



Published October 28th, 2011



David Grainger
TCP Innovations

David Grainger is an academic in the Department of Medicine, [Cambridge University](#), researching mechanisms underpinning chronic inflammatory diseases. He is also a leading consultant to the pharmaceutical industry through [TCP Innovations Ltd](#), and is the Chief Scientific Officer of [Funxional Therapeutics Ltd](#), a Cambridge-based biotechnology company he founded in 2005, which develops novel anti-inflammatory drugs. He delivers his often iconoclastic opinions on recent trends in life sciences industries through the Drug Baron blog.

Rotten Foundations: How defects in academic publishing are the biggest threat to the commercial life sciences sector

The cherished principles governing the publication of academic science haven't changed in a hundred years, with concept of peer review at its core – but this ancient system is failing, with catastrophic consequences for life science innovators and investors alike. It might serve academia well enough, but they are not the only constituency who rely on the publicly funded body of knowledge that is the scientific literature. Commercial innovation in the life sciences has its foundations in the public knowledge base – and the needs of these innovators, the people who convert knowledge into products that benefit us all, are no longer being served by the current system. It is time to demand a change.

It is not just a lack of competence that leads to faulty information being incorporated into the edifice of the scientific literature. The few proven cases of malpractice are surely just the tip of the iceberg. The likelihood of being caught out is so low, particularly in biological science, that it stretches credulity to breaking point to assume that scientists are inherently more honest than the rest of humanity. But whether the cause is error or malice, the fact remains that most of the literature in biological sciences is irrelevant, much is plain wrong and a surprising fraction of the remainder is deliberately misleading.

“the billions spent on academic research in biology yields precious little value beyond the self-serving metrics of ever-rising citation indices”

A radical shake up of how we publish biological science (and the problems are much more severe in biology than in 'harder' sciences such as chemistry or physics) is sorely needed to restore faith in the system, and to ensure the vast array of data that is generated at public expense is accessible and useful. If we do not, our innovators will drown in irrelevant information, be led down blind alleys by the factually inaccurate and eventually see their productivity lowered to the point where its economically suicidal to continue. You may not even have noticed yet, but a crisis of confidence in the life sciences, as deep as the global economic crisis, is already in full swing.

So what's the evidence that there is a problem in the first place? Why does it matter so much? And assuming we are convinced by the answers to these two questions, what can we possibly do about it?

For many in the industry it has for many years been an unspoken rule that at least 50% of studies from academic laboratories cannot be repeated in an industrial setting (it is only DrugBaron whose pessimism set that figure at 90%). But such opinions, however strongly held, are just opinions. More recently, though, data – difficult to obtain though it is – has started to back-up these opinions: John Ioannidis and his colleagues ([Ioannidis et al 2009](#) *Nature Genetics* **41:149**) largely failed to reproduce the findings from published gene array analyses despite working from the original datasets; Alexander Bell and colleagues ([Bell et al \(2009\)](#) *Nature Methods* **6:423**) didn't get the same results when sending the same sample to a number of independent proteomics laboratories; and Florian Prinz and colleagues at Bayer ([Prinz et al \(2011\)](#) *Nature Rev Drug Discovery* **10:712**) reproduced only 14 out of 67 (20%) of the biological studies they attempted to replicate. So maybe DrugBaron's 90% failure rate is so pessimistic. But if it isn't pessimistic its certainly depressing.

The fundamental problem is that biology IS inherently difficult to reproduce experimentally because it involves the properties of complex systems (cells, animals, humans). So this gives biologists an "excuse" – they can publish essentially anything and if someone else cannot replicate it, they just shrug their shoulders and say "Well, their cells must have been different from my cells". You cannot do that in chemistry or physics or computing or engineering – which is why the crisis in confidence in the scientific literature is essentially unique to biology.

In most cases, the consequence of bad science is "just another" failed clinical trial

The problem is made worse by the incentive structure we ask academic scientists to work under (with citation indices as the measure of productivity, or with the need to get grants based on exciting publications). Inevitably, this creates an incentive to bend or even break the rules to get ahead. But the inherent difficulty in reproducing experiments with complex biological systems means the unscrupulous know that they are very unlikely to get caught red-handed. They always have the "well that's what happened in MY experiment" argument, and the inherent variability in biology means its hard for the reader to pick the genuinely irreproducible from the plain wrong. Its a bit like the number of times people blame faulty email communications when they just didn't get round to doing something! After all, they know you cannot PROVE they hadn't sent an email that somehow got lost in the ether...

So at one extreme, there is undoubtedly outright fraud. This then merges into the data manipulation (dropping outliers to improve statistical significance, selecting the best rather than the typical experiment), and then into the widespread incompetence (most biologists understand virtually no statistics, so they use the wrong statistical frameworks for their experiments and so, genuinely believing the outcome, they publish false conclusions).

Neither of these factors would matter much, but for the almost complete failure of peer review as a mechanism to check for quality. Clearly, the reviewers can be on average no more competent than the authors (they are the same people after all), and worse still they never get to see the raw data. It can be nearly impossible to see what ACTUALLY went on behind the scenes when you get a nice polished manuscript that purports to tell an interesting tale.

The fundamental problem is that biology IS inherently difficult to reproduce experimentally – which gives biologists an "excuse"

DrugBaron, in his fifteen years as head of a [lab](#) studying cytokines in inflammation at one of the world's top academic institutions, Cambridge University, has seen his fair share of examples of this kind of bias in operation.

One group of scientists published over and over that a particular protein, called thrombospondin, activates the cytokine TGF-beta – eventually published even in *Cell*, one of the most well respected life science journals. Our group then discovered an artifact that underpinned their observations and *Cell*, perhaps predictably, politely declined to publish our demonstration. So did several other journals (without any technical criticism of our work). It was eventually published in the lowly *Biochemical Journal*. Today, nearly ten years later, if you review the literature, thrombospondin is universally cited as an activator of TGF-beta in reviews and the like even though the original observations have not been replicated and a definitive demonstration of the opposite has been published. Bottom line: you don't make a career out of proving other people wrong (and even if you do, no-one takes any notice of you anyway).

It doesn't even help if you point out your own errors. DrugBaron and his colleagues published, in *Nature Medicine*, that TGF-beta levels in blood are altered in people with heart disease. A few years later, we discovered that we had overstated the magnitude of the effect because of an artifact in our assay system we hadn't realised was there. We published a follow-on paper pointing out the artifact and moderating our original conclusions. Not only does almost everyone ignore the second paper and cite the first one at regular intervals, there are now more than a dozen separate replications of our original observation even though we know (and have published) that it was emphasized by an artifact. It seems likely these dozen apparent replications are also flawed.

You don't even have to do experiments to introduce bias: DrugBaron proposed, in 2006, that the protein PAI-1 was an inhibitor of the enzyme furin (based on bioinformatics) and that this could be a major pathogenic mechanism in diabetes. We published a substantial paper in *Bioessays* setting out what is universally recognised as a beautiful hypothesis. In fact the only trouble with it as a hypothesis is that its NOT RIGHT! In five years (and with hundreds of experiments) we gathered hardly a shred of evidence that PAI-1 really does inhibit furin. Of course, no-one wants to publish a negative experiment, so the hypothesis remained apparently untested. That alone would be bad enough – but then at the end of 2010 a French group stunned us, and most likely the rest of the diabetes world as well, with a publication apparently PROVING our hypothesis! They had done the same experiments as our lab, but whereas we saw little or no evidence that PAI-1 inhibits furin, they got a completely different answer – every experiment they reported showed a strong effect, and they concluded our hypothesis is correct. There must be a real risk that someone somewhere is treating this as a validated target for developing new therapeutics (DrugBaron certainly would if the published experiments were the only data he had access to).

**The people who convert knowledge into products that
benefit us all are no longer being served by the current system.
It is time to demand a change**

To muddy the waters even further, though, a failure to replicate an experiment can point to interesting new science, rather to incompetence or fraud. Several labs couldn't replicate our findings with broad-spectrum chemokine inhibitors (BSCIs) on human neutrophil migration. Here, though, there was a "real" reason – one that surprised and interested us. Responsiveness to BSCIs among neutrophils depends on the activation status of the cells (which differed between the labs).

You just can't tell when A cannot replicate B's experiments who got it wrong: A or B. Indeed, you have to be really careful that you really were replicating the experiment before you can judge if there is an issue at all. If you set out to replicate something, you really must replicate all the key details. You cannot do something a bit different because you want to extend the observations (or rather, you can and should do the latter, but you cannot call it an attempt at replication).

Whatever the cause of the lack of reproducibility of the biological sciences literature, the consequence is clear: you cannot just assume anything that you read in the literature is correct. The impact of the faulty reports spreads wider than the actual prevalence of error: it is analogous to the benefits of vaccination against infectious disease, where the benefit accrues not from any individual being vaccinated but from the whole population doing so. Once a significant minority are unvaccinated, then the disease can take hold (as we saw following the flawed concerns about the MMR vaccine – another illustration of how peer review failed to stop bad science being published with devastating consequences for those families whose children were blinded by measles). Similarly, once a significant minority of published studies are known to be wrong (whether through malice or incompetence it matters not) then the ability to trust any of it is lost. With no means to tell the fantastic from the fanciful, you are left with no choice but to become a career skeptic.

Yet our entire industry is built on scientific discovery, and is founded in the public knowledge base that is the scientific literature. If you dare not trust any of it, where does that leave life science investors?

As the examples above show, you can't even rely on replication within the literature – once something is published, it's quite likely similar findings will appear in the literature even if the original observations were flawed. That's pretty scary.

The only solution then is “wet diligence” – doing experiments yourself, or with people you trust. Ironically, attempting replication might not be the best way to do it: as our experience with BSCIs and human neutrophils showed us, it can be very hard to achieve perfect replication. An easier approach can be to make some predictions based on the published observations: predictions that need to be true if the underlying concept has any value. Then test some of those predictions experimentally. If the results mostly turn out consistent, you have the strongest available green light to proceed. But if you start to see inexplicable or inconsistent outcomes, then the risk is probably too high, no matter how promising the initial observations seem on the surface.

Without a root-and-branch overhaul of the publishing system, though, the huge public expenditure on biological research will continue to be largely wasted

The key is internal consistency. Most “made up” or “incompetent” science can be readily spotted by the well-trained eye. DrugBaron has used the ‘internal consistency’ test a lot on scientific publications: so much so that when lecturing Cambridge undergraduates many moons ago, he used to use a particular example of a *Cell* paper (again, I chose *Cell* because of its reputation for quality), a *Cell* paper by a Nobel laureate (who shall remain nameless to avoid embarrassment). It was, as you might guess, riddled with internal consistency errors: things like having the control levels in each group of experiments being statistically significantly different from each other, despite apparently being replications of the same experiment (the test arm differed in each group of experiments, but the control arms should not have been different). If the experimental system is valid, then while the control level will vary from experiment to experiment, simply because biological systems do vary, they will not show similarity in one group of experiments and then similarity around a different mean in a separate group of experiments. Such behaviour is the hallmark of data manipulation (while not understanding what you are doing) or else incompetence in executing the experiments. Either way, the conclusions drawn from such experiments should be treated with suspicion.

If you rely on published science for your living (in a pharmaceutical company or as a life sciences investor, for example) there are three simple tips DrugBaron can offer:

(1) Assume all biological research you see published is wrong unless you can find evidence to the contrary (rather than the other way round)

(2) Check every study you care about with a fine-toothed comb for internal inconsistencies, and if you find any, be very wary of the conclusions

(3) Make predictions based on the claimed conclusions and test the easiest ones experimentally (this is usually easier than trying to set up a perfect replication) and if the easy, cheap data you generate to test the predictions is inconsistent with what they are claiming, the original observations should be considered unreliable

In a few cases, enough evidence is gathered to demonstrate that the published science underlying clinical trials was false (or falsified). Duke University had to suspend cancer clinical trials using biomarkers to personalize treatment when it became clear that the underlying science most likely had been faked. Within weeks of the original publications, in the highly respected *New England Journal of Medicine* and then in *Nature Medicine*, some independent biostatisticians had started to find problems. They obtained the original data from the authors, but still found the approach was little better than chance (and much worse than claimed in the original manuscripts). A host of errors were found that eventually led to the suspension of the clinical trials based on the original studies. The whole house of cards eventually came tumbling down when it was reported that the senior author on the paper had lied repeatedly about his own experience and training. Quoted in *The Economist*, the biostatistician that originally uncovered the errors said “I find it ironic that we have been yelling for years about the science, which has the potential to be very damaging to patients, but that was not what has [eventually got the trial stopped].”

But in most cases, the outcome is “just another” failed clinical trial. Time and resources were wasted, the productivity of R&D investment declined further, but there was no way to distinguish a failure to extrapolate from high quality basic research on cells in culture and in animal models to the human condition, from a failure that was due to rotten foundations in incompetent or fraudulent scientific research.

The falling rate of Phase 2 trial successes suggests the problem is getting worse and worse (if the growing knowledge base were reliable, given that our collective skill and experience at conducting trials is improving all the while, the trend should be going in the opposite direction).

**Once a significant minority of published studies are known to be wrong
(whether through malice or incompetence it matters not)
then the ability to trust any of the literature is lost**

DrugBaron’s handy hints on how to “qualify” published scientific reports may help in the short term but systemic changes are needed in the way biological science data is published. “Open innovation” platforms, where a pre-competitive consortium assesses the reproducibility of reported findings will certainly emerge. The advantage that accrues from consigning the worst examples of irreproducible science to the dustbin should more than offset any loss of competitive edge. If such approaches became common-place they may even instigate a virtuous circle – if scientists felt their more interesting observations might be subject to a collective assessment of reproducibility, they may be more cautious about what they publish in the first place.

Without a root-and-branch overhaul of the publishing system, though, the huge public expenditure on biological research will continue to be largely wasted. Until the changes restore the confidence of the true “customers” for this public knowledge base, the commercial innovators and entrepreneurs who create from the building-blocks of blue-skies science the products that benefit us all, the billions spent on academic research in biology will continue to yield precious little value beyond the self-serving metrics of ever-rising citation indices.