

The Importance of Reproducibility in High-Throughput Biology: Case Studies in Forensic Bioinformatics

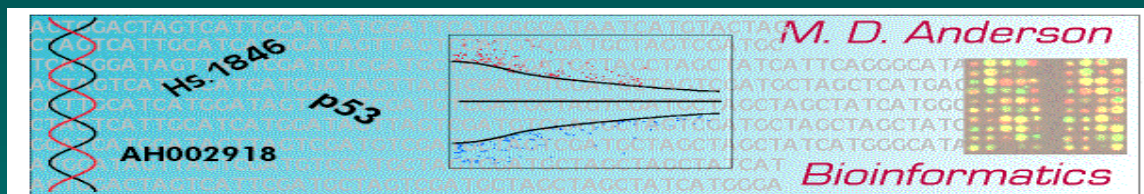
Keith A. Baggerly

Bioinformatics and Computational Biology

UT M. D. Anderson Cancer Center

kabagg@mdanderson.org

UNMC COPH Grand Rounds, Feb 15, 2012



Why is Reproducibility Important in H-T-B?

Our intuition about what “makes sense” is very poor in high dimensions. To use “genomic signatures” as biomarkers, we need to know they’ve been assembled correctly.

Without documentation, we may need to employ *forensic bioinformatics* to infer what was done to obtain the results.

Let’s examine some case studies involving an important clinical problem: *can we predict how a given patient will respond to available chemotherapeutics?*

Using Cell Lines to Predict Sensitivity

nature.com/naturemedicine

Genomic signatures to guide the use of chemotherapeutics

Anil Potti^{1,2}, Holly K Dressman^{1,3}, Andrea Bild^{1,3}, Richard F Riedel^{1,2}, Gina Chan⁴, Robyn Sayer⁴,
Janiel Cragun⁴, Hope Cottrill⁴, Michael J Kelley², Rebecca Petersen⁵, David Harpole⁵, Jeffrey Marks⁵,
Andrew Berchuck^{1,6}, Geoffrey S Ginsburg^{1,2}, Phillip Febbo¹⁻³, Johnathan Lancaster⁴ &
Joseph R Nevins¹⁻³

Potti et al (2006), *Nature Medicine*, 12:1294-1300.

The main conclusion is that we can use microarray data from cell lines (the NCI60) to define drug response “signatures”, which can be used to predict whether patients will respond.

They provide examples using 7 commonly used agents.

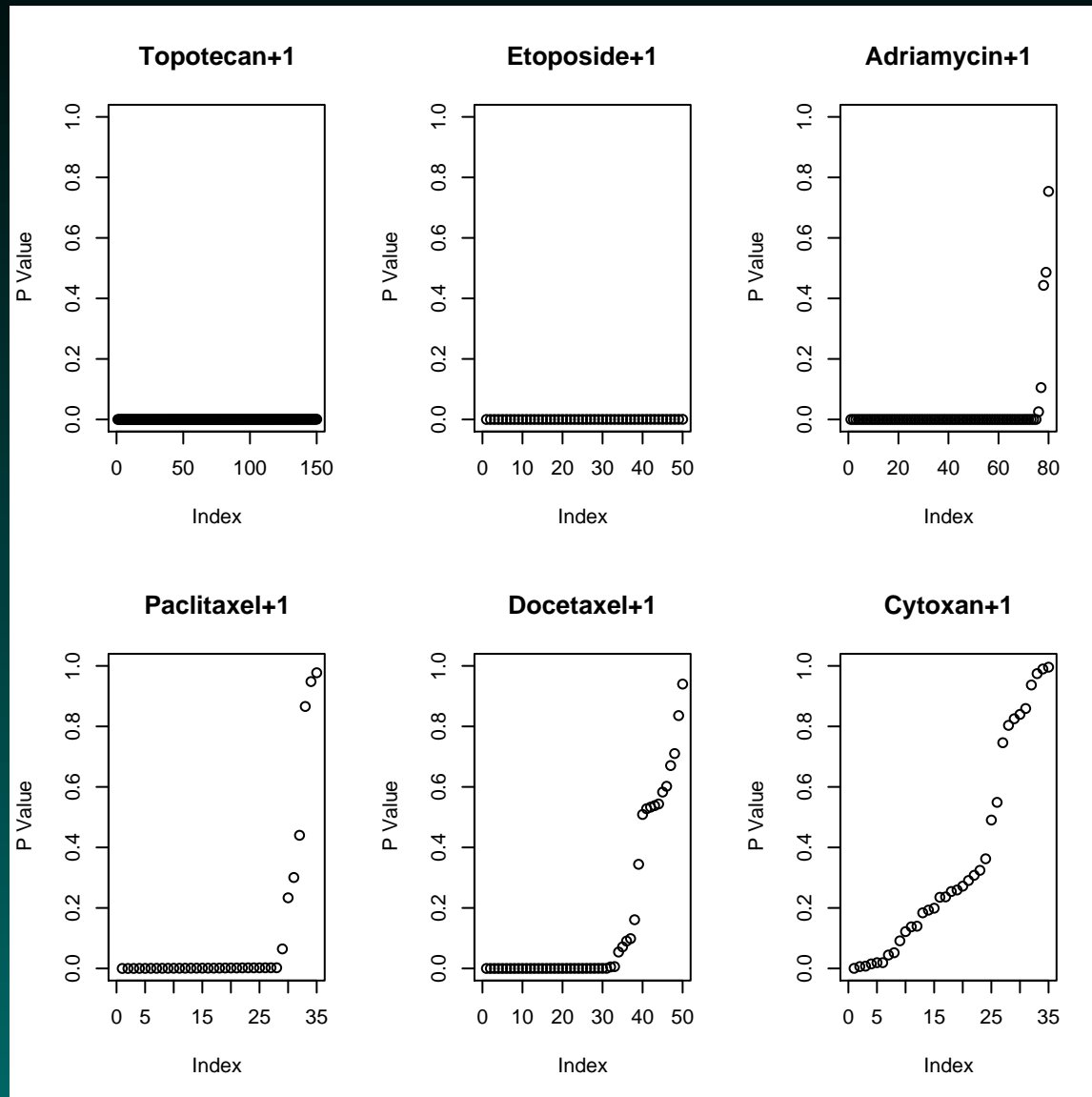
This got people at MDA very excited.

Their Gene List and Ours

```
> temp <- cbind(
  sort(rownames(pottiUpdated)[fuRows]),
  sort(rownames(pottiUpdated)[
    fuTQNorm@p.values <= fuCut]));
> colnames(temp) <- c("Theirs", "Ours");
> temp
```

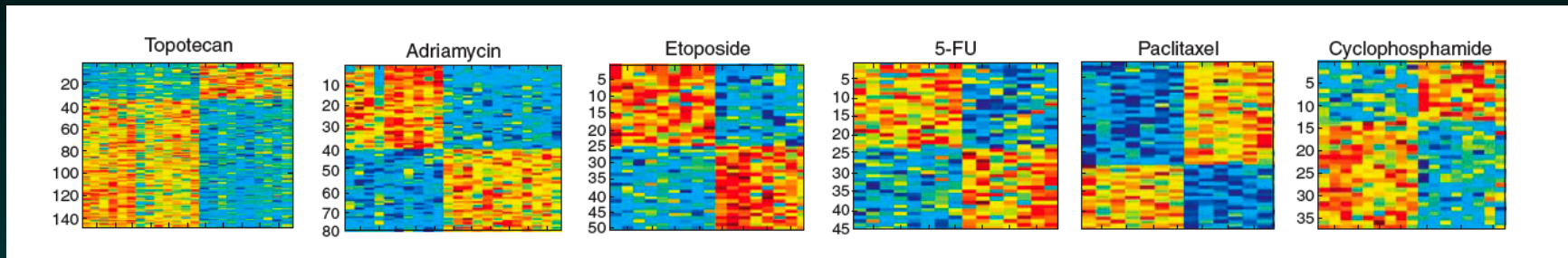
	Theirs	Ours
...		
[3,]	"1881_at"	"1882_g_at"
[4,]	"31321_at"	"31322_at"
[5,]	"31725_s_at"	"31726_at"
[6,]	"32307_r_at"	"32308_r_at"
...		

Offset P-Values: Other Drugs

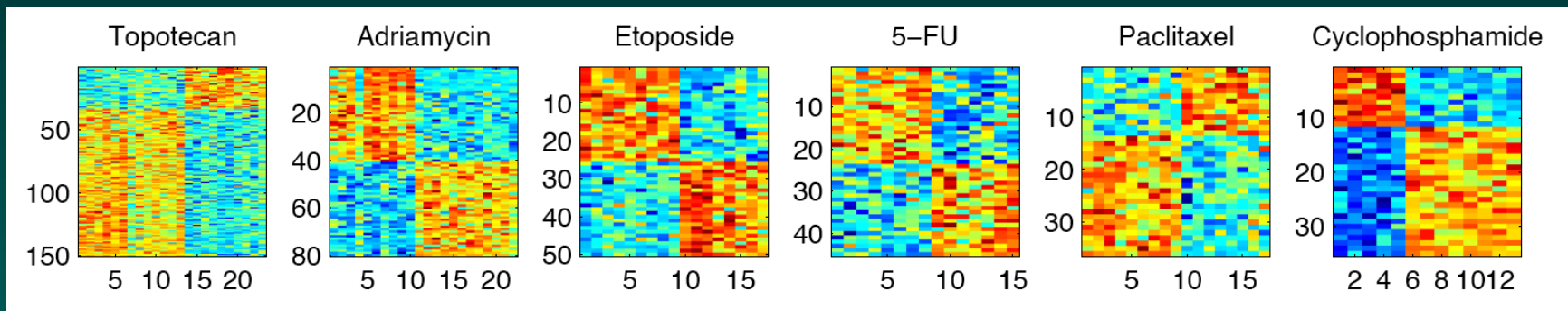


Heatmaps Match Exactly for Most Drugs

From the **paper**:



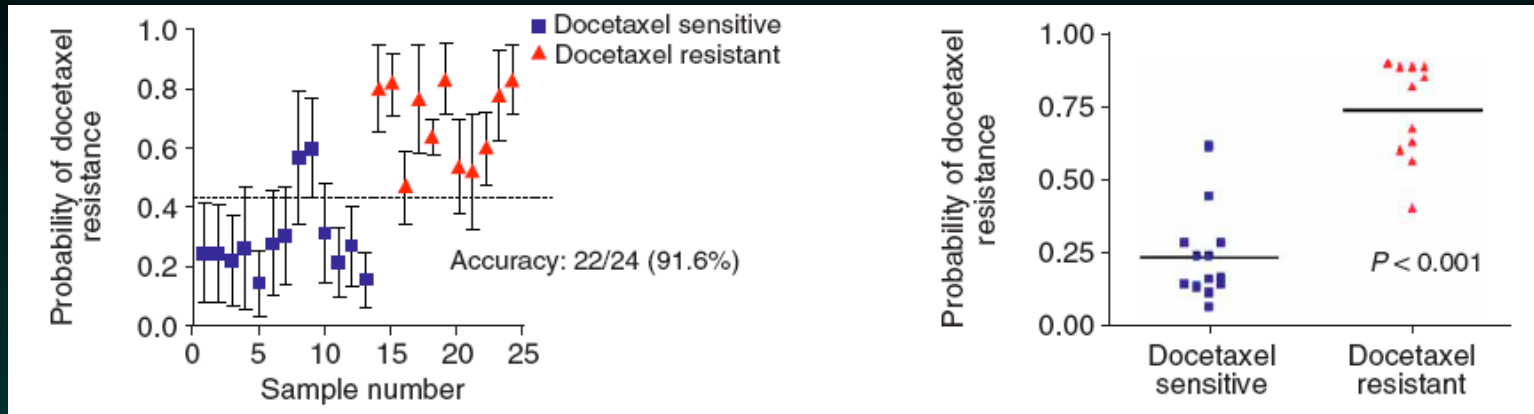
From the **software**:



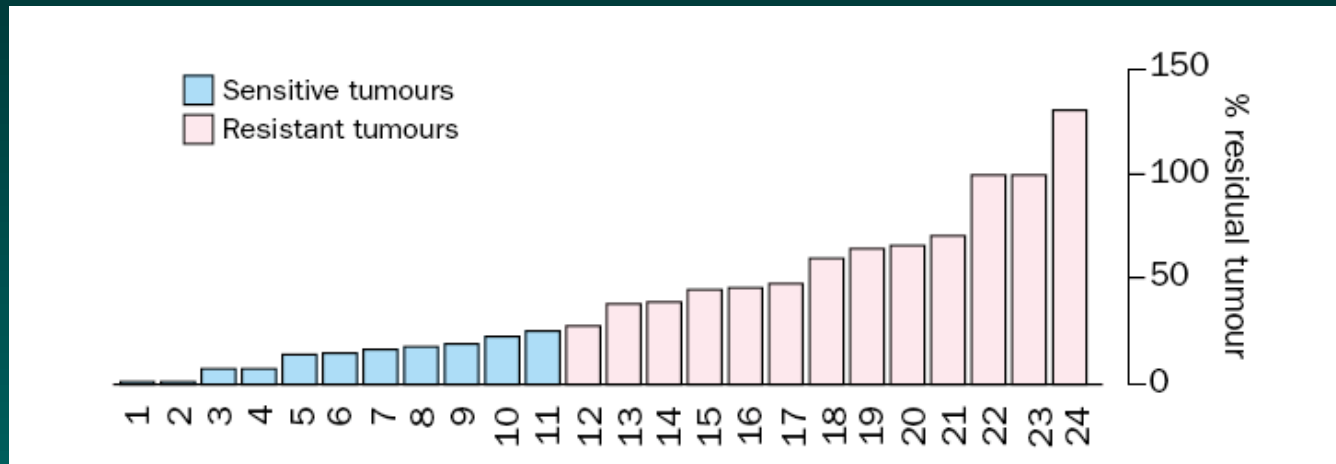
We match heatmaps but not gene lists?

Their software also gives *predictions*.

Predicting Docetaxel Response

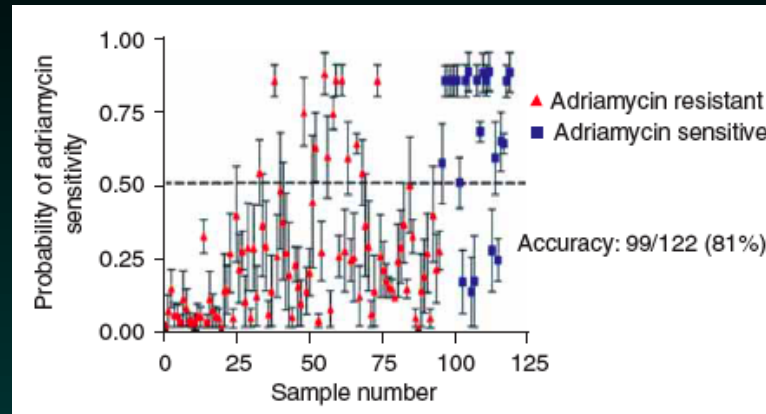


Potti et al, Nat Med 2006, 12:1294-300, Fig 1d

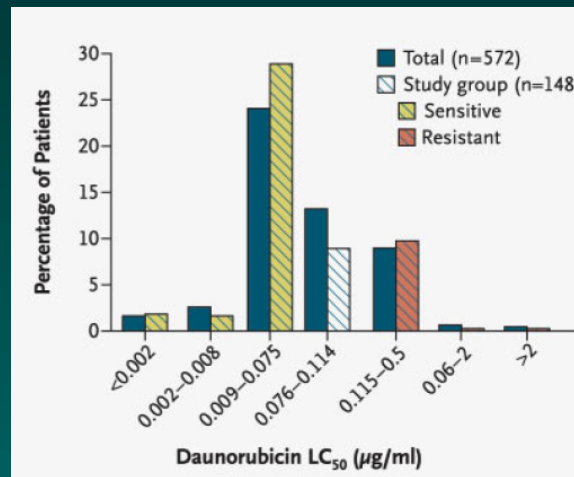


Chang et al, Lancet 2003, 362:362-9, Fig 2 top

Predicting Adriamycin Response

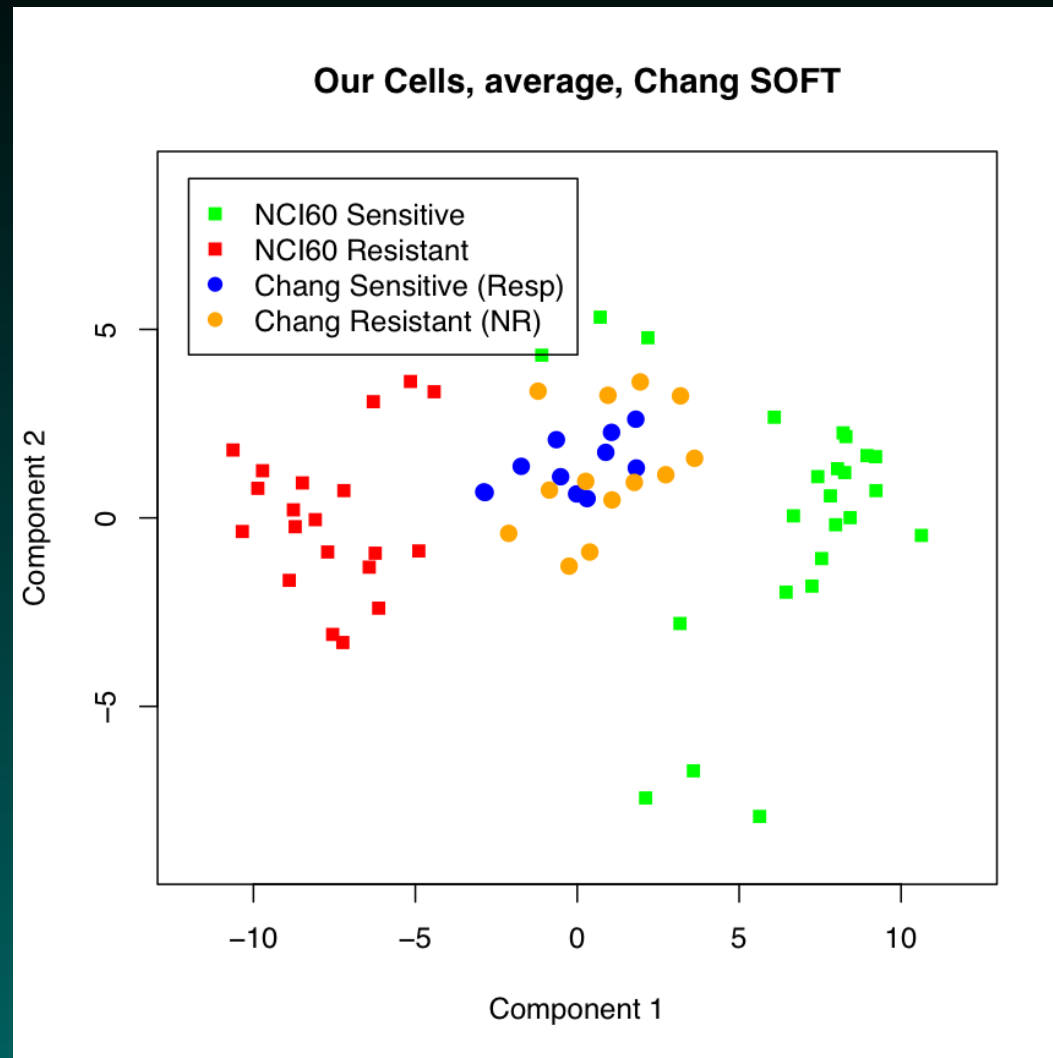


Potti et al, Nat Med 2006, 12:1294-300, Fig 2c



Holleman et al, NEJM 2004, 351:533-42, Fig 1

Trying it Ourselves



When we try it, *it doesn't work.*

Partial Timeline

2006:

- * **Nov 8**: Our first questions to Potti and Nevins.
- * **Nov 21**: Our first report describing errors.
- * **Nov-Dec**: More reports/questions: Nov 27, Dec 4, 13, 27.

2007:

- * **Jan 24**: We meet with Nevins at M.D. Anderson. We urge him to review the data.
 - * **Feb-Apr**: New data and code are posted. Some numbers change. We tell them we don't think it works.
 - * **Apr 25**: We send Potti and Nevins a draft for comment.
 - * **May**: We find problems with outliers. Potti and Nevins continue to insist it works, and want to **"bring this to a close"**.
-

A Repro Theme: Don't Take My Word For It!

Read the paper: Coombes, Wang & Baggerly, Nat Med, Nov 6, 2007, 13:1276-7, author reply 1277-8.

Try it yourselves: All of the raw data, documentation*, and code* is available from our web site (*and from Nat Med):

<http://bioinformatics.mdanderson.org/Supplements/ReproRsch-Chemo>.

Prompted by this example, we imposed new rules for writing reports in our department.

Potti/Nevins Rebuttal (Nat Med 13:1277-8)

Labels for Adria are correct – details on their web page.

They've gotten the approach to work again. (Twice.)

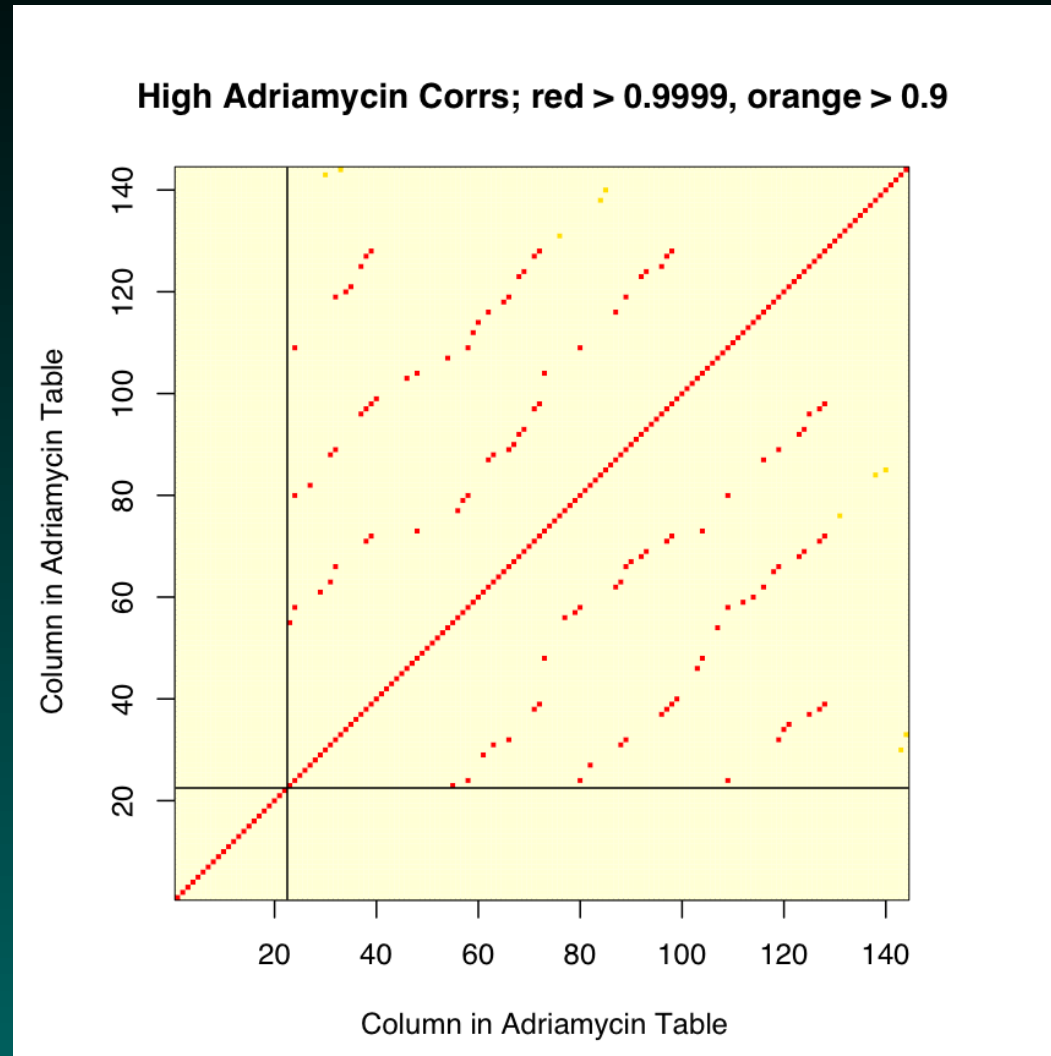
Pharmacogenomic Strategies Provide a Rational Approach to the Treatment of Cisplatin-Resistant Patients With Advanced Cancer

David S. Hsu, Bala S. Balakumaran, Chaitanya R. Acharya, Vanja Vlahovic, Kelli S. Walters, Katherine Garman, Carey Anders, Richard F. Riedel, Johnathan Lancaster, David Harpole, Holly K. Dressman, Joseph R. Nevins, Phillip G. Febbo, and Anil Potti

Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG 00-01 clinical trial

Hervé Bonnefoi, Anil Potti, Mauro Delorenzi, Louis Mauriac, Mario Campone, Michèle Tubiana-Hulin, Thierry Petit, Philippe Rouanet, Jacek Jassem, Emmanuel Blot, Véronique Becette, Pierre Farmer, Sylvie André, Chaitanya R Acharya, Sayan Mukherjee, David Cameron, Jonas Bergh, Joseph R Nevins, Richard D Iggo

Adriamycin 0.9999+ Correlations (Reply)



Redone Aug 08, “using ... 95 unique samples”.

Validation 1: Hsu et al

Pharmacogenomic Strategies Provide a Rational Approach to the Treatment of Cisplatin-Resistant Patients With Advanced Cancer

David S. Hsu, Bala S. Balakumaran, Chaitanya R. Acharya, Vanja Vlahovic, Kelli S. Walters, Katherine Garman, Carey Anders, Richard F. Riedel, Johnathan Lancaster, David Harpole, Holly K. Dressman, Joseph R. Nevins, Phillip G. Febbo, and Anil Potti

J Clin Oncol, Oct 1, 2007, 25:4350-7.

Same approach, using **Cisplatin** and **Pemetrexed**.

For cisplatin, U133A arrays were used for training. **ERCC1**, **ERCC4** and **DNA repair** genes are identified as “important”.

With some work, we matched the heatmaps. (Gene lists?)

The 4 We Can't Match (Reply)

203719_at, ERCC1,
210158_at, ERCC4,
228131_at, ERCC1, and
231971_at, FANCM (DNA Repair).

Another problem –

The last two probesets aren't on the U133A arrays that were used. They're on the U133B.

Validation 2: Bonnefoi et al

Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG 00-01 clinical trial

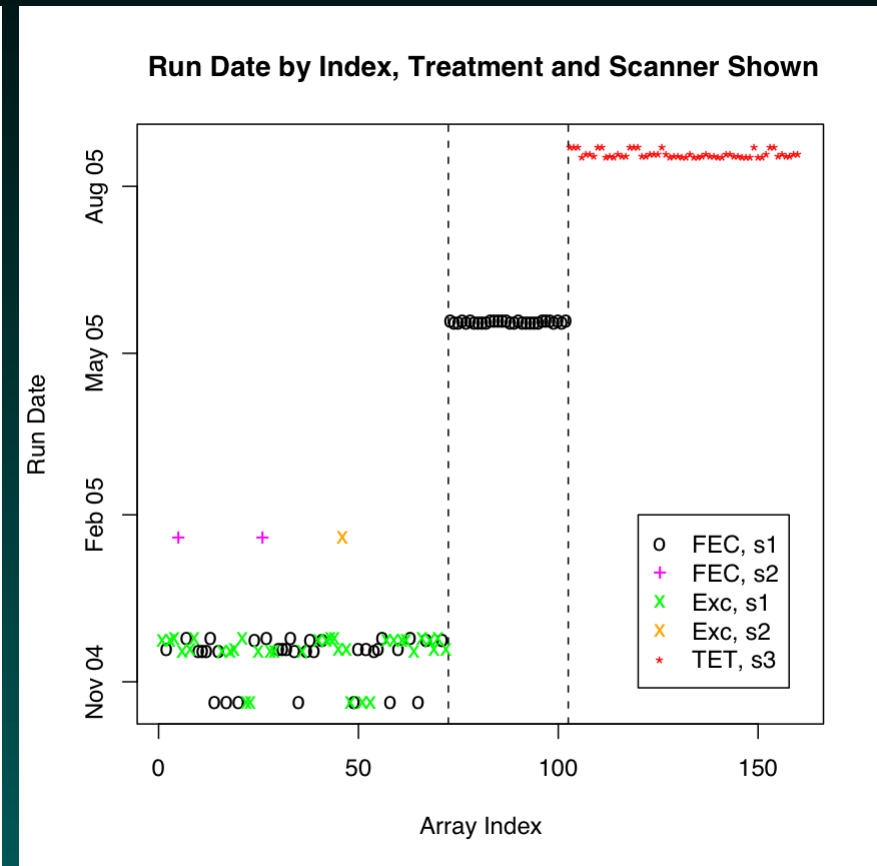
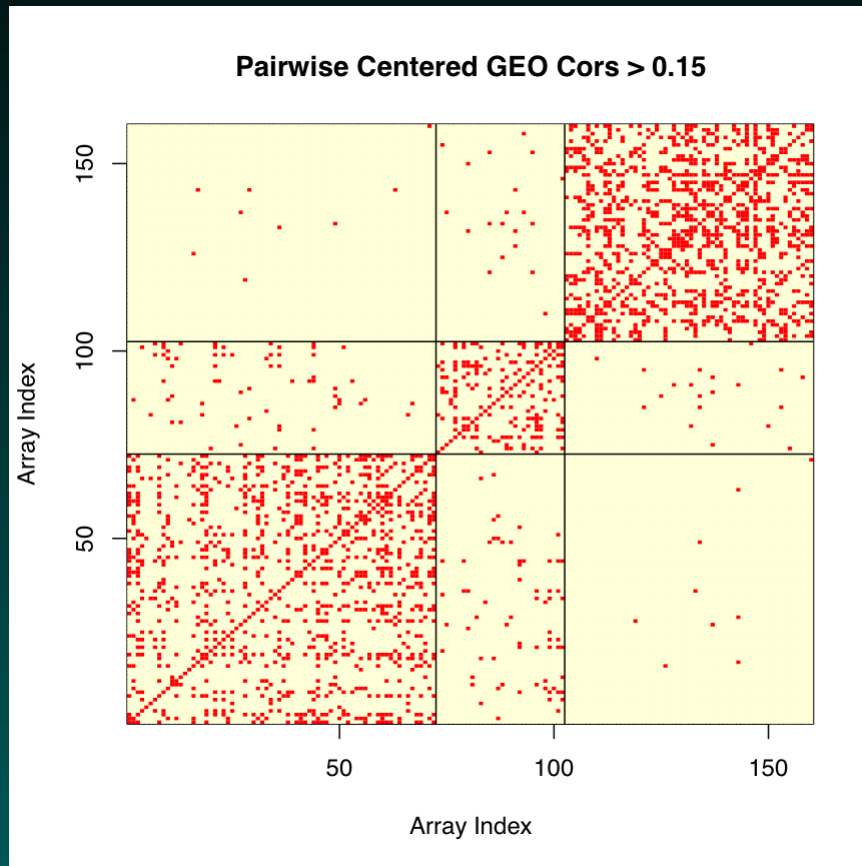
Hervé Bonnefoi, Anil Potti, Mauro Delorenzi, Louis Mauriac, Mario Camponé, Michèle Tubiana-Hulin, Thierry Petit, Philippe Rouanet, Jacek Jassem, Emmanuel Blot, Véronique Becette, Pierre Farmer, Sylvie André, Chaitanya R Acharya, Sayan Mukherjee, David Cameron, Jonas Bergh, Joseph R Nevins, Richard D Iggo

Lancet Oncology, Dec 2007, 8:1071-8. (early access Nov 14)

Similar approach, using signatures for Fluorouracil, Epirubicin, Cyclophosphamide, and Taxotere to predict response to combination therapies: **FEC** and **TET**.

Potentially improves ER- response from 44% to 70%.

Experimental Design Matters



High Sample Correlations
after Centering by Gene

Array Run Dates

How Are Results Combined?

Potti et al predict response to TFAC, Bonnefoi et al to TET and FEC. Let $P()$ indicate prob sensitive. The rules used are as follows.

How Are Results Combined?

Potti et al predict response to TFAC, Bonnefoi et al to TET and FEC. Let $P()$ indicate prob sensitive. The rules used are as follows.

$$P(TFAC) = P(T) + P(F) + P(A) + P(C) - P(T)P(F)P(A)P(C).$$

How Are Results Combined?

Potti et al predict response to TFAC, Bonnefoi et al to TET and FEC. Let $P()$ indicate prob sensitive. The rules used are as follows.

$$P(TFAC) = P(T) + P(F) + P(A) + P(C) - P(T)P(F)P(A)P(C).$$

$$P(ET) = \max[P(E), P(T)].$$

How Are Results Combined?

Potti et al predict response to TFAC, Bonnefoi et al to TET and FEC. Let $P()$ indicate prob sensitive. The rules used are as follows.

$$P(TFAC) = P(T) + P(F) + P(A) + P(C) - P(T)P(F)P(A)P(C).$$

$$P(ET) = \max[P(E), P(T)].$$

$$P(FEC) = \frac{5}{8}[P(F) + P(E) + P(C)] - \frac{1}{4}.$$

How Are Results Combined?

Potti et al predict response to TFAC, Bonnefoi et al to TET and FEC. Let $P()$ indicate prob sensitive. The rules used are as follows.

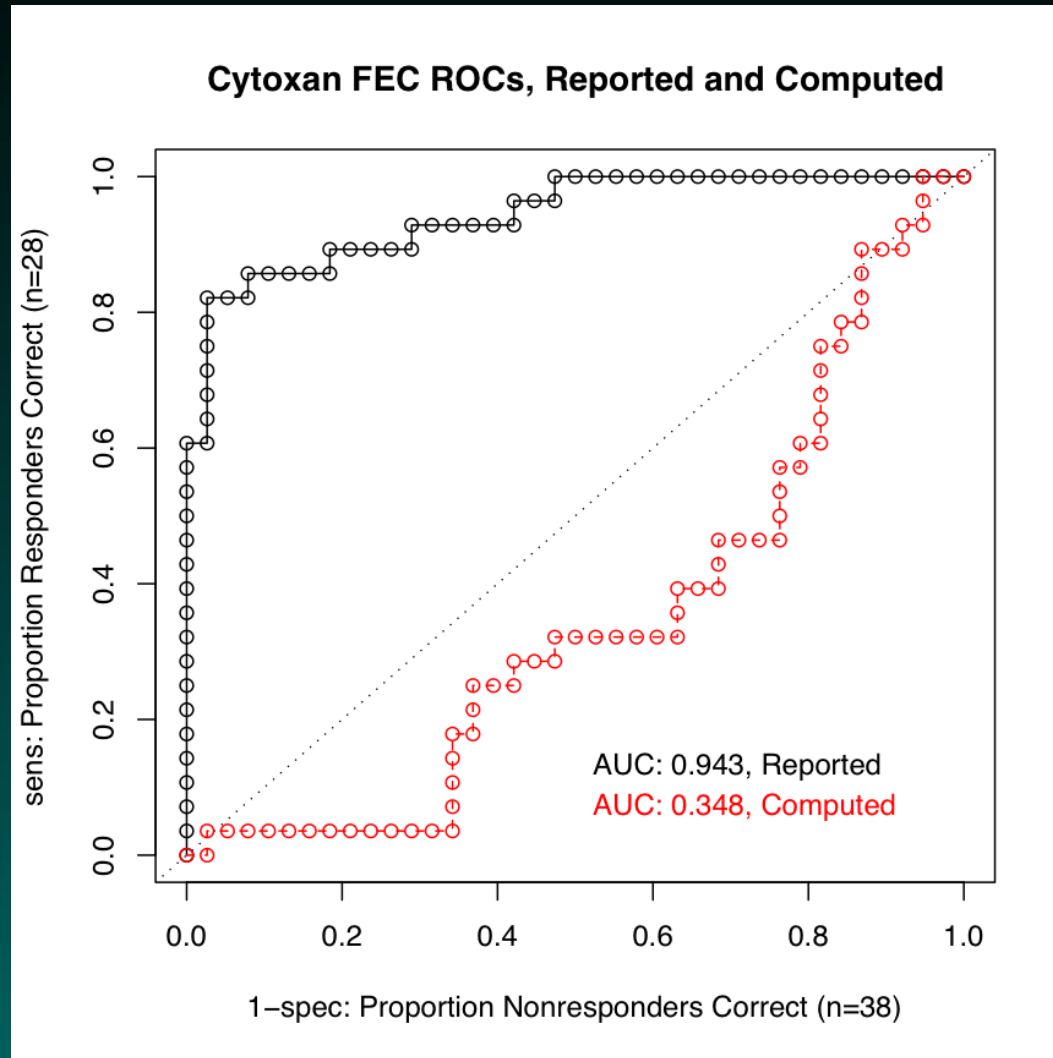
$$P(TFAC) = P(T) + P(F) + P(A) + P(C) - P(T)P(F)P(A)P(C).$$

$$P(ET) = \max[P(E), P(T)].$$

$$P(FEC) = \frac{5}{8}[P(F) + P(E) + P(C)] - \frac{1}{4}.$$

Each rule is different.

Predictions for Individual Drugs?



Does cytoxan make sense?

The Story Takes A Darker Turn

Jun 2009: we learn clinical trials had begun.

2007: pemetrexed vs cisplatin, pem vs vinorelbine.

2008: docetaxel vs doxorubicin, topotecan vs dox (Moffitt).

The Story Takes A Darker Turn

Jun 2009: we learn clinical trials had begun.

2007: pemetrexed vs cisplatin, pem vs vinorelbine.

2008: docetaxel vs doxorubicin, topotecan vs dox (Moffitt).

Sep 1, 2009: We submit a paper describing case studies to the *Annals of Applied Statistics*.

Sep 14, 2009: Paper accepted and available online at the *Annals of Applied Statistics*.

Sep-Oct 2009: Story covered by *The Cancer Letter*. NCI raises concerns with Duke's IRB behind the scenes. Duke starts internal investigation, suspends trials.

Are Our Objections Moot?

“Data was made available to us, **blinded**. All we got was the gene expression data. We ran the predictions and sent it back to the EORTC investigators” – *Joe Nevins, Oct 2.*

Are Our Objections Moot?

“Data was made available to us, **blinded**. All we got was the gene expression data. We ran the predictions and sent it back to the EORTC investigators” – *Joe Nevins, Oct 2.*

Sample info supplied:

Arm, Composite label


A, npCR Ep P- T3 N1 HB01 ...

A, pCR Ep Pp T2 N1 HB04

The data weren't blinded.

“we would not be able to reproduce the reported probabilities with the information we have about how they were obtained.”
– *Mauro Delorenzi, Oct 23.*

Jan 29, 2010



PO Box 9905 Washington DC 20016 Telephone 202-362-1809

**Duke In Process To Restart Three Trials
Using Microarray Analysis Of Tumors**

By Paul Goldberg

Duke University said it is in the process of restarting three clinical trials using microarray analysis of patient tumors to predict their response to chemotherapy.

Their investigation's results *"strengthen ... confidence in this evolving approach to personalized cancer treatment."*

We Asked for the Data

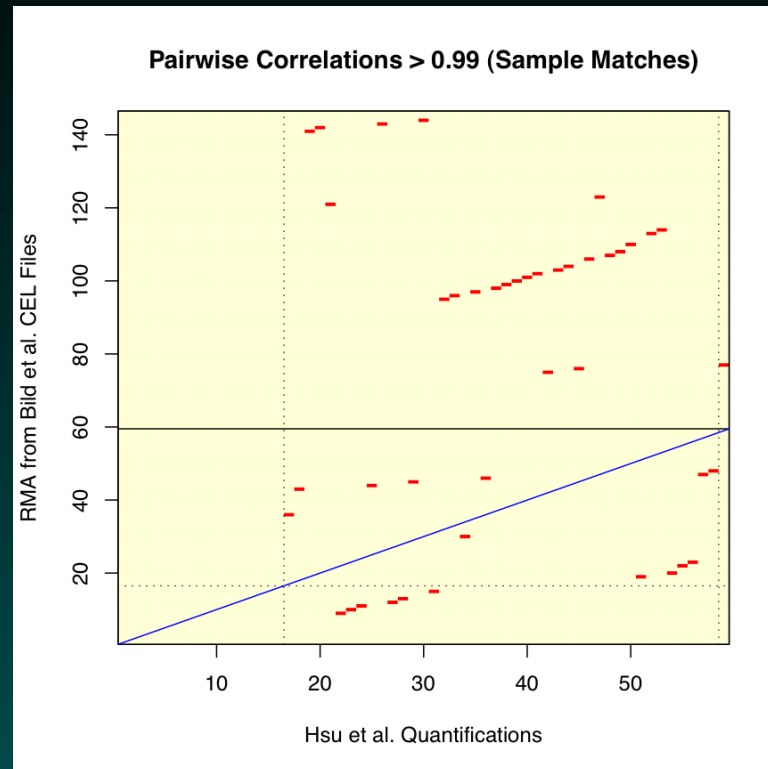
“While the reviewers approved of our sharing the report with the NCI, *we consider it a confidential document*” (Duke). A *future paper* will explain the methods.

There was, however, a major new development the restart announcement didn't mention.

In mid-Nov (mid-investigation), the Duke team posted new data for cisplatin and pemetrexed (in trials since '07).

These included quantifications for 59 ovarian cancer test samples (from GSE3149) used for predictor validation.

We Tried Matching The Samples

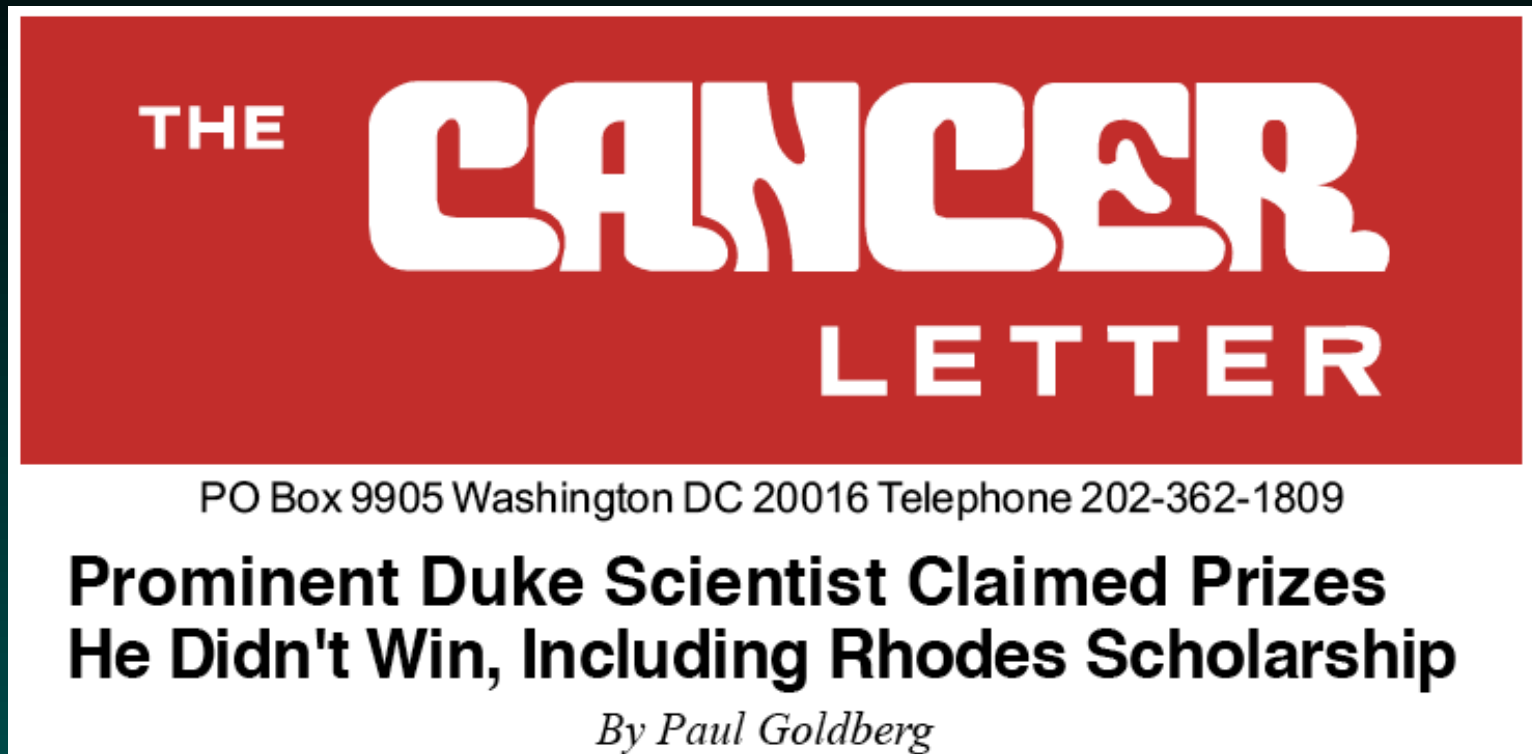


43 samples are mislabeled.

16 samples don't match because the genes are mislabeled.

All of the validation data are wrong.

A Catalyzing Event: July 16, 2010



Jul 19/20: Letter to Varmus; Duke resuspends trials.

Oct 22/9: First call for paper retraction.

Nov 9: Duke terminates trials.

Nov 19: call for Nat Med retraction, Potti resigns

Dec 20, 2010: the NCI Speaks

Sep 2009: Our paper received. Similar problems noted with CALGB 30702 application. Concerns sent to Duke IRB.

Nov 2009-Mar 2010: Data underlying the Lung Metagene Score (LMS) used in CALGB 30506 reexamined. Signature found unjustified and unstable. LMS pulled from trial.

April 2010: NCI learns it is partially funding NCT00509366. Data, code immediately requested.

May 2010: Problems found with cisplatin, pemetrexed signatures.

June 29, 2010: Duke team visits NCI. NCI directs that search for data justifying trials be conducted.

The IOM Reviews

Dec 20, 2010: NCI, FDA Presentations.

Mar 30-1, 2011: Case Studies. Joe Nevins presents.

I present. Duke historical document supplied.

Details clarify what happened with our Nov 2009 report.

Jun 30, 2011: NCI Presentation.

Aug 22, 2011: Duke Institutional Response.

Nov 4, 2011: Moffitt trial in *The Cancer Letter*.

Links to MP3 audio, documents, our annotations:

<http://bioinformatics.mdanderson.org/Supplements/ReproRsch-All/Modified/index.html>

Other Developments

Sep, 2011: Patient lawsuits filed.

Dec, 2011: NC Medical Board: reprimand, some settlements.

Dec, 2011: SC Medical Board: Potti application letter.

Jan, 2012: Acharya et al JAMA, Dressman et al JCO papers retracted.

Wikipedia

Online reputation manager

Press coverage

Feb 12, 2012: 60 Minutes.

http://www.cbsnews.com/8301-18560_162-57376073/deception-at-duke/

The Story's Not Over Yet...

"All the News That's Fit to Print"

The New York Times

Late Edition
Edwin Blackwell, 84, is the author of the book "The New York Times: A History" (2001).
NEW YORK, FRIDAY, JULY 8, 2011
\$2.00

Murdoch Closing Tabloid Linked to British Hacking

Former Editor Who Became Spokesman for the Prime Minister May Face Arrest

PARIS — Britain's media and political landscape shifted abruptly on the grounds of the London Daily Mail, owned by Rupert Murdoch, when the newspaper's former editor, James Murdoch, was linked to a series of e-mails that were part of a campaign to influence the 2010 general election in the United Kingdom. The e-mails, which were sent to the British Prime Minister, Gordon Brown, and other senior government officials, were part of a campaign to influence the election. The e-mails were sent to the British Prime Minister, Gordon Brown, and other senior government officials, were part of a campaign to influence the election. The e-mails were sent to the British Prime Minister, Gordon Brown, and other senior government officials, were part of a campaign to influence the election.

How Bright Promise in Cancer Testing Fell Apart

While scientists agree there is great promise in the science, the cost of the tests is too high for many patients.

NEW YORK — While scientists agree there is great promise in the science, the cost of the tests is too high for many patients. The tests, which are used to identify genetic mutations that can lead to cancer, are being used in a variety of ways. Some are used to identify genetic mutations that can lead to cancer, while others are used to identify genetic mutations that can lead to cancer. The tests, which are used to identify genetic mutations that can lead to cancer, are being used in a variety of ways. Some are used to identify genetic mutations that can lead to cancer, while others are used to identify genetic mutations that can lead to cancer.

Next Nation Full of Hope and Problems

Uganda's young people are optimistic about the future, but they are also facing many challenges.

NEW YORK — Uganda's young people are optimistic about the future, but they are also facing many challenges. The country has a young population, and many of its citizens are looking for better opportunities. However, the country is also facing many challenges, including poverty and corruption. The government is trying to address these challenges, but it is still a long way from solving them.

Not Providers Plan Penalties To Slow Piracy

Medical groups are looking for ways to protect their patients' information.

NEW YORK — Medical groups are looking for ways to protect their patients' information. The groups are concerned about the growing problem of medical piracy, which is the unauthorized use of patients' medical information. The groups are looking for ways to protect their patients' information, and they are planning to impose penalties on anyone who steals or leaks their patients' information.

Behind the Gentler Approach to Banks by U.S.

As Wall St. Policies Itself, Charges

NEW YORK — Behind the gentler approach to banks by the U.S. government, as Wall Street policies itself, charges. The government has been criticized for its handling of the financial crisis, and many people are questioning its approach to banks. However, the government has defended its approach, saying that it is necessary to stabilize the financial system.

Still 'Far Apart' on Debt, 2 Sides Will Seek \$4 Trillion in Savings

Republicans and Democrats are still far apart on how to deal with the national debt.

WASHINGTON — Republicans and Democrats are still far apart on how to deal with the national debt. The two sides have been unable to reach an agreement on how to reduce the deficit, and they are still far apart on how to deal with the national debt. The two sides are still far apart on how to deal with the national debt, and they are still far apart on how to deal with the national debt.

Founders Days

Many cities are celebrating the founding of the United States.

NEW YORK — Many cities are celebrating the founding of the United States. The cities are holding various events, including parades and fireworks, to commemorate the occasion. The cities are celebrating the founding of the United States, and they are celebrating the founding of the United States.

Mohawk Allow Are Charged

Two men are charged with the murder of a young girl.

NEW YORK — Two men are charged with the murder of a young girl. The men are accused of killing the girl, and they are facing charges of murder. The men are charged with the murder of a young girl, and they are facing charges of murder.

An Olympic Transformation

The city is preparing for the 2012 Summer Olympics.

NEW YORK — The city is preparing for the 2012 Summer Olympics. The city is making various changes to its infrastructure, including building new stadiums and improving its transportation system. The city is preparing for the 2012 Summer Olympics, and it is preparing for the 2012 Summer Olympics.

More to Perform Gay Marriage

Some states are still debating the issue.

NEW YORK — Some states are still debating the issue of gay marriage. The issue is still a controversial one, and many people are still debating it. Some states are still debating the issue, and some states are still debating the issue.

The Crisis of Values

Many people are questioning the values of the American dream.

NEW YORK — Many people are questioning the values of the American dream. The American dream is a concept that has been central to American culture, but many people are questioning it. They are questioning the values of the American dream, and they are questioning the values of the American dream.

Leading Figure in N.F.L. Dies

A former player and coach has passed away.

NEW YORK — A former player and coach has passed away. The man was a well-known figure in the NFL, and his death is a significant loss to the sport. The man was a well-known figure in the NFL, and his death is a significant loss to the sport.

Paul Krugman

The author of "The Great Unraveling" is back in the news.

NEW YORK — The author of "The Great Unraveling" is back in the news. The book is a critique of the American economy, and it has been widely discussed. The author is back in the news, and he is back in the news.

For now, we have a soapbox.

This is not an unmixed blessing.

What can we learn, and how are things moving forward?

Some Cautions/Observations/Lessons

This case is pathological.

But we've seen similar problems before.

The most common mistakes are simple.

Confounding in the Experimental Design

Mixing up the sample labels

Mixing up the gene labels

Mixing up the group labels

(Most mixups involve simple switches or offsets)

This simplicity is often hidden.

Incomplete documentation

What Should the Reproducibility Norm Be?

In our group we've prepared reports in *Sweave* since 2007.

For papers? (Baggerly + lots, *Nature*, Sep 22, 2010)

Things we look for:

1. Data (often mentioned, given MIAME)
2. Provenance
3. Code
4. Descriptions of Nonscriptable Steps
5. Descriptions of Planned Design, if Used.

For clinical trials?

Reasons for Hope

1. Our Own Experience
2. Duke's Trans Med Qual Framework (TMQF) Team
3. The NCI and Trials it Funds
4. The IOM, the FDA, and IDEs
5. Journals, Code and Data

Acknowledgements

Kevin Coombes

Shannon Neeley, Jing Wang

David Ransohoff, Gordon Mills

Jane Fridlyand, Lajos Pusztai, Zoltan Szallasi

M.D. Anderson Ovarian, Lung and Breast SPOREs

Baggerly and Coombes (2009), *Annals of Applied Statistics*, 3(4):1309-34. <http://bioinformatics.mdanderson.org/Supplements/ReproRsch-All>

For updates: <http://bioinformatics.mdanderson.org/Supplements/ReproRsch-All/Modified>.

Index

Title

Cell Line Story

Trying it Ourselves

The Reply

Adriamycin Followup

Hsu et al. (Cisplatin)

Bonnefoi et al.

Temozolomide Heatmaps

Timeline, Trials, Cancer Letter

Trial Restart and Objections

Final Lessons