

# Drug development: Raise standards for preclinical cancer research

C. Glenn Begley & & Lee M. Ellis

*Nature* **483**, 531–533 (29 March 2012)   doi:10.1038/483531a

Published online 28 March 2012

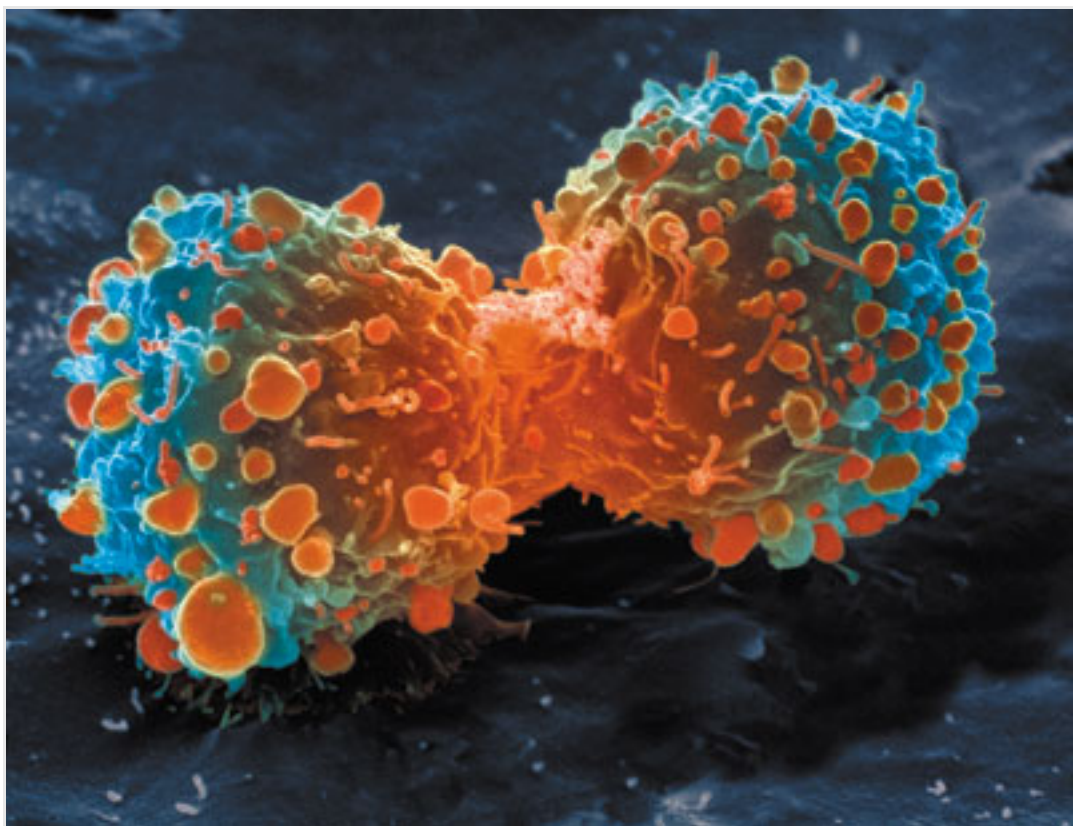
**Clarification (May, 2012)**

**C. Glenn Begley and Lee M. Ellis propose how methods, publications and incentives must change if patients are to benefit.**

**Subject terms:** Cancer   Drug discovery   Publishing

Efforts over the past decade to characterize the genetic alterations in human cancers have led to a better understanding of molecular drivers of this complex set of diseases. Although we in the cancer field hoped that this would lead to more effective drugs, historically, our ability to translate cancer research to clinical success has been remarkably low<sup>1</sup>. Sadly, clinical trials in oncology have the highest failure rate compared with other therapeutic areas. Given the high unmet need in oncology, it is understandable that barriers to clinical development may be lower than for other disease areas, and a larger number of drugs with suboptimal preclinical validation will enter oncology trials. However, this low success rate is not sustainable or acceptable, and investigators must reassess their approach to translating discovery research into greater clinical success and impact.

Many factors are responsible for the high failure rate, notwithstanding the inherently difficult nature of this disease. Certainly, the limitations of preclinical tools such as inadequate cancer-cell-line and mouse models<sup>2</sup> make it difficult for even the best scientists working in optimal conditions to make a discovery that will ultimately have an impact in the clinic. Issues related to clinical-trial design — such as uncontrolled phase II studies, a reliance on standard criteria for evaluating tumour response and the challenges of selecting patients prospectively — also play a significant part in the dismal success rate<sup>3</sup>.



S. GSCHMEISSNER/SPL

Many landmark findings in preclinical oncology research are not reproducible, in part because of inadequate cell lines and animal models.

Unquestionably, a significant contributor to failure in oncology trials is the quality of published preclinical data. Drug development relies heavily on the literature, especially with regards to new targets and biology. Moreover, clinical endpoints in cancer are defined mainly in terms of patient survival, rather than by the intermediate endpoints seen in other disciplines (for example, cholesterol levels for statins). Thus, it takes many years before the clinical applicability of initial preclinical observations is known. The results of preclinical studies must therefore be very robust to withstand the rigours and challenges of clinical trials, stemming from the heterogeneity of both tumours and patients.

**Confirming research findings**

The scientific community assumes that the claims in a preclinical study can be taken at face value — that although there might be some errors in detail, the main message of the paper can be relied on and the data will, for the most part, stand the test of time. Unfortunately, this is not always the case. Although the issue of irreproducible data has been discussed between scientists for decades, it has recently received greater attention (see [go.nature.com/q7i2up](https://go.nature.com/q7i2up)) as the costs of drug development have increased along with the number of late-stage clinical-trial failures and the demand for more effective therapies.

Over the past decade, before pursuing a particular line of research, scientists (including C.G.B.) in the haematology and oncology department at the biotechnology firm Amgen in Thousand Oaks, California, tried to confirm published findings related to that work. Fifty-three papers were deemed 'landmark' studies (see 'Reproducibility of research findings'). It was acknowledged from the outset that some of the data might not hold up, because papers were deliberately selected that described something completely new, such as fresh approaches to targeting cancers or alternative clinical uses for existing therapeutics. Nevertheless, scientific findings were confirmed in only 6 (11%) cases. Even knowing the limitations of preclinical research, this was a shocking result.

**Table 1: Reproducibility of research findings**

**Preclinical research generates many secondary publications, even when results cannot be reproduced.**

Of course, the validation attempts may have failed because of technical differences or difficulties, despite efforts to ensure that this was not the case. Additional models were also used in the validation, because to drive a drug-development programme it is essential that findings are sufficiently robust and applicable beyond the one narrow experimental model that may have been enough for publication. To address these concerns, when findings could not be reproduced, an attempt was made to contact the original authors, discuss the discrepant findings, exchange reagents and repeat experiments under the authors' direction, occasionally even in the laboratory of the original investigator. These investigators were all competent, well-meaning scientists who truly wanted to make advances in cancer research.

In studies for which findings could be reproduced, authors had paid close attention to controls, reagents, investigator bias and describing the complete data set. For results that could not be reproduced, however, data were not routinely analysed by investigators blinded to the experimental versus control groups. Investigators frequently presented the results of one experiment, such as a single Western-blot analysis. They sometimes said they presented specific experiments that supported their underlying hypothesis, but that were not reflective of the entire data set. There are no guidelines that require all data sets to be reported in a paper; often, original data are removed during the peer review and publication process.

Unfortunately, Amgen's findings are consistent with those of others in industry. A team at Bayer HealthCare in Germany last year reported<sup>4</sup> that only about 25% of published preclinical studies could be validated to the point at which projects could continue. Notably, published cancer research represented 70% of the studies analysed in that report, some of which might overlap with the 53 papers examined at Amgen.

Some non-reproducible preclinical papers had spawned an entire field, with hundreds of secondary publications that expanded on elements of the original observation, but did not actually seek to confirm or falsify its fundamental basis. More troubling, some of the research has triggered a series of clinical studies — suggesting that many patients had subjected themselves to a trial of a regimen or agent that probably wouldn't work.

These results, although disturbing, do not mean that the entire system is flawed. There are many examples of outstanding research that has been rapidly and reliably translated into clinical benefit. In 2011, several new cancer drugs were approved, built on robust preclinical data.

However, the inability of industry and clinical trials to validate results from the majority of publications on potential therapeutic targets suggests a general, systemic problem. On speaking with many investigators in academia and industry, we found widespread recognition of this issue.

**Improving the preclinical environment**

How can the robustness of published preclinical cancer research be increased? Clearly there are fundamental problems in both academia and industry in the way such research is conducted and reported. Addressing these systemic issues will require tremendous commitment and a desire to change the prevalent culture. Perhaps the most crucial element for change is to acknowledge that the bar for reproducibility in performing and presenting preclinical studies must be raised.

An enduring challenge in cancer-drug development lies in the erroneous use and misinterpretation of preclinical data from cell lines and animal models. The limitations of preclinical cancer models have been widely reviewed and are largely acknowledged by the field. They include the use of small numbers of poorly characterized tumour cell lines that inadequately recapitulate human disease, an inability to capture the human tumour environment, a poor appreciation of pharmacokinetics and pharmacodynamics, and the use of problematic endpoints and testing strategies. In addition, preclinical testing rarely includes predictive biomarkers that, when advanced to clinical trials, will help to distinguish those patients who are likely to benefit from a drug.

Wide recognition of the limitations in preclinical cancer studies means that business as usual is no longer an option. Cancer researchers must be more rigorous in their approach to preclinical studies. Given the inherent difficulties of mimicking the human micro-environment in preclinical research, reviewers and editors should demand greater thoroughness.

As with clinical studies, preclinical investigators should be blinded to the control and treatment arms, and use only rigorously validated reagents. All experiments should include and show appropriate positive and negative controls. Critical experiments should be repeated, preferably by different investigators in the same lab, and the entire data set must be represented in the final publication. For example, showing data from tumour models in which a drug is inactive, and may not completely fit an original hypothesis, is just as important as showing models in which the hypothesis was confirmed.

Studies should not be published using a single cell line or model, but should include a number of well-characterized cancer cell lines that are representative of the intended patient population. Cancer researchers must commit to making the difficult, time-consuming and costly transition towards new research tools, as well as adopting more robust, predictive tumour models and improved validation strategies. Similarly, efforts to identify patient-selection biomarkers should be mandatory at the outset of drug development.

Ultimately, however, the responsibility for design, analysis and presentation of data rests with investigators, the laboratory and the host institution. All are accountable for poor experimental design, a lack of robust supportive data or selective data presentation. The scientific process demands the highest standards of quality, ethics and rigour.

**“The scientific process demands the highest standards of quality, ethics and rigour.”**

**Building a stronger system**

What reasons underlie the publication of erroneous, selective or irreproducible data? The academic system and peer-review process tolerates and perhaps even inadvertently encourages such conduct<sup>5</sup>. To obtain funding, a job, promotion or tenure, researchers need a strong publication record, often including a first-authored high-impact publication. Journal editors, reviewers and grant-review committees often look for a scientific finding that is simple, clear and complete — a 'perfect' story. It is therefore tempting for investigators to submit selected data sets for publication, or even to massage data to fit the underlying hypothesis.

But there are no perfect stories in biology. In fact, gaps in stories can provide opportunities for further research — for example, a treatment that may work in only some cell lines may allow elucidation of markers of sensitivity or resistance. Journals and grant reviewers must allow for the presentation of imperfect stories, and recognize and reward reproducible results, so that scientists feel less pressure to tell an impossibly perfect story to advance their careers.

Although reviewers, editors and grant-committee members share some responsibility for flaws in the system, investigators must be accountable for the data they generate, analyse and submit. We in the field must remain focused on the purpose of cancer research: to

improve the lives of patients. Success in our own careers should be a consequence of outstanding research that has an impact on patients.

The lack of rigour that currently exists around generation and analysis of preclinical data is reminiscent of the situation in clinical research about 50 years ago. The changes that have taken place in clinical-trials processes over that time indicate that changes in prevailing attitudes and philosophies can occur (see 'Improving the reliability of preclinical cancer studies').

**Box 1: Recommendations: *Improving the reliability of preclinical cancer studies***

Full box

Improving preclinical cancer research to the point at which it is reproducible and translatable to clinical-trial success will be an extraordinarily difficult challenge. However, it is important to remember that patients are at the centre of all these efforts. If we in the field forget this, it is easy to lose our sense of focus, transparency and urgency. Cancer researchers are funded by community taxes and by the hard work and philanthropic donations of advocates. More importantly, patients rely on us to embrace innovation, make advances and deliver new therapies that will improve their lives. Although hundreds of thousands of research papers are published annually, too few clinical successes have been produced given the public investment of significant financial resources. We need a system that will facilitate a transparent discovery process that frequently and consistently leads to significant patient benefit.

**References**

- 1. Hutchinson, L. & Kirk, R. *Nature Rev. Clin. Oncol.* **8**, 189–190 (2011).
- 2. Francia, G. & Kerbel, R. S. *Nature Biotechnol.* **28**, 561–562 (2010).
- 3. Rubin, E. H. & Gilliland, D. G. *Nature Rev. Clin. Oncol.* <http://dx.doi.org/10.1038/nrclinonc.2012.22> (2012).
- 4. Prinz, F., Schlange, T. & Asadullah, K. *Nature Rev. Drug Discov.* **10**, 712 (2011).
- 5. Fanelli, D. *PLoS ONE* **5**, e10271 (2010).

Download references

**Author information**

**Affiliations**

C. Glenn Begley is a consultant and former vice-president and global head of Hematology and Oncology Research at Amgen, Thousand Oaks, California 91359, USA.

Lee M. Ellis is at the University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030, USA.

**Corresponding author**

Correspondence to: Lee M. Ellis

**Comments**

2012-03-28 06:47 AM

**Sander Heinsalu said:** Publication bias is a problem in all fields of research. The results of a paper should actually receive zero weight in the evaluation of its quality, otherwise there is the motivation to cherry-pick the data that give the most impressive result. The measure of quality should be the way the results were obtained – size of sample, experimental procedure, endpoints used. Ideally the

reviewers of a paper should not see its results at all, only the description of the experiment.

2012-03-29 10:50 AM

**Thomas Hanscheid said:** Publication bias of preclinical and basic science in high ranking journals may well have more reasons, perhaps, often the lack of some useful and necessary clinical input. I am always astonished to see that many of my (highly regarded) research colleagues in malaria have never seen a case of malaria, never diagnosed one case, or possibly thought about all the implications of treatment. In fact, quite a few have never been in a malaria endemic country, let alone set foot into a hospital or health care post there. I want to believe that many of the reviewers of their submitted papers are not like this, but some doubts remain. I am not even so sure about the Editorial Boards of these high ranking life-science journals. Thus, highly elegant studies (often in "models" of some form) presenting some impressive new mechanisms or new insight, usually based on cutting edge methodology in molecular biology, genetics or some intricate immunology pathways are written, edited and reviewed by "peers", in the literal sense of this word. I often wonder, if not quite a few of these scientific papers might look different (in a positive sense) if an experienced clinician/researcher would have also been involved (in the sense of being a "slightly different" peer) in the publication process.

2012-03-30 08:18 AM

**Alex Reynolds said:** If it is not possible to see what research was reviewed by Begley and Ellis, does this not lead one to ask if their own work is reproducible?

2012-04-02 03:39 AM

**Greg M said:** The claims presented here are pretty outlandish. Particularly relevant to "Hematology and Oncology" we now know that mice housed under different conditions with different microflora can have vastly different outcomes in any model, not just cancer. To suggest academic incompetence or outright unethical behavior is offensive, and is a particularly narrow view of why experiments are difficult to reproduce. Further, as indicated in Table 1, the entire definition of not-reproducible hinges on a priori profit motive of "robust" differences (whatever that means). There is always room for improvement in science, but this entire article is disingenuous and belittling to those of us who are on the front lines.

2012-04-06 12:45 PM

**Marcelo Behar said:** At first I thought this was an April Fool's joke: an article complaining about non-reproducible results and poor publishing practices that did not show the data underlying their own "results". I laughed at loud at the claim "The scientific community assumes that the claims in a preclinical study can be taken at face value"... thought it was pretty hilarious.

But I am not so sure this is a prank... so just in case here are my 2 cents.

I will not deny that cherry-picked results, poor controls, inadequate number of repeats, non-publishable negative results, or bad experimental habits in general are real problems in all scientific disciplines including biomedical research. However, this article is just sensationalism at it worst: making over-generalizing, grandiose claims without providing any supporting evidence. Which specific articles were picked, what criteria was used to categorize something as a Landmark finding, how were the claims tested, what reproducibility criteria were used, etc... speaking of cherry picked results, lack of controls, and poor publishing standards!

I am not familiar with the internal decision-making process in big pharma but if this article is serious, perhaps they should consider hiring scientists from a community that **does not** "assumes that the claims in a preclinical study can be taken at face value". Leaving aside dishonest data manipulation, problems arising from incomplete data, bad controls, poor practices, or limited applicability of the

results are usually evident from a critical review of the protocols/methods.

Cheers

2012-04-13 05:40 AM

**Uli Jung said:** While I applaud every kind of whistle blowing that helps improve transparency and decrease fraud and such in science – this comment is rather ridiculous. Half-way whistle blowing does not work, sorry. Who finds fraud, write about it in a reputable magazine but does not make sure the people committing it are actually exposed ... What is that, a self-serving publicity boost while supporting the network of the fraudulent actions through silence?

Usually there is that "conflict of interest" declaration in publications. A comment where someone labelled as "consultant" acts as the "good guy" who gets "halfway" exposes fraud but diligently avoids supporting the credibility (and maybe avoid personal attacks?) by not exposing the data of the claim ... is not needing a conflict of interest declaration?

2012-04-23 08:08 AM

**Bjoern Brembs said:** I'd concur with the previous commenters on the need for full disclosure of the research results. For instance, was reproducibility in any way correlated with journal rank as defined by IF? Thanks you.

2012-12-14 12:58 PM

**Milorad Milun said:** Reproducibility is the crucial part of scientific method. If indeed "mice housed under different conditions with different microflora can have vastly different outcomes in any model" (quot. from one of the comments above) then I wonder if mice presents a useful model-study that can be used outside a particular lab and on which further investigations can be built upon outside that one lab.

2015-02-18 08:02 AM

**Will kartr said:** Drug development is not an easy job to be done. Only the equipment use to create the drugs is need a good amount of investment. Any how for the better results it's quiet important to raise the standards. The perfection and the accuracy is first step towards the bright future of drug development.

William from <http://onlinewpmttest.org>

Subscribe to comments

*Nature*    ISSN 0028-0836    EISSN 1476-4687

SPRINGER NATURE

© 2012 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.  
partner of AGORA, HINARI, OARE, INASP, ORCID, CrossRef, COUNTER and COPE

