The “Garbage In” Problem in Cancer Research

Carolyn Compton, MD, PhD
Professor of Life Sciences, ASU
Professor of Laboratory Medicine and Pathology, Mayo Medical School
Adjunct Professor of Pathology, Johns Hopkins Medical Institutes
CMO, National Biomarker Development Alliance
CMO, Complex Adaptive Systems Institute

AACR Education Session
Philadelphia, PA
April 18, 2015
I have the following financial relationships to disclose:
Consultant for Indiuvmed
Board of Directors of HealthTell

- and -

I will not discuss off label use and/or investigational use in my presentation.
Disappearing Line Of Demarcation Between Biomarker Discovery, Development And Clinical Use

- Molecular Data
- Clinical Care / Research
- Diagnosis / Therapy
- New Diagnostics
- New Therapeutics

**PRECISION MEDICINE**

**Biospecimen Processing and Stabilization**

**Biospecimen Collection**

**QUALITY HERE**

**Biospecimen Analysis**

**QUALITY HERE**

**DETERMINES QUALITY HERE**
Biospecimen Quality Impacts Both Clinical And Research Outcomes

Effects on Clinical Outcomes

- Potential for incorrect diagnosis
- Potential for incorrect treatment
  - Therapy linked to diagnostic test on a biospecimen

Effects on Research Outcomes

- Irreproducible results
  - Variation in mutation data
  - Variation in gene expression data
- Misinterpretation of artifacts as biomarkers
Amgen attempts to verify results of 53 landmark studies in oncology and hematology; Only 6 (11%) could be reproduced.

*Nature* 483, 531-533 doi:10.1038/483531a, 2012
Irreproducibility in Biomedical Research: A Crisis in Confidence (Public View)

Unreliable research

Trouble at the lab

Scientists like to think of science as self-correcting. To an alarming degree, it is not.

Oct 19th 2013 | From the print edition

Annals of Science

The Truth Wears Off

Is there something wrong with the scientific method?

By Jonah Lehrer

December 13, 2010

On September 18, 2007, a few dozen neuroscientists, psychiatrists, and drug-company executives gathered in a hotel conference room in Brussels to hear some startling news. It had to do with a class of drugs known as atypical or second-generation antipsychotics, which came on the market in the early nineties. The drugs, sold under brand names such as Abilify, Seroquel, and Zypprexa, had been tested on schizophrenics in several large clinical trials, all of which had demonstrated a dramatic decrease in the subjects’ psychiatric symptoms. As a result, second-generation antipsychotics had become one of the fastest-growing and most profitable pharmacological classes. By 2001, Eli Lilly’s Zypprexa was generating more revenue than Prozac. It remains the company’s top-selling drug.

Why Most Published Research Findings Are False

John P. A. Ioannidis

Published: August 30, 2005  DOI: 10.1371/journal.pmed.0020124

Abstract

Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to false relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller, when effect sizes are smaller, when there is a greater number and lesser preselection of tested relationships, where there is greater flexibility in designs, definitions, outcomes, and analytical modes, when there is greater financial and other interest and prejudice, and when more teams are involved in a scientific field in chase of statistical significance. Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias. In this essay, I discuss the implications of these problems for the conduct and interpretation of research.
How Widespread Are Failures to Reproduce Published Reports?

- Mass spec diagnostic for ovarian cancer – results due to experimental artifact and bias – control and experimental specimens collected differently and run separately (Lancet, 2002)
- Five of 7 largest molecular epidemiology cancer studies did not classify patients better than chance (JNCI, 96:2004)
- Microarray drug sensitivity signatures – from cell lines – to predict patient response (named one of top100 breakthroughs in 2006) could not be reproduced in large clinical trial in 2009 (Nature Medicine, 2006)
- Assessment of 18 published microarray studies: 2 were reproducible (Science, 2011)
- Bayer Healthcare reported reproducibility rates of 25% in its attempt to reproduce discovery research (Nature Reviews Drug Discovery 10, 712 doi:10.1038/nrd3439-c1, 2011)
Quality Analytical Data Begins with Quality Analytes

Garbage in...

Purgamentum init, exit purgamentum.

Diamonds in...

Modified from Jerry Thomas

...Garbage out
Biomarker Development: What’s the Problem Here?

Estimated number of papers documenting thousands of claimed biomarkers

150,000

Estimated number of biomarkers routinely used in the clinic

100

Source: Poste G. Nature 469, 156-157 13 Jan 2011
Biomarker: A measurable characteristic used as an indicator of a biological state or condition

- Drug development – markers of efficacy, toxicity and surrogate endpoints for clinical trials
- Early detection (broad or specific detection/ corroboration of specific disease stage)
- Rational choice of treatments (patient stratification)
- Assessment of treatment effectiveness
- Prognosis, prediction
- Prevention, surveillance
- Treatment, disease monitoring

No Biomarkers, No Precision Oncology
# Sources of Bias in Molecular Marker Research in Cancer - Ransohoff and Gourlay, 2010

<table>
<thead>
<tr>
<th>Source of Bias</th>
<th>Location of Bias: Before or After Specimens Are Received in the Laboratory</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Features of subjects, determined in selection:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>X</td>
<td>Cancer subjects are male, whereas control subjects are mainly female. Bias: Assay results may depend on sex.</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comorbid conditions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specimen collection</td>
<td>X</td>
<td>Cancer specimens come from one clinic, whereas controls come from a different clinic. Bias: Assay results may depend on conditions that differ between clinics.</td>
</tr>
<tr>
<td>Specimen storage and handling</td>
<td>X X</td>
<td>Cancer specimens are stored for 10 years because it takes longer to collect them, whereas control specimens are collected and stored over 1 year. Bias: Assay results may vary with duration of storage, or with different numbers of thaw-freeze cycles.</td>
</tr>
<tr>
<td>Specimen analysis</td>
<td>X</td>
<td>Cancer specimens are run on one day, whereas control specimens are run on a different day. Bias: Assay results may depend on day of analysis in a machine that “wanders” over time.</td>
</tr>
</tbody>
</table>

**NOTE.** The table shows examples of different sources of bias and the location of the bias before or after specimens are received in the laboratory. The list is not exhaustive; other biases may be important, and the biases listed may or may not be important in any given research study, depending on details of biology and technology (i.e., what is being measured and how it might be influenced).
The Vision Of Precision Oncology Cannot Be Realized Without Biomarkers

Biomarker

- A measurable characteristic serving as an indicator of a biological state or condition
- Most often measured from biospecimens
- Required characteristics:

  - Quantifiable
  - Reproducible
  - Clinically relevant

All of these can be distorted by pre-analytical variation
Pervasive Standards Deficits Contribute to the Lack of Progress in Biomarker Development

- Poor access to rigorously annotated, fit-for-purpose biospecimens from stringently phenotyped sources
- Insufficient control of pre-analytical parameters
- Low reproducibility of academic publications
- Variable analytical standards
- Idiosyncratic ‘lab-specific’ analytical methods
- Small studies lacking statistical power
- Chaotic data reporting formats and poor database interoperability
- Poor compliance with journal policies on reporting standards
- Non-existent quality management systems
Requirements for biospecimen quality are related to:

- The stringency of the analysis to be performed
- The requirements of the specific platform used
- The lability/stability of the molecular species to be analyzed
Pre-analytical Factors Affect Both Molecular Quality And Molecular Composition

Specimen is **viable** and biologically reactive

Molecular composition subject to further alteration/ degradation

Factors (examples):
- Antibiotics
- Time at room temperature

Factors (examples):
- Time 0
- Temperature of room
- Type of fixative
- Time in fixative
- Rate of freezing
- Size of aliquots

Pre-acquisition

Post-acquisition

Does it matter???
The facts:

- Between 32 and 75% of all laboratory test errors occur in the pre-analytical phase.
- Insufficient specimen quality (or quantity) may account for over 60% of pre-analytical errors.
- Genomic tests are not exempt from this issue.

Pre-analytical Variables: Impact on Test Results

HER2 IHC and FISH in Breast Cancer: Loss of Signal with Time to Fixation

pMAPK IHC of Colon Cancer: Gain of Signal with Time to Fixation


Hartmut Juhl, Indivumed GmbH, BRN
Pre-analytical Variables: Surgery and Pathology Contributions

Number of Genes Showing >2-Fold Change in Expression Level Pre vs. Post Surgery

Percentage of Patients with >2-Fold Change in Selected Protein Expression Level Pre vs. Post Surgery

Expression of >15% of genes and up 60% of selected proteins change >2-fold during surgery and postsurgical processing time

K David et al, Oncotarget November 2014
Blood Collection And Plasma Processing: Circulating Genomic Biomarkers And Tumor Cells
# Plasma Biomarkers: Pre-analytical Variations With Known Effects On Analyte Assays

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Variations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Venipuncture</td>
<td>Needle gauge&lt;br&gt;Preming volumes</td>
</tr>
<tr>
<td>Phlebotomy</td>
<td>Patient position (seated /reclining)&lt;br&gt;Tourniquet time&lt;br&gt;Tube orders&lt;br&gt;Venipuncture sites</td>
</tr>
<tr>
<td>Collection device</td>
<td>Tube types</td>
</tr>
<tr>
<td>Blood derivatives and processing</td>
<td>Anticoagulant types&lt;br&gt;Temperatures&lt;br&gt;Centrifugation speeds&lt;br&gt;Processing time</td>
</tr>
<tr>
<td>Time between collection and storage</td>
<td>Variable or unknown times</td>
</tr>
<tr>
<td>Storage and shipping</td>
<td>Temperature Duration</td>
</tr>
</tbody>
</table>
Evolution Of Biomarker Testing In The “Omics Era”

- Multianalyte Tests
- Whole Genome Sequencing

Unianalyte Tests

Raising the Bar for Specimen Quality!

Networks and Systems
The technological capacity exists to produce low-quality data from low-quality analytes with unprecedented efficiency.

We now have the ability to get the wrong answers with unprecedented speed.

Starting materials of known, consistent quality are required to assure analysis data of known, consistent quality.
NGS Is One Of Those Powerful Tools Moving Rapidly Into Clinical Application

- Rational choice of treatments (patient stratification)
- Assessment of treatment effectiveness / disease evolution
- Treatment/disease monitoring
- Risk assessment
- Prognosis (outcome)
- Early detection
It all starts with the “Right Stuff”.

- Biospecimens and Analysis of Molecular Pathway/Network Perturbations
- Multiplex Assays and Complex Signal Deconvolution Algorithms
- Novel Instrumentation, Automation and Large Scale Informatics
- Patient Profiling, Rational Rx and Health Monitoring

Courtesy of G. Poste
NBDA: Realizing an End-To-End, Standards-Based Approach to Biomarker Development

Early Discovery (Biology Verified Patient Samples)
Translatable Discovery (Clinical Measure Established)
Assay Development (Analyte - Reagents - Technology – Robust)
Assay Performance (Analytical Validation)
Biomarker Qualification (“Fit for Clinical Purpose)
Biomarker Validation (Clinical Validation)

Standards are needed at every step and across the continuum
Biospecimens Flank End-To-End Biomarker Development

The Continuous Feedback Loop of Quality

NBDA
National Biomarker Development Alliance
- Academic discovery scientists
- Clinical investigators
- Industry (pharma, biotech, diagnostics)
- CROs
- Clinicians
- Regulators
- Accreditation organizations
- Payors
- Patients

- Better science, greater efficiency, cost savings, better medicine – because there are patients waiting
The goal:

- Converge on the pre-analytical steps of the biospecimen lifecycle that most compromise the quality of tissue and blood for NGS and mass spec
  - “Top 10 List”

- Identify where the greatest value can be delivered in the control of pre-analytical variation (*biggest quality bang for the buck*)
  - “Top 3 List”

- Define the performance metrics required to achieve control of the highest-value variables

- Define a cost-effective strategy for implementation and compliance with those metrics
Think: Pareto Principle (20/80 rule)

For many events 80% of the effects come from 20% of the cause
1. Time to stabilization
   Tissue: Fixation within 1 hour
   Blood: N/A to blood extraction

2. Method of processing
   Tissue: Time in formalin 6-24 hours
   - Section thickness <3 mm
   Blood: Room temp (15-25° C)
   - Maintained in transport

3. Method of stabilization
   Tissue: Standardize formalin and tissue - fixative volume ratio
   Blood: 3 tubes: RNA, DNA
   optional specialty tube
   - Minimum 10 inversions

4. Metadata collected
   Tissue: Time to fixation
   Deviations
   Fixative QC
   Blood: Site (vein or line)
   Tourniquet
   Draw order
   Volume of tube fill

5. Storage conditions
   Tissue blocks: room temp (15-25° C)
   Blood analytes -80° C
The CAP Isoving Ahead

Goal:

- Implementation of the Top 5 through the College of American Pathologists (CAP) Laboratory Accreditation Program checklists
- New reimbursements codes sought, if needed
- Reinforcement through FDA guidance, funder requirements, etc.

Next steps:

- MOU between the NBDA and the CAP in process
- Personalized Healthcare Committee (PHC) of CAP begins education and implementation through the CAP Laboratory Accreditation Program
- PHC further develops, refines and updates key pre-analytics
Historic transformation of practice with far-reaching impact:

- Variably variable and unknown quality to uniform, known quality that is consistent with molecular analysis
- Simultaneous impact on both clinical and research results
- “Convenience samples” will be fit for purpose!
If your research involves human biospecimens, think sample quality.

“If you don’t have the time to do it right, when will you have the time to do it over?”

- John Wooden, Coach UCLA
The “Garbage In” Problem in Cancer Research

Carolyn Compton, MD, PhD
Professor of Life Sciences, ASU
Professor of Laboratory Medicine and Pathology, Mayo Medical School
Adjunct Professor of Pathology, Johns Hopkins Medical Institutes
CMO, National Biomarker Development Alliance