The Industry Experience Attempting to Validate Published Data

or

How to Identify the Papers Where the Data Just Won't Stand Up

PCAST

Jan 31, 2014

"Conflict of Interest" Statement

C. Glenn Begley, MB BS, PhD, FRACP, FRCPA, FRCPath

CSO & SVP TetraLogic, Malvern, PA
Non-Executive Director at Oxford BioTherapeutics, UK
SAB member for several, non-public biotech companies

Studies performed at Amgen Inc., 2002-2012
Vice President and Global Head, Hematology and Oncology Research
- continue to hold Amgen stock

Prior to Amgen: >20 years medical oncologist, hemato-pathologist
research scientist and department head

I will not discuss off-label use and/or investigational use of drugs

Drug development: Raise standards for preclinical cancer research
C. Glenn Begley & Lee M. Ellis, 28 March 2012
Nature 483, 531-533 doi:10.1038/483531a
Six Red Flags for Suspect Work
C. Glenn Begley
Nature 497, 433-434, 23 May 2013

Begley's position statement

- These results do not challenge the validity or legitimacy of the scientific method
- Not talking about fraud: the subject is scientific-laziness, sloppiness, exaggeration, desperation
- The vast majority of investigators want to do the right thing
- That this debate is occurring in public confirms the strength our scientific system

Cancer is Evolution

We don't get to "set the bar"

The challenge and opportunity for cancer research

1) **Challenge: Inherent to the disease, and extremely difficult to address**
   a. Cell lines are artificial
   b. Animal models are artificial
   c. Human cancers are heterogeneous in practically every dimension we care to think about:
      i. Tissue of origin
      ii. Genetic alterations
      iii. Host adaptations and interactions
      iv. Genetic instability
      v. Pathways of escape from growth control
      vi. Invasiveness and metastasis potential….etc…

2) **Opportunity: Inherent to our system, and <u>more easily addressed</u>**
   a. Poor experimental design
      i. Lack of blinding
      ii. Lack of adequate controls
      iii. Lack of prospective hypothesis statement
      iv. Lack of appropriate statistical power
   b. Poor reagents
   c. Poor analysis e.g. Inappropriate statistical methods
   d. Failure to reject hypothesis after observing discordant, valid experimental results
   e. Deliberate bias in selecting positive rather than negative results to report, publish, cite, fund…
      i. Scientists
      ii. Journal editors
      iii. Funding agencies
      iv. Press
   f. Failure to ask and follow through on ***"Why is this result not what I expected?"***

An unappreciated challenge to oncology drug discovery: publication bias

Sometimes we **<u>can</u>** "set the bar"

But we get what we incentivize

Industry relies heavily upon targets and pathways identified by academic groups

Between 2002-2012, Amgen was not able to reproduce 47of 53 seminal publications. These were publications that reported something completely "new" (not "binary" publications)

The spectrum of irreproducibility

- data could not reproduced by the original investigators with their reagents in their lab
- specific data reproduced, but not a general finding
- data selection bias: a single, non-representative experiment was reported

Investigators often required Amgen to sign a Confidentiality Agreement to allow an exchange of reagents and to allow Amgen scientists to work in their labs

These studies have had substantial impact

Wasted effort: multiple investigators,
multiple companies,
opportunity cost

Some papers have spawned a whole field with hundreds of secondary publications

Clinical studies initiated

Amgen's experience is, unfortunately, not unique

"…only in approx. 20–25% of the projects were the relevant published data completely in line with our in-house findings…..in most cases (this) resulted in termination of the projects because the evidence ….was insufficient to justify further investments into these projects"

Prinz, F., Schlange, T. & Asadullah, K. Believe it or not: how much can we rely on published data on potential drug targets? *Nat Rev Drug Discov* 10, 712,

Preclinical irreproducibility is a "systemic" problem, driven by current incentives

Careers are built on publications in "top-tier" journals:
drive grants, fame, promotion

Top-tier journals want "the best" story
simple, clear, compelling

"Positive" studies are rewarded:
little recognition of value of negative studies,
reward for "finding" the answer a Reviewer wants/expects

Although the Investigator and host institution are ultimately responsible and accountable, the greatest likelihood for change will come from journals and granting agencies

Although reducing incentives for publication would likely improve scientific rigor and quality, it is unrealistic

Highest probability for change will come from:
raising the standards of publication
encouraging publication of confirmatory data
rewarding findings that refute high-profile studies

High-profile studies typically fail at multiple levels:

Begley's six criteria for judging scientific reports:

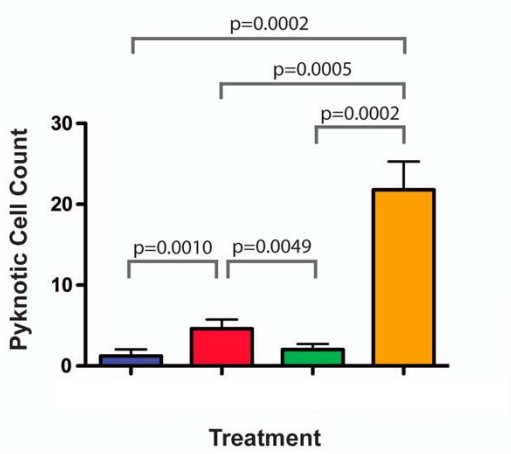
1) Were studies blinded?
   - Almost never
2) Were all results shown?
   - Typically not — "representative examples" & data selection bias; western blots that show only a slice; no size markers
3) Were experiments repeated?
   - Often not — westerns/immuno-precipitation usually only performed once; typically only use 1/2 siRNAs and in 1/2 cell lines; confusion between replicates and independent experiments
4) Were positive and negative controls shown?
   - Typically not
5) Were reagents validated?
   - Frequently not — IHC with a polyclonal anti-peptide Ab; small molecule inhibitors
6) Were the statistical tests appropriate?
   - Typically not

Nature 497, 433-434, 23 May 2013

Reviewers, editors of "top-tier" journals repeatedly tolerate poor quality science

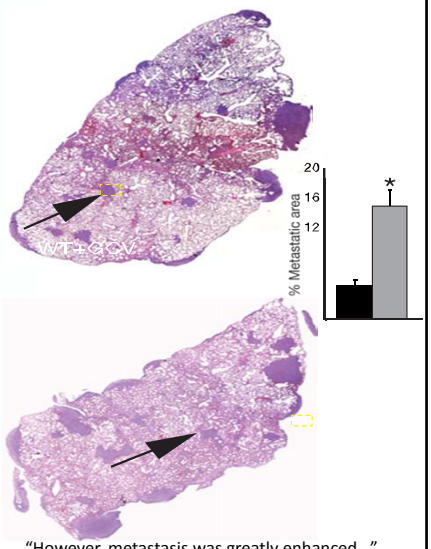
Some selected examples

1) Were studies blinded?



"Five separate fields of H&E stained paraffin tumor sections were quantified for the presence of pyknotic nuclei per square millimeter by visual inspection."

Nature Genetics, 2009

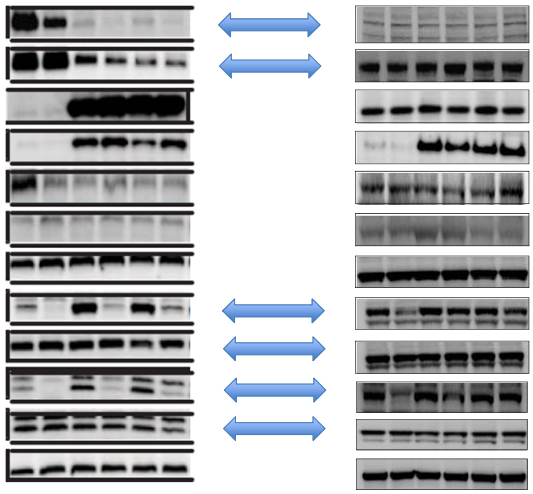


"However, metastasis was greatly enhanced…"

Cancer Cell, 2012

2) Were all the results shown?

The same antibodies from two different Figures show different patterns of bands

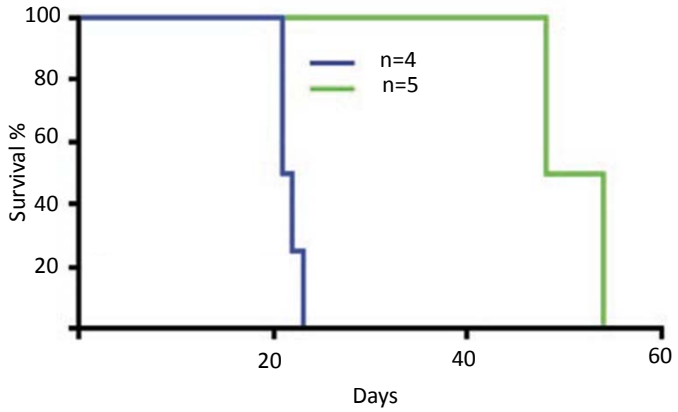


Nature, 2012

Were experiments ever repeated?
Were size standards used?

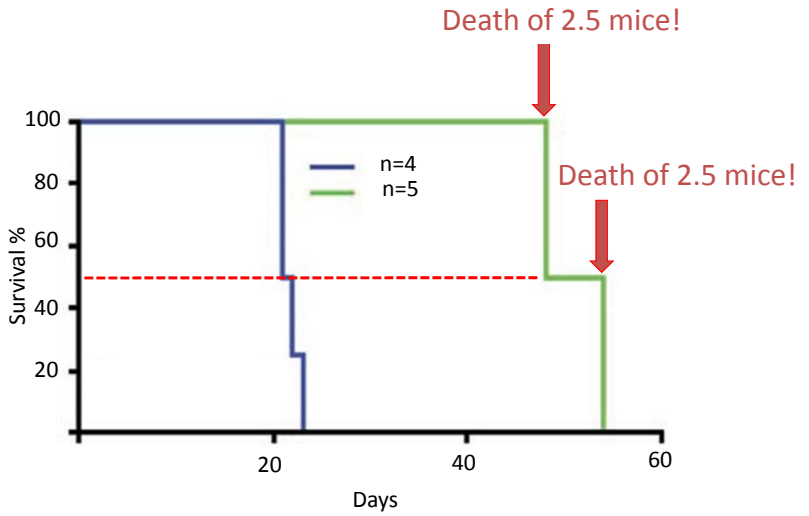
Were antibodies validated?
Were exposures in the 'linear range'?

2) Were experiments repeated?



Nature Genetics, 2009

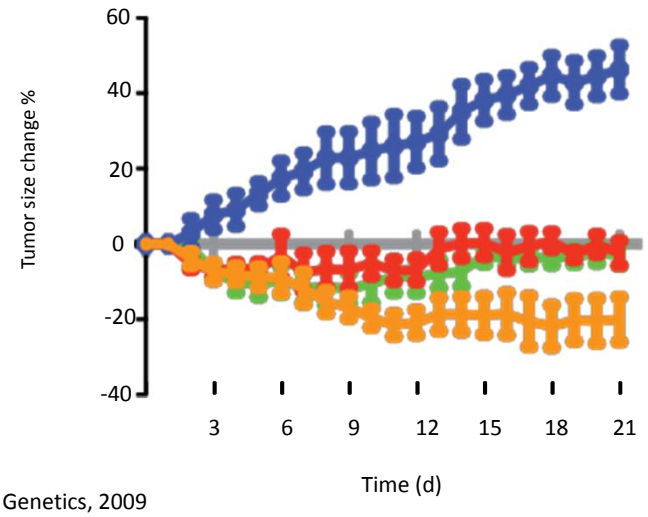
2) Were experiments repeated?



Was the paper actually read by the co-authors (n=10)? (Reviewers? Editors?)

Nature Genetics, 2009

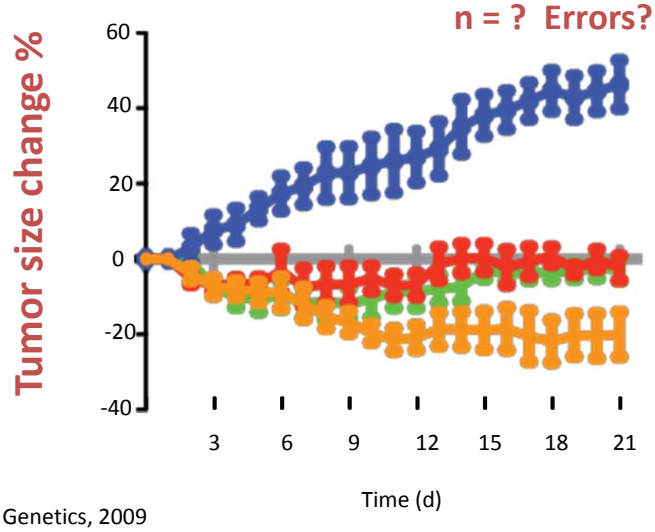
3) Were experiments repeated?



Nature Genetics, 2009

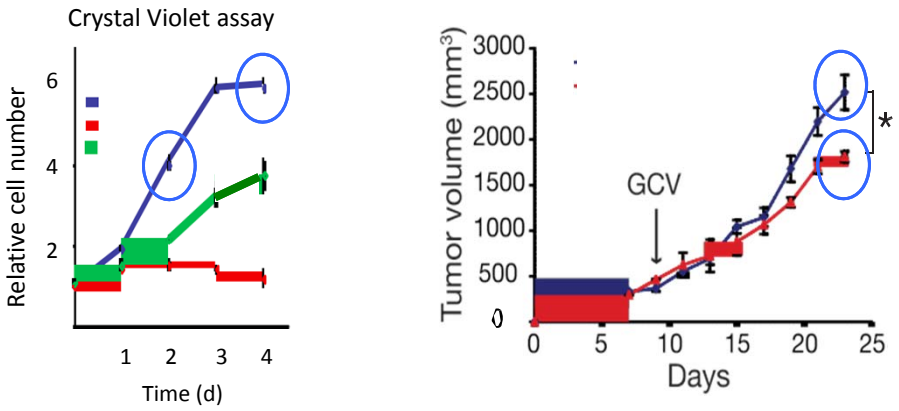
3) Were experiments repeated?

n = ? Errors? not stated



Nature Genetics, 2009

3) Were experiments repeated?



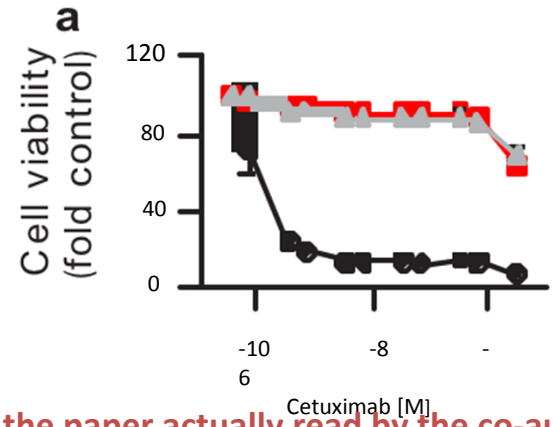
n=? Errors? Replicates? Repeats? **Not stated**

Cell proliferation is an exponential, not a linear function!

Nature Genetics, 2009 Cancer Cell, 2012

3) Were experiments repeated?

4 similar Figures showing <u>100-FOLD</u> increase over control



Was the paper actually read by the co-authors (n=26)? (Reviewers? Editors?)

Nature, 2012

4) Were positive and negative controls shown?

"…targeting of Met using PF2341066 inhibited the EMT program shift and suppressed metastasis…Although PF2341066 is also an inhibitor of ALK, quantitative RT-PCR (**data not shown**) and immunostaining revealed that ALK was not expressed in 4T1 tumors with or without PF2341066 treatment.."

AND

Ignoring the fact that PF2341066 targets 16 kinases

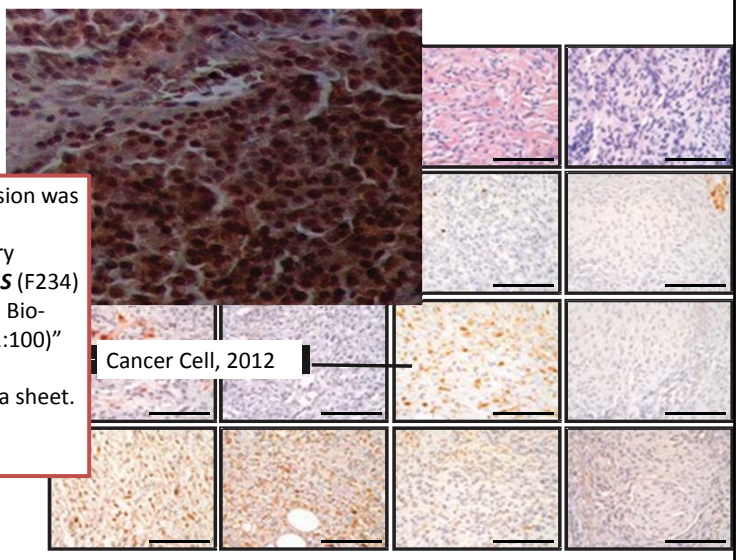
Cancer Cell, 2012

5) Were reagents validated?

"KRAS protein expression was evaluated by immunohistochemistry …using ***a specific KRAS*** (F234) antibody. (Santa Cruz Bio-technology; dilution 1:100)"

Versus Santa Cruz data sheet.

Nature, 2012



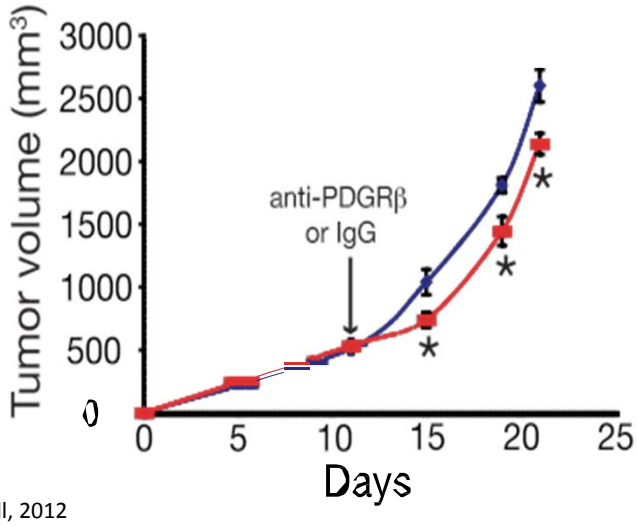
Cancer Cell, 2012

Nature Genetics, 2009

6) Were statistical tests appropriate?

n = ? not stated

Error bars display SEM; *p < 0.05.



Cancer Cell, 2012

Conclusions

We have a systemic problem

Our system tolerates (encourages?) these behaviors

The principal responsibility rests with the Investigator and the Host Institution

Patients deserve, and certainly expect, more

Acknowledgements

- The outstanding group of scientists in Hematology/Oncology Research at Amgen
- Lee Ellis
- Frank Calzone, Dan Hicklin, Fran Visco, Lanny Kirsch, Margaret Kripke, George Morstyn, Bill Sheridan

Recommendation:

Investigators, Institutions,
Reviewers, Funding Agencies, Consumers, Advocates, Press demand:

1) **STUDIES ARE BLINDED**
2) All results are shown
3) Experiments are repeated
4) Positive and negative controls are shown
5) Reagents are validated
6) Appropriate statistical tests are applied