotecan | 0.79 | 0.574 | -0.61 | 0.7 | 0.82 | 0.62 | -0.89 | 0.01 | 0.9 | 0.46 |
| L-685458 | 0.73 | 0.451 | -0.53 | -0.67 | 0.72 | 0.03 | 0.78 | 0.891 | 0.81 | 0.251 |
| LBW242 | 0.93 | 0.544 | 0.564 | 0.65 | 0.71 | 0.584 | 0.814 | 0.69 | 0.77 | 0.287 |
| Lapatinib | -0.01 | 0.467 | 0.557 | 0.69 | 0.03 | 0.634 | 0.73 | 0.77 | -0.77 | 0.373 |
| Nilotinib | 0.09 | 0.45 | 0.65 | 0.65 | 0.79 | 0.52 | 0.77 | 0.926 | 0.83 | -0.24 |
| Nutlin-3 | 0.859 | 0.36 | -0.51 | 0.68 | 0.805 | 0.48 | 0.01 | 0.73 | -0.79 | 0.2 |
| PD-0325901 | 0.809 | 0.419 | 0.663 | 0.68 | 0.77 | 0.504 | -0.09 | 0.8 | 0.85 | 0.365 |
| PD-0332991 | -0.06 | 0.43 | 0.54 | -0.72 | 0.856 | 0.49 | 0.79 | 0.873 | 0.8 | 0.23 |
| PF2341066 | 0.73 | 0.424 | 0.55 | -0.05 | 0.74 | 0.45 | 0.72 | 0.71 | -0.35 | 0.26 |
| PHA-665752 | 0.803 | 0.39 | -0.58 | -0.17 | 0.801 | 0.561 | 0.74 | 0.69 | 0.77 | 0.22 |
| PLX4720 | 0.74 | 0.36 | 0.64 | 0.15 | 0.05 | 0.42 | 0.06 | -0.44 | 0.8 | -0.2 |
| Paclitaxel | 0.77 | -0.475 | 0.47 | 0.712 | 0.77 | -0.511 | 0.7 | 0.73 | 0.78 | 0.337 |
| Panobinostat | -0.71 | 0.472 | -0.58 | 0.63 | -0.77 | 0.49 | -0.79 | -0.78 | -0.77 | -0.304 |
| RAF265 | 0.78 | -0.42 | -0.642 | 0.6 | -0.75 | -0.48 | -0.06 | -0.74 | -0.2 | 0.29 |
| Sorafenib | 0.11 | 0.35 | 0.609 | 0.64 | -0.72 | 0.44 | 0.16 | -0.01 | 0.79 | 0.247 |
| TAE684 | 0.17 | -0.445 | 0.626 | 0.62 | 0.77 | -0.18 | -0.34 | 0.75 | -0.79 | 0.232 |
| TKI258 | 0.45 | 0.38 | -0.566 | 0.6 | 0.71 | 0.46 | -0.77 | 0.78 | 0.81 | 0.242 |
| Topotecan | 0.73 | 0.457 | 0.48 | 0.66 | -0.75 | 0.502 | 0.17 | -0.7 | 0.8 | 0.426 |
| ZD-6474 | 0.77 | 0.38 | 0.53 | -0.45 | 0.77 | 0.549 | 0.21 | 0.75 | 0.13 | 0.278 |
| Placebo | 0.78 | 0.35 | 0.47 | -0.62 | 0.76 | 0.48 | -0.01 | 0.71 | -0.5 | 0.18 |
| Placebo | 0.7 | 0.06 | -0.48 | 0.13 | -0 | -0.08 | -0.71 | 0.11 | 0.11 | -0.19 |

TABLE 5
*Benjamini-Hochberg FDR corrected p-values of largest Spearman correlation. Red values signify 1% significance.*

| | cent: | othe: | haem: | skin: | ovar: | lung: | panc: | brea: | larg: | all: |
|---|---|---|---|---|---|---|---|---|---|---|
| 17-AAG | 40.94 | 0.062 | 12.56 | 32.32 | 11.78 | 21.44 | 93.86 | 48.53 | 22.75 | 0 |
| AEW541 | 28.3 | 1.1 | 0.047 | 56.58 | 3.75 | 100 | 6.59 | 100 | 100 | 0.039 |
| AZD0530 | 23.65 | 14.98 | 0.228 | 0.263 | 6.98 | 2.46 | 100 | 100 | 100 | 0.903 |
| AZD6244 | 22.17 | 0.036 | 0.003 | 0.25 | 100 | 16.08 | 22.46 | 1.14 | 10.27 | 0 |
| Erlotinib | 100 | 0.007 | 1.05 | 100 | 41.94 | 0.017 | 46.05 | 18.51 | 25.42 | 0 |
| Irinotecan | 26.11 | 0.487 | 0.586 | 3.25 | 1.95 | 4.28 | 0.855 | 95.14 | 18.81 | 0 |
| L-685458 | 33.64 | 0.001 | 14.29 | 14.46 | 28.56 | 43.23 | 40.98 | 23.37 | 12.72 | 0.06 |
| LBW242 | 100 | 1.77 | 29.37 | 71.55 | 100 | 1.91 | 100 | 100 | 5.56 | 3.47 |
| Lapatinib | 100 | 0 | 4.25 | 100 | 100 | 0.094 | 2.88 | 0.864 | 92.9 | 0 |
| Nilotinib | 70.72 | 0.472 | 0.117 | 100 | 21.5 | 19.77 | 10.44 | 100 | 26.75 | 0.152 |
| Nutlin-3 | 1.41 | 13.53 | 30.38 | 16.1 | 53.33 | 18.76 | 25.31 | 7.13 | 100 | 5.98 |
| PD-0325901 | 100 | 11.12 | 0.002 | 0.12 | 20.04 | 0.885 | 100 | 5.93 | 1.71 | 0 |
| PD-0332991 | 57.68 | 1.15 | 5.34 | 8.31 | 34.79 | 100 | 25.42 | 100 | 100 | 2.16 |
| PF2341066 | 100 | 0.023 | 0.727 | 65.27 | 9.25 | 100 | 41.45 | 19.61 | 78.73 | 0.197 |
| PHA-665752 | 19.77 | 100 | 0.49 | 100 | 100 | 19.58 | 80.23 | 100 | 100 | 1.81 |
| PLX4720 | 100 | 2.02 | 1.36 | 9.91 | 100 | 100 | 8.27 | 50.66 | 100 | 4.52 |
| Paclitaxel | 0.212 | 0.06 | 100 | 0.431 | 0.853 | 0.201 | 1.4 | 27.2 | 0.178 | 0 |
| Panobinostat | 60.6 | 0.006 | 0.197 | 8.52 | 0.124 | 2.23 | 23.5 | 0.334 | 7.39 | 0 |
| RAF265 | 72.41 | 0.35 | 0.02 | 39.83 | 25.8 | 7.66 | 80.01 | 14.14 | 18.19 | 0 |
| Sorafenib | 9.54 | 75.13 | 5.16 | 0.289 | 33 | 56.67 | 6.59 | 100 | 32.53 | 0.027 |
| TAE684 | 100 | 0.021 | 0.034 | 100 | 58.07 | 100 | 100 | 7.8 | 44.83 | 0.036 |
| TKI258 | 6.81 | 2.25 | 0.092 | 100 | 42.94 | 3.42 | 6.42 | 45.01 | 25.67 | 0.003 |
| Topotecan | 100 | 0.101 | 14.9 | 1.62 | 7.82 | 0.439 | 87.97 | 44.17 | 6.99 | 0 |
| ZD-6474 | 17.06 | 16.75 | 1.3 | 100 | 22.93 | 13.88 | 42.94 | 79.42 | 46.88 | 0.017 |
| Placebo | 36 | 100 | 39.81 | 51.13 | 23.94 | 100 | 100 | 19.82 | 70.3 | 100 |
| Placebo | 39.41 | 100 | 100 | 100 | 35.49 | 100 | 45.09 | 100 | 14.1 | 8.71 |

TABLE 6
*Largest Spearman correlations, appearing in red if it is significant at 1% using Benjamini-Hochberg FDR correction.*

| | cent: | othe: | haem: | skin: | ovar: | lung: | panc: | brea: | larg: | all: |
|---|---|---|---|---|---|---|---|---|---|---|
| 17-AAG | 0.7 | 0.431 | 0.51 | -0.62 | 0.75 | -0.44 | 0.69 | 0.7 | 0.8 | 0.359 |
| AEW541 | -0.71 | -0.39 | 0.604 | -0.61 | -0.78 | 0.01 | 0.75 | -0.06 | -0.26 | 0.251 |
| AZD0530 | 0.72 | 0.35 | -0.579 | 0.723 | 0.76 | 0.48 | 0.15 | -0.39 | 0.33 | -0.226 |
| AZD6244 | 0.72 | -0.438 | -0.644 | 0.724 | 0.03 | 0.45 | -0.73 | -0.78 | -0.81 | 0.304 |
| Erlotinib | 0.14 | 0.458 | -0.55 | -0.17 | -0.72 | 0.558 | 0.71 | -0.72 | 0.79 | 0.321 |
| Irinotecan | -0.81 | 0.519 | -0.646 | 0.74 | -0.84 | 0.64 | -0.886 | 0.84 | -0.92 | 0.451 |
| L-685458 | -0.71 | 0.487 | 0.51 | 0.65 | 0.73 | 0.44 | -0.71 | -0.72 | 0.82 | -0.251 |
| LBW242 | 0.01 | -0.39 | -0.49 | -0.6 | 0.2 | 0.49 | 0.24 | -0.34 | 0.82 | -0.21 |
| Lapatinib | 0.06 | 0.551 | -0.53 | 0.1 | 0 | 0.534 | 0.77 | 0.785 | -0.75 | 0.386 |
| Nilotinib | 0.74 | 0.466 | 0.597 | 0.2 | 0.75 | -0.49 | 0.82 | -0.06 | 0.82 | -0.263 |
| Nutlin-3 | 0.78 | 0.36 | -0.49 | 0.64 | 0.71 | 0.45 | -0.72 | -0.74 | -0.34 | 0.21 |
| PD-0325901 | -0.19 | 0.36 | -0.65 | -0.736 | -0.74 | 0.499 | -0.1 | -0.75 | 0.84 | 0.33 |
| PD-0332991 | 0.74 | -0.44 | 0.53 | 0.7 | 0.73 | 0.08 | 0.79 | -0.18 | -0.18 | 0.24 |
| PF2341066 | -0.07 | -0.443 | -0.56 | -0.61 | -0.76 | -0.06 | 0.71 | 0.72 | 0.76 | 0.238 |
| PHA-665752 | -0.72 | 0.09 | -0.567 | -0.13 | 0.03 | 0.45 | -0.69 | -0.05 | 0.29 | -0.22 |
| PLX4720 | 0.1 | -0.39 | -0.55 | -0.65 | 0.02 | -0.01 | -0.76 | -0.7 | -0.15 | -0.21 |
| Paclitaxel | 0.817 | 0.432 | -0.37 | 0.714 | 0.804 | 0.523 | -0.79 | -0.71 | -0.876 | 0.344 |
| Panobinostat | 0.69 | 0.46 | -0.582 | 0.66 | 0.834 | -0.49 | -0.73 | -0.81 | 0.81 | 0.31 |
| RAF265 | 0.74 | -0.435 | -0.623 | 0.63 | 0.73 | -0.49 | 0.75 | -0.74 | 0.8 | 0.296 |
| Sorafenib | 0.74 | 0.33 | 0.52 | 0.721 | -0.73 | -0.42 | -0.75 | 0.1 | 0.78 | 0.253 |
| TAE684 | 0.13 | -0.444 | 0.609 | 0.04 | 0.71 | -0.19 | -0.25 | -0.74 | -0.77 | 0.251 |
| TKI258 | 0.74 | 0.38 | -0.594 | -0.19 | 0.72 | 0.48 | -0.76 | 0.7 | 0.79 | 0.268 |
| Topotecan | 0.03 | 0.425 | 0.5 | -0.69 | -0.76 | 0.511 | 0.69 | -0.7 | -0.82 | 0.418 |
| ZD-6474 | -0.73 | 0.36 | 0.55 | -0.42 | 0.78 | -0.45 | -0.72 | 0.68 | -0.77 | 0.259 |
| Placebo | -0.71 | -0.14 | 0.48 | 0.61 | 0.74 | -0.01 | -0.12 | -0.72 | -0.76 | -0.09 |
| Placebo | -0.7 | 0.09 | -0.06 | 0.19 | 0.73 | -0.08 | -0.71 | 0.04 | 0.8 | -0.21 |

TABLE 7
*Westfall-Young corrected p-values of largest Pearson correlation. Red values signify 1% significance.*

| | cent: | othe: | haem: | skin: | ovar: | lung: | panc: | brea: | larg: | all: |
|---|---|---|---|---|---|---|---|---|---|---|
| 17-AAG | 80.3 | 0 | 5.2 | 47.1 | 8.6 | 1.9 | 75.9 | 87.9 | 51.4 | 0 |
| AEW541 | 43.6 | 0 | 0 | 87.7 | 7.6 | 67.9 | 11.8 | 87.3 | 85 | 0 |
| AZD0530 | 44.5 | 48.7 | 15.7 | 1.2 | 6.4 | 2.3 | 52.2 | 6.6 | 84.2 | 1.1 |
| AZD6244 | 2.4 | 13 | 0.3 | 6.9 | 9.4 | 2.8 | 75.6 | 0.3 | 8 | 0 |
| Erlotinib | 10.3 | 0 | 17.3 | 18.6 | 4.6 | 0.6 | 12.4 | 0.2 | 29.1 | 0 |
| Irinotecan | 19.3 | 0 | 3.8 | 5.8 | 2.5 | 3.4 | 1.2 | 47.9 | 23.9 | 0 |
| L-685458 | 56 | 3.5 | 18.9 | 49.7 | 35.5 | 71 | 28 | 0.9 | 13.1 | 0.3 |
| LBW242 | 1.4 | 12.8 | 19.6 | 22.9 | 63.7 | 6.6 | 6.3 | 83 | 29 | 0.2 |
| Lapatinib | 95 | 8 | 1 | 30.1 | 36.4 | 0 | 22.1 | 2.4 | 71.8 | 0 |
| Nilotinib | 51.8 | 31.3 | 1.9 | 68.7 | 16.6 | 71.7 | 62.5 | 1.5 | 20.4 | 6 |
| Nutlin-3 | 0 | 13.8 | 46.4 | 11.1 | 10.6 | 60.4 | 76.5 | 49.5 | 27.1 | 13.5 |
| PD-0325901 | 2.8 | 0.3 | 0 | 1 | 12.1 | 0.7 | 96.2 | 1.6 | 0.9 | 0 |
| PD-0332991 | 56.3 | 9.6 | 1.4 | 83 | 0.2 | 62.6 | 34.4 | 18.3 | 23.9 | 4.9 |
| PF2341066 | 21.2 | 26.2 | 2.2 | 71.2 | 17.8 | 98.8 | 25.4 | 12.2 | 41.7 | 0.6 |
| PHA-665752 | 20.5 | 7.4 | 5.8 | 65.9 | 12.2 | 13.6 | 23.8 | 62.8 | 71.4 | 5.8 |
| PLX4720 | 67.2 | 43.6 | 77.8 | 47.2 | 84.3 | 36.9 | 95.2 | 97.5 | 63.7 | 21.2 |
| Paclitaxel | 1.3 | 0 | 26.7 | 0.2 | 1.5 | 0.2 | 19.8 | 4.2 | 13.2 | 0 |
| Panobinostat | 49.9 | 0 | 2.4 | 8.4 | 2.1 | 0.5 | 27.8 | 3.4 | 23.9 | 0 |
| RAF265 | 10.5 | 2.6 | 0.6 | 30.7 | 39.6 | 16.9 | 37 | 13.3 | 45.5 | 0.2 |
| Sorafenib | 74.9 | 48.1 | 29.5 | 10.5 | 90.2 | 34.6 | 38.3 | 93.4 | 8.2 | 0.5 |
| TAE684 | 58.9 | 0.2 | 0 | 47.6 | 5.9 | 98.8 | 56.5 | 4.4 | 90.4 | 0.6 |
| TKI258 | 39.6 | 5.5 | 64.1 | 38.2 | 45.8 | 14.8 | 16.7 | 1.3 | 7.7 | 0.2 |
| Topotecan | 5 | 0 | 12.7 | 1.9 | 5.2 | 0.4 | 41.3 | 44.3 | 4.7 | 0 |
| ZD-6474 | 14.8 | 3.2 | 2 | 39.9 | 9.6 | 1.2 | 90.3 | 4.4 | 96.7 | 0 |
| Placebo | 17.3 | 79.9 | 57.1 | 40.5 | 19.6 | 22.2 | 99.1 | 46.8 | 79.3 | 56.1 |
| Placebo | 16 | 51.9 | 61.2 | 76.1 | 83.3 | 39.8 | 24.4 | 66.2 | 64.5 | 77.4 |

TABLE 8
*Largest Pearson correlations, appearing in red if it is significant at 1% using a Westfall-Young permutation method.*

| | cent: | othe: | haem: | skin: | ovar: | lung: | panc: | brea: | larg: | all: |
|---|---|---|---|---|---|---|---|---|---|---|
| 17-AAG | 0.63 | 0.446 | 0.51 | 0.58 | 0.74 | 0.47 | 0.64 | 0.6 | 0.74 | 0.375 |
| AEW541 | 0.67 | 0.442 | 0.61 | 0.53 | 0.78 | 0.39 | 0.71 | 0.63 | 0.7 | 0.252 |
| AZD0530 | 0.67 | 0.33 | 0.51 | 0.72 | 0.74 | 0.52 | 0.66 | 0.73 | 0.7 | 0.23 |
| AZD6244 | 0.86 | 0.39 | 0.67 | 0.64 | 0.78 | 0.5 | 0.64 | 0.86 | 0.79 | 0.343 |
| Erlotinib | 0.75 | 0.512 | 0.48 | 0.63 | 0.75 | 0.59 | 0.73 | 0.85 | 0.8 | 0.338 |
| Irinotecan | 0.79 | 0.574 | 0.58 | 0.7 | 0.82 | 0.62 | 0.87 | 0.83 | 0.9 | 0.46 |
| L-685458 | 0.73 | 0.45 | 0.48 | 0.62 | 0.72 | 0.41 | 0.78 | 0.89 | 0.81 | 0.25 |
| LBW242 | 0.93 | 0.54 | 0.56 | 0.65 | 0.71 | 0.58 | 0.81 | 0.69 | 0.77 | 0.29 |
| Lapatinib | 0.6 | 0.47 | 0.56 | 0.69 | 0.69 | 0.634 | 0.73 | 0.77 | 0.73 | 0.373 |
| Nilotinib | 0.71 | 0.45 | 0.65 | 0.65 | 0.79 | 0.52 | 0.77 | 0.93 | 0.83 | 0.24 |
| Nutlin-3 | 0.859 | 0.36 | 0.49 | 0.68 | 0.81 | 0.48 | 0.69 | 0.73 | 0.77 | 0.2 |
| PD-0325901 | 0.81 | 0.42 | 0.663 | 0.68 | 0.77 | 0.5 | 0.6 | 0.8 | 0.85 | 0.365 |
| PD-0332991 | 0.71 | 0.43 | 0.54 | 0.7 | 0.86 | 0.49 | 0.79 | 0.87 | 0.8 | 0.23 |
| PF2341066 | 0.73 | 0.42 | 0.55 | 0.58 | 0.74 | 0.45 | 0.72 | 0.71 | 0.74 | 0.26 |
| PHA-665752 | 0.8 | 0.39 | 0.52 | 0.58 | 0.8 | 0.56 | 0.74 | 0.69 | 0.77 | 0.22 |
| PLX4720 | 0.74 | 0.36 | 0.64 | 0.58 | 0.68 | 0.42 | 0.64 | 0.59 | 0.8 | 0.2 |
| Paclitaxel | 0.77 | 0.45 | 0.47 | 0.71 | 0.77 | 0.49 | 0.7 | 0.73 | 0.78 | 0.337 |
| Panobinostat | 0.65 | 0.472 | 0.52 | 0.63 | 0.76 | 0.49 | 0.72 | 0.75 | 0.76 | 0.287 |
| RAF265 | 0.78 | 0.4 | 0.57 | 0.6 | 0.71 | 0.47 | 0.73 | 0.71 | 0.74 | 0.29 |
| Sorafenib | 0.66 | 0.35 | 0.61 | 0.64 | 0.62 | 0.44 | 0.67 | 0.59 | 0.79 | 0.25 |
| TAE684 | 0.64 | 0.44 | 0.626 | 0.62 | 0.77 | 0.35 | 0.67 | 0.75 | 0.69 | 0.23 |
| TKI258 | 0.67 | 0.38 | 0.52 | 0.6 | 0.71 | 0.46 | 0.7 | 0.78 | 0.81 | 0.24 |
| Topotecan | 0.73 | 0.457 | 0.48 | 0.66 | 0.75 | 0.5 | 0.67 | 0.65 | 0.8 | 0.426 |
| ZD-6474 | 0.77 | 0.38 | 0.53 | 0.59 | 0.77 | 0.55 | 0.63 | 0.75 | 0.67 | 0.278 |
| Placebo | 0.78 | 0.35 | 0.47 | 0.6 | 0.76 | 0.48 | 0.63 | 0.71 | 0.73 | 0.18 |
| Placebo | 0.7 | 0.31 | 0.42 | 0.56 | 0.65 | 0.4 | 0.69 | 0.64 | 0.74 | 0.17 |

TABLE 9
*Westfall-Young corrected p-values of largest Spearman correlation. Red values signify 1% significance.*

| | cent: | othe: | haem: | skin: | ovar: | lung: | panc: | brea: | larg: | all: |
|---|---|---|---|---|---|---|---|---|---|---|
| 17-AAG | 21.1 | 0 | 13 | 48.8 | 3 | 5.8 | 77.9 | 69.1 | 47.1 | 0 |
| AEW541 | 36.1 | 13.5 | 0 | 50.3 | 3 | 41.9 | 20.9 | 32.2 | 71.9 | 0 |
| AZD0530 | 9.2 | 8.4 | 17.4 | 3.7 | 1.8 | 1 | 36.7 | 51.3 | 33.4 | 1.4 |
| AZD6244 | 2.8 | 4.5 | 0 | 4.7 | 90 | 6.5 | 70.2 | 5.4 | 5.6 | 0 |
| Erlotinib | 88.4 | 0 | 10.2 | 82.2 | 11.3 | 0 | 43.2 | 3.4 | 43.8 | 0 |
| Irinotecan | 38.9 | 0 | 5.4 | 4.9 | 1.1 | 6.9 | 1 | 87.2 | 39.7 | 0 |
| L-685458 | 83.6 | 0 | 3.5 | 13.2 | 43 | 17.7 | 27.8 | 59.2 | 36 | 0 |
| LBW242 | 9.7 | 17.7 | 32.8 | 26.6 | 80.4 | 1.4 | 30 | 96.2 | 22.8 | 1.9 |
| Lapatinib | 96.6 | 0 | 3.9 | 48 | 42.5 | 0 | 0 | 2.7 | 85.2 | 0 |
| Nilotinib | 64.7 | 1.8 | 0.5 | 95.2 | 25.6 | 30.7 | 57 | 72.7 | 46.3 | 0.8 |
| Nutlin-3 | 1.6 | 8.3 | 6.2 | 32.4 | 34.8 | 6.2 | 95.4 | 29.1 | 36.9 | 3 |
| PD-0325901 | 18.8 | 2.1 | 0 | 1.3 | 19.4 | 0.7 | 90.6 | 4.3 | 4.1 | 0 |
| PD-0332991 | 3.9 | 3.1 | 2.1 | 8.8 | 27.4 | 49.3 | 36.4 | 83.9 | 90.9 | 1.2 |
| PF2341066 | 69.2 | 0 | 5.3 | 64.2 | 52.1 | 22.7 | 21.4 | 7.7 | 31.4 | 0 |
| PHA-665752 | 76.4 | 76.2 | 3.3 | 46.9 | 92.8 | 12.1 | 82.8 | 48 | 65.3 | 17.7 |
| PLX4720 | 65.7 | 76.6 | 11.8 | 53.2 | 96.1 | 26.5 | 96.9 | 67.8 | 75.5 | 53.3 |
| Paclitaxel | 0.7 | 0.2 | 63.1 | 0.5 | 3.9 | 0 | 31.8 | 7.7 | 8 | 0 |
| Panobinostat | 25.2 | 0 | 9.9 | 4.4 | 0 | 1.3 | 23.9 | 1.4 | 25.6 | 0 |
| RAF265 | 32.3 | 2.2 | 0.7 | 10.2 | 5.5 | 20 | 12.7 | 28.3 | 20.9 | 0 |
| Sorafenib | 37.9 | 37.2 | 1.5 | 0.7 | 33.3 | 35.5 | 45.8 | 30.1 | 27.2 | 0 |
| TAE684 | 74.5 | 0.4 | 0 | 28.1 | 7.5 | 61.7 | 21.7 | 27.8 | 81.2 | 0 |
| TKI258 | 19.2 | 1 | 2.5 | 84.7 | 19.9 | 2.7 | 45 | 7.3 | 22.6 | 0 |
| Topotecan | 20.4 | 0 | 11.9 | 5.7 | 9.1 | 1.2 | 17.1 | 30.8 | 17.4 | 0 |
| ZD-6474 | 11.3 | 5.8 | 1.8 | 85.8 | 69.2 | 1.1 | 23.8 | 39.5 | 72.1 | 0 |
| Placebo | 85.4 | 74.9 | 30.1 | 21.5 | 8 | 77.5 | 61.5 | 30.7 | 98.6 | 74.3 |
| Placebo | 76.1 | 76.9 | 50.4 | 67 | 22 | 61.5 | 30.2 | 37.9 | 42.6 | 84.2 |

TABLE 10
*Largest Spearman correlations, appearing in red if it is significant at 1% using a Westfall-Young permutation method.*

| | cent: | othe: | haem: | skin: | ovar: | lung: | panc: | brea: | larg: | all: |
|---|---|---|---|---|---|---|---|---|---|---|
| 17-AAG | 0.71 | 0.442 | 0.49 | 0.58 | 0.79 | 0.46 | 0.65 | 0.65 | 0.77 | 0.359 |
| AEW541 | 0.69 | 0.34 | 0.623 | 0.58 | 0.78 | 0.41 | 0.72 | 0.68 | 0.72 | 0.251 |
| AZD0530 | 0.73 | 0.35 | 0.48 | 0.68 | 0.8 | 0.49 | 0.7 | 0.67 | 0.77 | 0.22 |
| AZD6244 | 0.76 | 0.36 | 0.65 | 0.66 | 0.63 | 0.46 | 0.65 | 0.75 | 0.82 | 0.304 |
| Erlotinib | 0.62 | 0.468 | 0.5 | 0.55 | 0.76 | 0.589 | 0.69 | 0.75 | 0.75 | 0.321 |
| Irinotecan | 0.78 | 0.543 | 0.59 | 0.73 | 0.86 | 0.62 | 0.91 | 0.8 | 0.89 | 0.452 |
| L-685458 | 0.63 | 0.486 | 0.53 | 0.64 | 0.7 | 0.43 | 0.71 | 0.65 | 0.78 | 0.246 |
| LBW242 | 0.72 | 0.33 | 0.46 | 0.61 | 0.65 | 0.49 | 0.7 | 0.6 | 0.78 | 0.21 |
| Lapatinib | 0.6 | 0.549 | 0.51 | 0.58 | 0.69 | 0.567 | 0.837 | 0.77 | 0.71 | 0.386 |
| Nilotinib | 0.71 | 0.45 | 0.58 | 0.57 | 0.73 | 0.46 | 0.75 | 0.65 | 0.8 | 0.24 |
| Nutlin-3 | 0.78 | 0.35 | 0.51 | 0.6 | 0.71 | 0.46 | 0.61 | 0.69 | 0.77 | 0.21 |
| PD-0325901 | 0.7 | 0.37 | 0.653 | 0.7 | 0.73 | 0.51 | 0.62 | 0.74 | 0.83 | 0.33 |
| PD-0332991 | 0.81 | 0.41 | 0.53 | 0.68 | 0.71 | 0.44 | 0.76 | 0.64 | 0.73 | 0.24 |
| PF2341066 | 0.65 | 0.426 | 0.52 | 0.57 | 0.69 | 0.42 | 0.72 | 0.74 | 0.77 | 0.238 |
| PHA-665752 | 0.64 | 0.29 | 0.52 | 0.59 | 0.63 | 0.44 | 0.64 | 0.67 | 0.72 | 0.19 |
| PLX4720 | 0.66 | 0.29 | 0.49 | 0.58 | 0.66 | 0.42 | 0.62 | 0.65 | 0.71 | 0.18 |
| Paclitaxel | 0.81 | 0.42 | 0.43 | 0.72 | 0.78 | 0.531 | 0.71 | 0.74 | 0.82 | 0.344 |
| Panobinostat | 0.7 | 0.464 | 0.5 | 0.66 | 0.841 | 0.49 | 0.7 | 0.79 | 0.77 | 0.31 |
| RAF265 | 0.73 | 0.4 | 0.58 | 0.65 | 0.76 | 0.45 | 0.79 | 0.71 | 0.79 | 0.296 |
| Sorafenib | 0.68 | 0.32 | 0.54 | 0.7 | 0.71 | 0.41 | 0.69 | 0.7 | 0.78 | 0.253 |
| TAE684 | 0.65 | 0.39 | 0.606 | 0.6 | 0.75 | 0.39 | 0.71 | 0.7 | 0.71 | 0.251 |
| TKI258 | 0.71 | 0.39 | 0.53 | 0.54 | 0.72 | 0.47 | 0.69 | 0.74 | 0.77 | 0.268 |
| Topotecan | 0.71 | 0.435 | 0.49 | 0.65 | 0.75 | 0.49 | 0.72 | 0.69 | 0.79 | 0.418 |
| ZD-6474 | 0.74 | 0.36 | 0.54 | 0.54 | 0.71 | 0.49 | 0.73 | 0.68 | 0.72 | 0.259 |
| Placebo | 0.63 | 0.29 | 0.46 | 0.62 | 0.75 | 0.37 | 0.66 | 0.68 | 0.67 | 0.17 |
| Placebo | 0.64 | 0.29 | 0.44 | 0.56 | 0.72 | 0.39 | 0.71 | 0.69 | 0.75 | 0.16 |

TABLE 11
*p-value of type independence (see $H_{0,3}$) for drug AZD0530*

| | hema.& ly. | skin |
|---|---|---|
| hema.& ly. | - | 0.035 |
| skin | - | - |

TABLE 12
*p-value of type independence (see $H_{0,3}$) for drug AZD6244*

| | hema.& ly. | skin |
|---|---|---|
| hema.& ly. | - | 0.037 |
| skin | - | - |

TABLE 13
*p-value of type independence (see $H_{0,3}$) for drug Irinotecan*

| | hema.& ly. | pancreas |
|---|---|---|
| hema.& ly. | - | 0.038 |
| pancreas | - | - |

TABLE 14
*p-value of type independence (see $H_{0,3}$) for drug Lapatinib*

| | lung | breast |
|---|---|---|
| lung | - | 0.036 |
| breast | - | - |

TABLE 15
*p-value of type independence (see $H_{0,3}$) for drug Panobinostat*

| | hema.& ly. | ovary | breast |
|---|---|---|---|
| hema.& ly. | - | 0.035 | 0.781 |
| ovary | - | - | 0.025 |
| breast | - | - | - |

TABLE 16
*p-value of type independence (see $H_{0,3}$) for drug PD-0325901*

| | hema.& ly. | skin | lung |
|---|---|---|---|
| hema.& ly. | - | 0.046 | 0.793 |
| skin | - | - | 0.027 |
| lung | - | - | - |

TABLE 17
*Using a model fitted with (row) to predict the response in type (column) for drug Lapatinib ActArea.*

| | central n.s. | other | hema.& ly. | skin | ovary | lung | pancreas | breast | l. intestine |
|---|---|---|---|---|---|---|---|---|---|
| central n.s. | 1.928 | 1.551 | 1.772 | 1.133 | 2.091 | 1.496 | 1.657 | 1.782 | 1.678 |
| other | 3.292 | 0.924 | 1.104 | 1.145 | 1.063 | 0.871 | 1.452 | 0.941 | 1.070 |
| hema.& ly. | 22.062 | 1.381 | 0.929 | 1.816 | 1.159 | 1.175 | 0.966 | 1.421 | 1.124 |
| skin | 1.608 | 1.238 | 1.295 | 1.167 | 1.277 | 1.206 | 1.164 | 1.511 | 1.325 |
| ovary | 3.553 | 1.035 | 1.029 | 1.135 | 1.065 | 1.030 | 1.004 | 1.293 | 1.118 |
| lung | 2.847 | 0.840 | 1.084 | 1.030 | 1.118 | 0.861 | 1.370 | 0.608 | 0.947 |
| pancreas | 3.179 | 1.053 | 1.051 | 1.097 | 1.006 | 1.045 | 1.806 | 1.322 | 1.144 |
| breast | 5.352 | 1.064 | 3.671 | 1.858 | 3.101 | 1.094 | 4.454 | 1.039 | 1.287 |
| l. intestine | 8.497 | 1.043 | 1.111 | 1.790 | 1.555 | 1.043 | 1.480 | 1.088 | 1.297 |

TABLE 18
*Using a model fitted with (row) to predict the response in type (column) for drug Topotecan ActArea.*

| | central n.s. | other | hema.& ly. | skin | ovary | lung | pancreas | breast | l. intestine |
|---|---|---|---|---|---|---|---|---|---|
| central n.s. | 1.017 | 1.038 | 2.658 | 1.058 | 1.034 | 0.949 | 1.064 | 1.289 | 1.601 |
| other | 0.733 | 0.764 | 1.518 | 1.153 | 0.927 | 0.917 | 1.012 | 1.090 | 1.195 |
| hema.& ly. | 2.631 | 1.717 | 1.058 | 3.952 | 1.535 | 2.124 | 3.846 | 2.518 | 1.521 |
| skin | 1.092 | 2.050 | 7.185 | 1.092 | 1.950 | 1.238 | 1.778 | 1.274 | 8.130 |
| ovary | 0.827 | 0.939 | 1.712 | 1.063 | 0.988 | 1.042 | 0.874 | 1.305 | 1.073 |
| lung | 0.848 | 1.030 | 2.466 | 0.843 | 0.994 | 1.064 | 0.822 | 0.929 | 0.942 |
| pancreas | 1.013 | 1.150 | 3.657 | 1.005 | 1.100 | 1.000 | 1.034 | 1.039 | 1.055 |
| breast | 1.102 | 1.355 | 4.466 | 1.036 | 1.344 | 1.044 | 1.067 | 1.090 | 1.228 |
| l. intestine | 0.993 | 1.234 | 1.838 | 1.133 | 1.119 | 0.987 | 1.367 | 1.152 | 0.998 |

## ACKNOWLEDGEMENTS

I would like to thank to my supervisor Sach Mukherjee and my colleagues at the NKI-AVL in Amsterdam, Frank Dondelinger, Steven Hill, Anas Rana and Nicholas Städler.

## REFERENCES

[1] BENJAMINI, Y. and HOCHBERG, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57** 289–300.

[2] FRIEDMAN, J. H., HASTIE, T. and TIBSHIRANI, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software* **33** 1–22.

[3] HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. *Springer Series in Statistics*. Springer Verlag.

[4] MEINSHAUSEN, N., MAATHUIS, M. H. and BÜHLMANN, P. (2011). Asymptotic optimality of the Westfall-Young permutation procedure for multiple testing under dependence. *Ann. Stat.* **39** 3369–3391.

[5] MEINSHAUSEN, N. and RICE, J. (2006). Estimating the proportion of false null hypotheses among a large number of independently tested hypotheses. *ANN. STAT* **34** 373–393.

[6] MENDEN, M. P., IORIO, F., GARNETT, M., MCDERMOTT, U., BENES, C. H., BALLESTER, P. J. and SAEZ-RODRIGUEZ, J. (2013). Machine Learning Prediction of Cancer Cell Sensitivity to Drugs Based on Genomic and Chemical Properties. *PLoS ONE* **8** e61318.

TABLE 19
*Using a model fitted with (row) to predict the response in type (column) for drug AZD6244 ActArea.*

| | central n.s. | other | hema.& ly. | skin | ovary | lung | pancreas | breast | l. intestine |
|---|---|---|---|---|---|---|---|---|---|
| central n.s. | 1.254 | 1.195 | 1.328 | 2.981 | 1.060 | 1.151 | 3.201 | 1.006 | 3.247 |
| other | 0.992 | 1.006 | 0.882 | 1.825 | 0.705 | 0.994 | 1.514 | 1.903 | 1.650 |
| hema.& ly. | 10.266 | 3.411 | 0.737 | 1.346 | 4.779 | 5.089 | 2.685 | 10.076 | 1.188 |
| skin | 6.344 | 2.680 | 1.382 | 1.218 | 4.367 | 3.585 | 1.405 | 8.325 | 1.017 |
| ovary | 1.073 | 1.087 | 1.295 | 2.710 | 1.014 | 1.049 | 2.684 | 1.112 | 2.861 |
| lung | 1.207 | 1.020 | 1.114 | 2.887 | 1.112 | 1.122 | 1.893 | 1.020 | 2.107 |
| pancreas | 3.632 | 1.609 | 1.070 | 1.188 | 2.491 | 2.029 | 1.217 | 4.739 | 1.288 |
| breast | 1.227 | 1.184 | 1.334 | 3.847 | 1.530 | 0.991 | 2.119 | 0.965 | 2.103 |
| l. intestine | 7.429 | 3.142 | 1.538 | 1.020 | 5.139 | 4.242 | 1.711 | 9.747 | 0.974 |

[7] TIBSHIRANI, R., SAUNDERS, M., ROSSET, S., ZHU, J. and KNIGHT, K. (2005). Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society Series B* 91–108.