The Importance of Reproducible Research: Lessons from Train Wrecks

Keith A. Baggerly
Bioinformatics and Computational Biology
UT M. D. Anderson Cancer Center
kabagg@mdanderson.org

Chicago ASA, Apr 1, 2016
Why is Reproducible Research Important?

Datasets are getting bigger and more complex.

Our intuition about what “makes sense” is very poor in high dimensions.

To use “omics-based signatures” as biomarkers, we need to know they’ve been assembled correctly. (This applies to results from simpler assays as well.)

Without documentation, we may need to employ (lengthy!) *forensic bioinformatics* to infer what was done.

Let’s look at a few case studies from oncology...
A Proteomics Case Study

MECHANISMS OF DISEASE

Mechanisms of disease

Use of proteomic patterns in serum to identify ovarian cancer


100 ovarian cancer patients
100 normal controls
16 patients with “benign disease”

Use 50 cancer and 50 normal spectra to train a classification method; test the algorithm on the remaining samples.
Which Group is Different?
Really?
Processing Can Trump Biology: Design!
Some Timeline

2004:
* Early Jan: Correlogic, Quest and LabCorp advertise the forthcoming “OvaCheck” assay at SGO.
* Jan 29: Critiques available online
* Feb 3: New York Times coverage
* Feb-Mar: Letters from FDA to companies involved
* July: FDA rules omics signatures are medical devices and will be regulated accordingly.

2006:
* FDA releases draft guidance on IVDMIAs
* NCI Clinical Proteomic Technologies for Cancer (CPTAC)
Are Things Better Now?

New York Times, 2.3.04

New Cancer Test Stirs Hope and Concern

By ANDREW POLLACK

Jill Daimer’s mother died in 2002 from ovarian cancer, detected too late to be effectively treated. So Ms. Daimer is eagerly awaiting the introduction of a new test that holds the promise of detecting early-stage ovarian cancer far more accurately than any test available now, using only blood from a finger prick.

Not only does she plan to be tested, but an advocacy group she helped found, Ovarian Awareness of Kentucky, also intends to spread the word to women and doctors.

“If it’s going to happen to me or anyone I know, I want it to be caught at an early stage,” said Ms. Daimer, who lives in Louisville.

The new test, expected to be available in the next few months, could have a big effect on public health if it works as advertised. That is because when ovarian cancer is caught early, when it is treatable by surgery, more than 90 percent of women live five years or longer. But right now, about three-quarters of cases are detected after the cancer has advanced, and then only 30 percent of women survive five years.

The test is also the first to use a new technology that some believers say could revolutionize diagnostics. It looks not for a single telltale protein — like the prostate-specific antigen, or P.S.A., used to diagnose prostate cancer — but rather for a complex fingerprint formed by all the proteins in the blood. Similar tests are being developed for prostate, pancreatic, breast and other cancers. The technique may work for other diseases as well.

“I’ve been in cancer research for 40 years and I think it’s the most important breakthrough in those years,” said Dr.

Continued on Page 6
Are Things Better Now?

New York Times, 2.3.04

New Cancer Test Stirs Hope and Concern

By ANDREW POLLACK

Jill Doiner's mother died in 2002 from ovarian cancer, described too late to be effectively treated.

So Ms. Doiner is eagerly awaiting the introduction of a new test that holds the promise of detecting early-stage ovarian cancer far more accurately than any test available now, using only blood from a finger prick.

Not only does she plan to be tested, but an advocacy group she helped found, Ovarian Awareness of Kentucky, also intends to spread the word to women and doctors.

"If it's going to happen to me or anyone I know, I want it to be caught at an early stage," said Ms. Doiner, who lives in Louisville.

The new test, expected to be available in the next few months, could have a big effect on public health if it works as advertised. That is because while ovarian cancer is caught early, when it is treatable by surgery, more than 60 percent of women live five years or longer. But right now, about three-quarters of cases are detected after the cancer has advanced, and then only 30 percent of women survive five years.

The test is also the first to use a new technology that some believers say could revolutionize diagnostics. It looks for a single telltale protein — like the prostate-specific antigen, or P.S.A., used to diagnose prostate cancer — but rather for a complex fingerprint formed by all the proteins in the blood. Similar tests are being developed for prostate, pancreatic, breast and other cancers. The technology may work for other diseases as well.

"I've been in cancer research for 40 years and I think it's the most important breakthrough in these years," said Dr. 

Continued on Page 6

Cancer Test For Women Raises Hope, And Concern

By ANDREW POLLACK

A new blood test aimed at detecting ovarian cancer at an early, still treatable stage is stirring hopes among women and their physicians. But the Food and Drug Administration and some experts say the test has not been proved to work.

Is This Unique to Proteomics?

High Sample Correlations

Array Run Dates

See Leek et al, Nat Rev Gen 2010 for more examples.
Using Cell Lines to Predict Sensitivity

Genomic signatures to guide the use of chemotherapeutics

Anil Potti¹,², Holly K Dressman¹,³, Andrea Bild¹,³, Richard F Riedel¹,², Gina Chan⁴, Robyn Sayer⁴, Janiel Cragun⁴, Hope Cottrill⁴, Michael J Kelley², Rebecca Petersen⁵, David Harpole⁵, Jeffrey Marks⁵, Andrew Berchuck¹,⁶, Geoffrey S Ginsburg¹,², Phillip Febbo¹,³, Johnathan Lancaster⁴ & Joseph R Nevins¹,³


The main conclusion: we can use microarray data from cell lines (the NCI60) to define drug response “signatures”, which can 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

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

<table>
<thead>
<tr>
<th></th>
<th>Theirs</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[3,]</td>
<td>&quot;1881_at&quot;</td>
<td>&quot;1882_g_at&quot;</td>
</tr>
<tr>
<td>[4,]</td>
<td>&quot;31321_at&quot;</td>
<td>&quot;31322_at&quot;</td>
</tr>
<tr>
<td>[5,]</td>
<td>&quot;31725_s_at&quot;</td>
<td>&quot;31726_at&quot;</td>
</tr>
<tr>
<td>[6,]</td>
<td>&quot;32307_r_at&quot;</td>
<td>&quot;32308_r_at&quot;</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```
Predicting Response: Docetaxel

Potti et al (2006), Nature Medicine, 12:1294-300, Fig 1d

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

Potti et al (2006), Nature Medicine, 12:1294-300, Fig 2c

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

2006:
* Nov 8: Our first questions to Potti and Nevins.
* Nov 21: Our first report describing errors.

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”.
Adriamycin 0.9999+ Correlations

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


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

203719 at, ERCC1,
210158 at, ERCC4,
228131 at, ERCC1, and
231971 at, FANCM (DNA Repair).
The 4 We Can’t Match

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

The last two probesets aren’t on the U133A arrays that were used. They’re on the U133B.
The Reason We Really Care

Jun 2009: we learn clinical trials had begun.

2007: pemetrexed vs cisplatin, pem vs vinorelbine.

2008: docetaxel vs doxorubicin, topotecan vs dox (Moffitt).
The Reason We Really Care

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:

NCI raises concerns with Duke’s IRB behind the scenes.
Duke starts internal investigation, suspends trials.
New Data

Early-Nov ’09 (mid-investigation), the Duke team posted new data for cisplatin and pemetrexed (in lung trials since ’07).

These included quantifications for the 59 ovarian cancer test samples (from GSE3149, which has 153 samples) they used to validate their predictor.
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.

We reported this to Duke and to the NCI in mid-November.
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

“WeWhile 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.

This did give us one more option...
We Asked for the Data

“We 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.

This did give us one more option...

In May 2010, we obtained a copy of the reviewers’ report from the NCI under FOIA (Cancer Letter, May 14).

In our assessment (and others’), it did not justify restarting trials.

*There was no mention of our Nov 2009 report.*
A Catalyzing Event: July 16, 2010

Prominent Duke Scientist Claimed Prizes He Didn't Win, Including Rhodes Scholarship

By Paul Goldberg

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
Other Developments

117 patients were enrolled in the trials.

Sep, 2011: Patient lawsuits filed (11+ settlements).

Misconduct investigation (ongoing).
10 retractions, 6+ “partial retractions”
FDA Review, Discussions with Duke IRB


Feb 12, 2012: 60 Minutes.

Some Cautions/Observations

These cases are 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 about non-genomic tests?
Screening for Ovarian Cancer

Using a biomarker for early detection, can we reduce mortality?

We hope to see three things:

1. A **High PPV** (few unnecessary surgeries)
2. A **Stage Shift** (the early/late stage ratio increases)
3. A **Lagged Improvement in Outcomes** (overall survival)

Some results from **UKCTOCS** (the largest such trial to date) were released at the end of 2015.

The test had a high PPV, and they saw a stage shift.
Does Screening Improve Survival? (1/2)

Primary Outcome: Log-rank test
Does Screening Improve Survival? (1/2)

Primary Outcome: Log-rank test $p = 0.10$ (wrong test)
Does Screening Improve Survival? (2/2)

Web Figure 3a: Cumulative ovarian cancer deaths by randomisation group with RP models overlaid - MMS versus no screening (C).

Post-hoc **Weighted Log Rank**
Does Screening Improve Survival? (2/2)

Post-hoc Weighted Log Rank $p = 0.023$
But Wait...

<table>
<thead>
<tr>
<th></th>
<th>Number of women (n)</th>
<th>Deaths (n)</th>
<th>Mortality reduction 0–14 years (%)</th>
<th>p value</th>
<th>Mortality reduction 0–7 years (%)</th>
<th>Mortality reduction 7–14 years (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ovarian cancer (primary analysis)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cox model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMS</td>
<td>50,624</td>
<td>148</td>
<td>15% (-3 to 30)</td>
<td>0.10</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>USS</td>
<td>50,623</td>
<td>154</td>
<td>11% (-7 to 27)</td>
<td>0.21</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>No screening</td>
<td>101,299</td>
<td>347</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Royston- Parmar model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMS</td>
<td>50,624</td>
<td>148</td>
<td>16% (-1 to 33)</td>
<td>0.11</td>
<td>8% (-20 to 31)</td>
<td>23% (1 to 46)</td>
</tr>
<tr>
<td>USS</td>
<td>50,623</td>
<td>154</td>
<td>12% (-6 to 29)</td>
<td>0.18</td>
<td>2% (-27 to 26)</td>
<td>21% (-2 to 42)</td>
</tr>
<tr>
<td>No screening</td>
<td>101,299</td>
<td>347</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Royston-Parmar model (excluding prevalent cases)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMS</td>
<td>50,561</td>
<td>120</td>
<td>20% (-2 to 40)</td>
<td>0.021</td>
<td>8% (-27 to 43)</td>
<td>28% (-3 to 49)</td>
</tr>
<tr>
<td>No screening</td>
<td>101,183</td>
<td>281</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Weighted log-rank (post-hoc)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMS</td>
<td>50,624</td>
<td>148</td>
<td>22% (3 to 38)*</td>
<td>0.023</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>USS</td>
<td>50,623</td>
<td>154</td>
<td>20% (0 to 35)*</td>
<td>0.049</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>No screening</td>
<td>101,299</td>
<td>347</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Ovarian cancer (primary analysis)</td>
<td>Number of women (n)</td>
<td>Deaths (n)</td>
<td>Mortality reduction 0-14 years (%)</td>
<td>p-value</td>
<td>Mortality reduction 0-7 years (%)</td>
<td>Mortality reduction 7-14 years (%)</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------------</td>
<td>------------</td>
<td>----------------------------------</td>
<td>---------</td>
<td>----------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Cox model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMS</td>
<td>50 624</td>
<td>148</td>
<td>15% (-3 to 30)</td>
<td>0.10</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>USS</td>
<td>50 623</td>
<td>154</td>
<td>11% (-7 to 27)</td>
<td>0.21</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>No screening</td>
<td>101 1299</td>
<td>347</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>Royston-Parmar model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMS</td>
<td>50 624</td>
<td>148</td>
<td>16% (-1 to 33)</td>
<td>0.11</td>
<td>8% (-20 to 31)</td>
<td>23% (1 to 46)</td>
</tr>
<tr>
<td>USS</td>
<td>50 623</td>
<td>154</td>
<td>12% (-6 to 29)</td>
<td>0.18</td>
<td>2% (-27 to 26)</td>
<td>21% (-2 to 42)</td>
</tr>
<tr>
<td>No screening</td>
<td>101 1299</td>
<td>347</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>Royston-Parmar model (excluding prevalent cases)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMS</td>
<td>50 561</td>
<td>120</td>
<td>20% (-2 to 40)</td>
<td>0.021</td>
<td>8% (-27 to 43)</td>
<td>28% (-3 to 49)</td>
</tr>
<tr>
<td>No screening</td>
<td>101 183</td>
<td>281</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>Weighted log-rank (post-hoc)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMS</td>
<td>50 624</td>
<td>148</td>
<td>22% (3 to 38)*</td>
<td>0.023</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>USS</td>
<td>50 623</td>
<td>154</td>
<td>20% (0 to 35)*</td>
<td>0.049</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>No screening</td>
<td>101 1299</td>
<td>347</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
</tbody>
</table>

P-values < 0.05 with CI’s including 0?
We Asked for the Relevant Raw Data

We’d like a table with one row per patient, with

1. which arm the patient was in
2. the time of last followup
3. whether the patient did or didn’t die of ovarian cancer

The “implicit promise” assumption: if you show a statistical result, you should be able to back it up if asked
We Asked for the Relevant Raw Data

We’d like a table with one row per patient, with

1. which arm the patient was in
2. the time of last followup
3. whether the patient did or didn’t die of ovarian cancer

The “implicit promise” assumption: if you show a statistical result, you should be able to back it up if asked

UKCTOCS is not releasing these data this time
We Asked for the Relevant Raw Data

We’d like a table with one row per patient, with

1. **which arm** the patient was in
2. the **time of last followup**
3. whether the patient **did or didn’t die** of ovarian cancer

The “implicit promise” assumption: if you show a statistical result, you should be able to back it up if asked

UKCTOCS is not releasing these data this time

I qualified my earlier statements accordingly
Should We Be Surprised?

Is there a community expectation of data sharing?

Folks running trials at MD Anderson

SGO

Companies

Can we get trials reported at all?

NCI - Since Jan 1, per NCTN guidelines (p.52)

Journals? Well, my timing was good...
ICMJE, Jan 26, 2016

“the ICMJE proposes to require authors to share with others the deidentified individual-patient data (IPD) underlying the results presented in the article (including tables, figures, and appendices or supplementary material) no later than 6 months after publication.”

This proposal is open for comment now.
"A second concern held by some is that a new class of research person will emerge - people who had nothing to do with the design and execution of the study but use another group’s data for their own ends, possibly stealing from the research productivity planned by the data gatherers, or even use the data to try to disprove what the original investigators had posited. There is concern among some front-line researchers that the system will be taken over by what some researchers have characterized as "research parasites." "
I’m a Research Parasite...

Retraction Watch

Intl Soc Comp Biol (ICSB)

Simply Statistics

hashtag ResearchParasite on Twitter

hashtag dataparasite

“Translation” into plain English

ASA Statement planned

Citation yes, ceding of control, no
I Don’t Handle Frustration Well

They won’t supply the data
I Don’t Handle Frustration Well

They won’t supply the data

But I’m a research parasite
I Don’t Handle Frustration Well

They won’t supply the data

But I’m a research parasite

Do we have to wait?
I Don’t Handle Frustration Well

They won’t supply the data

But I’m a research parasite

Do we have to wait?

Let’s think about image processing...
This is a KM Curve (Sort of)

Doesn’t this have the information we need?
In the MMS group, we observed an average of 63 (74%) prevalent ovarian cancers from the 338 cases, and we included an average of 108 (76%) prevalent cases from the 638 cases in the no-screening group, showing the expected equal proportion of prevalent cases in the two groups. Hazards were not proportional (p=0.03), so we fitted separate RP models to the survival data for each group, showing excellent overlap with the non-parametric Kaplan-Meier cumulative mortality curves (figure 4). The hazards between the MMS and no-screening groups were significantly different (p<0.01), showing that ovarian cancer mortality was significantly lower in the MMS group (20% vs 2 to 40%) than in the no-screening group (50%). The mortality reduction was also higher for years 7-14 (28%) vs 16-49% than for years 1-7 (p=0.01-0.04). Because the hazard ratios were not constant over time and the Cox model has low power for detection of a low effect of this type, we did a single post hoc analysis using the Wajeb test inspired by the rationale in the ESCO Cancer Screening Trial report (appendix p 15). The median from randomization to death in the no-screening group was over 8 years (randomization to cancer diagnosis is 6 years, diagnosis to death 2.5 years). This analysis suggested a significant reduction in ovarian cancer mortality in the MMS group compared with the no-screening group, but not in the US group (table 3).

In addition to the ovarian and fallopian tube cancers, we noted further 12 deaths due to primary peritoneal cancer in the no-screening group, 12 in the MMS group, and none in the US group (table 3). This resulted in smaller mortality reductions with the MMS versus no screening and US versus no screening compared with the ovarian cancer mortality reductions (figures 2, table 3) As with the primary outcome, the no-screening group hazard rate continues to rise, whereas the MMS and US hazard rates start leveling off (appendix p 72). The RP model also yielded higher mortality reductions for years 7-14 than for years 0-7, with the MMS reduction increasing if prevalent cases are excluded (figure 4).
Curves Extracted With R

We want curves 35-37, and axes. No overplotting!
We Can Get Most of What We Need

We can get death times.
We can get rough numbers at risk.
We can get numbers of deaths at each time.

We have to guess censoring times within an interval, but these shouldn’t markedly affect inferences.

We sent the data we extracted on to the UKCTOCS folks.

Should ICMJE be waiting a year?

So, what can we do with the data?
Still looks real; I like these more than the parametric tests.
These are not Isolated Problems


NCI focus meeting Sep 2012.


NAS meeting Feb 26-7, 2015.

METRICS Nov 19-20, 2015

ENAR Webinar Nov 20, 2015
Some Cost Breakdowns

Reasons for Hope

1. Our Own (Evolving!) Experience & Sanity Checks

2. Better tools (knitr, markdown, GitHub)

3. Journals, Code and Data

4. The IOM, the FDA, and IDEs*

5. The NCI and Trials it Funds

6. OSTP, Congress, Science, Nature
Some Places to Learn More

Karl Broman’s Tools for RR Course

Roger Peng’s Coursera course and notes (2013)

Christopher Gandrud’s book (2e, 2015)

Yihui Xie’s book (2e, 2015)


NAS meeting, Feb 26-7, 2015

SISBID Reproducible Research Short Course, July 2015
Our Own Stuff


http://bioinformatics.mdanderson.org/Supplements/ReproRsch-All/Modified/StarterSet

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

1st Summer Institute in Statistics for Big Data
Module 5: Reproducible Research Short Course, July 2015
Things We Consider Today

**Tools** for Making Reproducibility Easier
- Mindset: Begin with the End in Mind
- Mindset: Review with Expectation
- Literate Programming
- Code Packaging/Versioning/Sharing

**Expansion** from Reproduction to Replication
- Cutoffs and P-Values
- Getting Basic Stats Right
- Batch Effects

**Design and Big Data**

**Sanity Checks** and Controls

**Report Structure**

**Plotting Data**
Acknowledgments

Kevin Coombes

Steve Skates & UKCTOCS

Yang Zhao, Ying Wang, Shelley Herbrich
Shannon Neeley, Jing Wang
David Ransohoff, Gordon Mills
Jane Fridlyand, Lajos Pusztai, Zoltan Szallasi

M.D. Anderson Ovarian, Lung, Breast, Bladder SPOREs
Thanks!