
The “See”–Value App: Visual Decision Making for Drug Development

H. Diehl¹, A. M. Stein², N. Roy Chowdhury², T. Broderick¹

¹ MIT, ² Novartis Pharmaceuticals

Scenario: Genomics of Drug Sensitivity

You are looking for top associations between genetic features and drug response using the GDSCTools package (Cokelaer et al. 2017)

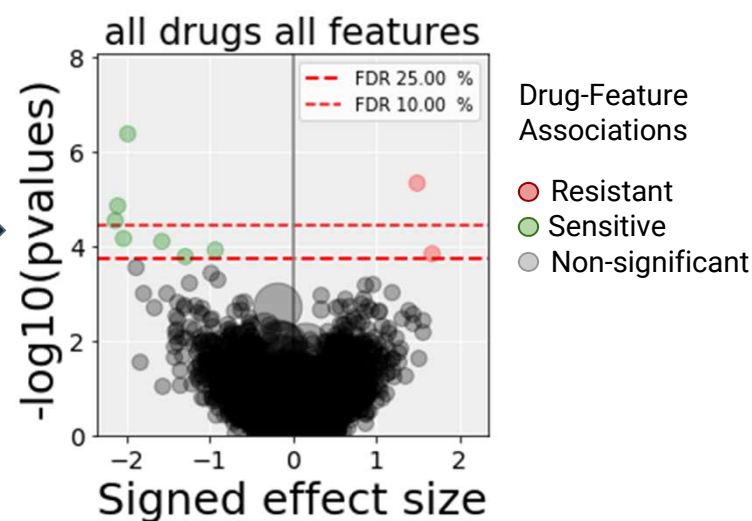
Genetic Features

Cell Line	Tissue	Mut 1	...	Mut 47
1	Lung	1	..	0
.....
998	Renal	0	...	1

Drug Response

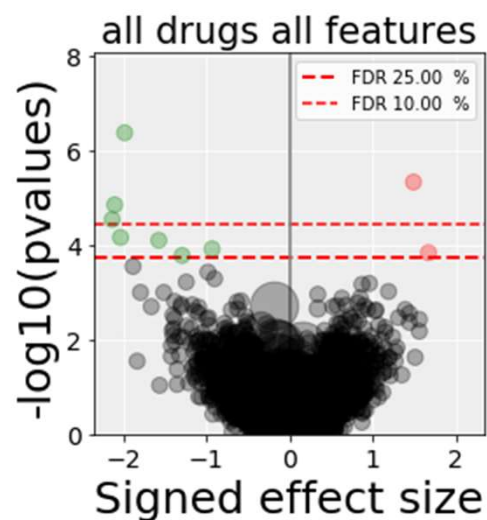
Cell Line	Drug 1 IC50	...	Drug 256 IC50
1	1.107	..	-5.950
.....
998	2.755	...	-4.603

ANOVA

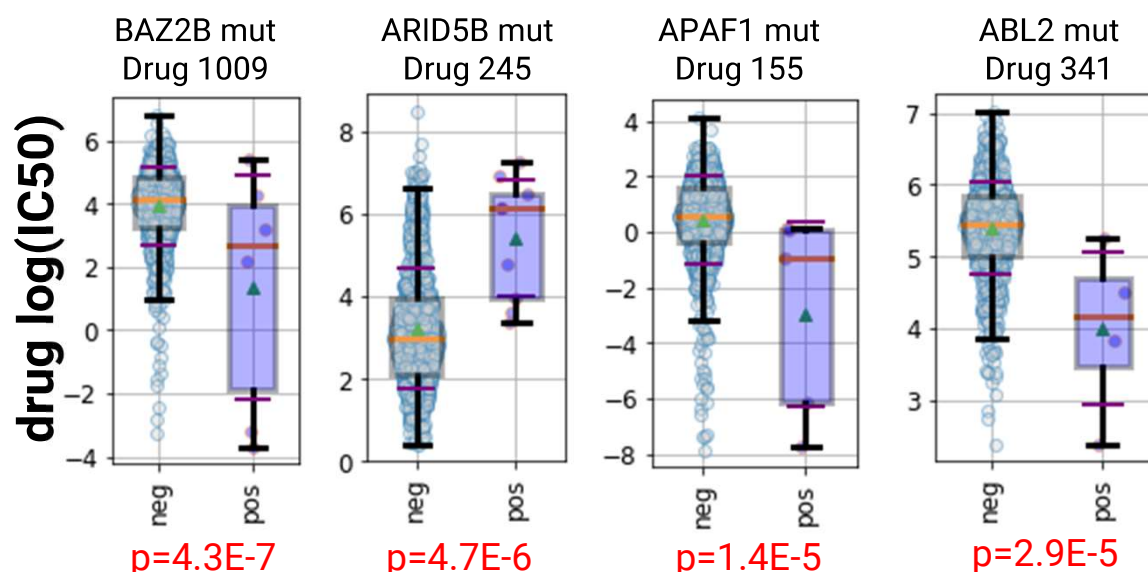


Scenario: Genomics of Drug Sensitivity

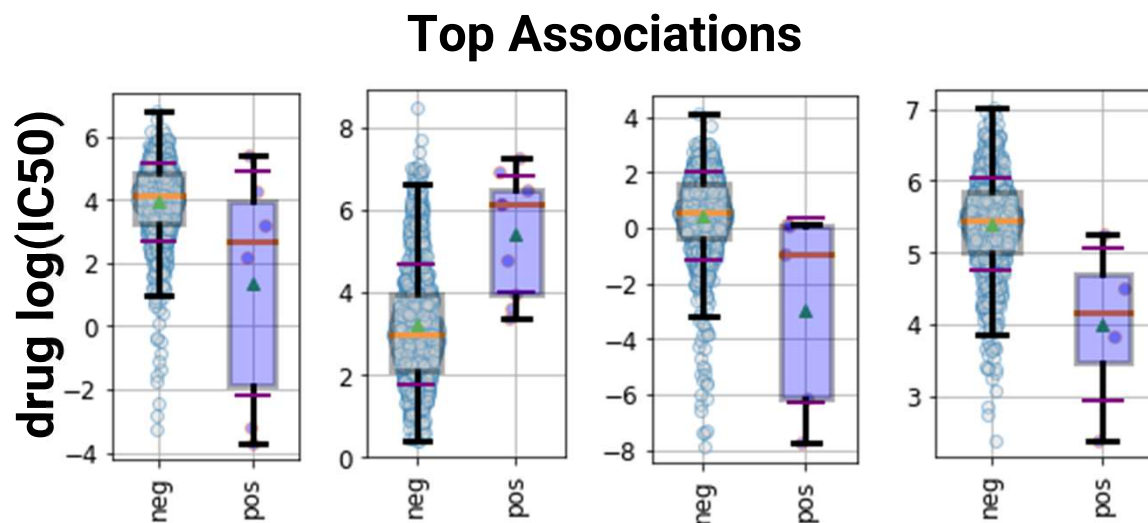
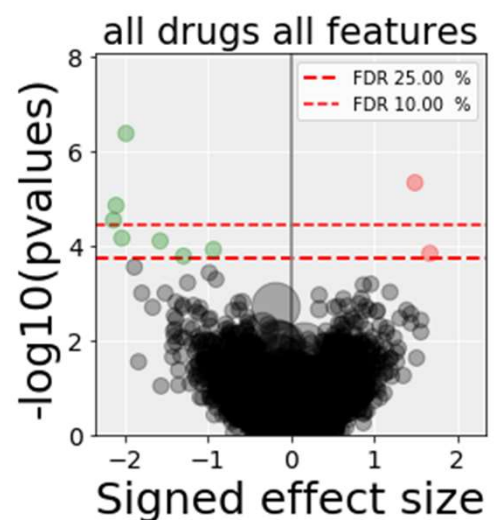
You are looking for top associations between genetic features and drug response using the GDSCTools package (Cokelaer et al. 2017)



Top Associations



Scenario: Genomics of Drug Sensitivity



Further trials would reveal no real effects - in fact this is synthetic data with drug response synthetically generated independent of genetic features.

What went wrong?

Problem

Signals seen in exploratory data analysis inform trial design.



But humans are great at seeing signals even when they aren't there.

Solution

A visual check for significance that is...

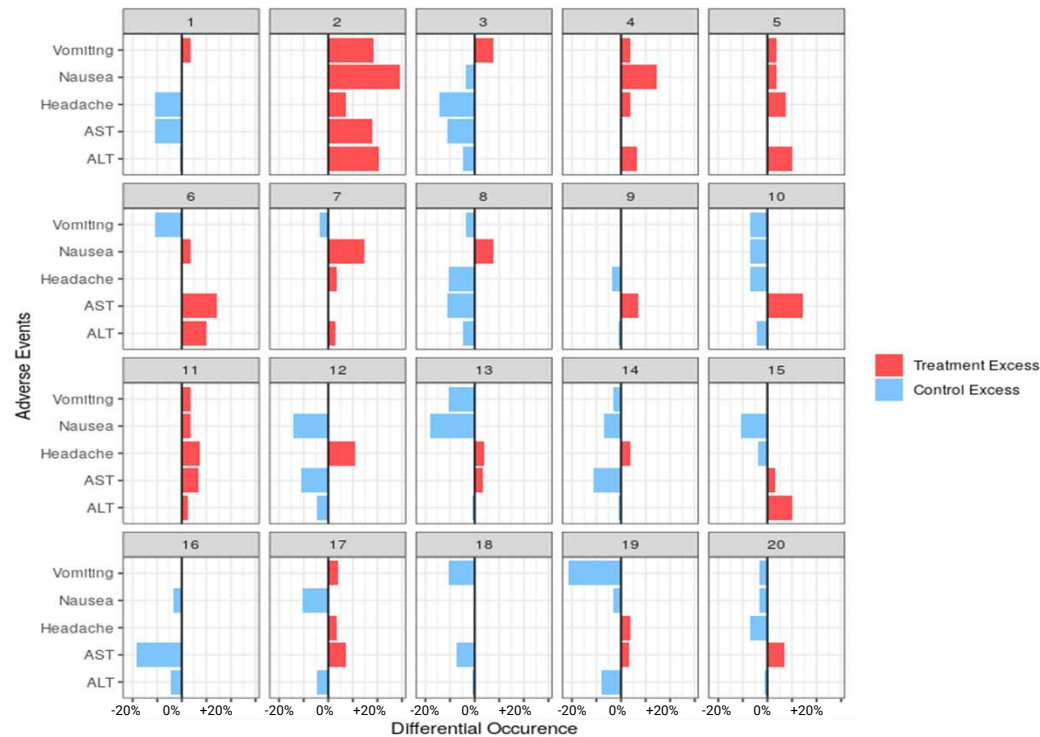
- **Fast, easy, and intuitive**
- **Convincing even when users are incentivized to see a signal**
- **Easy to integrate into current workflows**

What to Expect

- ☑ **Motivate** the need for more visual inference tools
- ☐ **Introduce** the Visual Lineup Test
- ☐ **Re-visit** the 'Genomics of Drug Sensitivity' Example
- ☐ **Present** the See-Value App
- ☐ **Illustrate** Usage of the App

A Lineup Protocol

Protocol described by Buja et al (2009), Majumder et al (2015), Roy Chowdhury et al (2015), and others.



Which plot looks most different from the others?

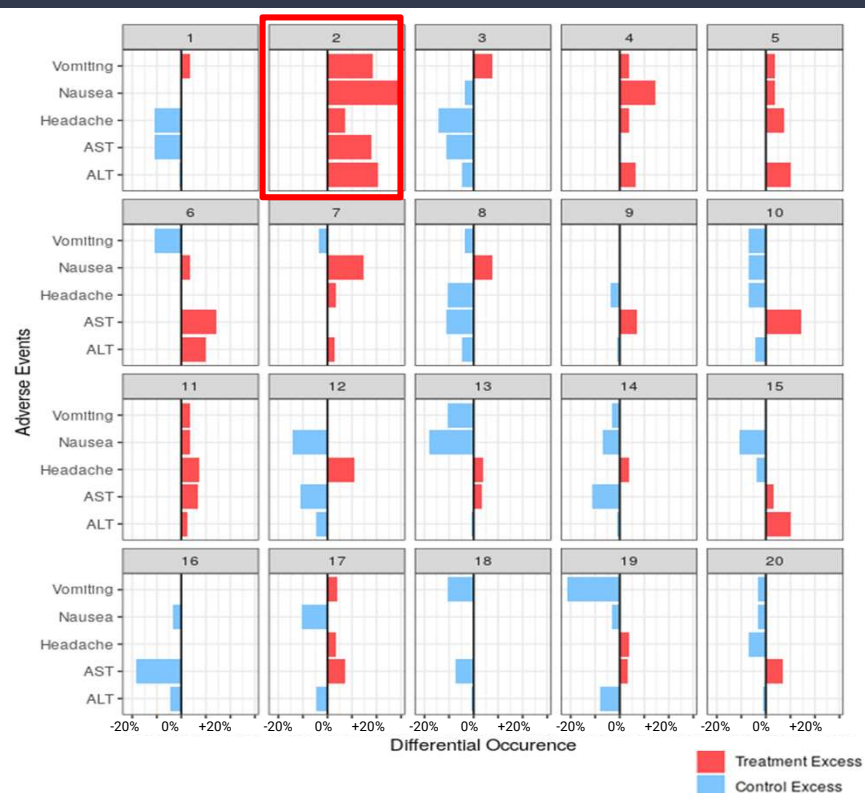
A Lineup Protocol

Null Hypothesis:

Control and treatment have same adverse event occurrence rate.

Permutation:

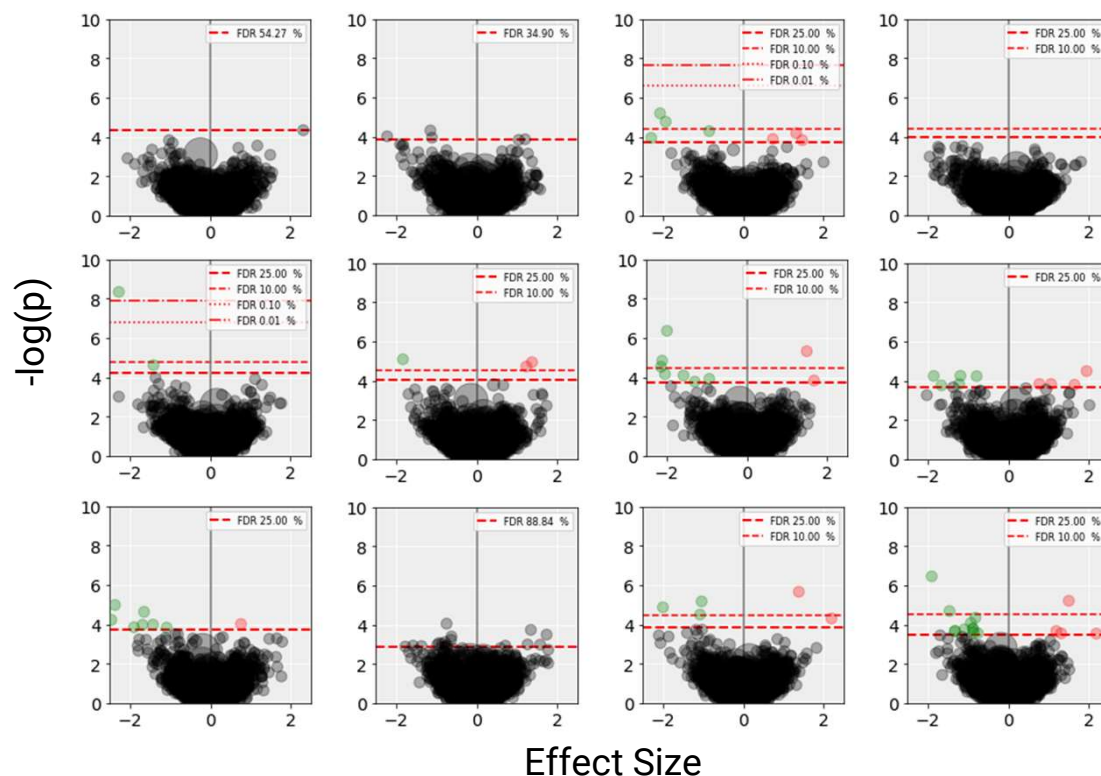
If group assignment is independent of adverse event risk, group label can be freely permuted



Lineup Test:

1. Plot the real data along with the scrambled data.
2. Pick the “most unusual” plot.
3. If you can pick out your real data, your result may be significant.

Genetics of Drug Sensitivity – Lineup Method



Permutation:

Scramble the cell line ID in the drug response table.

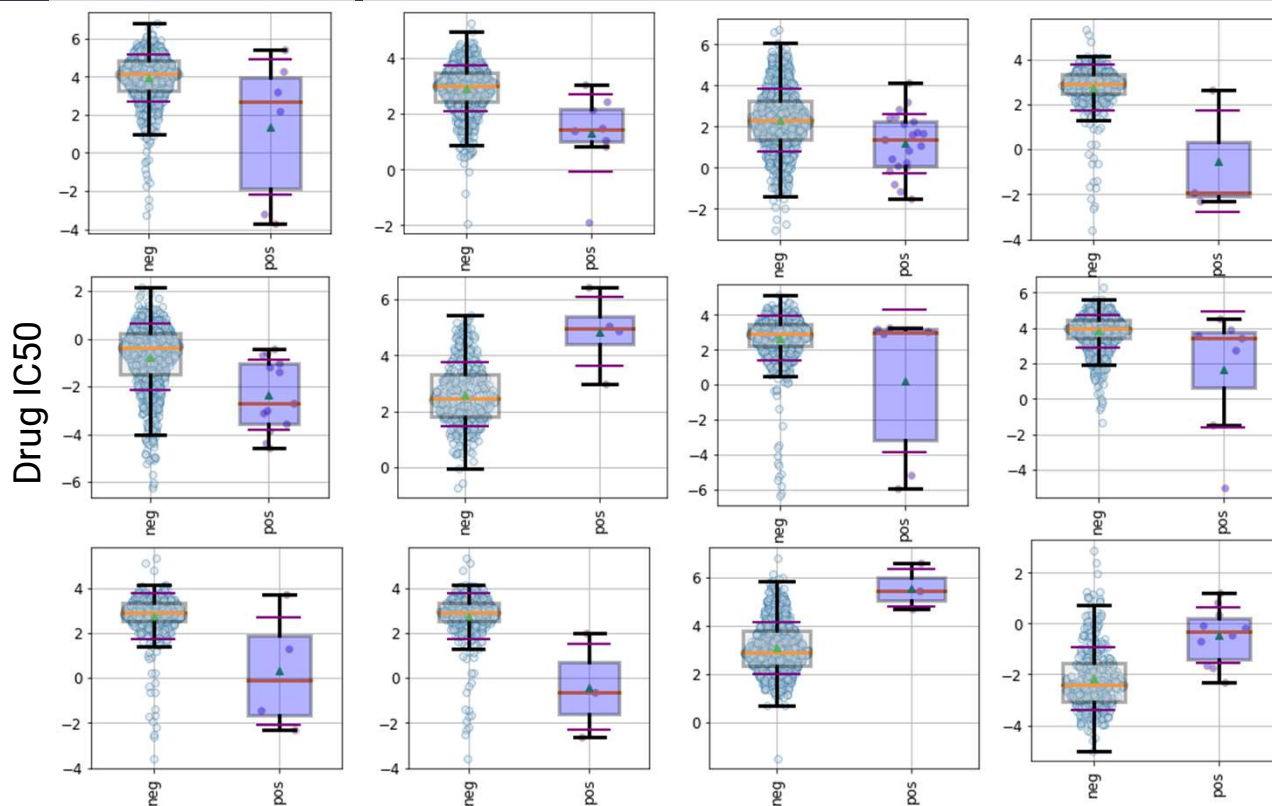
Observation:

No facet that is compellingly distinct from the rest.

Takeaway:

The synthetic data's volcano plot that originally looked promising doesn't look that unusual for null data.

Genetics of Drug Sensitivity – Lineup Method



Permutation:

Scramble the cell line ID in the drug response table.

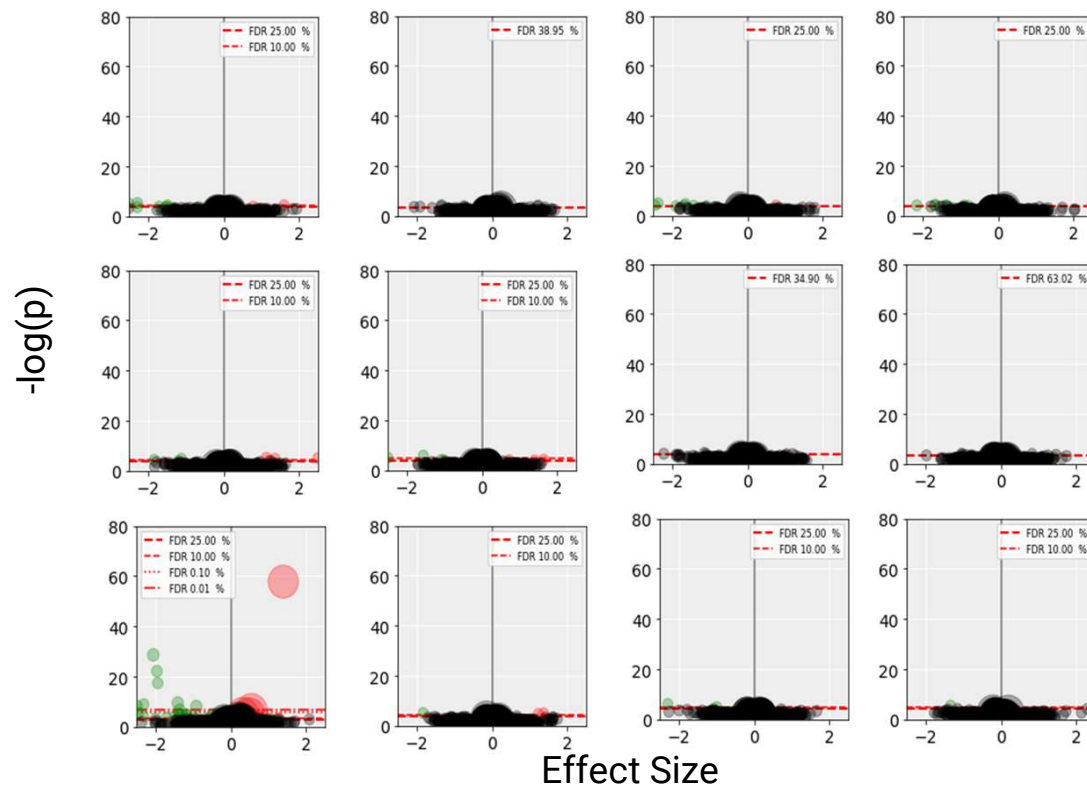
Observation:

Again, no single facet stands out as obviously distinct.

Takeaway:

The box plot illustrating the 'most significant' association we saw on our synthetic data was fairly typical for null generated data.

Genetics of Drug Sensitivity – Lineup Method



Permutation:

Scramble the cell line ID in the drug response table.

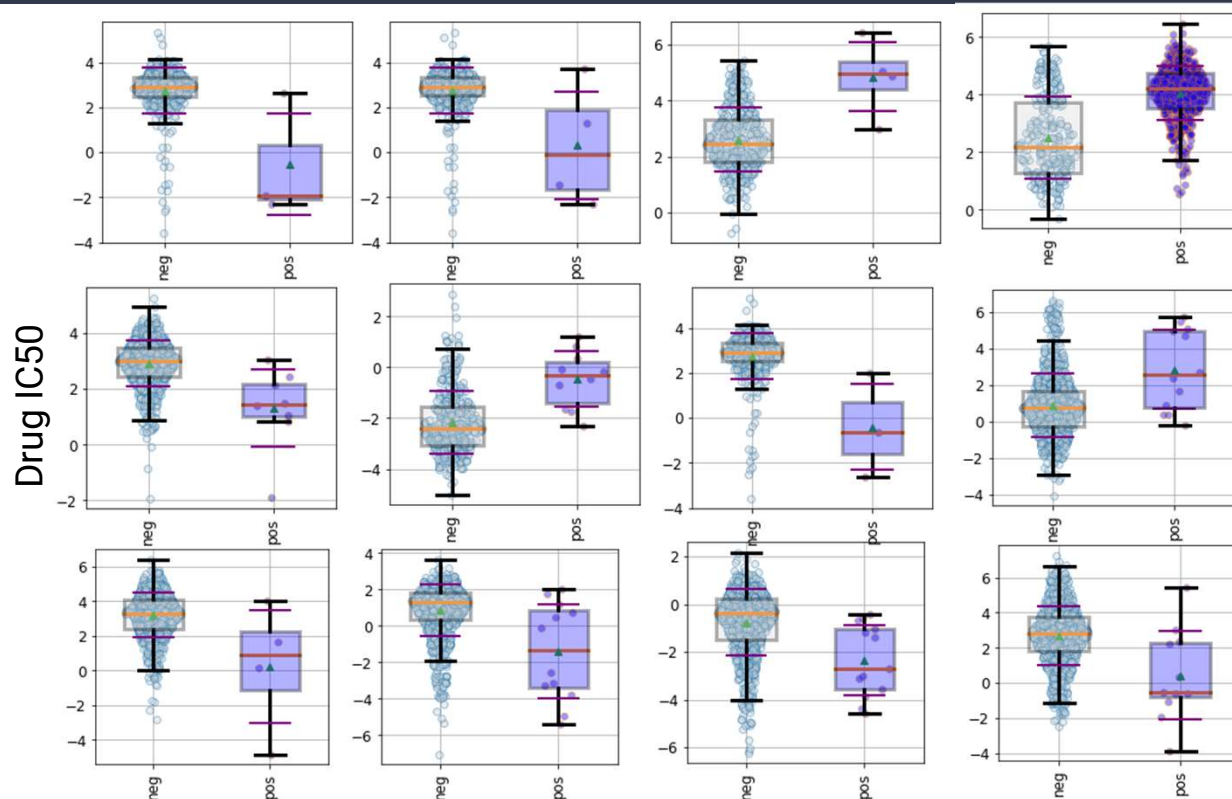
Observation:

Facet 9 notably is the only facet with numerous associations above 15 on the y-axis.

Takeaway:

A volcano plot containing published data with realistic associations does stand out against null plots suggesting real, meaningful associations.

Genetics of Drug Sensitivity – Lineup Method



Null Hypothesis:
Drug response is independent of all genetic features.

Observation:
Facet 4 stands out as it has a much more balanced number of cell lines in the + and - groups.

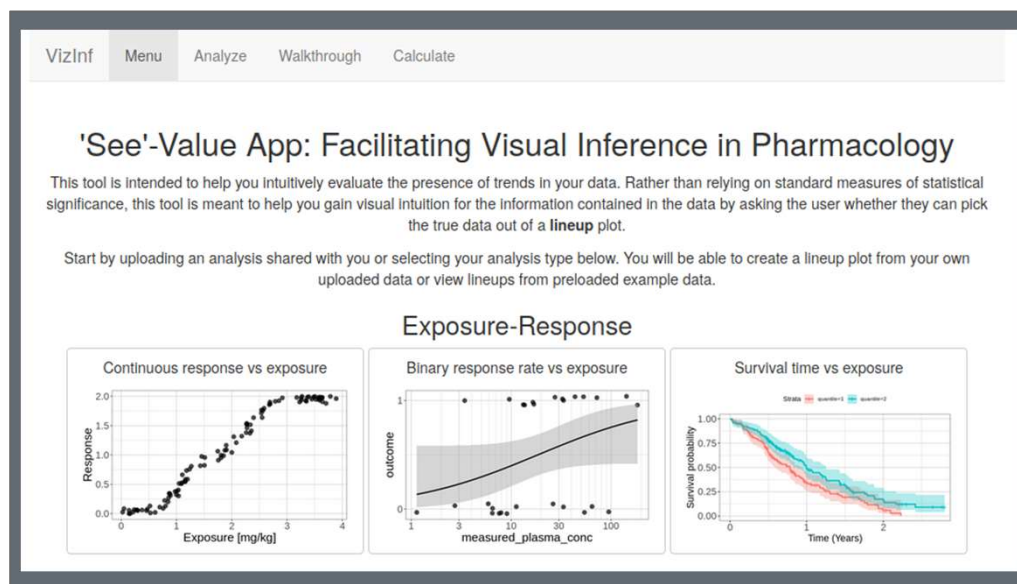
Takeaway:
It is likely many people would see facet 4 as notably distinct. This is the true data, suggesting there is a signal.

Contribution: “See”-Value App

www.see-value.org

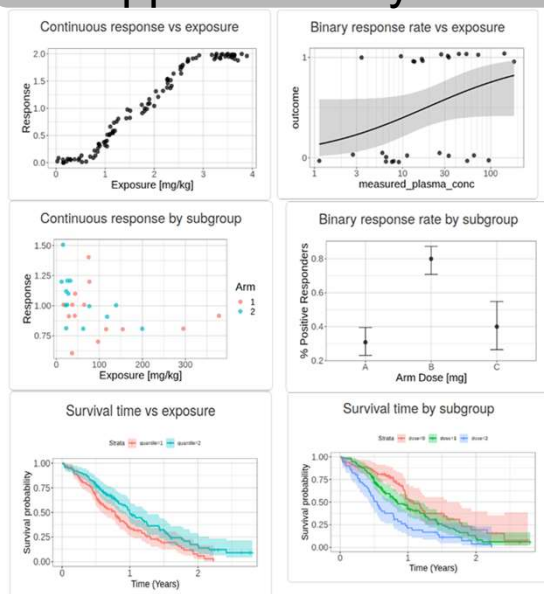
We provide a Shiny app that..

- Facilitates visual permutation tests as described by Buja et al (2009), Majumder et al (2015), Roy Chowdhury et al (2015).
- Supports common inference tasks used in drug development.
- Enables team-based voting and significance calculations.



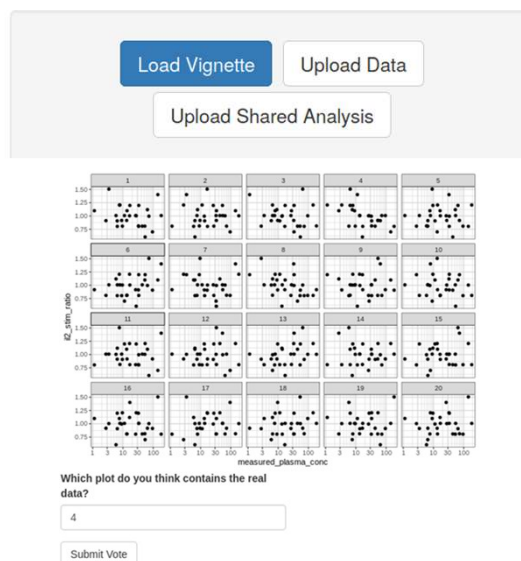
App workflow: Team Based Voting

Step 1: Select a supported analysis.



With options growing

Step 2: Explore pre-loaded vignette



Learn and practice with lineups

Step 3: Upload and configure your data

Load Vignette

Upload Data

Upload Shared Analysis

Upload Data

Browse... fake_continuous_exposure_binary_response.csv

Upload complete

Exposure

measured_plasma_conc

Outcome

il2_stim_ratio

Plot Settings

☒ Log X

☒ Log Y

☒ Logistic Overlay

Generate Lineup

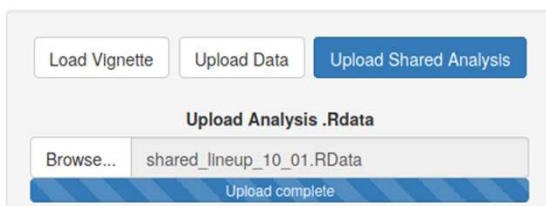
Save your analysis setup to share with team members (file includes your uploaded data).

Download Analysis Setup

Download sharable configuration

App workflow: Team Based Voting

Step 4: Teammates vote and report



...

Which plot do you think contains the real data?

You selected plot 1, but the true data was in plot 14.

Analysis leader tallies
correct/incorrect votes

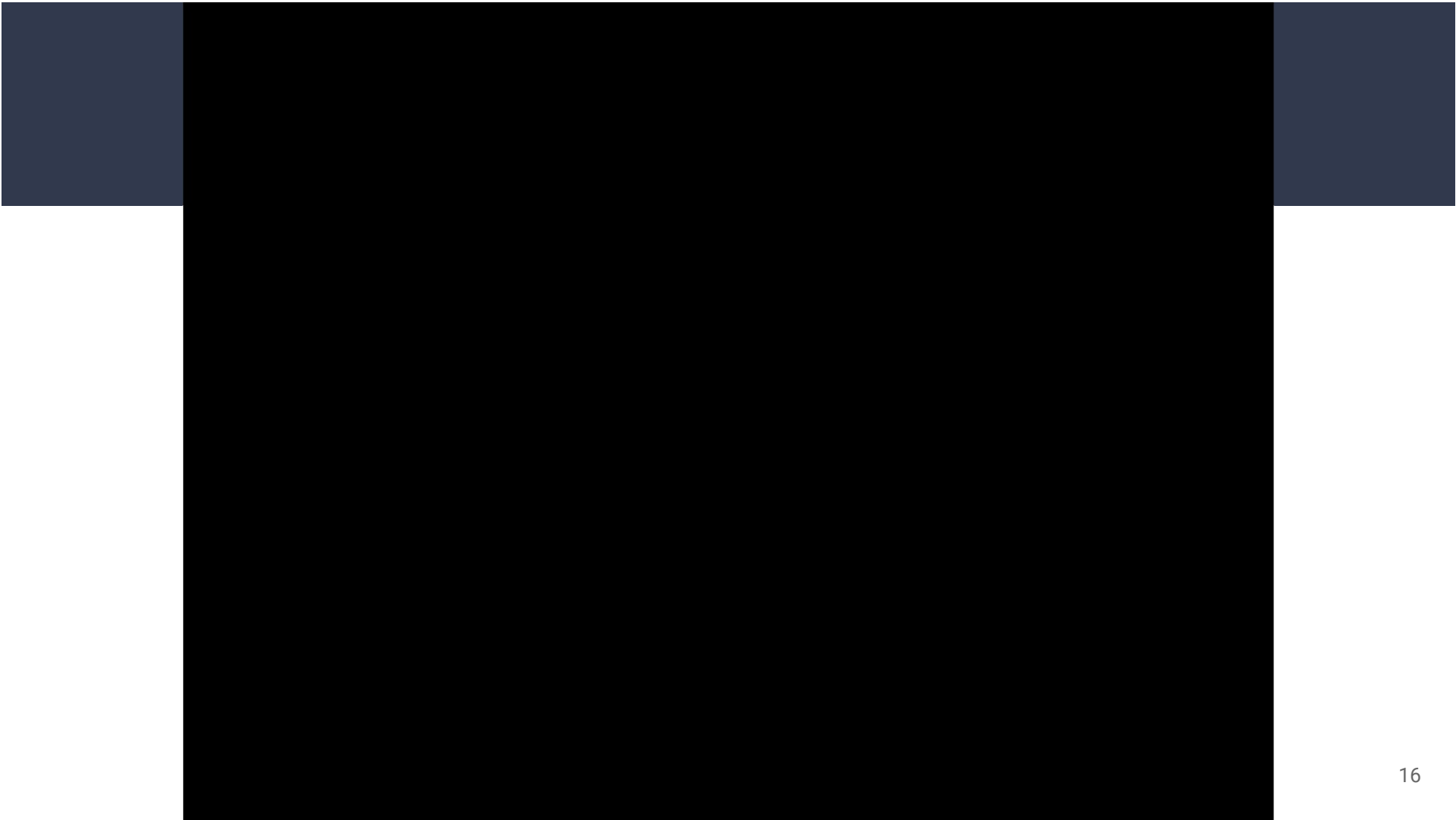
Step 5: 'See'-Value calculated in-app

Total Participants Performing Lineup Test

of Participants Correctly Identifying Data

'See'-Value: $2e-11$

Interpretation similar to
traditional p-value



Conclusion

Help your teams understand their data and develop an intuition for distinguishing a strong signal from random variability

Try out the app at:

Feedback to:

www.see-value.org

<at> mit.edu

hdiuhl

References

Buja, Cook, Hofmann, Lawrence, Lee, Swayne, and Wickham. 2009 [Statistical inference for exploratory data analysis and model diagnostics](#)

Majumder, Hofmann and Cook. 2013 [Validation of Visual Statistical Inference, Applied to Linear Models](#)

Roy Chowdhury, Cook, Hofmann, Majumder, Lee, and Toth. 2015 [Using visual statistical inference to better understand random class separations in high dimension, low sample size data](#)

Cokelaer et al. GDSCTools for mining pharmacogenomic interactions in cancer. Bioinformatics, 2017, <https://doi.org/10.1093/bioinformatics/btx744>

DSAIRM by [Handel Group at UGA](#) for app template

[nullabor](#) package by Cook et al. for lineup generation