flatiron® | ResearchX

WE WILL BE STARTING AT 1PM ET

EP 04

How Novel Methodologies are Powering Integrated Evidence

flatiron | ResearchX

Integrated evidence

Evidence that is more robust as a result of bringing together multiple sources of data.

GENERATE

COMBINE

ANALYZE

flatiron | Research$x$

Agenda

Pooling multiple real-world comparators — how to quantify heterogeneity

Daniel Backenroth, PhD
Scientific Director
Janssen

Post-prediction inference

Jeff Leek, PhD
Professor & Director of Johns Hopkins Data Science Lab
Johns Hopkins Bloomberg School of Public Health

BMS / Flatiron hybrid control designs

David Paulucci, MS
Associate Director of Data Science
BMS

Sanhita Sengupta, PhD
Senior Manager, Data Science
BMS

Katherine Tan, PhD
Senior Quantitative Scientist
Flatiron Health

Which type of methodological use cases does your organization use? *Select all that apply*

A. **Pooling** RWD to increase power or generalizability

B. **Hybrid control** methodologies to integrate trial and RWD

C. Analysis of **Machine Learning**-derived variables

Pooling multiple real-world comparators – how to quantify heterogeneity

Daniel Backenroth, PhD

Scientific Director, Janssen

03.30.2022

flatiron | ResearchX

Pooling multiple real-world comparators—how to quantify heterogeneity

Daniel Backenroth
Statistical Modeling & Methodology
April 13, 2022

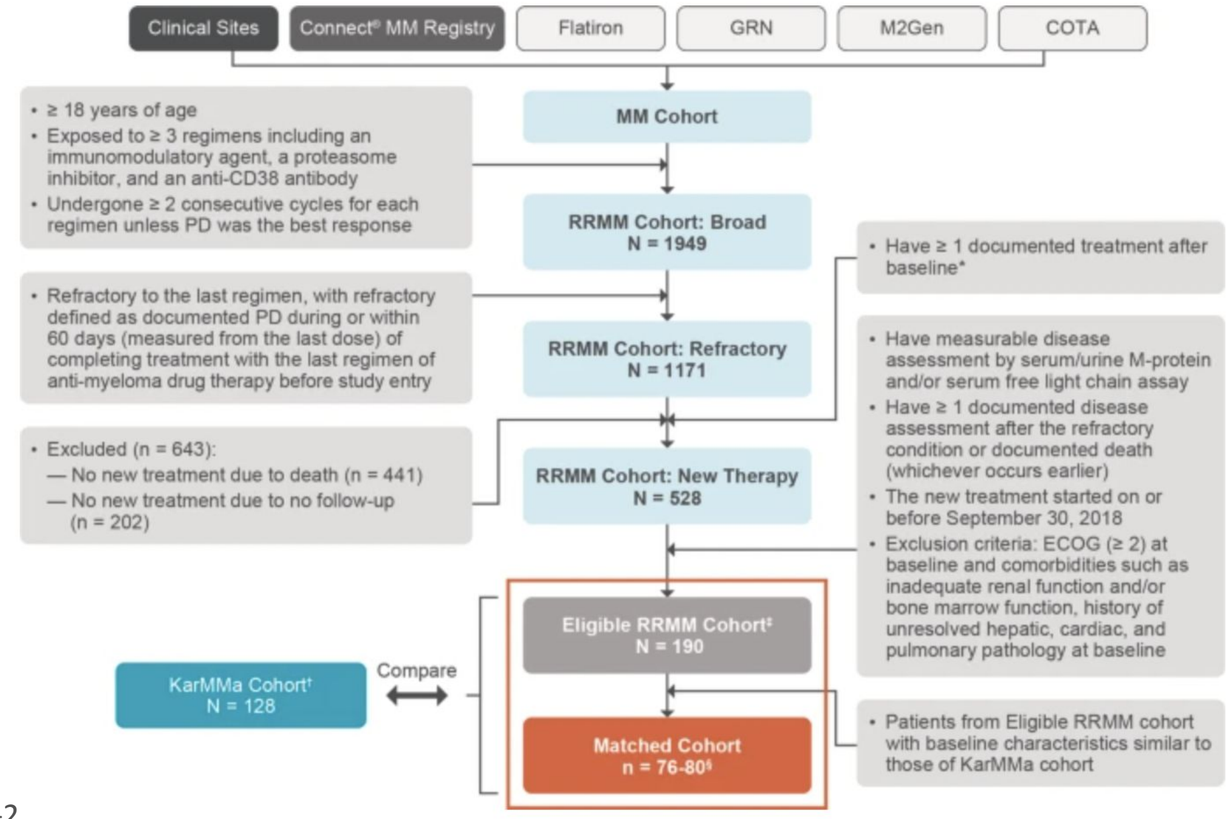
Judith Hinton Andrew, *Rock Composite 22*
Artwork from The Creative Center

Background: Why Pool?

For rare subpopulations, it may be necessary to **pool RWD** sources to serve as a real-world comparator cohort (rwCC) for a single arm clinical trial (SAT)

KarMMa-RW idecabtagene vicleucel to RWD comparison*

Fig. 1: Selection process for real-world Eligible Cohort.



*https://doi.org/10.1038/s41408-021-00507-2

How to Pool?

1. **Dataset selection**
   - Quality / fitness for use (Duke-Margolis, 2019)
   - Similarity of datasets (i.e., harmonizable values, definitions and populations)
2. **Data integration**
   - Deduplication
   - Data transformations and derivations
3. **Descriptive assessment of heterogeneity**
4. **Pooled data analysis**

Problem: how to assess heterogeneity?

We have multiple rwCC for a SAT, we want to evaluate their consistency.

Key rwCC assumption:

- After adjustment for confounders, SAT to rwCC comparison yields same estimate as would have gotten from RCT, up to sampling variability

- Implies no heterogeneity among rwCC-SAT comparison results

Warning

If we have multiple rwCC and all rwCC to SAT comparisons are consistent:

- This is better than if they are inconsistent
- But rwCC assumption could still be incorrect
  - All comparisons could be biased in the same way

Presenting evidence for homogeneity with a descriptive statistical test can raise confidence in analysis using multiple rwCC, even though it doesn't prove rwCC assumption holds

Where can heterogeneity come from?

- Unmeasured confounding—populations are different + information in datasets doesn’t account for confounding
  - Different baseline characteristics
  - Different supportive care after baseline
  - Different treatments received

- Collection of data on baseline information or outcomes differs between datasets
  - E.g., more deaths or progression events are missing from one dataset

Two methods to evaluate consistency

1. "Aggregate" method:
   - For each rwCC, calculate effect estimates vs. SAT
   - Then compare these effect estimates
2. "Individual patient data" method:
   - Compare rwCC datasets to each other, after matching/weighting to SAT

Focus here is aggregate method

- Can be calculated even if sponsor lacks access to all real-world datasets that are used
  - Common situation where disease registries are used/analyses are carried out in different jurisdictions with strict data protection rules

Aggregate method

Standard* method is Cochran's Q-test

- method for testing null hypothesis of homogeneity in meta-analysis using weighted sum of squared deviations around weighted mean

$$Q = \sum_{i=1}^{k} w_i (x_i - \bar{x}_w)^2$$

- Weights w_i are reciprocals of estimated variances of effect estimates x_i; weighted mean is $\bar{x}_w$

But see Hoaglin DC. Misunderstandings about Q and 'Cochran's Q test' in meta-analysis. *Stat Med*. 2016 Feb 20;35(4):485-95.

Aggregate method assumption

- Q test assumes effect estimates are **independent**
- *But here effect estimates are dependent since all rwCC are compared to the same single-arm trial*

- We present Adjusted Q test
  - Calculate estimated covariance matrix Σ of vector of statistics $\vec{x}$ from each rwCC-SAT comparison
  - Calculate $\vec{y} = \Sigma^{-1/2}\vec{x}$
  - Transformed statistics $\vec{y}$ can be assumed independent with identity covariance matrix, and standard Q test can be used

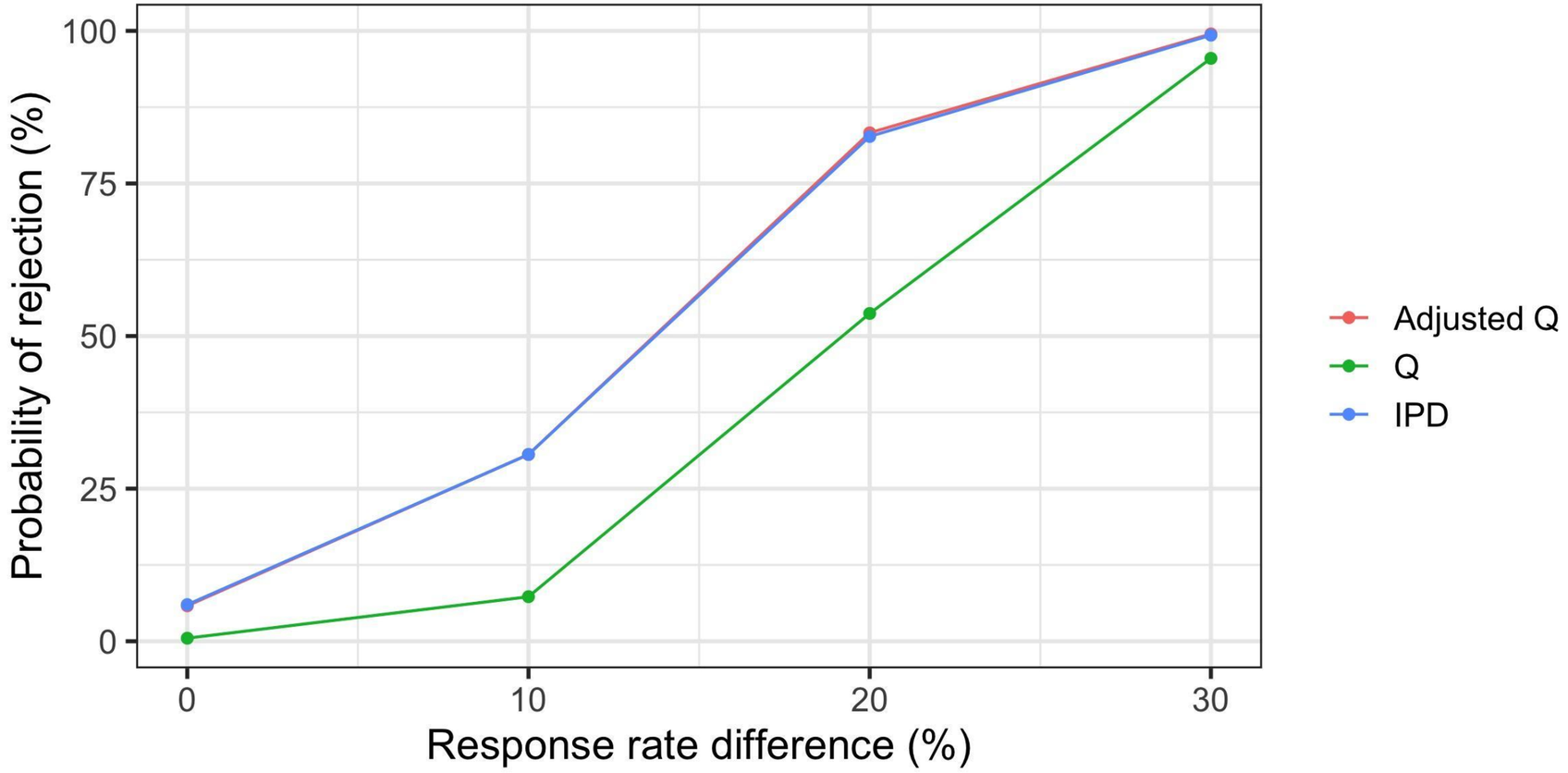
- How to calculate covariance matrix?
  - Suggest to use bootstrap
  - Resample from single-arm trial and rwCCs, calculate estimates and then calculate their covariance
  - As long as bootstrap samples for SAT are identical (shared random-number seed), can calculate covariance of estimates

Simulation

Simulated trials with binary endpoint

- 50% response rate in SAT and in one rwCC, same or different response rate in second rwCC (due to unmeasured confounding)
- Compare Q test, Adjusted Q test and IPD method (logistic regression) on probability of rejection of null hypothesis

Simulation results: Adjusted Q test works



Summary

- Pooling RWD can improve power and generalizability
- Rigorous data selection and integration are critical
- Presenting evidence for homogeneity with an appropriate descriptive statistical test can raise confidence in analysis using multiple rwCC
  - Q test is inappropriate because effect estimates from rwCC to SAT comparisons are not independent
  - Adjusted Q test can then be used instead
  - Can be used even if sponsor cannot access all rwCC, by using bootstrap to calculate covariance matrix of estimates

Acknowledgments/next steps

This is joint work and product of a Flatiron-Janssen collaboration to explore issues in pooling RWD sources.

Project team:

Daniel Backenroth PhD (Janssen)

Olivier Humblet ScD & Trevor Royce MD, MPH, MS (Flatiron)

Sponsors:

Meghna Samant PhD (Flatiron)

Jose Pinheiro PhD, Trilok Parekh PhD & Kiran Patel MD (Janssen)

Manuscript is in preparation, stay tuned!

BMS / Flatiron hybrid control designs

David Paulucci, MS
Associate Director of Data Science, BMS

Sanhita Sengupta, PhD
Senior Manager, Data Science, BMS

Katherine Tan, PhD
Senior Quantitative Scientist, Flatiron Health

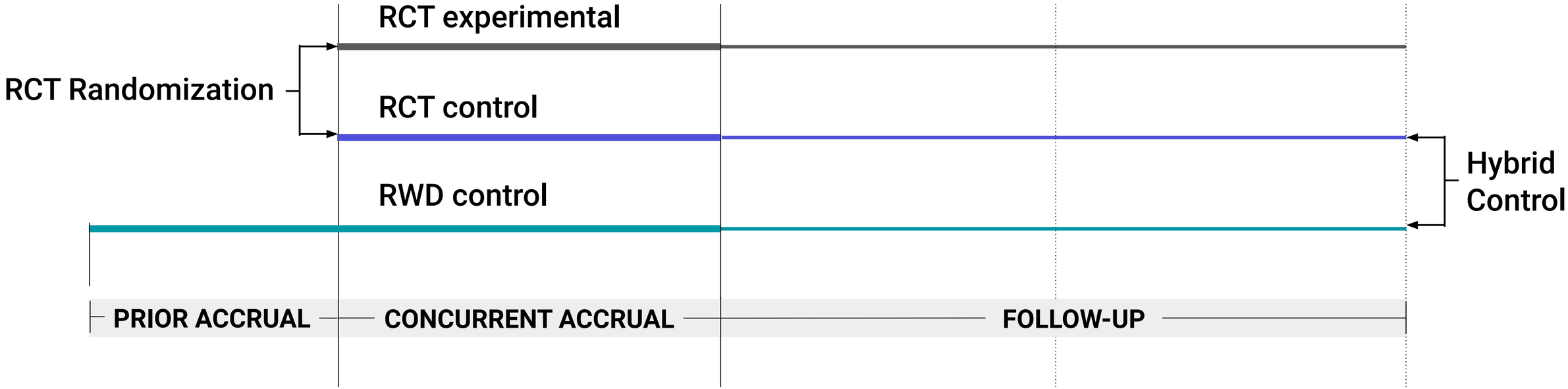
04.13.2022

flatiron. | ResearchX

Hybrid controlled designs using RWD can transform RCT and accelerate patient access to effective therapies

- Randomized controlled trials are the gold standard for evaluating the benefit-risk of new drugs and regulatory decision making in Pharma
- They are expensive, take a long time and at times hard to enroll (rare disease, avoiding control arm therapy due to ethics/patient preference), negatively impacting timely patient access to effective therapies
- Hybrid controlled designs address enrollment and timeline issues by supplementing RCT with external data

Example design of hybrid controlled trials



First patient in
Start accruing patients in trial and RWD & possibly select prior RWD patients

Last Patient in
Stop accruing patients both in trial and RWD

Assess borrowing
Compare outcomes and baseline characteristic; determine number of events that can be borrowed

Final analysis
Compare trial experimental vs hybrid control (control + RWD)

Examples of statistical borrowing methods

FREQUENTIST

1. Test-then-pool[1]
2. Two-step regression[4] (frequentist analog to modified power prior model)
   - Regression Step 1: Calculate hazard ratio comparing trial control to external control, to determine amount of downweighting
   - Regression Step 2: Calculate hazard ratio comparing treatment to hybrid control, to estimate treatment effect

BAYESIAN

1. Commensurate prior model[2]
2. Power prior model with fixed power parameter[3]
3. Meta-analytic predictive (MAP) prior models[5]
4. Modified power prior[6]

Bristol Myers Squibb™ flatiron

[1]Viele et al. (2014), [2]Hobbs et al. (2012), [3]Chen et al., 2000; Ibrahim et al., 2000, [4]Proposed by Flatiron as a simple alternative, [5]Neuenschwander et al. (2010), [6]Duan et al., 2006; Neuenschwander et al., 2009

Examples of statistical borrowing methods

FREQUENTIST

1. Test-then-pool[1]
2. Two-step regression[4] (frequentist analog to modified power prior model)
   - Regression Step 1: Calculate hazard ratio comparing trial control to external control, to determine amount of downweighting
   - Regression Step 2: Calculate hazard ratio comparing treatment to hybrid control, to estimate treatment effect

BAYESIAN

1. Commensurate prior model[2]
2. Power prior model with fixed power parameter[3]
3. Meta-analytic predictive (MAP) prior models[5]
4. Modified power prior[6]

Bristol Myers Squibb™ flatiron

[1]Viele et al. (2014), [2]Hobbs et al. (2012), [3]Chen et al., 2000; Ibrahim et al., 2000, [4]Proposed by Flatiron as a simple alternative, [5]Neuenschwander et al. (2010), [6]Duan et al., 2006; Neuenschwander et al., 2009

Examples of statistical borrowing methods

FREQUENTIST

1. Test-then-pool[1]
2. Two-step regression[4] (frequentist analog to modified power prior model)
   - Regression Step 1: Calculate hazard ratio comparing trial control to external control, to determine amount of downweighting
   - Regression Step 2: Calculate hazard ratio comparing treatment to hybrid control, to estimate treatment effect

BAYESIAN

1. Commensurate prior model[2]
2. Power prior model with fixed power parameter[3]
3. Meta-analytic predictive (MAP) prior models[5]
4. Modified power prior[6]

Bristol Myers Squibb™ flatiron

[1]Viele et al. (2014), [2]Hobbs et al. (2012), [3]Chen et al., 2000; Ibrahim et al., 2000, [4]Proposed by Flatiron as a simple alternative, [5]Neuenschwander et al. (2010), [6]Duan et al., 2006; Neuenschwander et al., 2009

Examples of statistical borrowing methods

FREQUENTIST

1. Test-then-pool[1]
2. Two-step regression[4] (frequentist analog to modified power prior model)
   - Regression Step 1: Calculate hazard ratio comparing trial control to external control, to determine amount of downweighting
   - Regression Step 2: Calculate hazard ratio comparing treatment to hybrid control, to estimate treatment effect

BAYESIAN

1. Commensurate prior model[2]
2. Power prior model with fixed power parameter[3]
3. Meta-analytic predictive (MAP) prior models[5]
4. Modified power prior[6]

Bristol Myers Squibb™ flatiron

[1]Viele et al. (2014), [2]Hobbs et al. (2012), [3]Chen et al., 2000; Ibrahim et al., 2000, [4]Proposed by Flatiron as a simple alternative, [5]Neuenschwander et al. (2010), [6]Duan et al., 2006; Neuenschwander et al., 2009

Hybrid controls are part of a larger project lifecycle

Evaluate fit-for-purpose of external source RWD

Data availability & criticality

Construct appropriate RWD analytic cohort to form hybrid control

Cohort selection from RWD

Balance baseline patient characteristics between RWD and RCT cohorts

Emulate RCT with hybrid controls

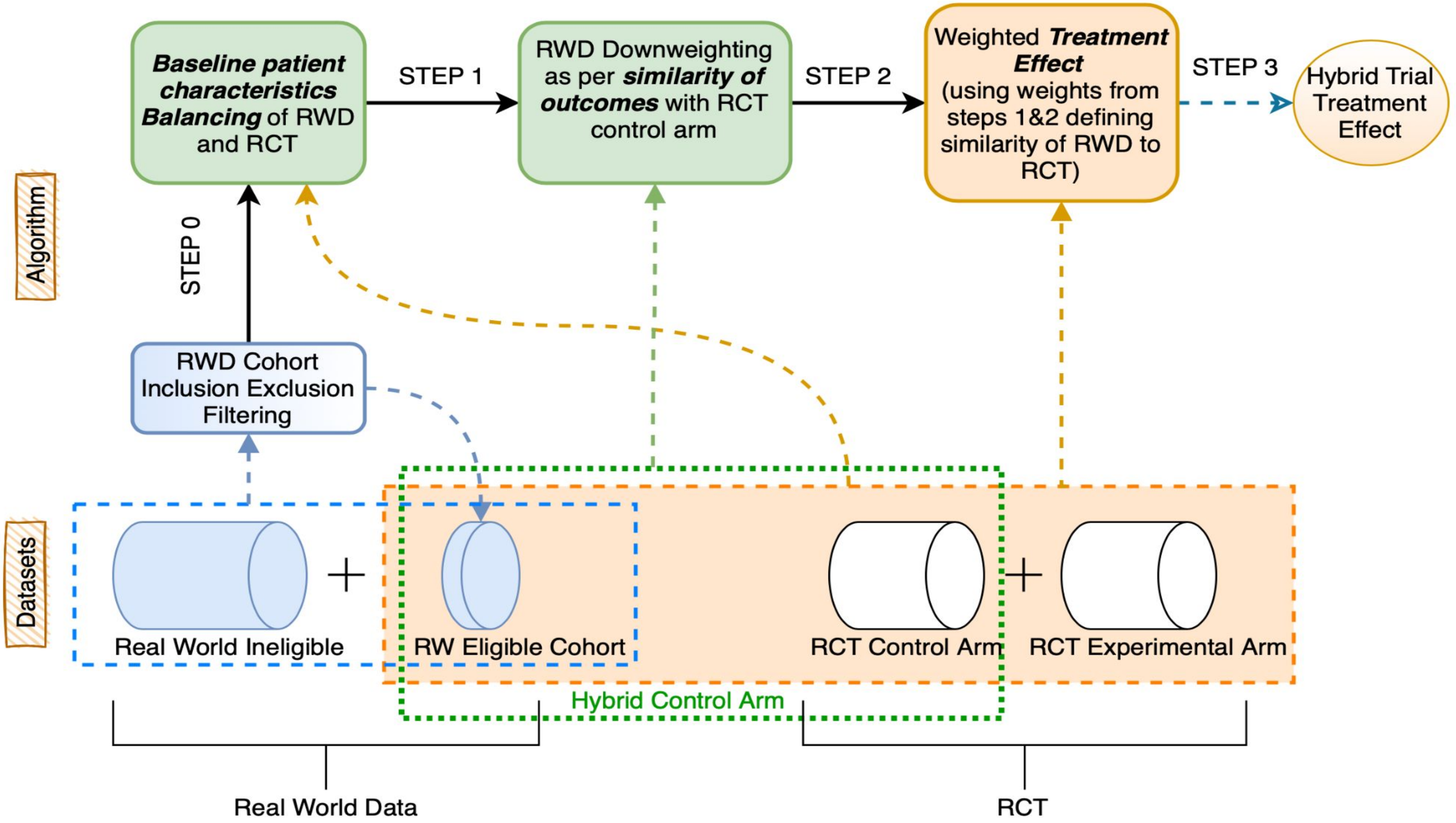
Estimate amount of borrowing from RWD to form hybrid control

Assess treatment effect and timeline savings from using hybrid control

Hybrid Control Arm case study: Research Objectives

- To describe differences in patient characteristics and overall survival (OS) between the control arm of a completed RCT and an emulated hybrid control arm derived from the control arm of that completed RCT and Flatiron real-world data (RWD).
    - Describe the data fitness and considerations of using the source RWD cohort for cohort selection, covariate balancing, and outcomes assessment.
- To evaluate bias in the estimated treatment effect of OS from the RCT under an emulated hybrid control design, using a single borrowing methodology
    - Evaluate impact on study duration under a hybrid control design

Methodology Details:



Hybrid Controlled Trial: Quantitative Evaluation of Emulation

Predefined success metrics for the emulated hybrid controlled trial were achieved

Treatment effect: OS Hazard Ratio in the same direction (<1) and within confidence interval of RCT

Impact on study duration: Based on the number of events that could be borrowed, the emulation demonstrated a ***hypothetical*** **trial reduction time of between 7–11 months**

However, real-world patients were observed to have poorer OS compared to trial control patients.

- Important to have high quality real-world data and strict alignment to trial I/E criteria, as well as utilizing analytical methods (e.g., ECOG imputation) to power baseline characteristic alignment

Hybrid Controlled Trial: Lessons Learned

Overall, this patient-level emulation exercise using hybrid control designs, via a RWD/drug sponsor collaboration, is an **important advancement in the area of integrating RWE and RCT data**.

Critical to maximize data availability and completeness for key prognostic and confounding variables, to successfully derive a real-world cohort that is closely aligned with RCT

Evaluation of multiple statistical borrowing approaches is likely needed to inform the optimal method in the study context

Robust application of other statistical methodologies that pertain to using RWD (not specific to hybrid controls), e.g. propensity score adjustment

Timing of interim analyses at which the borrowing decision is made needs to be carefully considered to optimize trial timeline savings

Post-prediction inference

Jeff Leek, PhD

Professor & Director of Johns Hopkins Data Science Lab,
Johns Hopkins Bloomberg School of Public Health

04.13.2022

Post-prediction inference

(what we do after we have machine learned everything)

@jtleek

jtleek.com

(look for “Talks”)

Disclosures

Current: Professor of Biostatistics, Director Data Science Lab, Johns Hopkins Bloomberg School of Public Health

Summer: Vice President, Chief Data Officer, and J Orin Edson Foundation Professor at Fred Hutchinson Cancer Research Center

Relationships:

- Instructor Coursera Programs
- Co-Founder Streamline Data Science (https://streamlinedatascience.io/healthcare)
- Co-Founder papr (https://www.papr.io/)
- Collaborator/Speaker/Advisory Board
    - Flatiron Health, Johnson and Johnson, Point Field Partners

Methods for correcting inference based on outcomes predicted by machine learning

Siruo Wang[a], Tyler H. McCormick[b,c], and Jeffrey T. Leek[a,1]

[a]Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205; [b]Department of Statistics, University of Washington, Seattle, WA 98195; and [c]Department of Sociology, University of Washington, Seattle, WA 98195

Edited by Robert Tibshirani, Stanford University, Stanford, CA, and approved October 6, 2020 (received for review January 24, 2020)

Many modern problems in medicine and public health leverage machine-learning methods to predict outcomes based on observable covariates. In a wide array of settings, predicted outcomes are used in subsequent statistical analysis, often without accounting for the distinction between observed and predicted outcomes. We call inference with predicted outcomes postprediction inference. In this paper, we develop methods for correcting statistical inference using outcomes predicted with arbitrarily complicated machine-learning models including random forests and deep neural nets. Rather than trying to derive the correction from first principles for each machine-learning algorithm, we observe that there is typically a low-dimensional and easily modeled representation of the relationship between the observed and predicted outcomes. We build an approach for postprediction inference that naturally fits into the standard machine-learning framework where the data are divided into training, testing, and validation sets. We train the prediction model in the training set, estimate the relationship between the observed and predicted outcomes in the testing set, and use that relationship to correct subsequent inference in the validation set. We show our postprediction inference (postpi) approach can correct bias and improve variance estimation and subsequent statistical inference with predicted outcomes. To show the broad range of applicability of our approach, we show postpi can improve inference in two distinct fields: modeling predicted phenotypes in repurposed gene expression data and modeling predicted causes of death in verbal autopsy data. Our method is available through an open-source R package: https://github.com/leekgroup/postpi.

known inheritance patterns for the disease. The predicted outcome can be used in place of the observed Alzheimer's status when performing a genome-wide association study (15).

This is just one example of the phenomenon of postprediction inference (postpi). Although common, this approach poses multiple statistical challenges. The predicted outcomes may be biased, or the predicted outcomes may have less variability than the actual outcomes. Standard practice in many applications is to treat predicted outcomes as if they were observed outcomes in subsequent regression models (6, 14–18). As we will show, uncorrected postprediction inference will frequently have deflated standard errors, bias, and inflated false positive rates.

Postprediction inference appears across fields and has been recognized as a potential source of error in recent work on prevalence estimation (see for example refs. 19 and 20 in the context of dataset shift and ref. 21 in document class prevalence estimation). Here, we focus on developing analytical and bootstrap-based approaches to correct regression estimates, SEs, and test statistics in inferential regression models using predicted outcomes. We examine settings where a predicted outcome becomes the dependent variable in the subsequent inferential regression analysis. We derive an analytical correction in the case of linear regression and bootstrap-based corrections for more general regression models, focusing on linear and logistic regression as they are the most common inferential models. Our bootstrap-based approach can, however, easily be extended to any generalized linear regression inference model.

Both our analytical and bootstrap-based corrections take

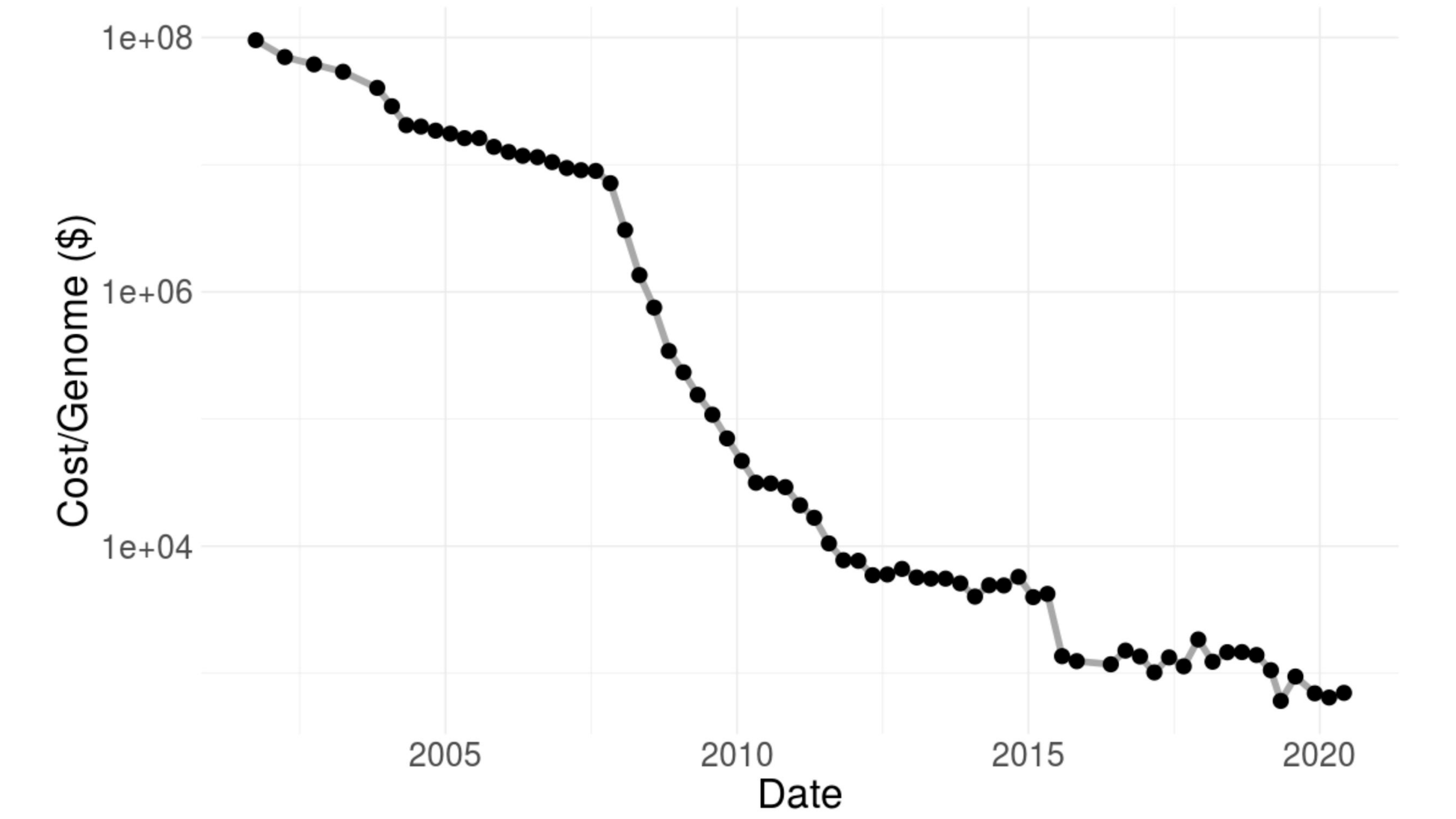
The real brains here

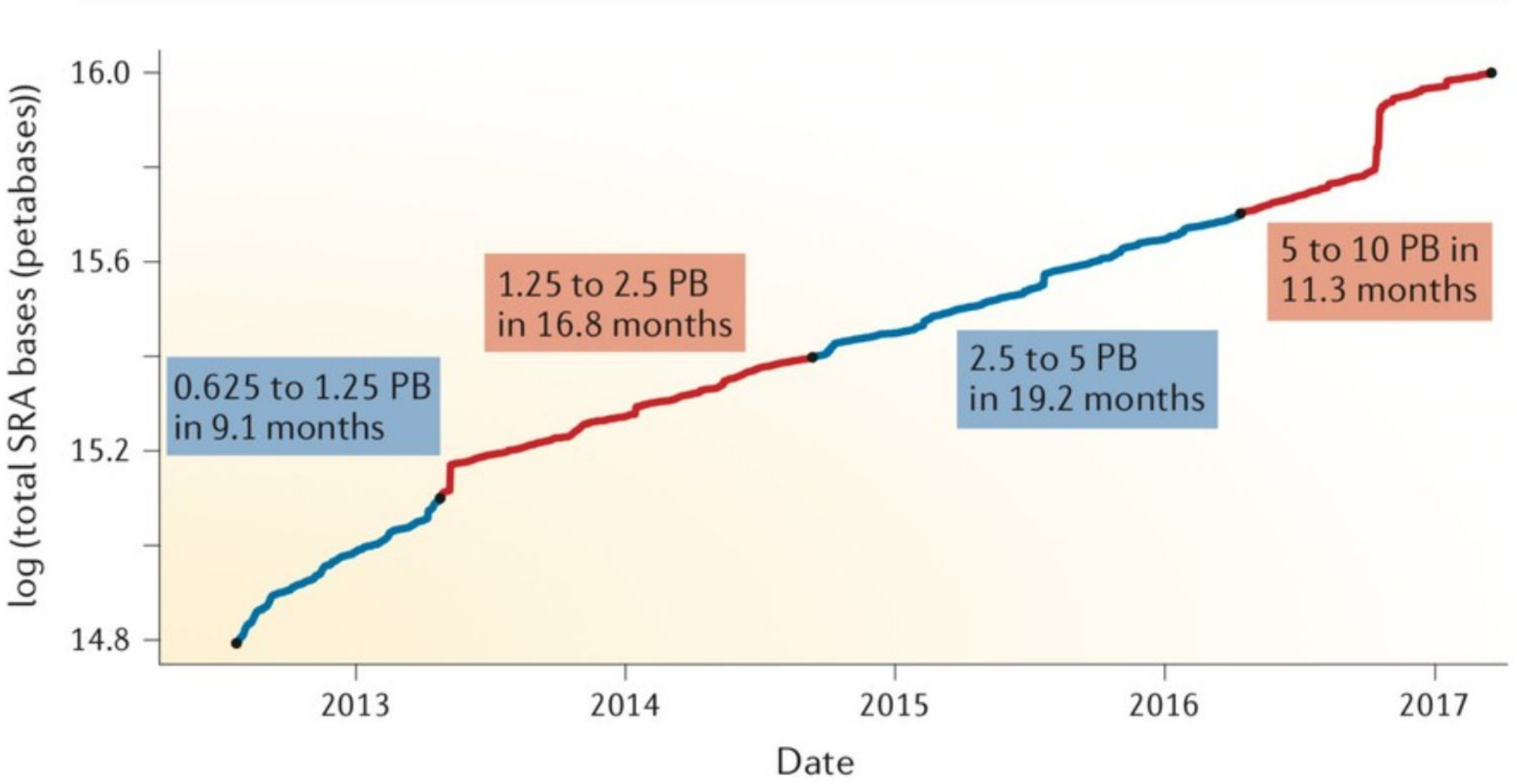
Sara Wang

Tyler McCormick

N = SAMPLE SIZE

$$N = \frac{(\$ \text{ YOU HAVE})}{(\$ \text{ PER SAMPLE})}$$


1e+08
1e+06
1e+04
Cost/Genome ($)
2005
2010
2015
2020
Date

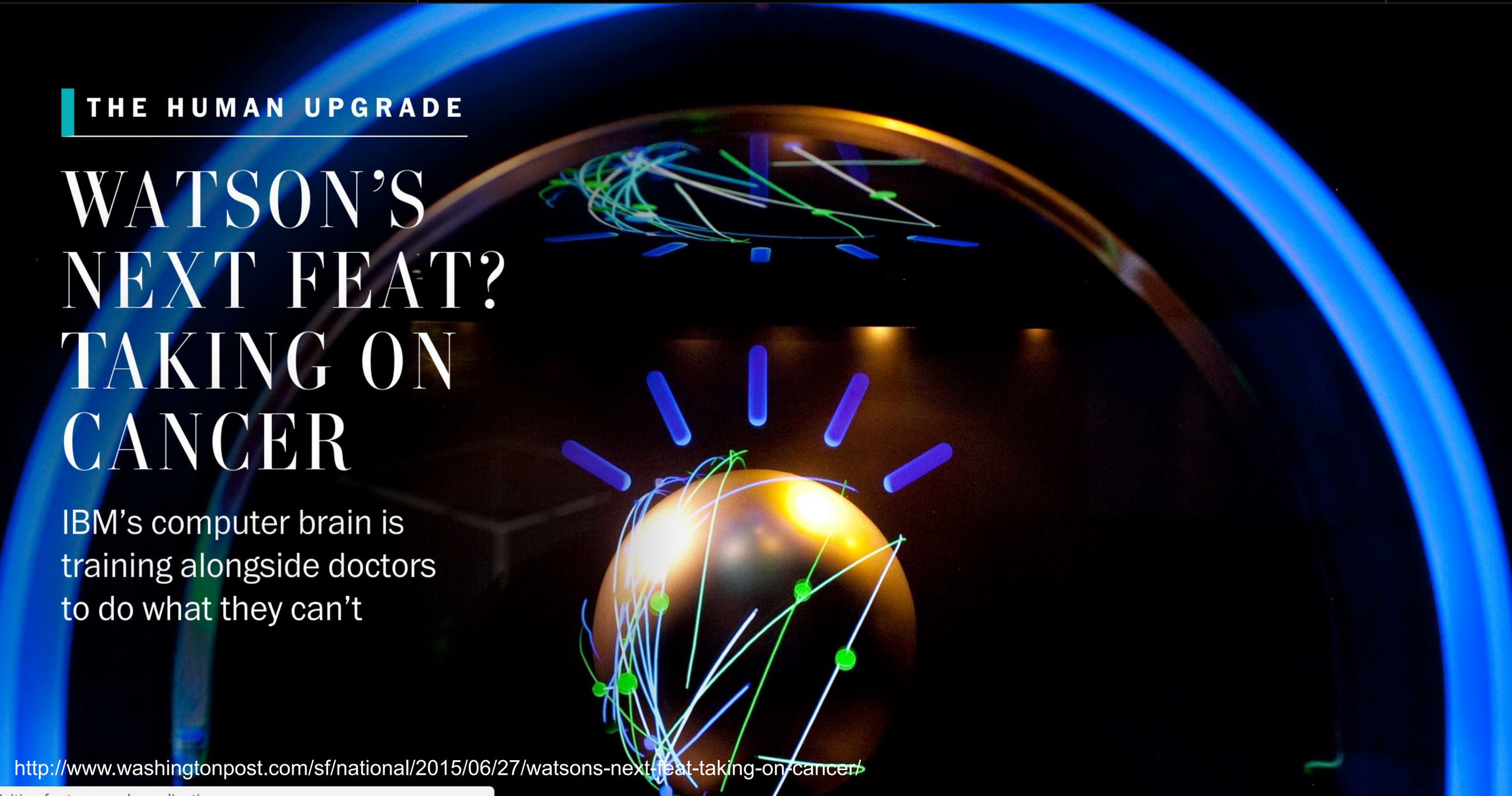


Langmead & Nellore, *Nat Rev. Genet.* 2018

THE HUMAN UPGRADE

WATSON'S NEXT FEAT? TAKING ON CANCER

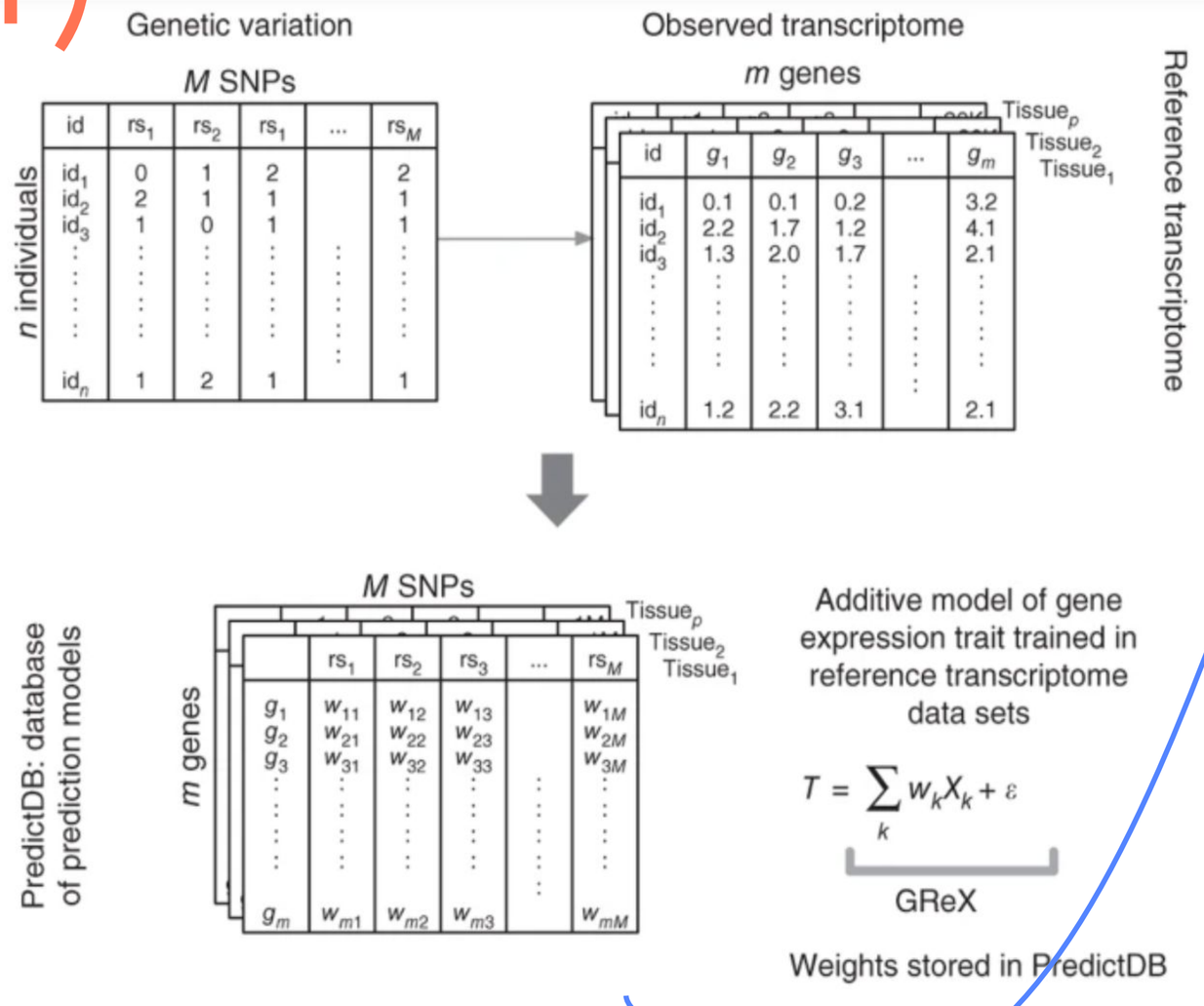
IBM's computer brain is training alongside doctors to do what they can't



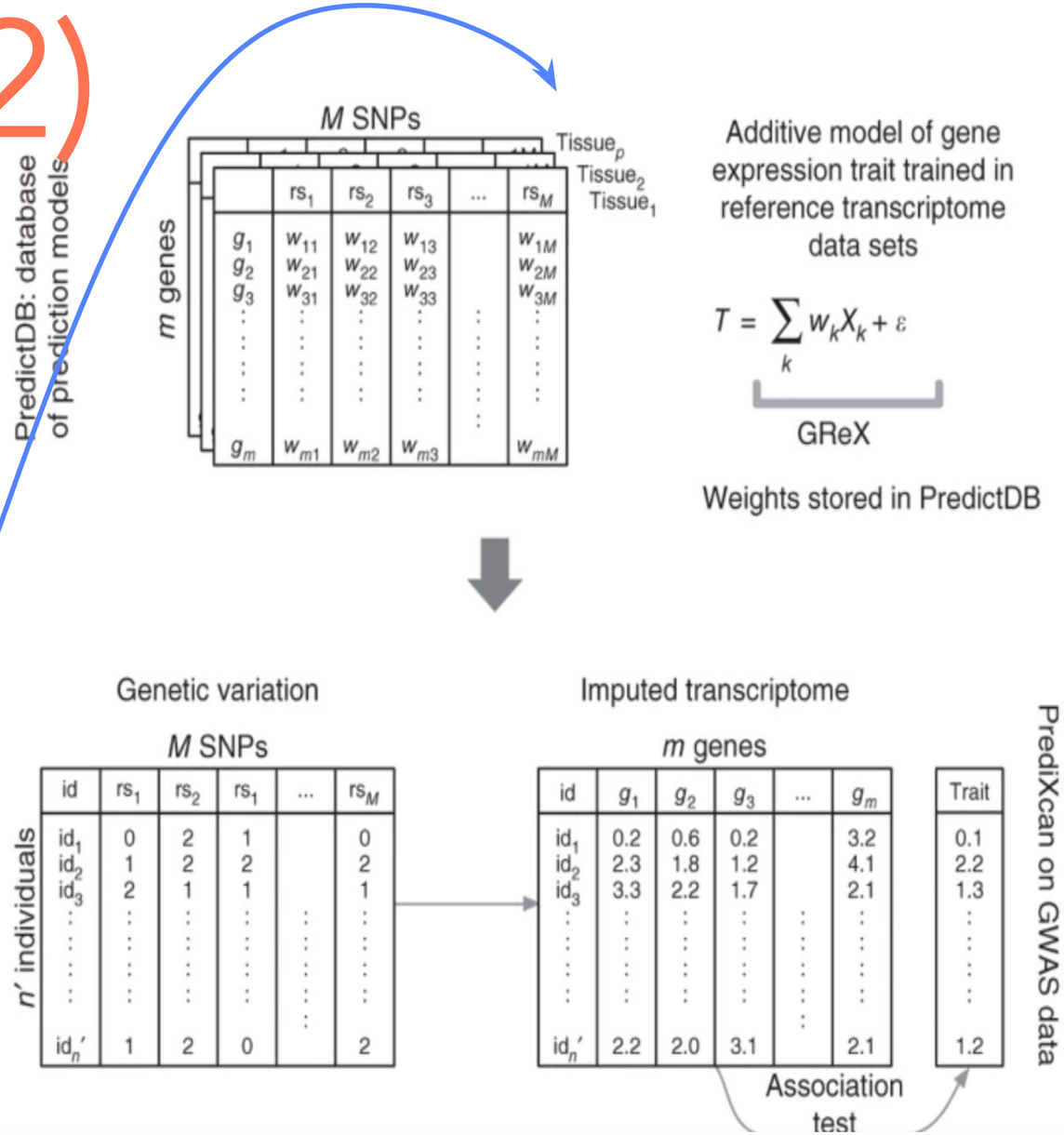
http://www.washingtonpost.com/sf/national/2015/06/27/watsons-next-feat-taking-on-cancer/

A new observation

1)

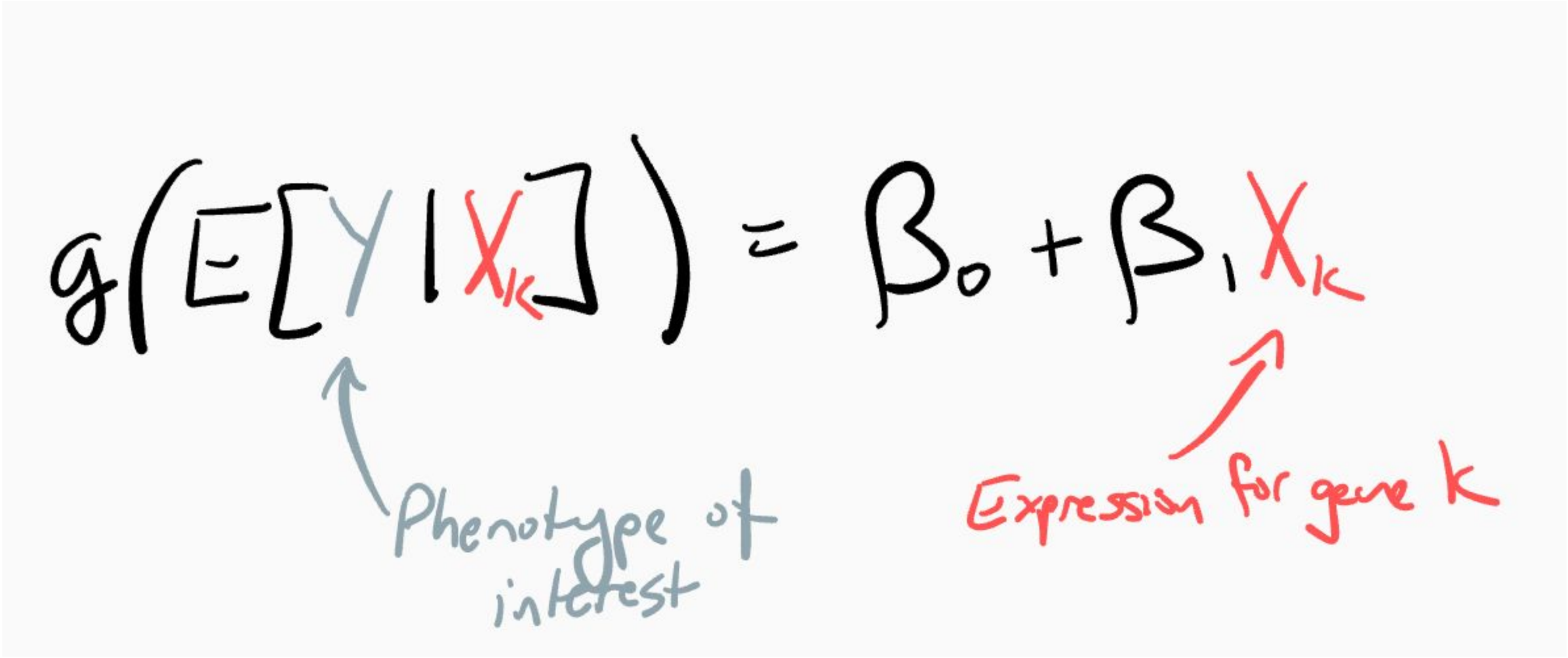


2)

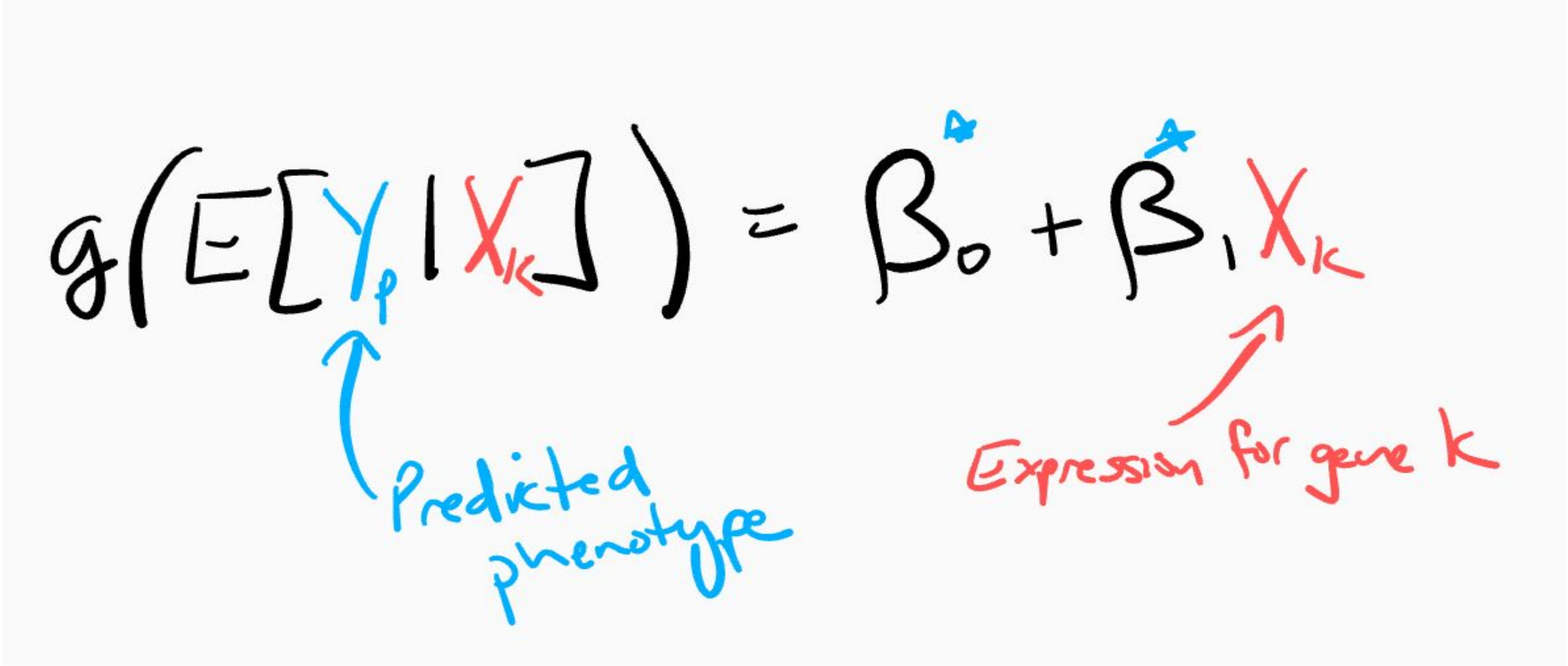


https://www.nature.com/articles/ng.3367

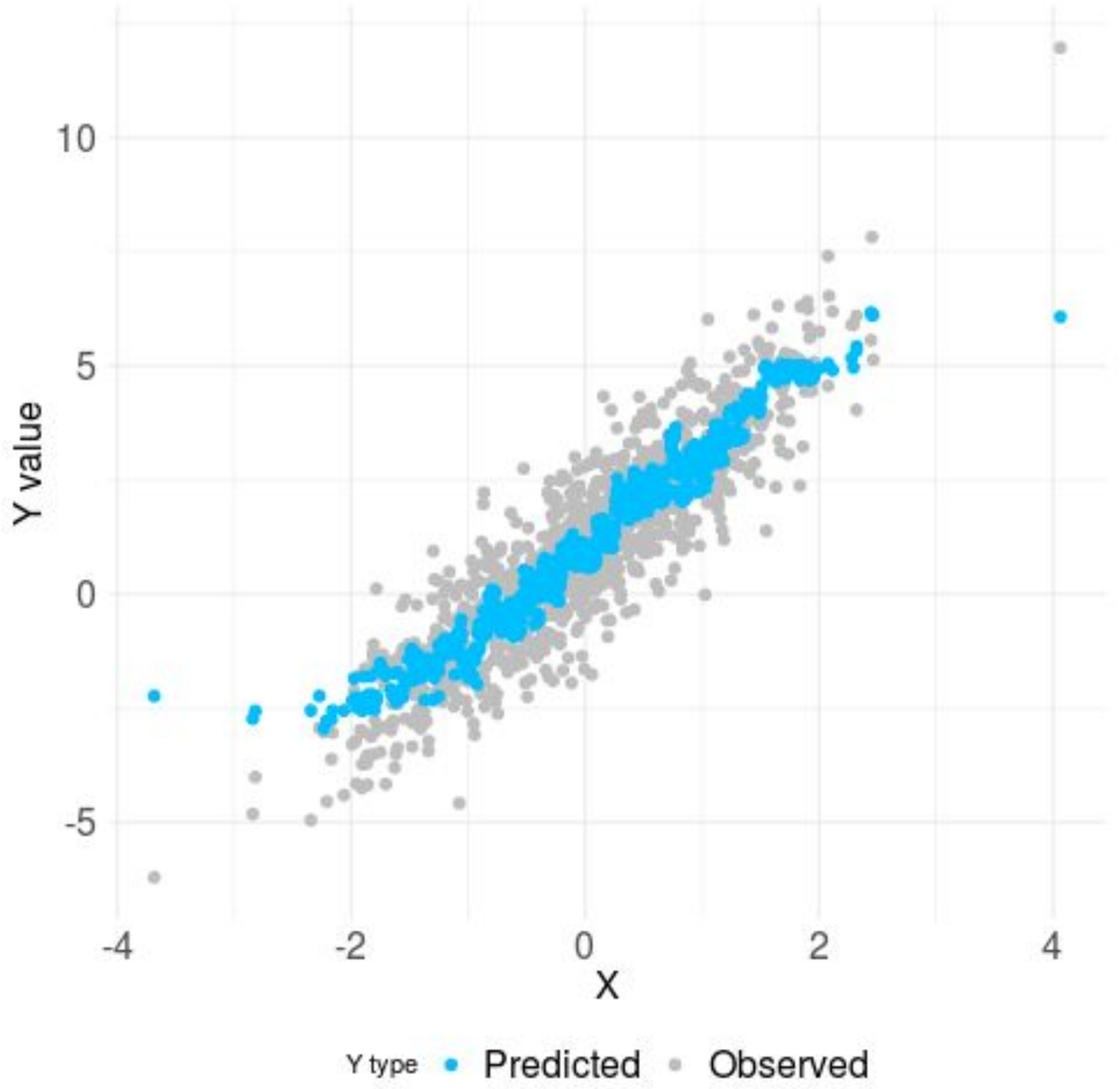
It can cause problems



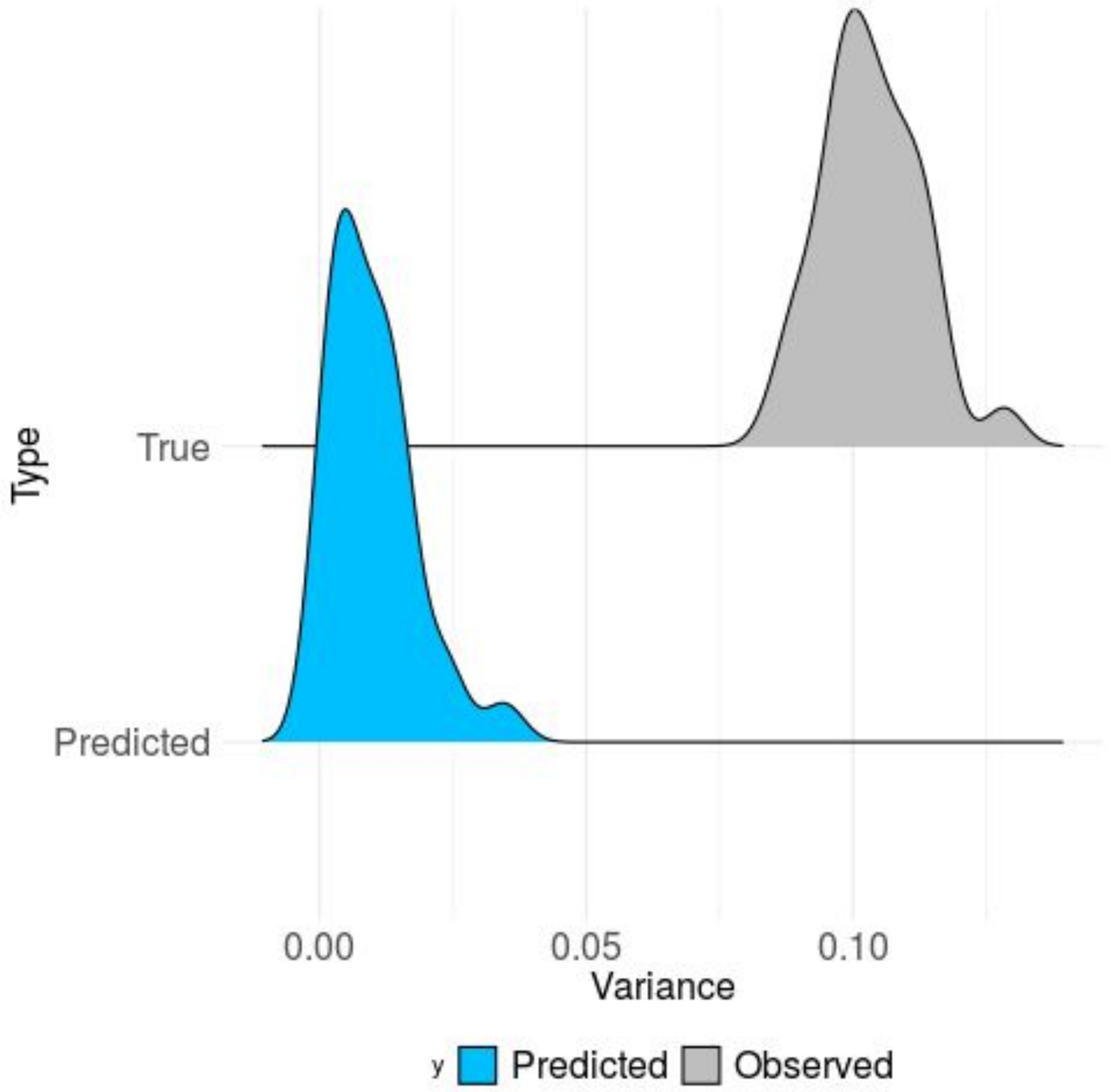
It can cause problems



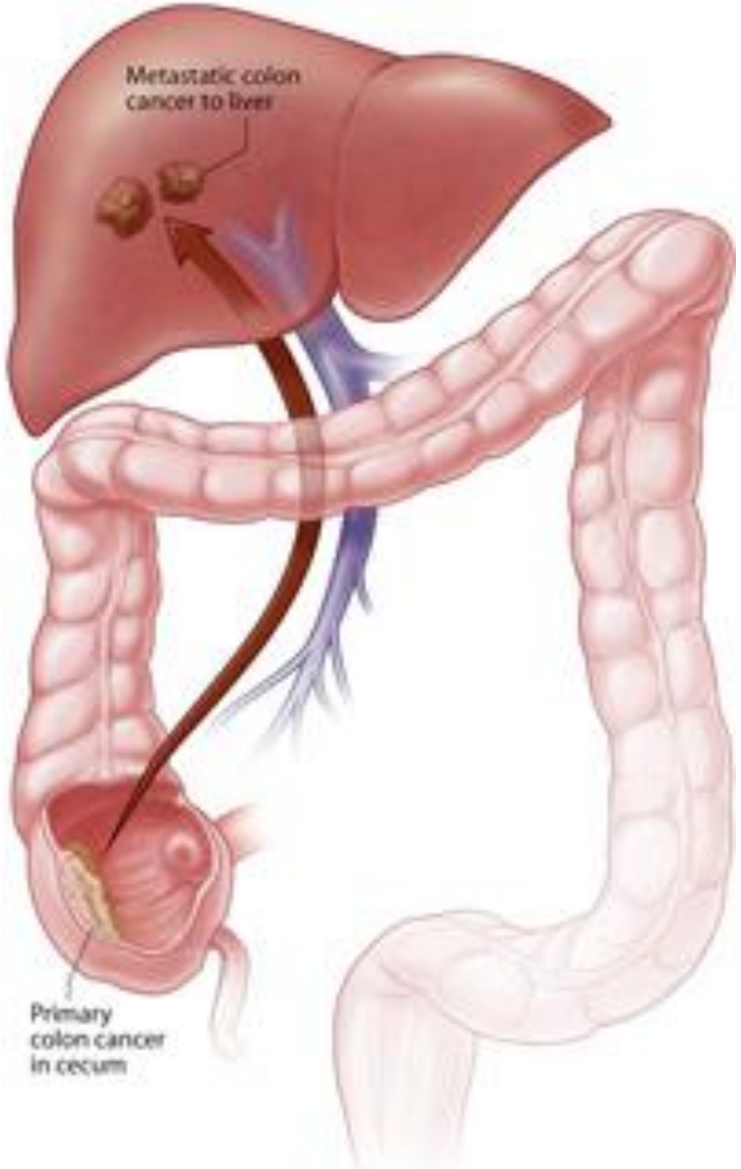
It can cause problems



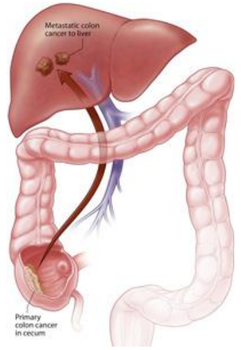
Underestimated variance



What genes are prognostic of colorectal cancer metastasis ?



www.hopkinsmedicine.org



What makes primary cancer different than metastatic cancer?

Collect patient samples and information

~6 months

Find a researcher with access to patient samples

1-2 wks

Extract DNA/RNA from samples

Sequence samples

2-4 wks

Process sequencing data

3-6 months

1-3 months

Data cleaning

Analyze data and answer biological question

1 month - 1+ years

Total: 2+ years

ACTACTTT

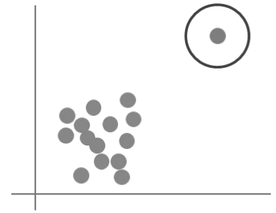
Sequence data

Process/quantify

recount3

What % expressed?

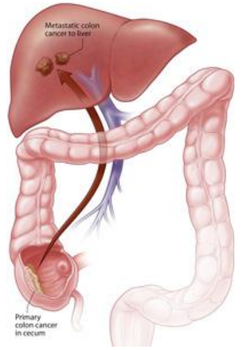
New genes?



Important outliers?

$\log\det \Theta - \mathrm{tr}\, S\Theta - \rho\|\Theta\|_1$

Best methods?



What makes primary cancer different than metastatic cancer?

Find a researcher with access to patient samples

Collect patient samples and information

3~6 months

Extract DNA/RNA from samples

1-2 wks

Sequence samples

2-4 wks

Process sequencing data

3-6 months

Data cleaning

1-3 months

Analyze data and answer biological question

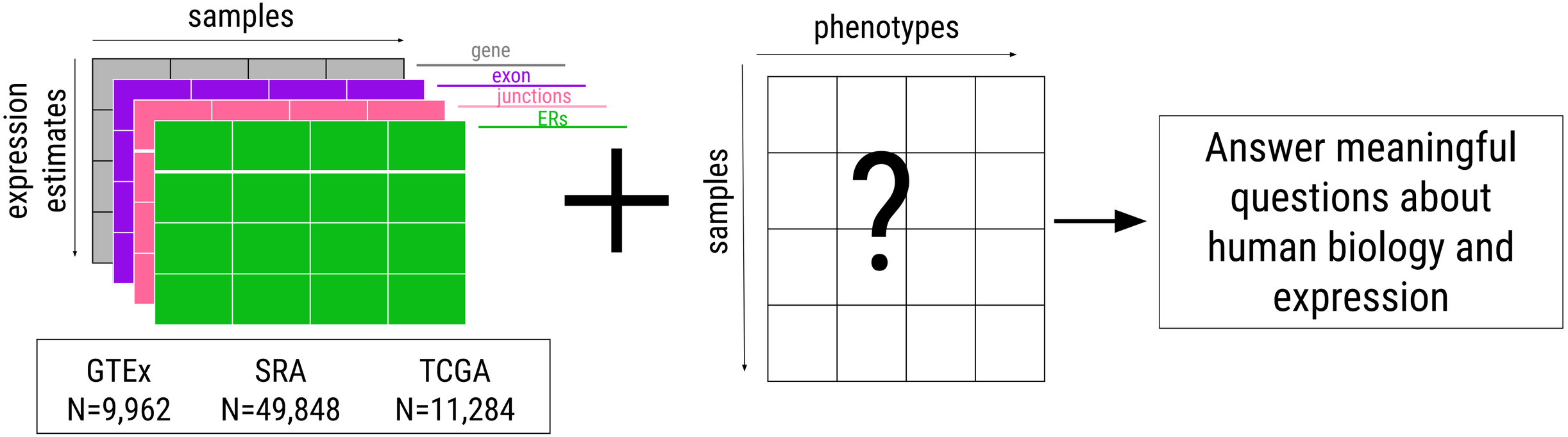
1 month - 1+ years

Total: 1+ years

A new problem

recount2

expression data for ~70,000 human samples



SRA phenotype information is far from complete

	Sex	Tissue	Race	Age
6620	female	liver	NA	NA
6621	female	liver	NA	NA
6622	female	liver	NA	NA
6623	female	liver	NA	NA
6624	female	liver	NA	NA
6625	male	liver	NA	NA
6626	male	liver	NA	NA
6627	male	liver	NA	NA
6628	male	liver	NA	NA
6629	male	liver	NA	NA
6630	male	liver	NA	NA
6631	NA	blood	NA	NA
6632	NA	blood	NA	NA
6633	NA	blood	NA	NA
6634	NA	blood	NA	NA
6635	NA	blood	NA	NA
6636	NA	blood	NA	NA

Even when information *is* provided, it's not always clear…

Sex across the SRA:

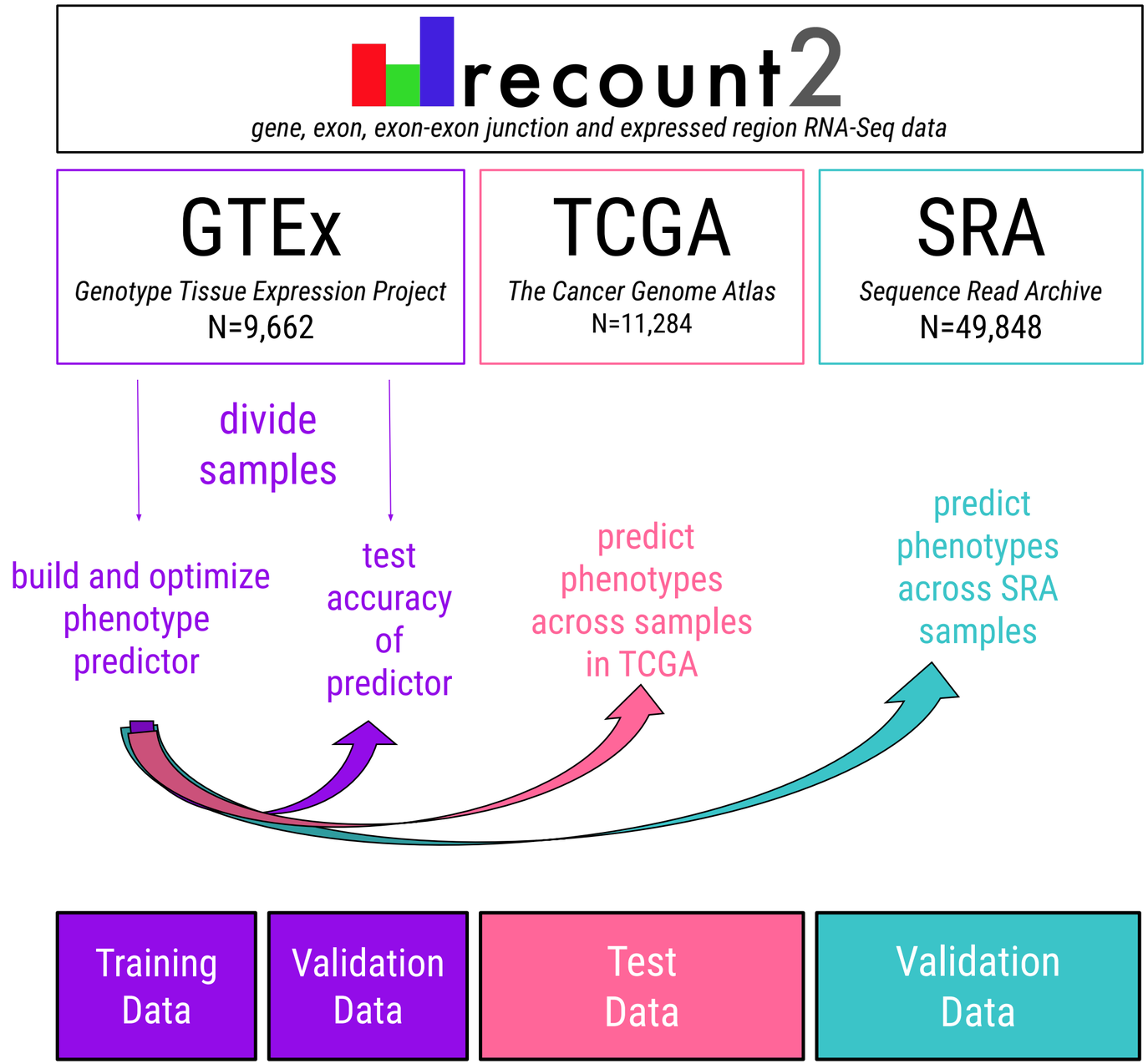
Level	Frequency
F	95
female	2036
Female	51
M	77
male	1240
Male	141
Total	**3640**

"1 Male, 2 Female", "2 Male, 1 Female", "3 Female", "DK", "male and female" "Male (note: ….)", "missing", "mixed", "mixture", "N/A", "Not available", "not applicable", "not collected", "not determined", "pooled male and female", "U", "unknown", "Unknown"

# of NAs	# w/sex assigned
44,957	4,700

Goal :

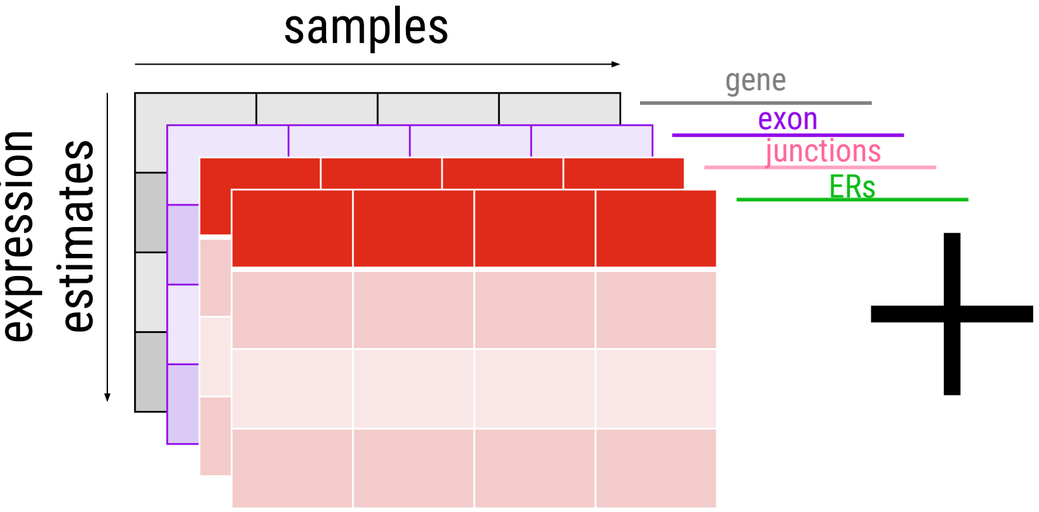
to accurately predict critical phenotype information for all samples in *recount2*



Problem solved (thanks Shannon!)

recount2

expression data for ~70,000 human samples



GTEx	SRA	TCGA
N=9,962	N=49,848	N=11,284

phenotypes

samples

sex	tissue	Cell line?
M	Blood	yes
F	Heart	no
F	Liver	no

Answer meaningful questions about human biology and expression

ACTACTTT

Sequence data

Process/quantify

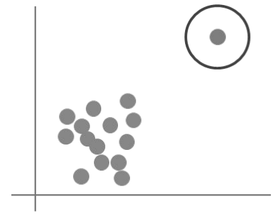
Clean/predict

Metadata

recount3

What % expressed?

New genes?



Important outliers?

$\log\det \Theta - \mathrm{tr}\, S\Theta - \rho\|\Theta\|_1$

Best methods?

It's happening everywhere

Published: 10 August 2015

A gene-based association method for mapping traits using reference transcriptome data

Eric R Gamazon, Heather E Wheeler, Kaanan P Shah, Sahar V Mozaffari, Keston Aquino-Michaels, Robert J Carroll, Anne E Eyler, Joshua C Denny, GTEx Consortium, Dan L Nicolae, Nancy J Cox & Hae Kyung Im

Nature Genetics **47**, 1091–1098(2015) | Cite this article

https://www.nature.com/articles/ng.3367

It’s happening everywhere

Computational and Systems Biology, Genetics and Genomics

Identifying gene expression programs of cell-type identity and cellular activity with single-cell RNA-Seq

Dylan Kotliar, Adrian Veres, M Aurel Nagy, Shervin Tabrizi, Eran Hodis, Douglas A Melton, Pardis C Sabeti

Harvard Medical School, United States; Broad Institute of MIT and Harvard, United States; Massachusetts Institute of Technology, United States; Harvard University, United States; Howard Hughes Medical Institute, United States

Tools and Resources · Jul 8, 2019

Cited 12 Views 12,639 Annotations 0

Cite as: eLife 2019;8:e43803 DOI: 10.7554/eLife.43803

https://elifesciences.org/articles/43803

It’s happening everywhere

Genetics and Genomics

An atlas of polygenic risk score associations to highlight putative causal relationships across the human phenome

Tom G Richardson, Sean Harrison, Gibran Hemani, George Davey Smith

University of Bristol, United Kingdom

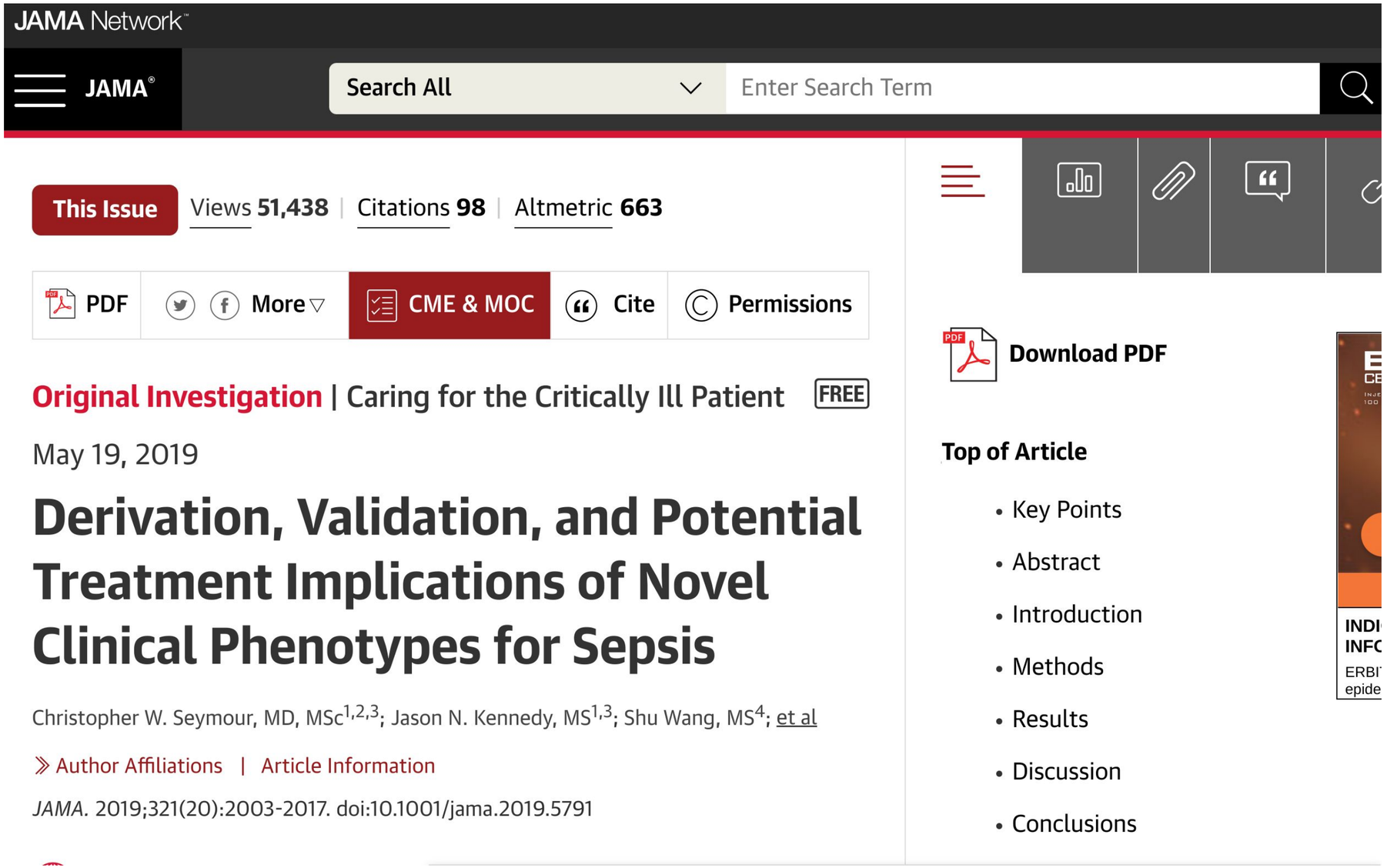
Tools and Resources · Mar 5, 2019

Cited 37 Views 11,298 Annotations 2

Cite as: eLife 2019;8:e43657 DOI: 10.7554/eLife.43657

https://elifesciences.org/articles/43657

It's happening everywhere



https://jamanetwork.com/journals/jama/fullarticle/2733996

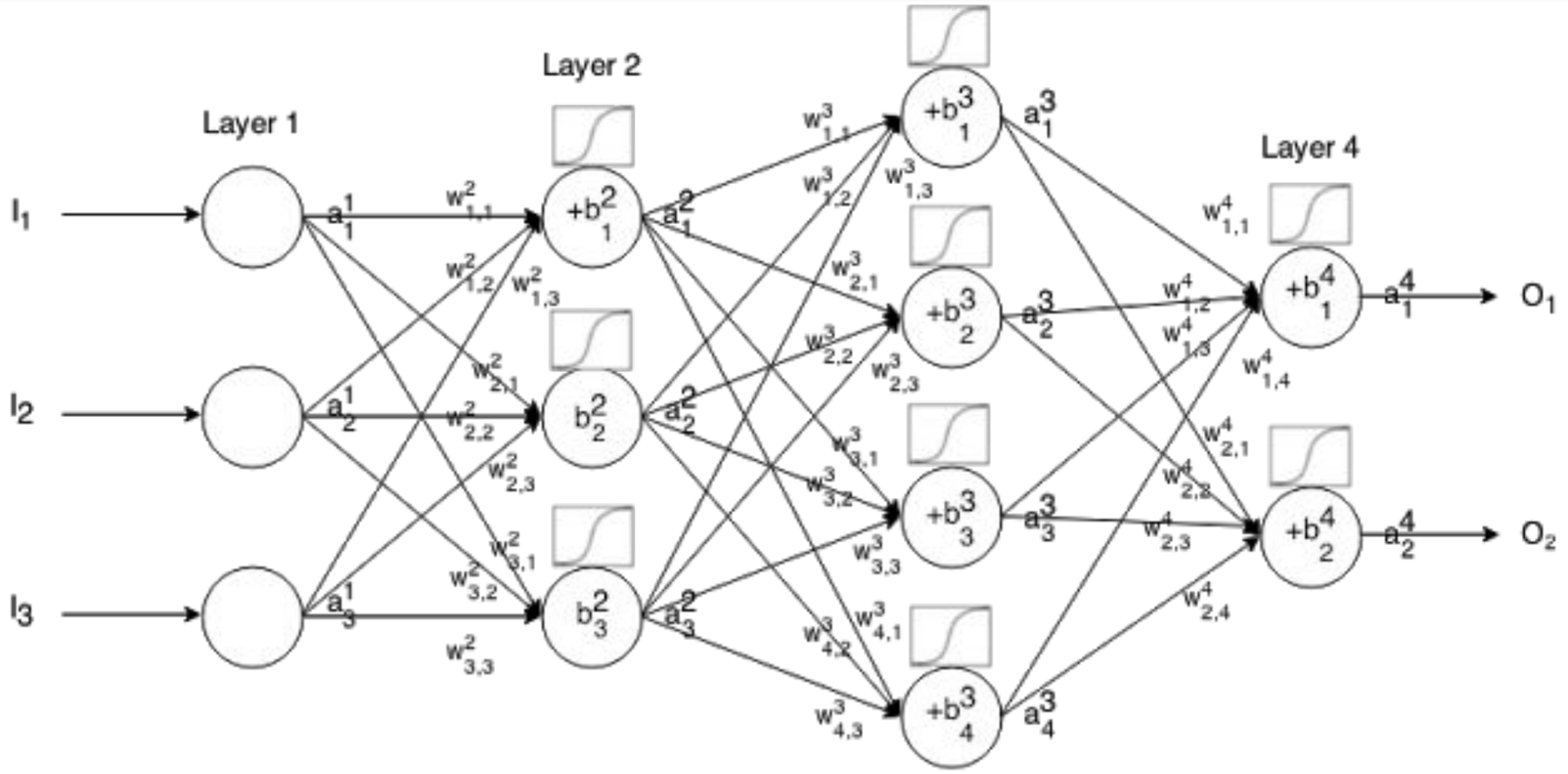
We need “post-prediction inference”

$$E[Y|X] = \beta_0 + \beta_1 X$$

$$E[Y_p|X] = \beta_0^{\star} + \beta_1^{\star} X$$

$$E[Y_p \mid X] = \beta_0^{\star} + \beta^{\star} X$$

This is hard to model



Matters arising

Transparency and reproducibility in artificial intelligence

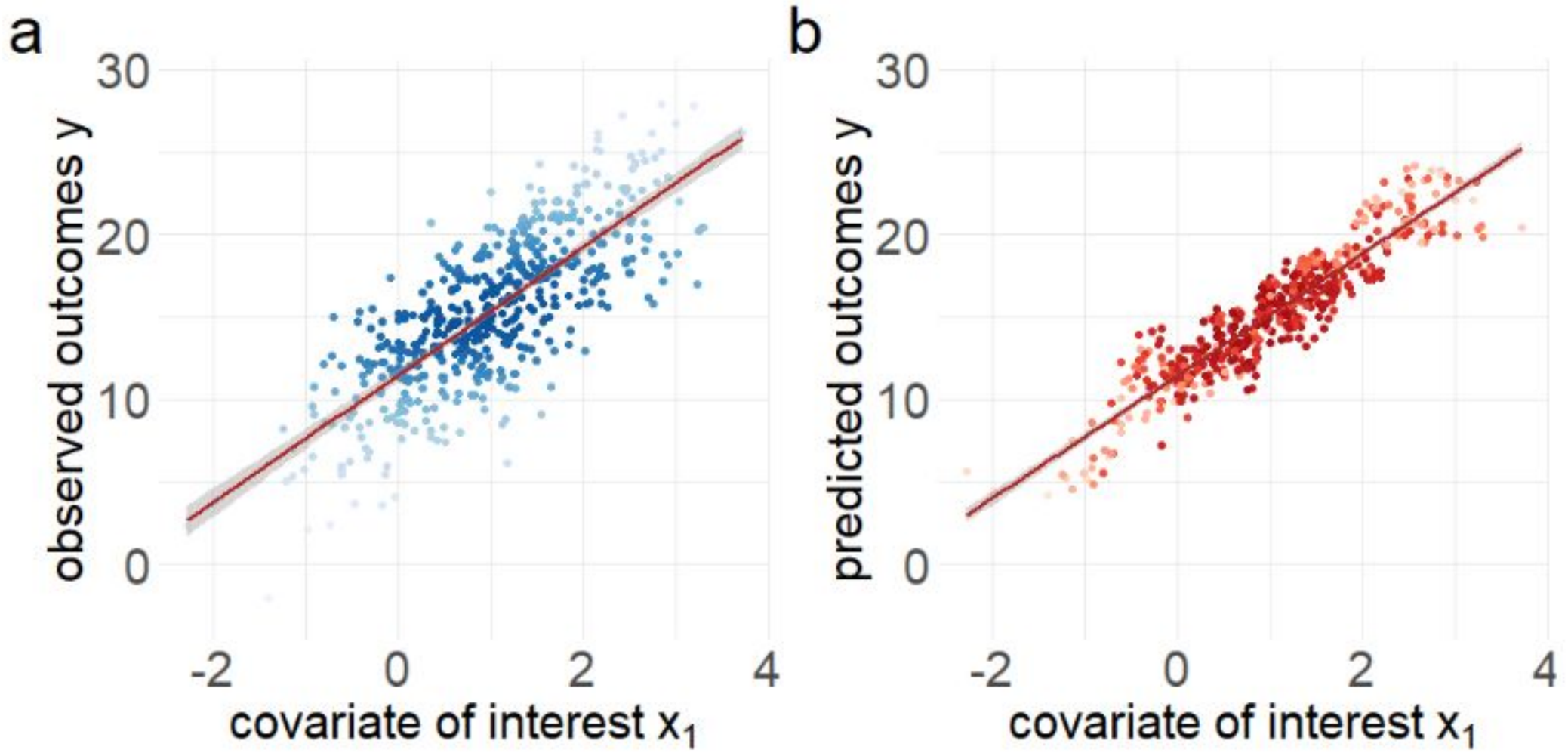
https://doi.org/10.1038/s41586-02
Received: 1 February 2020
Accepted: 10 August 2020
Check for updates

Table 1 | Essential hyperparameters for reproducing the study for each of the three models

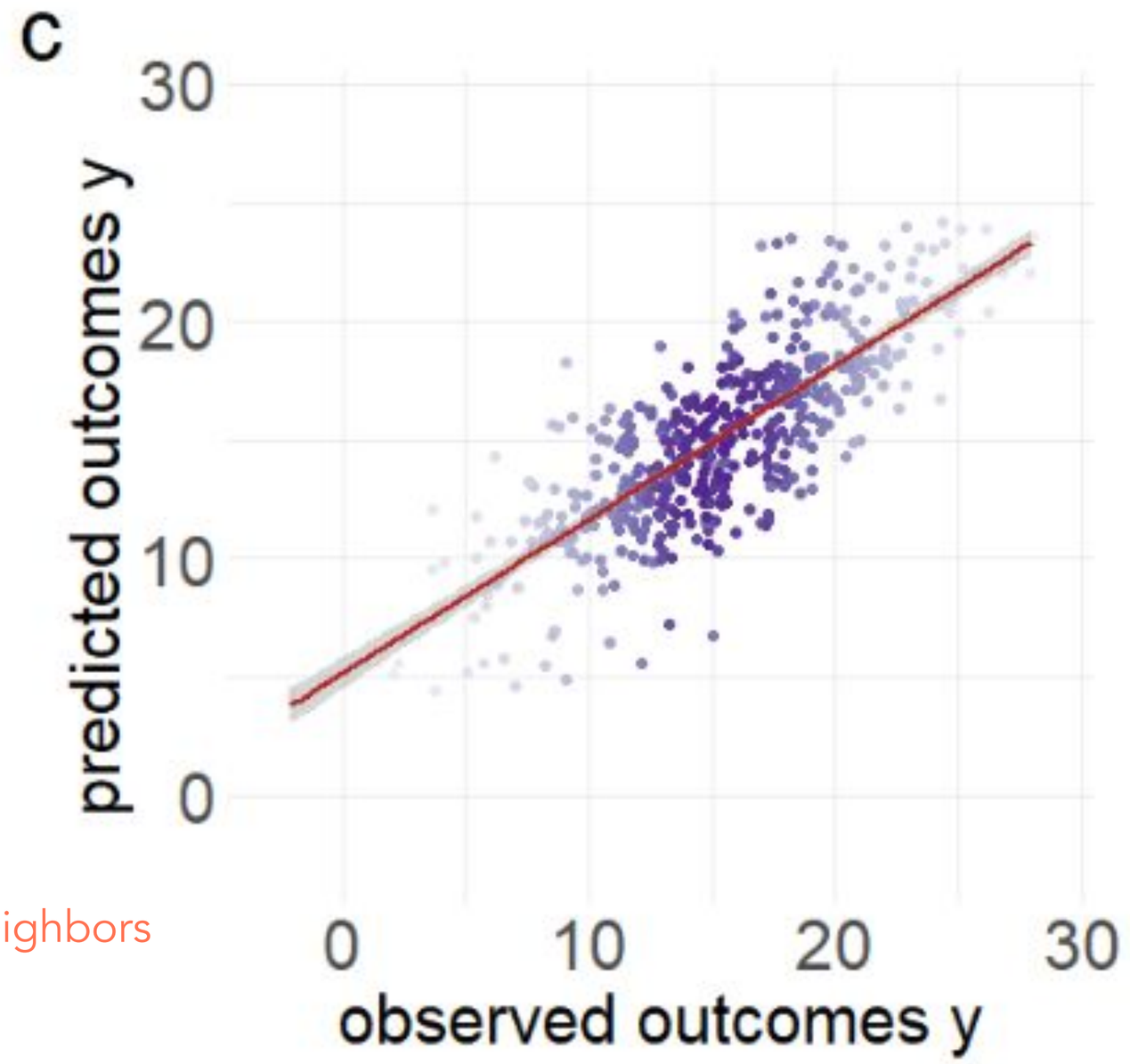
	Lesion	Breast	Case
Learning rate	Missing	0.0001	Missing
Learning rate schedule	Missing	Stated	Missing
Optimizer	Stochastic gradient descent with momentum	Adam	Missing
Momentum	Missing	Not applicable	Not applicable
Batch size	4	Unclear	2
Epochs	Missing	120,000	Missing

https://www.nature.com/articles/s41586-020-2766-y

A key observation

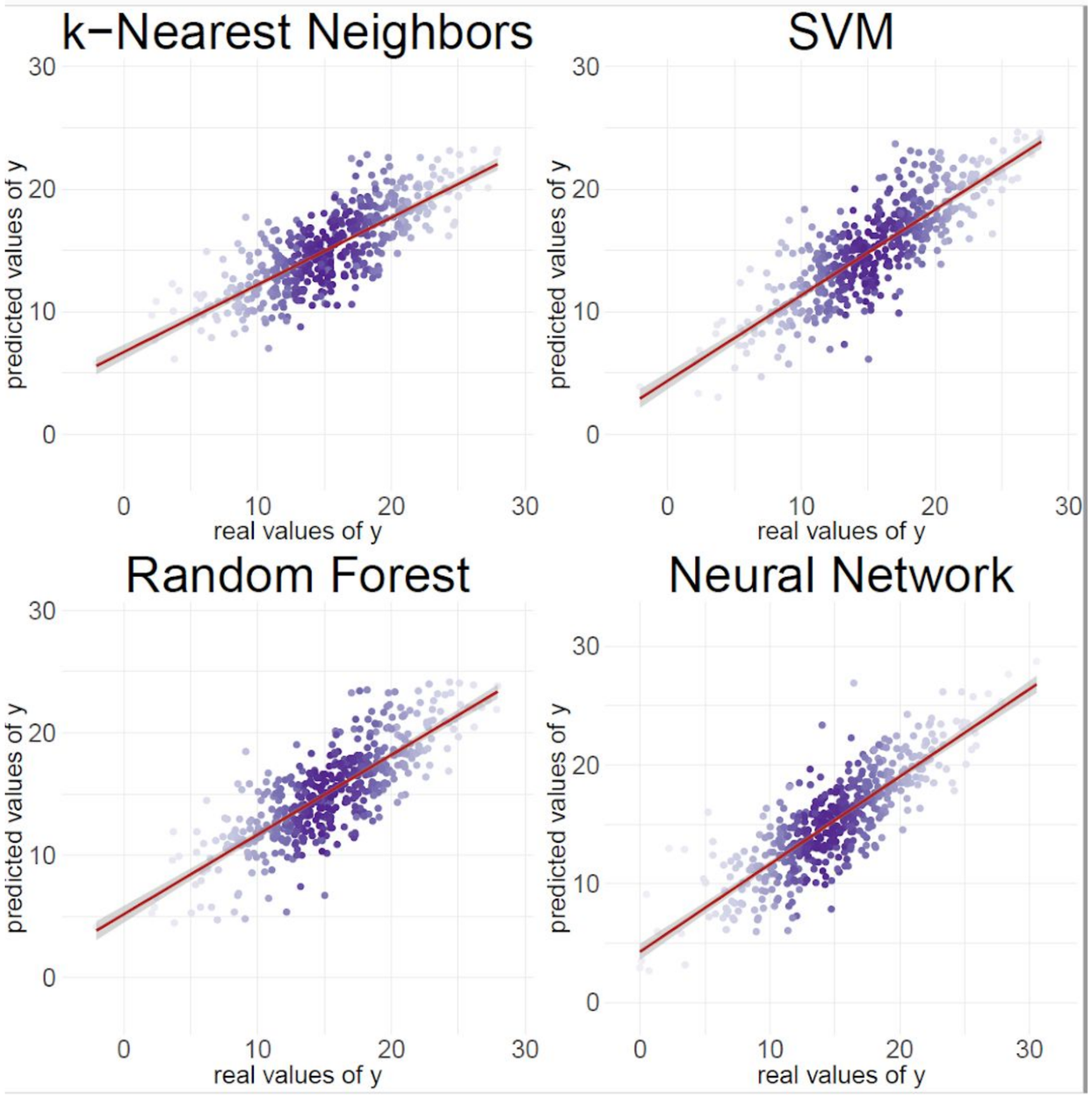


A key observation



Predicted with K-nearest neighbors

Doesn’t depend on the prediction model



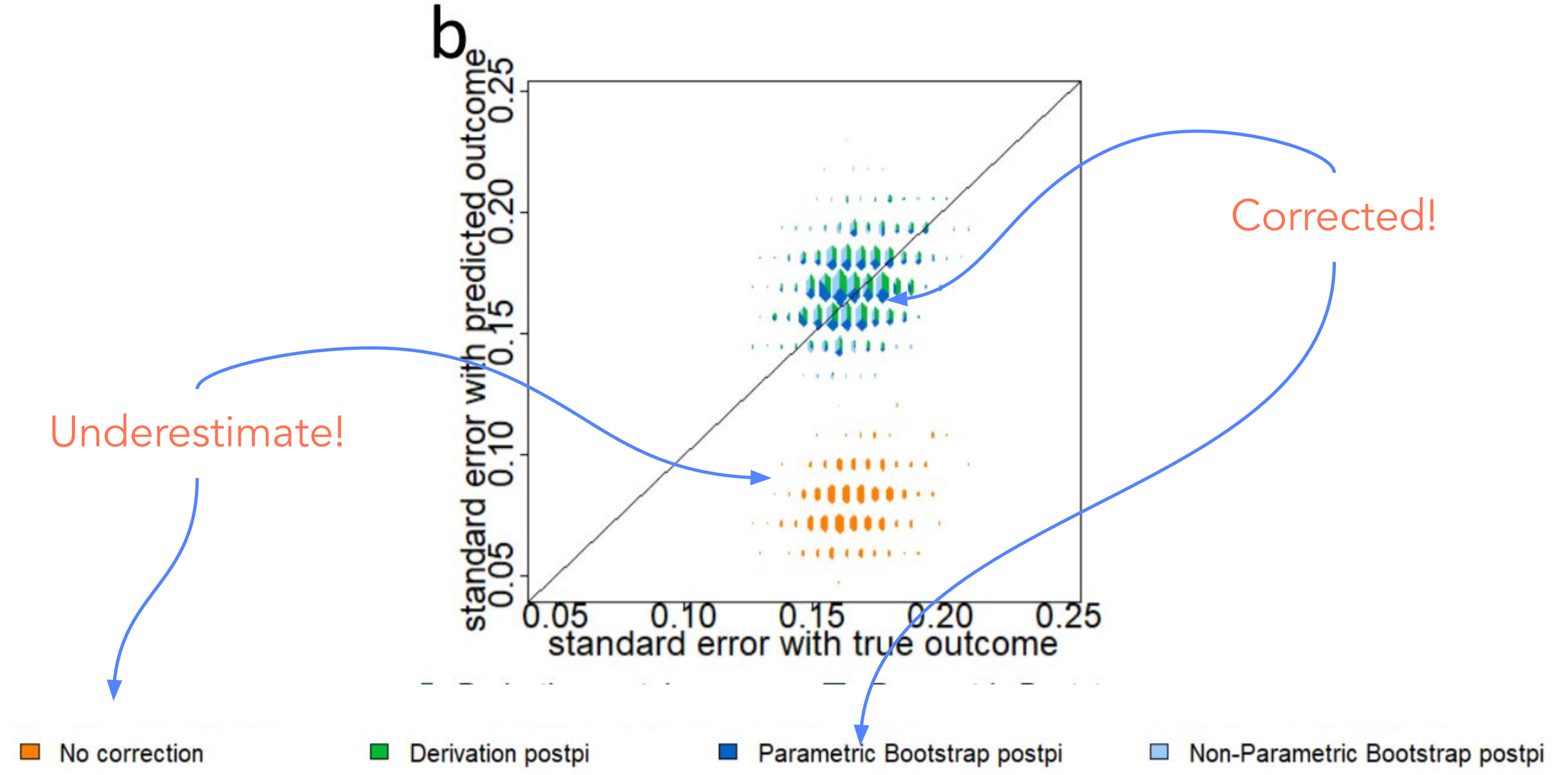
$$\tilde{E}[Y_p \mid X]$$

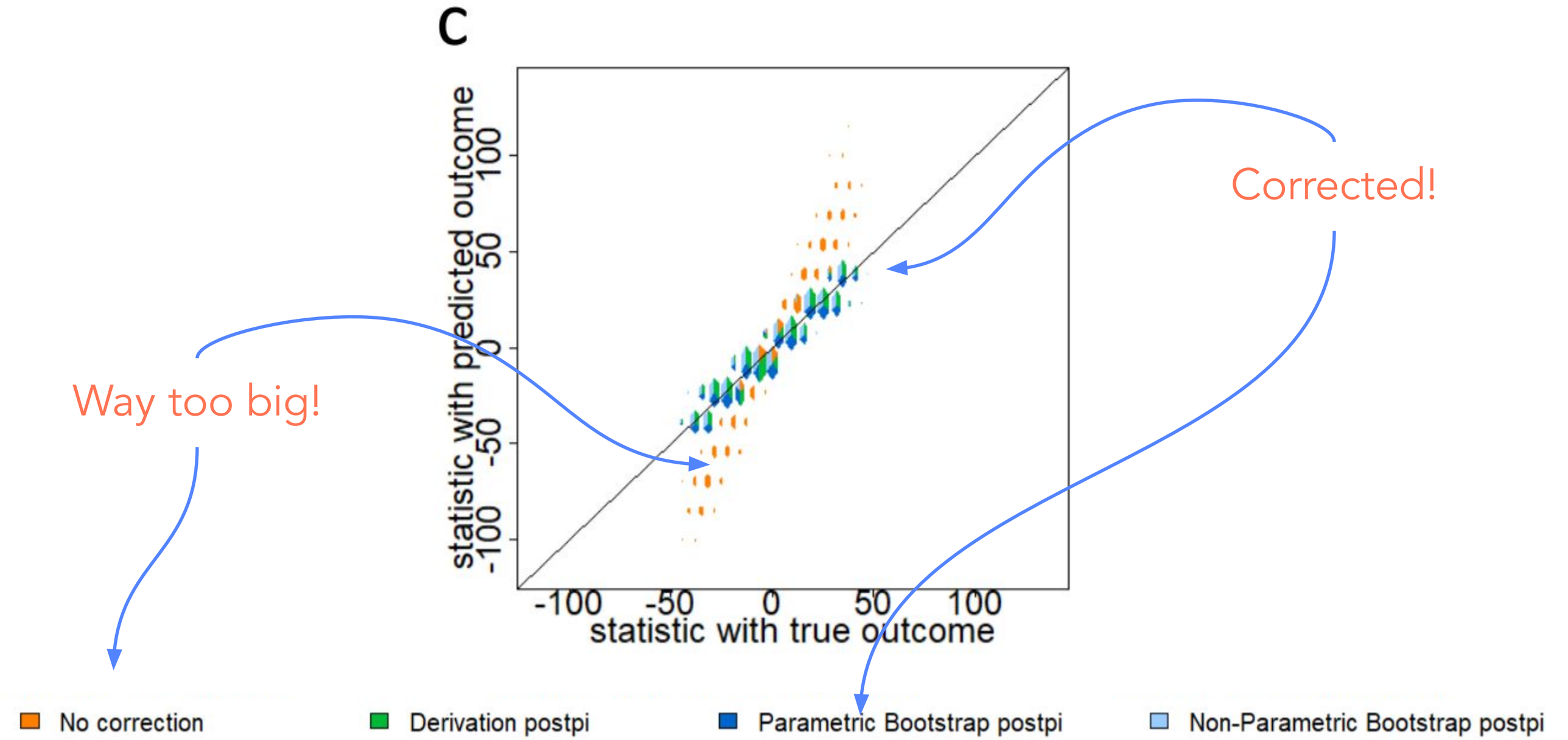
$$= \tilde{E}\{E[Y_p \mid Y, X] \mid X\}$$

$$\approx E\{E[Y_p \mid Y] \mid X\}$$

"Relationship model"

Simulation

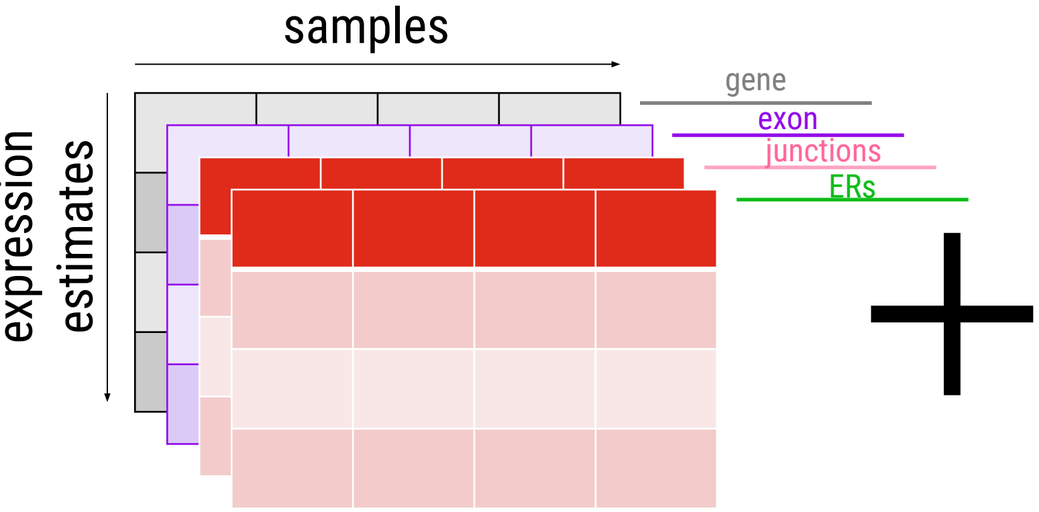

b
standard error with predicted outcome
0.05
0.10
0.15
0.20
0.25
standard error with true outcome
Underestimate!
Corrected!
No correction
Derivation postpi
Parametric Bootstrap postpi
Non-Parametric Bootstrap postpi


C
statistic with predicted outcome
-100
-50
0
50
100
statistic with true outcome
-100
-50
0
50
100
Way too big!
Corrected!
No correction
Derivation postpi
Parametric Bootstrap postpi
Non-Parametric Bootstrap postpi

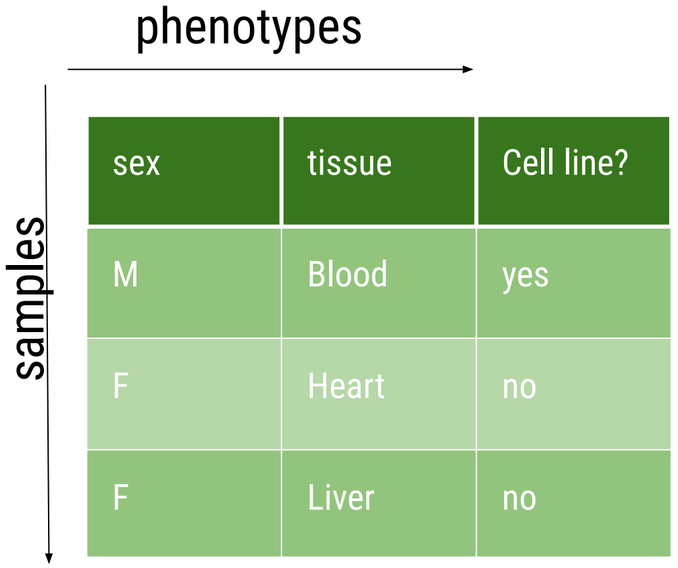
Back to recount2

recount2

expression data for ~70,000 human samples

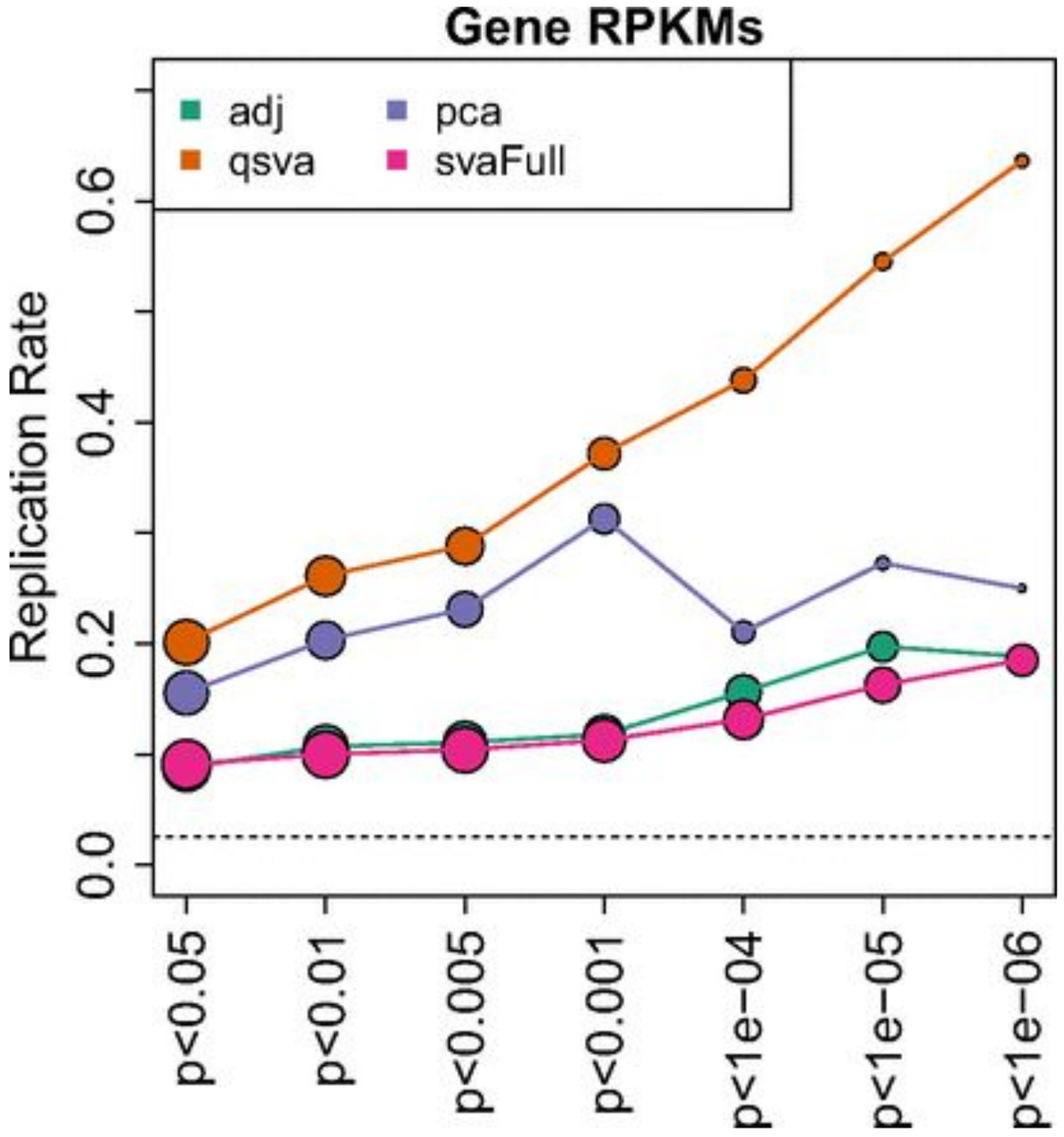


GTEx	SRA	TCGA
N=9,962	N=49,848	N=11,284



sex	tissue	Cell line?
M	Blood	yes
F	Heart	no
F	Liver	no

Answer meaningful questions about human biology and expression

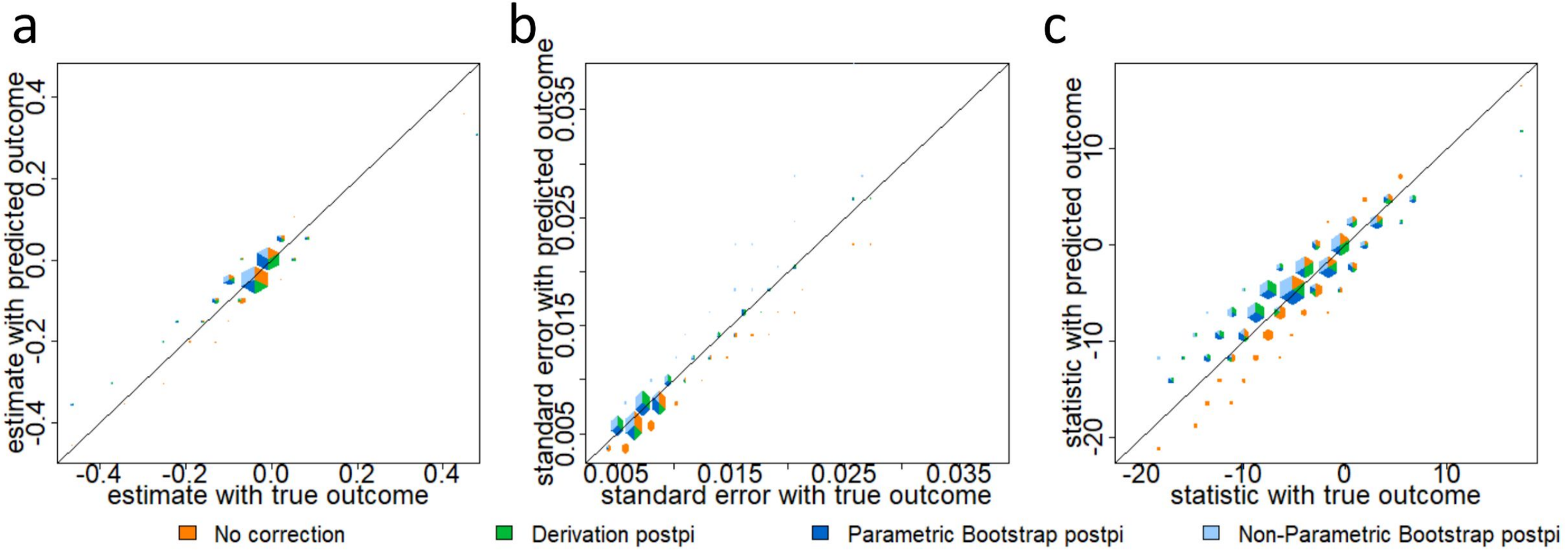


qSVA framework for RNA quality correction in differential expression analysis

Andrew E. Jaffe, Ran Tao, Alexis L. Norris, Marc Kealhofer, Abhinav Nellore, Joo Heon Shin, Dewey Kim, Yankai Jia, Thomas M. Hyde, Joel E. Kleinman, Richard E. Straub, Jeffrey T. Leek, and Daniel R. Weinberger

PNAS July 3, 2017 114 (27) 7130-7135; first published June 20, 2017 https://doi.org/10.1073/pnas.1617384114

Edited by Pasko Rakic, Yale University, New Haven, CT, and approved May 19, 2017 (received for review October 27,



https://academic.oup.com/nar/article/46/9/e54/4920847

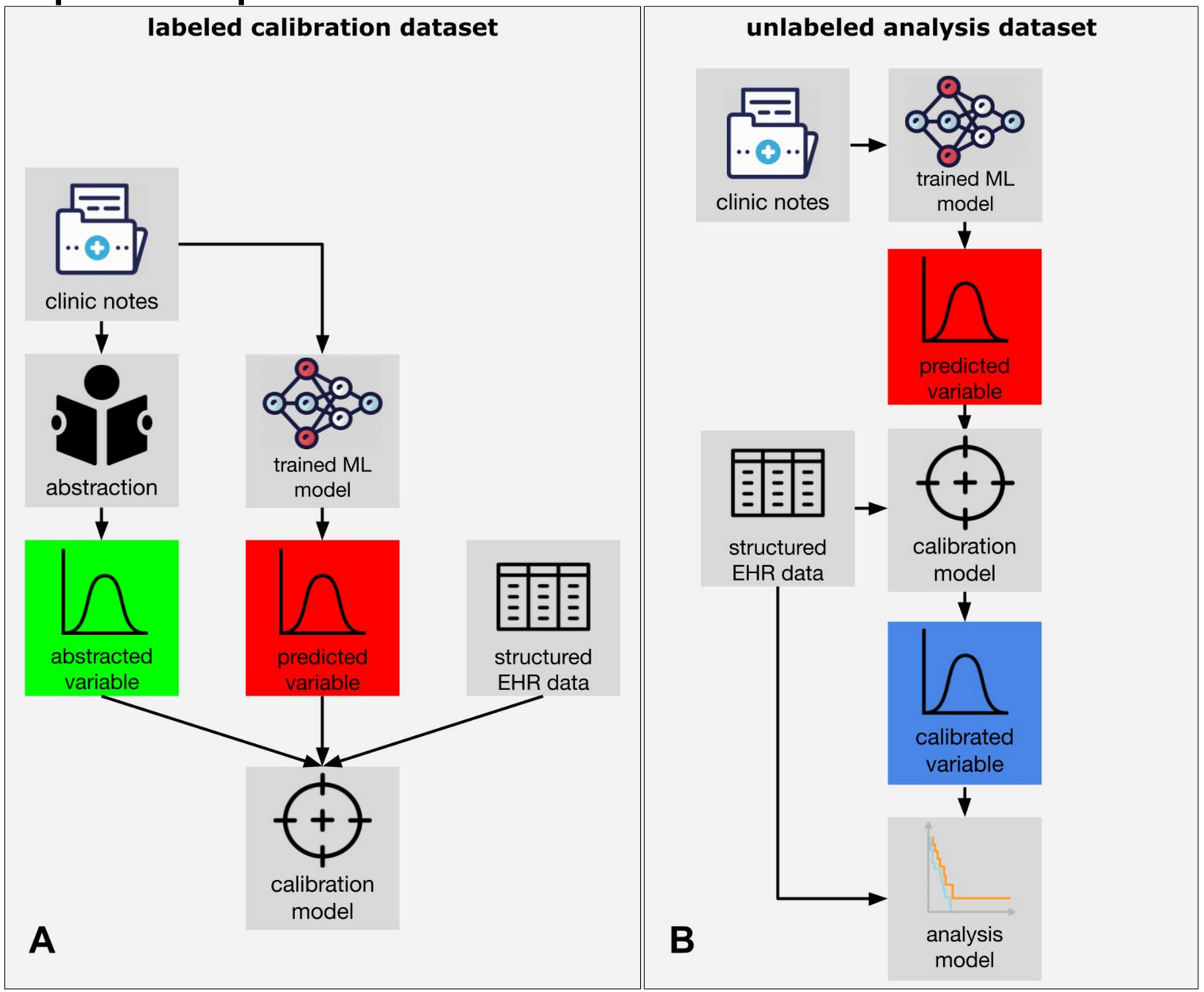
Post-prediction inference and RWE

Arjun Sondhi

Alex Rich

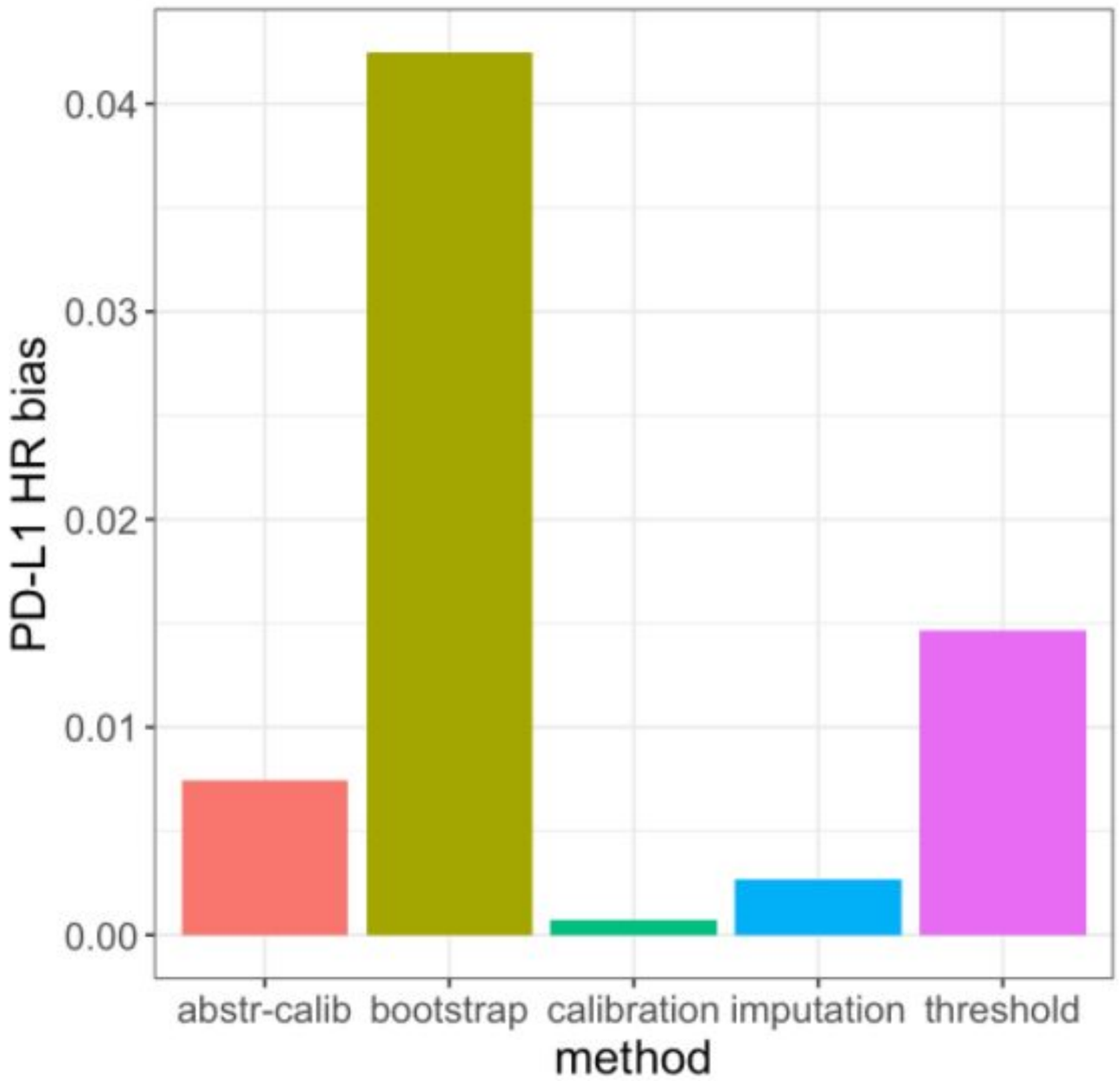
Sara Wang

How does post-prediction inference for RWE work?



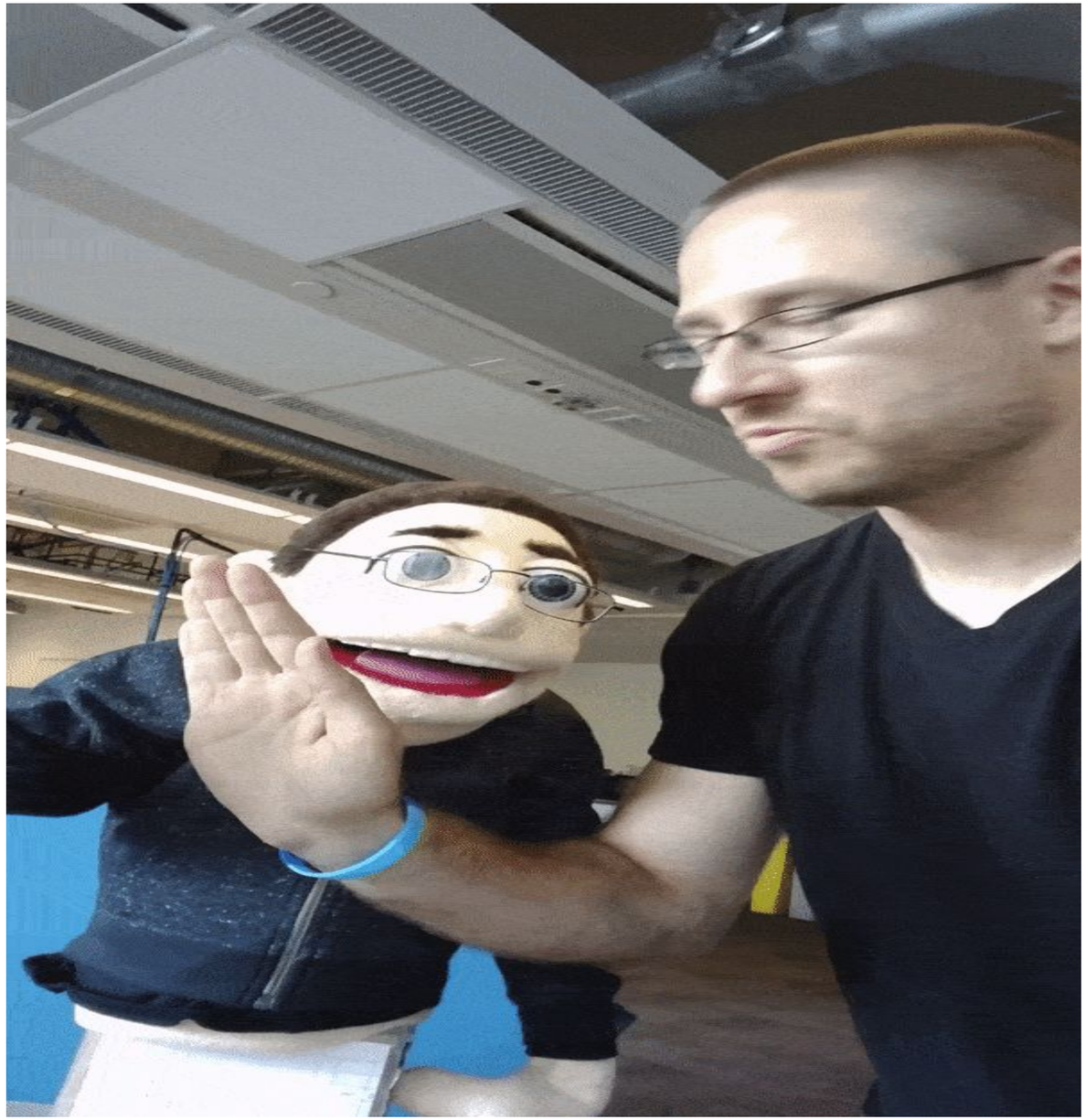
Sondhi et al. *in prep*

Calibration/Imputation can improve inference



Sondhi et al. *in prep*

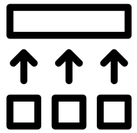
We can work with the machines!



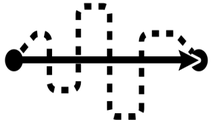
Thank you!

Tying it all together

With **thoughtfulness** and **methodologic rigor**, integrated evidence can increase…



Generalizability and power of RWD analyses



Efficiency of clinical trials

Depth and sophistication of RWD

Q&A

Please submit questions through the Q&A feature at the bottom of your screen.

Daniel Backenroth, PhD
Scientific Director
Janssen

David Paulucci, MS
Associate Director of Data Science
BMS

Katherine Tan, PhD
Senior Quantitative Scientist
Flatiron Health

Olivier Humblet, ScD
Moderator
Senior Quantitative Scientist
Flatiron Health

Sanhita Sengupta, PhD
Senior Manager, Data Science
BMS

Next on ResearchX

EP 05 | April 27

Life sciences case studies: Using RWE to support decision-making

EP 06 | May 11

Centering the patient’s voice: A discussion

flatiron® | ResearchX

rwe@flatiron.com