

Replicability in Science: 1: A Case Study

giovanni_parmigiani@dfci.harvard.edu

Padova, September 18, 2023

- Consulting Related to the Topic:
Martingale Labs, Delfi Diagnostics
- Speaker's Bureau: None
- Grant/Research support from: NIH-NCI, NSF
Licensing of BayesMendel software for genetic counseling.
Licensing of Ask2me database.
- Stockholder in: Phaeno Biotechnology
- Honoraria from: Academic Only
- co-Founder / Chief Scientific Officer: Phaeno Biotechnology
- Relevant patents: "Methods and systems for treatment of ovarian cancer" WO2014153442A2.
- Patents on diagnostic use of various genes
- Employee of: Dana Farber Cancer Institute

none of the analyses described involve licensed products

Reproducibility and Replicability in Science



An **ad hoc committee of the National Academies** of Sciences, Engineering, and Medicine explored the issues of reproducibility and replication in scientific and engineering research, focusing on defining reproducibility and replicability, and examining the extent of non-reproducibility and non-replicability.

Reproducibility and Replicability in Science



“We define reproducibility to mean computational reproducibility –obtaining consistent computational results using the same input data, computational steps, methods, and code, and conditions of analysis”.

Reproducibility and Replicability in Science



We define ... replicability to mean obtaining **consistent** results across **studies** aimed at **answering** the same scientific question, each of which has obtained its own **data**.

we need to define:

DATA

STUDY

ANSWER

CONSISTENCY

Today: A case study from precision medicine

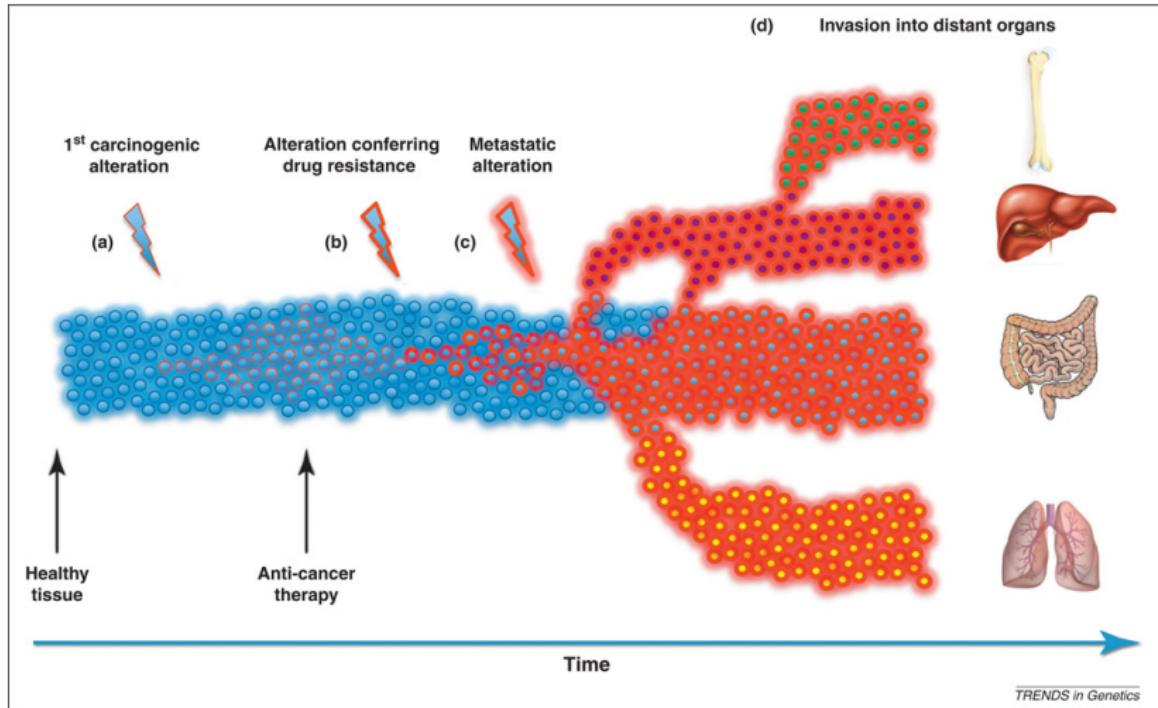
- Cancer Biomarkers
- High Throughput Gene Expression Measurements
- Comprehensive Review of Ovarian Cancer Studies
- Meta-analysis
- Exploring Replicability

Tomorrow: A closer look at the NAS report

- How to build collections of studies to address questions of replicability
- Aspects of replicability of parametric inference across studies and how it compares to meta-analysis
- Aspect of replicability of predictive inference across studies

Mind the gap!

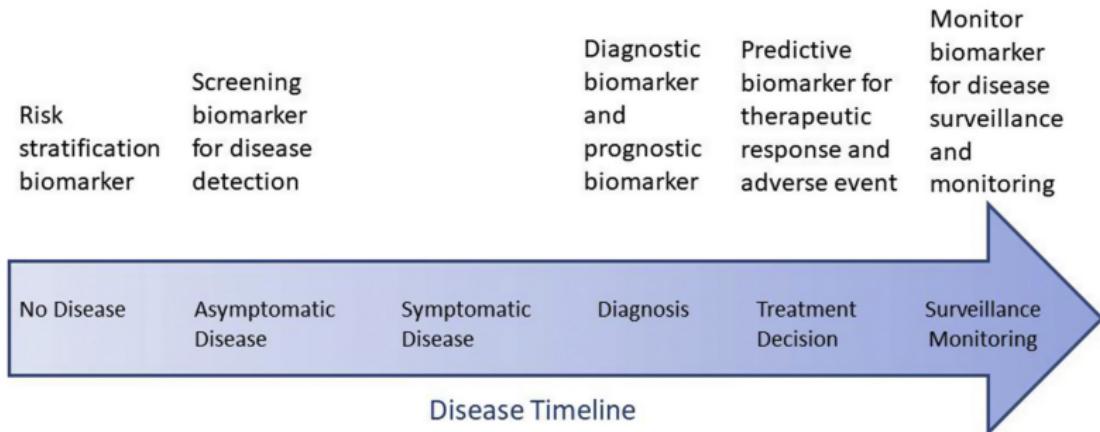
cancer evolution



biomarkers from a “system biology” perspective

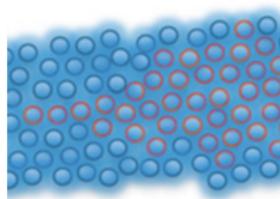
- Multiple, interconnected changes in cellular machinery override normal biological regulation, and lead to cells becoming neoplastic and invasive.
- Active genes, their respective products, and other organic chemicals made by the cell are identifiers that make up the **“molecular signature”** of a cell.
- Detection of such changes (e.g.“molecular signatures”) during initiation, transformation, and progression can be used for:
 - **Discovery** of both associations and mechanisms of evolution
 - **Development** of clinical management strategies

clinical roles of biomarkers in cancer



Ou et al 2021 [https://www.jto.org/article/S1556-0864\(21\)01663-4/fulltext](https://www.jto.org/article/S1556-0864(21)01663-4/fulltext)

high-throughput measurements in cancer



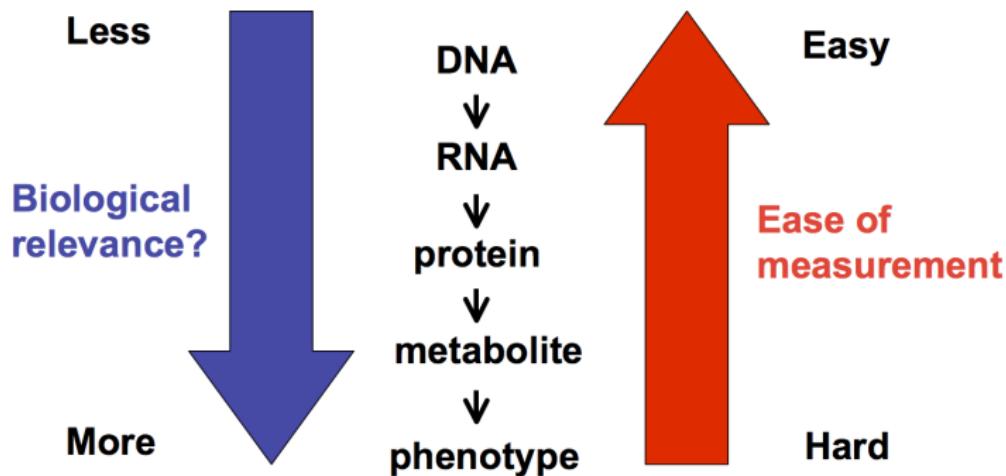
Changes in DNA compared to germline

DNA organization and accessibility

Levels of mRNA ("expression")

Level of Proteins and Metabolites

biological relevance and biotechnologies

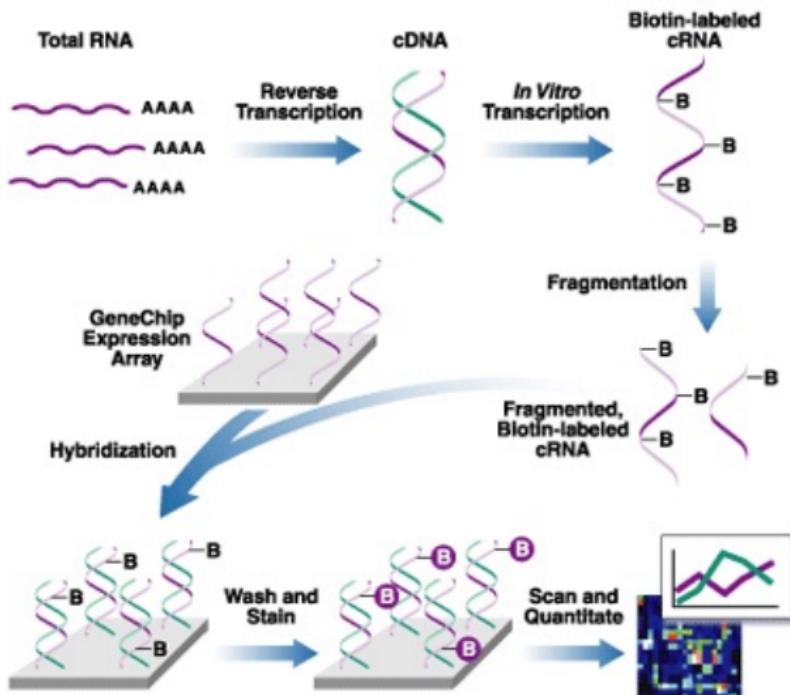


Transcription is a commonly studied form of cellular regulation

stages of gene expression measurement

Specimen Acquisition Preparation Protocols RNA or cDNA
quantification Raw Data Your Data

example: Affymetrix gene expression arrays



.CEL file format for Affymetrix arrays

probes	Symbols	Entrez_IDs	ma_0uM_r1
117_at	HSPA6	3310	2.2398
209552_at	PAX8	7849	5.1657
1255_g_at	GUCA1A	2978	2.1911
1316_at	THRA	7867	5.8759
226380_at	PTPN21	11099	8.3151
1405_i_at	CCL5	6352	2.1911
209975_at	CYP2E1	1571	2.2582
1438_at	EPHB3	2049	4.3058
203193_at	ESRRRA	2101	7.9473
1494_f_at	CYP2A6	1548	2.2238
1552256_a_at	SCARB1	949	7.8378
216251_s_at	TTLL12	23170	7.9864
1552258_at	LINC00152	112597	2.1911
203892_at	WFDC2	10406	8.4272
1552264_a_at	MAPK1	5594	6.4566
1552266_at	ADAM32	203102	2.2173
230763_at	SPATA17	128153	4.4470
1552271_at	PRR22	163154	2.1911
1552274_at	PXX	54899	3.2453
223346_at	VPS18	57617	6.5887
1552277_a_at	MSANTD3	91283	8.8748
1558703_at	SLC46A1	113235	6.9686
1552280_at	TIN04	91937	2.1911
1552281_at	SLC39A5	283375	3.6017
1552287_s_at	AFG3L1P	172	9.3857

Biomarker discovery: association analysis,(multiple) hypothesis testing

Biomarker-based Predictive Modeling: survival analysis, machine learning

Validation: Is a biomarker / prediction model adequate for a certain set of tasks;

Replicability: Are scientific results about a biomarker varies across multiple independent studies.

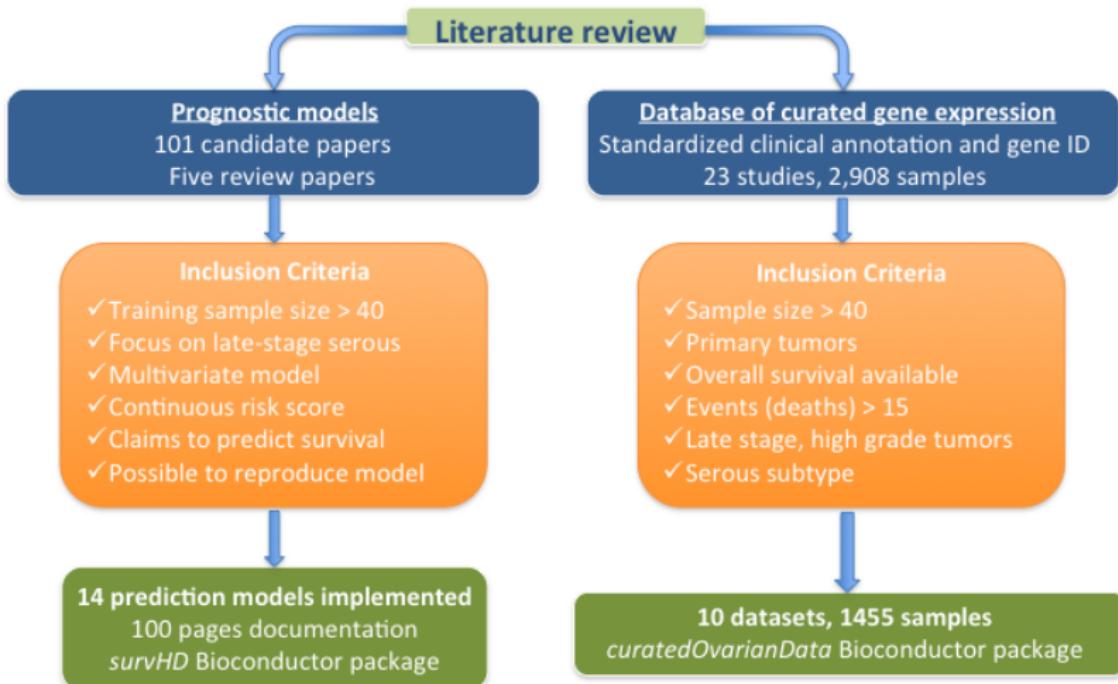
"Validating results".

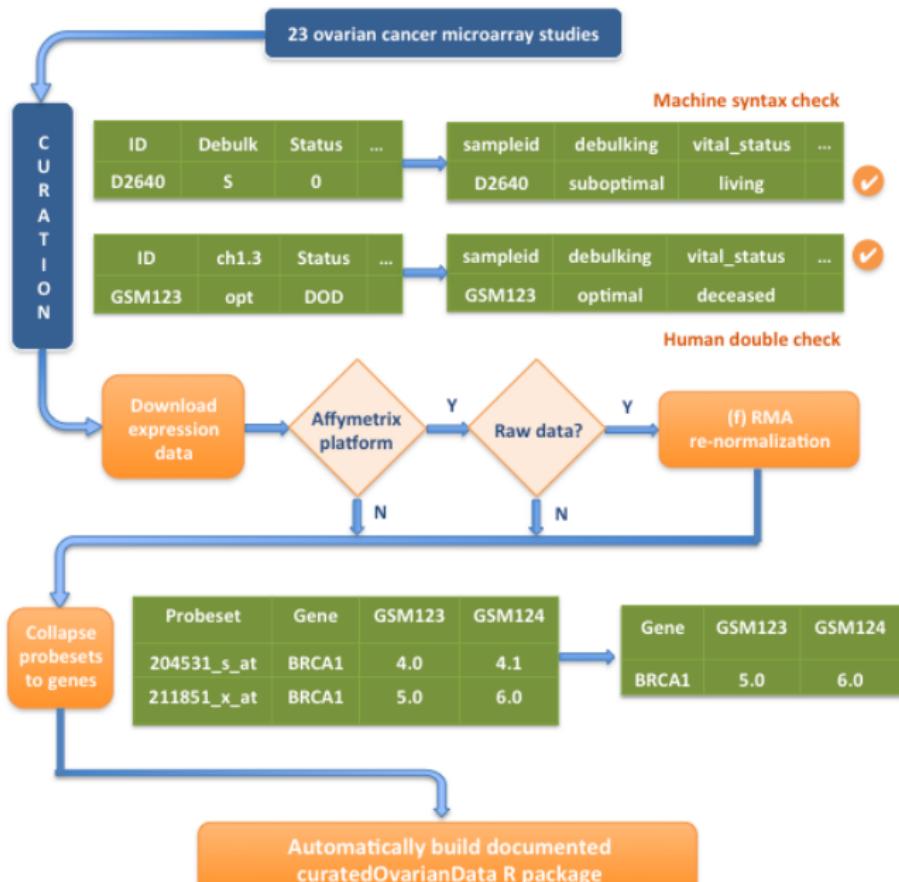
meta-analysis

Comprehensive Review: Collection of all relevant publications and dataset based on preset inclusion criteria

Meta-analysis: Joint analysis of the above. It can be based on summaries or "raw" data.

Meta-analysis overview





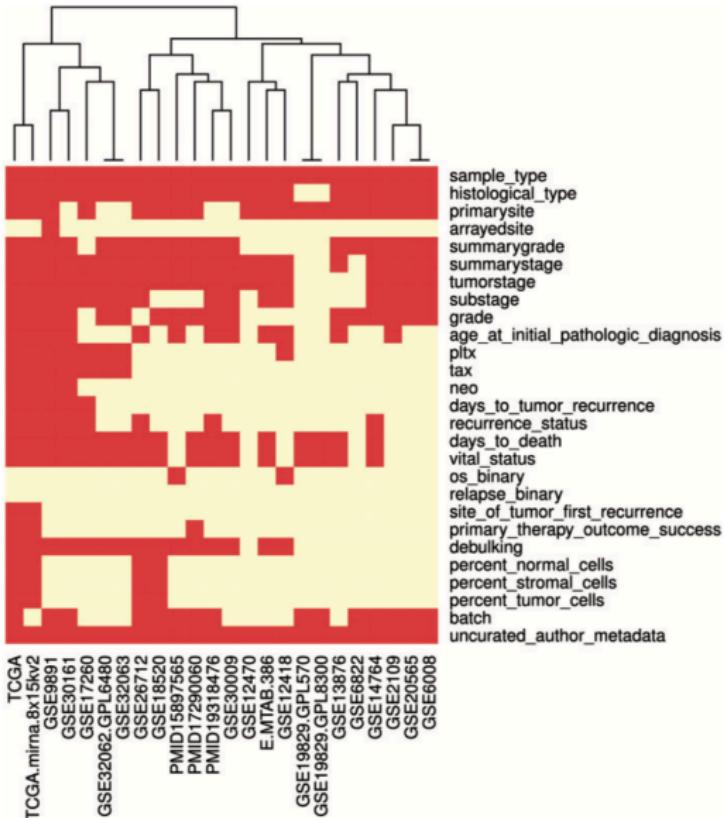
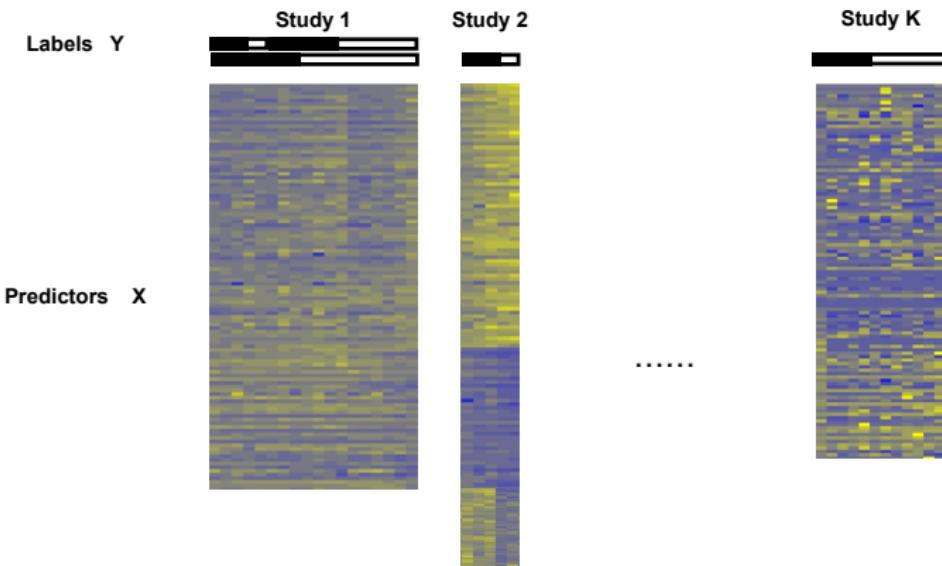


Figure 2. Available clinical annotation. This heatmap visualizes for each curated clinical characteristic (rows) the availability in each data set (columns). Red indicates that the corresponding characteristic is available for at least one sample in the data set. See Table 2 for descriptions of these characteristics.

data structure



meta-analysis and replicability: CXCL12

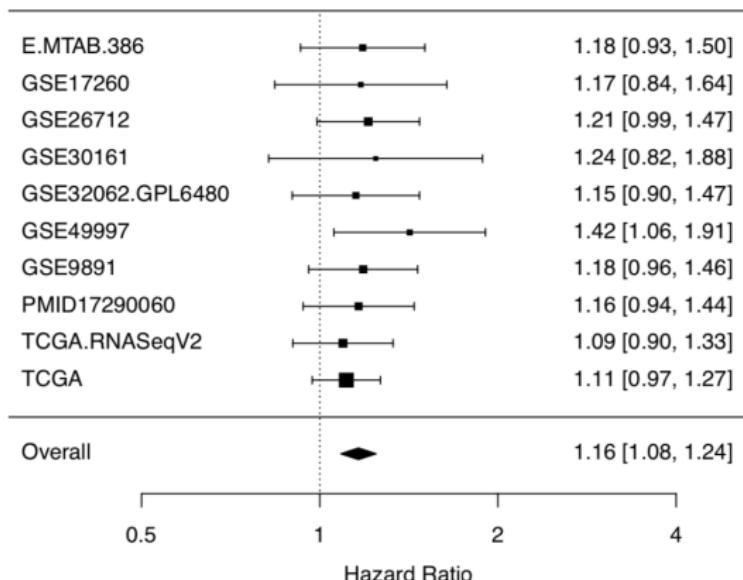


Figure 2: Validation of CXCL12 as an independent predictor of survival

This figure shows a forest plot as in Figure 1, but the CXCL12 expression levels were adjusted for debulking status (optimal versus suboptimal) and tumor stage. The p-value for the overall HR, found in `res$pval`, is `1.8e-05`.

meta-analysis and replicability: CXCL12 10 years after

Google Scholar CXCL12 ovarian cancer 

Articles About 2,500 results (0.08 sec)

Any time Try Harvard Library

Since 2020

Since 2019

Since 2016

Custom range...

Sort by relevance

Sort by date

include patents

include citations

 Create alert

CXCL12 promotes human ovarian cancer cell invasion through suppressing ARHGAP10 expression N Luo, D Chen, L Liu, L Li, Z Cheng - Biochemical and biophysical ..., 2019 - Elsevier
The CXCL12/CXCR4 axis is strongly implicated as key determinant of tumor invasion and metastasis in ovarian cancer. However, little is known about the potential downstream signals of the CXCL12/CXCR4 axis that contribute to ovarian cancer cell invasion and ...
 99 Cited by 1 Related articles All 3 versions

CXCR7 is not obligatory for CXCL12-CXCR4-induced epithelial-mesenchymal transition in human ovarian cancer N Zheng, W Liu, J Chen, B Li, J Liu... - Molecular ..., 2019 - Wiley Online Library
Although the CXCL12-CXCR4/CXCR7 chemokine axis is demonstrated to play an integral role in tumor progression, the controversy exists and the role of CXCL12-CXCR4/CXCR7 signaling axis in epithelial-mesenchymal transition (EMT) of human ovarian cancer has not ...
 99 Cited by 5 Related articles All 3 versions

Role of CXCL12-CXCR4 axis in ovarian cancer metastasis and CXCL12-CXCR4 blockade with AMD3100 suppresses tumor cell migration and invasion in vitro Y Liu, CC Ren, L Yang, YM Xu... - Journal of cellular ..., 2019 - Wiley Online Library
Ovarian cancer (OC) is a lethal gynecologic tumor, which brings its mortality to the head. CXCL12 and its receptor chemokine receptor 4 (CXCR4) have been found to be highly expressed in OC and contribute to the disease progression by affecting tumor cell ...
 99 Cited by 9 Related articles All 3 versions

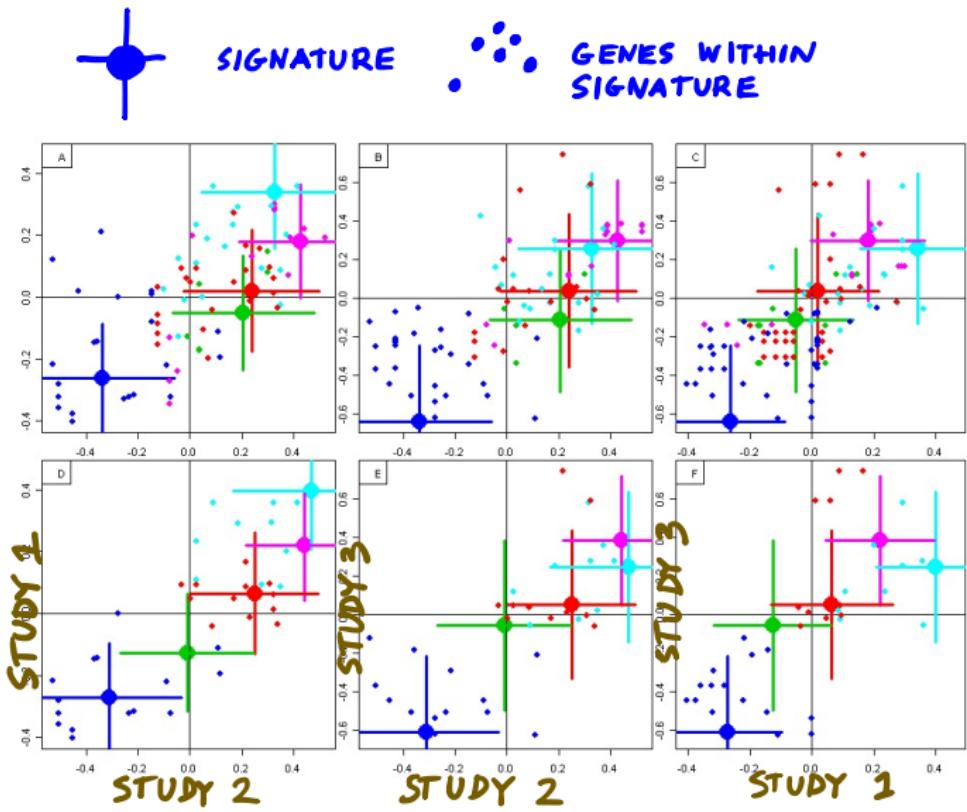
Dual blockade of CXCL12-CXCR4 and PD-1-PD-L1 pathways prolongs survival of ovarian tumor-bearing mice by prevention of immunosuppression in the tumor ... Y Zeng, B Li, Y Liang, PM Reeves, X Qu, C Ran... - The FASEB ..., 2019 - fasebj.org
... not only increased tumor apoptosis and necrosis by blockade of CXCL12 pathway but ... antitumor immune responses in immunocompetent murine tumor models of ovarian cancer (15) and ... groups have also explored antitumor effects of AMD3100 on various cancers that express ...
 99 Cited by 13 Related articles All 4 versions

[PDF] wiley.com Try Harvard Library

[PDF] fasebj.org Try Harvard Library

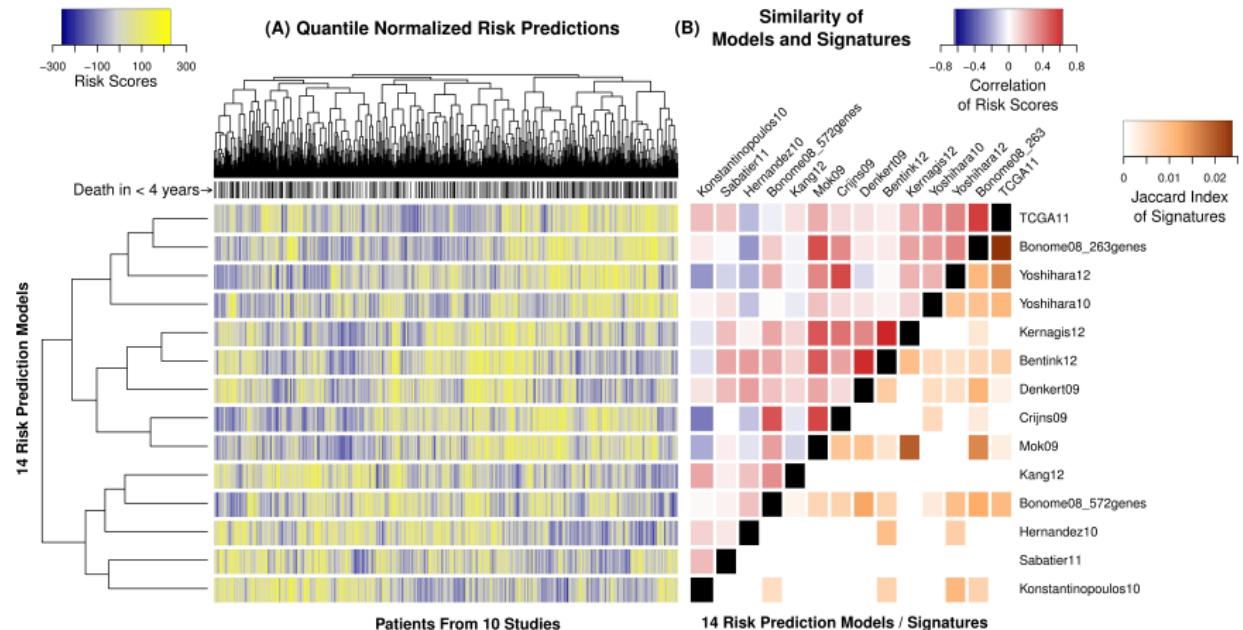
what is a study?

FILTERED UNFILTERED



SCALE : PREDICTION STRENGTH (COX EFFECT SIZE)

replicability of variable selection and individual scores



classifier cross-study validation matrix

Waldron et al 2014

Implemented Models

Validation Statistics for 14 Models in 10 Datasets

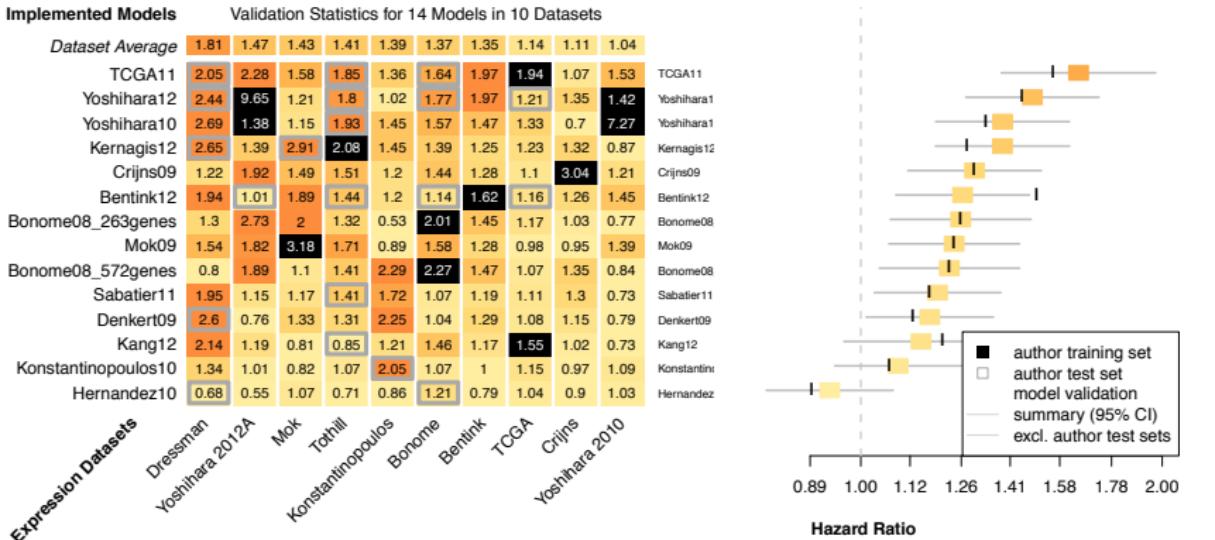
	Dataset Average	1.81	1.47	1.43	1.41	1.39	1.37	1.35	1.14	1.11	1.04	
TCGA11	2.05	2.28	1.58	1.85	1.36	1.64	1.97	1.94	1.07	1.07	1.53	TCGA11
Yoshihara12	2.44	9.65	1.21	1.8	1.02	1.77	1.97	1.21	1.35	1.35	1.42	Yoshihara1
Yoshihara10	2.69	1.38	1.15	1.93	1.45	1.57	1.47	1.33	0.7	0.7	7.27	Yoshihara1
Kernagis12	2.65	1.39	2.91	2.08	1.45	1.39	1.25	1.23	1.32	1.32	0.87	Kernagis12
Crijns09	1.22	1.92	1.49	1.51	1.2	1.44	1.28	1.1	3.04	3.04	1.21	Crijns09
Bentink12	1.94	1.01	1.89	1.44	1.2	1.14	1.62	1.16	1.26	1.26	1.45	Bentink12
Bonomo08_263genes	1.3	2.73	2	1.32	0.53	2.01	1.45	1.17	1.03	1.03	0.77	Bonomo08
Mok09	1.54	1.82	3.18	1.71	0.89	1.58	1.28	0.98	0.95	0.95	1.39	Mok09
Bonomo08_572genes	0.8	1.89	1.1	1.41	2.29	2.27	1.47	1.07	1.35	1.35	0.84	Bonomo08
Sabatier11	1.95	1.15	1.17	1.41	1.72	1.07	1.19	1.11	1.3	1.3	0.73	Sabatier11
Denkert09	2.6	0.76	1.33	1.31	2.25	1.04	1.29	1.08	1.15	1.15	0.79	Denkert09
Kang12	2.14	1.19	0.81	0.85	1.21	1.46	1.17	1.55	1.02	1.02	0.73	Kang12
Konstantinopoulos10	1.34	1.01	0.82	1.07	2.05	1.07	1	1.15	0.97	0.97	1.09	Konstantinopoulos10
Hernandez10	0.68	0.55	1.07	0.71	0.86	1.21	0.79	1.04	0.9	0.9	1.03	Hernandez10

Expression Datasets

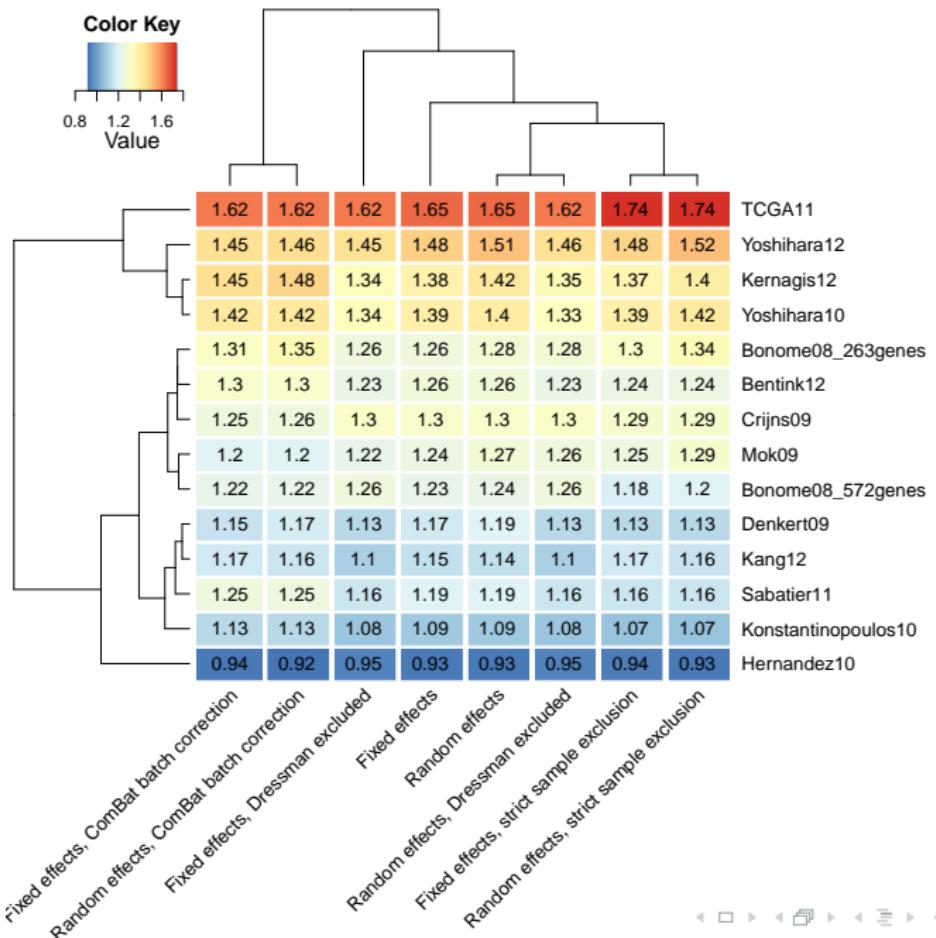
Dressman
Yoshihara 2012A
Mok
Tothill
Konstantinopoulos
Bonomo
Bentink
TCGA
Crijns
Yoshihara 2010

classifier cross-study validation

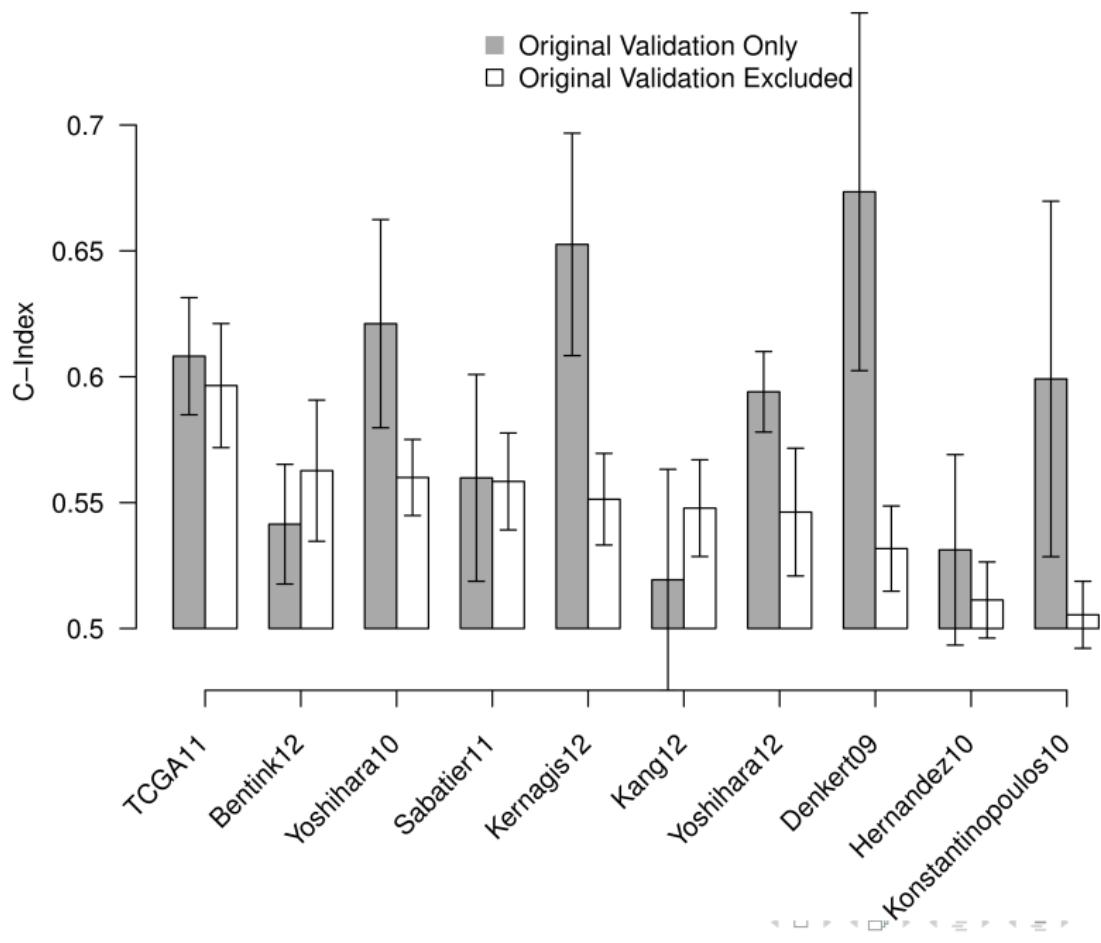
Waldron et al JNCI 2014



sensitivity analysis

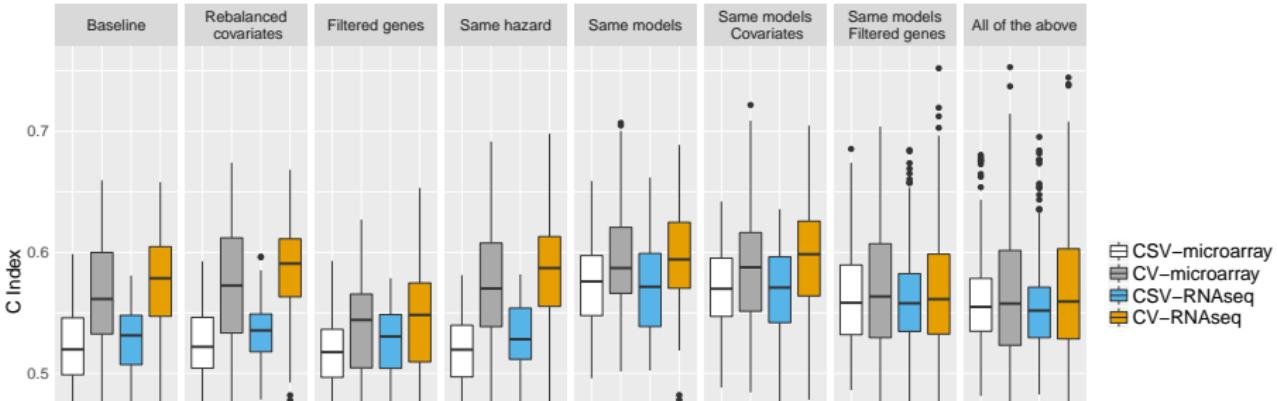


selection bias in choice of validation study?



determinants of CV/CSV gap

Zhang et al Biostatistics 2018



nas replicability: did we learn anything?

We define ... replicability to mean

obtaining **consistent** results

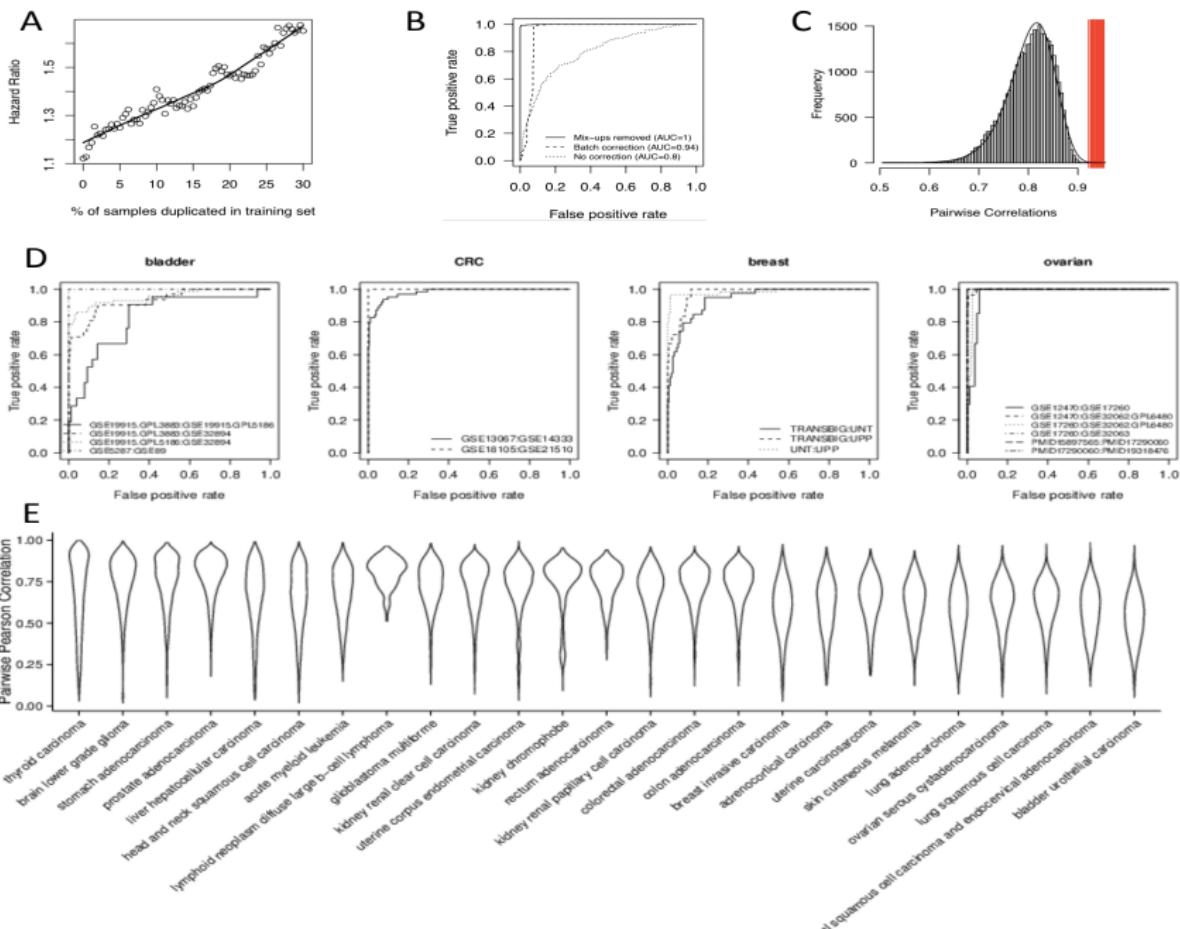
across **studies**

aimed at **answering** the same scientific question,
each of which has obtained its own **data**

credits and references

- Ganzfried, B.F., Riester, M., Haibe-Kains, B., Risch, T., Tyekucheva, S., Jazic, I., Wang, X.V., Ahmadifar, M., Birrer, M.J., Parmigiani, G., Huttenhower, C., Waldron, L., 2013. curatedOvarianData: clinically annotated data for the ovarian cancer transcriptome. Database 2013, bat013–bat013.
- Waldron, L., Haibe-Kains, B., Culhane, A.C., Riester, M., Ding, J., Wang, X.V., Ahmadifar, M., Tyekucheva, S., Bernau, C., Risch, T., Ganzfried, B.F., Huttenhower, C., Birrer, M., Parmigiani, G., 2014. Comparative meta-analysis of prognostic gene signatures for late-stage ovarian cancer. *JNCI* 106, dju049–dju049.
- Zhao, S.D., Parmigiani, G., Huttenhower, C., Waldron, L., 2014. Más-o-menos: a simple sign averaging method for discrimination in genomic data analysis. *Bioinformatics* 30, 3062–3069.
- Bernau, C., Riester, M., Boulesteix, A.-L., Parmigiani, G., Huttenhower, C., Waldron, L., Trippa, L., 2014. Cross-study validation for the assessment of prediction algorithms. *Bioinformatics* 30, i105–12.
- Waldron, L., Riester, M., Ramos, M., Parmigiani, G., Birrer, M., 2016. The Doppelgänger Effect: Hidden Duplicates in Databases of Transcriptome Profiles. *JNCI* 108, djw146.
- Zhang Y, Bernau C, Parmigiani G, Waldron L. The impact of different sources of heterogeneity on loss of accuracy from genomic prediction models. *Biostatistics*. 2018;6(e1002358):701.

DIGRESSIONS



aspects of reproducibility

Waldron et al 2014

Model	Reproducibility†			
	Model Provided	Training Data Available	Validation Data Available	Verified Implementation
TCGA11 [12]	YES	YES	YES	YES
Denkert09 [13]	YES	YES	YES	YES
Bonomo08_263genes [14]	YES	YES	YES	YES
Bonomo08_572genes [14]	YES	YES	YES	YES
Mok09 [15]	NO	YES	YES	PARTIALLY
Yoshihara12 [16]	YES	-	YES	YES
Yoshihara10 [17]	YES	-	YES	YES
Bentink12 [18]	YES	-	YES	YES
Kang12 [19]	YES	YES	YES	PARTIALLY
Crijns09 [20]	NO	YES	NO	NO
Kernagis12 [21]	PARTIALLY	YES	YES	PARTIALLY
Sabatier11 [22]	PARTIALLY	NO	NO	NO
Konstantinopoulos10 [23]	YES	-	YES	PARTIALLY
Hernandez10 [24]	PARTIALLY	-	YES	PARTIALLY

Compare to Ioannidis 08:

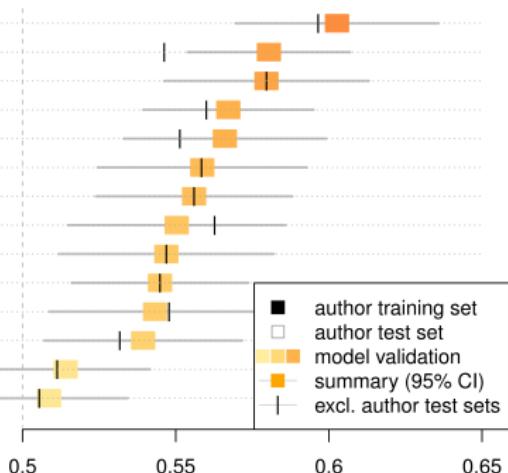
"We reproduced two analyses in principle and six partially or with some discrepancies; ten could not be reproduced."

using c-stat instead

(A) Implemented Models Validation Statistics for 14 Models in 10 Datasets

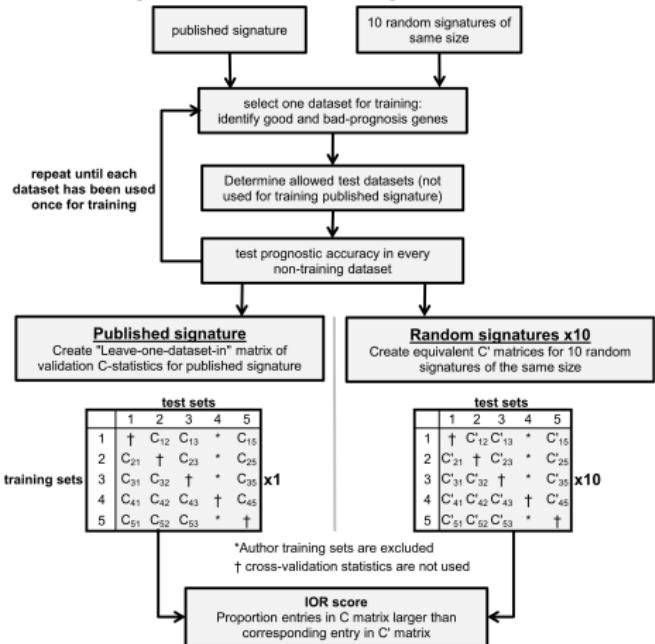
Dataset	Average	0.61	0.58	0.57	0.56	0.56	0.55	0.55	0.54	0.54	0.53
TCGA11		0.62	0.69	0.6	0.63	0.61	0.47	0.57	0.6	0.64	0.55
Yoshihara12		0.63	0.81	0.64	0.6	0.62	0.51	0.5	0.58	0.57	0.55
Bonome08_263genes		0.57	0.68	0.58	0.6	0.62	0.53	0.6	0.54	0.56	0.52
Yoshihara10		0.7	0.55	0.62	0.53	0.55	0.53	0.54	0.8	0.56	0.52
Kernagis12		0.66	0.58	0.63	0.56	0.55	0.55	0.65	0.57	0.55	0.54
Sabatier11		0.64	0.54	0.56	0.57	0.54	0.62	0.55	0.57	0.56	0.52
Crijns09		0.5	0.6	0.59	0.55	0.58	0.55	0.56	0.47	0.54	0.67
Bentink12		0.65	0.56	0.55	0.61	0.55	0.57	0.57	0.53	0.53	0.52
Bonome08_572genes		0.57	0.6	0.54	0.55	0.64	0.63	0.55	0.5	0.53	0.54
	Mok09	0.53	0.6	0.56	0.57	0.57	0.53	0.69	0.57	0.51	0.51
Kang12		0.63	0.54	0.52	0.54	0.57	0.54	0.49	0.54	0.58	0.52
Denkert09		0.67	0.52	0.54	0.53	0.53	0.58	0.53	0.51	0.52	0.55
Hernandez10		0.56	0.61	0.56	0.54	0.53	0.5	0.5	0.54	0.49	0.51
Konstantinopoulos10		0.57	0.5	0.52	0.48	0.49	0.6	0.5	0.51	0.53	0.5
Expression Datasets	Dressman	Yoshihara 2012A	Tothill	Bentink	Bonome	Konstantinopoulos	Mok	Yoshihara 2010	TCGA	Crijns	

(B) Summary Validation Statistics

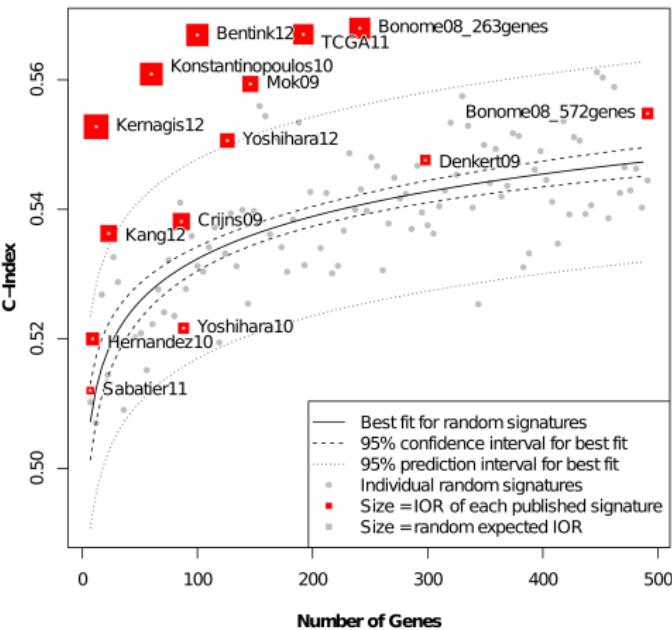


signatures

A) Methodology for comparing prognostic quality of gene sets to random gene sets



B) Gene set Improvement over Random Signatures



Generates collections of studies

Within and across study variation is closely matching empirical collections based on comprehensive reviews.

Major Sources of Bias in DNA prep

- Fragmentation of chromatin : Bias due to more efficient sonication of euchromatin (not so dense) than heterochromatin (more dense)
- Size selection: Bias due to heating agarose gel slices in chaotropic salt buffer
- PCR: Bias due to preferential amplification of templates with neutral GC%

<https://doi.org/10.1016/j.yexcr.2014.01.008>

Additional Sources of Bias in RNA prep

- RNA extraction using Trizol: selective loss of GC poor or highly structured small RNAs at low RNA concentrations
- Ribosomal RNA (rRNA) depletion/ mRNA enrichment: bias due to exonuclease targeting partially degraded mRNAs
- RNA fragmentation by RNase III: not completely random, leading to reduced complexity
- Random hexamer priming bias: not completely random
- Reverse transcription: antisense artefacts
- Adapter ligation bias due to substrate preferences of T4 RNA ligases
- Reduced ligation efficiency due to RNA modifications