thebmj

BMJ 2018;360:k959 doi: 10.1136/bmj.k959 (Published 20 March 2018)

ANALYSIS

Trials that say "maybe": the disconnect between exploratory and confirmatory testing after drug approval

Clinical trials that explore the repurposing of drugs for off-label uses are common. But without a commitment to rigorously testing the hypotheses generated by these exploratory trials, patients are put in harm's way, argue **Benjamin Carlisle and colleagues**

Benjamin Carlisle *PhD student*, Carole A Federico *PhD student*, Jonathan Kimmelman *associate professor*

Studies of Translation, Ethics and Medicine (STREAM), Biomedical Ethics Unit, McGill University, 3647 rue Peel, Montreal QC H3A 1X1, Canada

Key messages

- After a new drug receives approval, companies and public sponsors often run numerous small trials exploring the drug's activity in different indications
- The level of evidence produced in such trials is usually low, and drug companies and public sponsors often fail to follow up on promising exploratory findings by running large, confirmatory trials
- The poor uptake of confirmatory testing often results in prolonged periods where ineffective and costly drugs are used off-label; the sacrifice of patients who participated in the original small trials goes unrealised
- Ethics committees and policy makers should devise measures that encourage better coordination between exploratory and confirmatory trials after drugs are approved

Most clinical trials are directed not towards getting new drugs approved but at repurposing already licensed drugs for new applications.[1] When drugs are first licensed, companies, public funders, and medical centres often mount vigorous trial programmes exploring a drug's activity in other indications or in combination with other treatments. These studies generate myriad hypotheses about possible treatment options for patients.

This uncertainty about clinical value, which we call "clinical agnosticism,"[2] provides grounds for rigorous, hypothesis testing trials. But, as we describe below, confirmatory trials are often not swiftly performed. This delay has ethical implications: the sacrifice of research volunteers in exploratory studies goes unappreciated, and other patients and healthcare systems are at risk of using ineffective and unsafe treatments for many years.

How post-license trials create clinical agnosticism

The primary aim of exploratory trials (typically phase I and II trials) is to produce evidence that a new treatment might be effective and might compete with standard of care. Favourable results establish a moral basis for randomising patients in larger, confirmatory trials (typically phase III). But, because exploratory trials are based on surrogate endpoints, small sample sizes, limited follow up, or unplanned subgroup analyses, they are rarely sufficient to guide clinical practice.[3 4] About 60% of drugs fail to maintain their promise in confirmatory trials,[1] making these trials critical for measuring a drug's value.[5]

For unlicensed drugs, regulators generally require confirmatory trials before they grant full marketing approval. Regulations provide powerful incentives for drug companies to resolve uncertainties quickly.[6] They also make it likely that ineffective treatments are intercepted before reaching clinical practice.

After approval, however, physicians and clinical practice guidelines can recommend off-label prescription based on the evidence generated in exploratory trials, and encouraging findings are often taken up into practice without proof of value.

Sorafenib and sunitinib

We looked at the cancer drugs sorafenib and sunitinib, which have over 10 years of exploratory trial activity since approval (in 2005 and 2006, respectively). Of all monotherapy trials launched within five years of their US approval (n=132),[7 8] we found 41 (31%) that met their primary efficacy endpoint with acceptable toxicity. Five of these trials were the basis of four recommendations for off-label use in clinical practice guidelines by the National Comprehensive Cancer Network (NCCN), an organisation whose guidelines in the US are used to set reimbursement policies. Seven other recommendations were based on four negative trials that reported subgroups of patients that responded to the drugs. But in only one of the 13 total NCCN recommendations of sunitinib or sorafenib for off-label use was the hypothesis generated in exploratory studies followed by a completed randomised trial using either overall survival or progression free survival as primary endpoint (table 1).

Correspondence to: J Kimmelman jonathan.kimmelman@mcgill.ca

Pregabalin

Pregabalin was approved in 2004 for the treatment of partial seizures, diabetic peripheral neuropathy, and post-herpetic neuralgia (box 1). It is widely used off-label for acute, sub-acute, and chronic non-cancer pain (such as low back pain) on the basis of exploratory evidence.[12] A meta-analysis published in 2017—seven years after the first test of pregabalin in chronic low back pain—found that the existing evidence was limited for gabapentinoid activity and that large, high quality trials were needed.[13] To our knowledge, no large, randomised trials testing pregabalin in low back pain have been published to date.

Box 1: Exploratory evidence for pregabalin

In 2004, the FDA approved pregabalin for the treatment of partial seizures, diabetic peripheral neuropathy, and post-herpetic neuralgia

On the basis of exploratory evidence, pregabalin is widely used off-label; in one Canadian study, 75% (268/355) of patients taking pregabalin for chronic non-cancer pain did so for unapproved, off-label pain conditions[9]

In 2009, Pfizer paid $2.3bn (£1.7bn; €1.9bn) to settle claims relating to the off-label promotion of pregabalin and other drugs[10]; these promotions often drew on exploratory evidence, and cost the US healthcare system hundreds of millions of dollars[11]

No large, randomised trials testing pregabalin in low back pain have been published to date

Without confirmatory evidence, we don't know whether pregabalin is effective for low back pain or whether healthcare resources would be better directed at proven treatments

Unresolved clinical agnosticism is harmful

The moral basis for exposing patients to unproven treatments in exploratory trials derives from the expectation that, if positive, they will lead to confirmatory testing. Researchers and sponsors have obligations to build on encouraging findings deriving from such exposures.

Prolonged clinical agnosticism can harm patients if doctors use treatments that are ineffective, even with the best of intentions. In debilitating disease, where standard of care options are exhausted, suspicion (rather than proof) of efficacy is often sufficient to influence treatment. In cancer, for example, up to one third of treatments are given off-label14; in office based care, 20% of prescriptions are for off-label indications.[15] About 94% of recommendations in NCCN guidelines are based on "lower level" evidence—generally small, non-randomised, early phase trials.[16] Harms to patients accrue over years if prescription recommendations are based on spurious exploratory findings.

These practices have effects on healthcare systems. When treatments are reimbursed for long periods based on exploratory evidence, healthcare systems must either charge greater premiums or allocate fewer resources to proven treatments (such as palliation for advanced disease). At the same time, drug companies have little incentive to fund large trials that might disprove a drug's value. Patients might be less willing to participate in a confirmatory clinical trial if a therapy is available to them off-label. With pregabalin's annual global sales approaching $5bn, Pfizer paid $2.3bn in 2009 to settle claims relating to the off-label promotion of it and other drugs.[10] These promotions often drew on exploratory evidence and cost the US healthcare system hundreds of millions of dollars.[11]

There are understandable reasons why favourable exploratory findings might sometimes go unconfirmed for extended periods. Getting funding and recruiting patients with rare diseases may be difficult. But drug companies are capable of meeting these challenges under regulatory pressure. Trials supporting FDA approval of orphan drugs have a median enrolment of 96 and are randomised 30% of the time,[17] but the evidence supporting recommendations for sorafenib and sunitinib is considerably weaker. Another mitigating factor is that exploratory trials sometimes yield important insights about pathophysiology, biomarkers, or research techniques. But this doesn't justify the patterns we have described—all trials supporting recommendations for sorafenib and sunitinib in table 1 stated their primary goal as testing efficacy, not exploring techniques or concepts. Finally, treatment effects in exploratory trials are occasionally very large and well supported by mechanistic knowledge; in which case, their findings can be sufficient to prove a drug's value without confirmatory trials.

Towards resolving clinical agnosticism

Several current initiatives could tackle prolonged clinical agnosticism. The US National Cancer Institute (NCI) has launched an "exceptional responder" programme that aims to determine whether outliers from negative clinical cancer trials can "open up new drug development avenues."[18] Public investments in comparative effectiveness programmes (such as the Agency for Healthcare Research and Quality's effective healthcare programme) could also identify the most important clinical hypotheses to test. But more should be done.

Firstly, the ethical evaluation of post-license exploratory trials should take into account that many positive findings never undergo confirmatory testing; similar arguments have been made about underpowered trials in general.[19] Ethics committees should consider whether there is evidence of a credible path forward. Regulators can play a similar role, based on their obligations to protect public health. Evidence that might be considered includes commitments and financial backing from companies or research consortia to pursue randomised trials if favourable findings emerge. Patients entering exploratory trials should also understand the flaws in medical research. They should be told that trials rarely prove a treatment's efficacy, and that, even when findings are promising, sponsors do not always follow up with definitive trials.

Secondly, financial incentives could be used to re-balance the broader research agenda. Policy makers should consider that slackening drug approval standards would probably exacerbate prolonged agnosticism. They should also consider measures that might encourage companies to run confirmatory studies—especially in the context of rare disorders. Paediatric exclusivity, which grants an extra six months of patent or data exclusively when companies test drugs in children, is one example of how policies can encourage companies to tackle evidence gaps. Public funding bodies and federal agencies should set aside a larger budget for confirmatory trials. These proposals would concentrate more trial resources in fewer research groups, but this may be worthwhile given the stakes for patients and healthcare systems. Currently, the NCI and the US National Institute of Neurological Disorders and Stroke (NINDS) allocated more funding towards phase II trials than the National Heart, Lung, and Blood Institute (NHLBI). Among clinicaltrials.gov records, confirmatory trials (defined as combined phase II-III or phase III status) represented 12.4% of NCI funded trials, 25.1% of NINDS funded trials, and 39.3% of NHLBI funded trials. This is in contrast to early phase, exploratory studies, which represent 87.6%, 74.9%, and 60.7% of phase II testing, respectively.

Thirdly, journal editors, referees, and doctors can do a better job of vetting manuscripts for "spin." In one study, 59% of 92 trials reporting a negative primary outcome claimed clinical promise on the basis of secondary outcomes.[20] In a follow-up randomised trial, experts rated treatments as more promising

when abstracts presented to them were "spun" in this manner.[21] Early phase trial reports should generally be presented as generating—not resolving—clinical hypotheses.

Having a large repertoire of clinical hypotheses is key to improving outcomes for patients—especially where effective management strategies are lacking. But, for this system to work, early phase exploratory testing must be tightly coupled with late phase, confirmatory trials. Studies of clinical development[7 8] indicate that drug companies and public funding agencies often generate prolonged clinical agnosticism, which can be sufficient to influence physicians and healthcare system decisions. Who could blame companies for not seeking to capitalise on this opportunity? Policy makers should recognise the ethical and healthcare system stakes of ensuring a reliable coordination of exploratory and confirmatory trials.

Contributors and sources: All authors participated in the conception, drafting, and revision of the present manuscript. All data collection was performed by BC and CF. JK studies ethics and drug development; CF and BC are students in his research group. The observations described in this essay originated from a series of projects aimed at measuring the burdens and clinical impact associated with drug development efforts in cancer and pain. JK is guarantor.

Funding: This work was funded by CIHR, the Rabinowitch Fellowship, and The Louise and Alan Edwards Foundation Studentship in Pain Research.

Competing interests: We have read and understood BMJ policy on competing interests and have no relevant interests to declare.

Provenance and peer review: Not commissioned, externally peer reviewed.

1 Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. Clinical development success rates for investigational drugs. *Nat Biotechnol* 2014;32:40-51. 10.1038/nbt.2786 24406927
2 London AJ. Clinical equipoise: foundational requirement or fundamental error? In: Steinbock B, ed. *Oxford handbook of bioethics*. Oxford University Press, 2007: 571-96.
3 Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled?*Ann Intern Med* 1996;125:605-13. 10.7326/0003-4819-125-7-199610010-00011 8815760
4 Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005-2012. *JAMA* 2014;311:368-77. 10.1001/jama.2013.282034 24449315
5 Doshi P. FDA unease about faster drug approval. *BMJ* 2017;357:j169110.1136/bmj.j1691.
6 Naci H, Smalley KR, Kesselheim AS. Characteristics of preapproval and postapproval studies for drugs granted accelerated approval by the US Food and Drug Administration. *JAMA* 2017;318:626-36. 10.1001/jama.2017.9415 28810023
7 Carlisle B, Demko N, Freeman G, etal . Benefit, risk, and outcomes in drug development: a systematic review of sunitinib. *J Natl Cancer Inst* 2015;108:djv292-292.26547927
8 Mattina J, Carlisle B, Hachem Y, Fergusson D, Kimmelman J. Inefficiencies and patient burdens in the development of the targeted cancer drug sorafenib: a systematic review. *PLoS Biol* 2017;15:e2000487. 10.1371/journal.pbio.2000487 28158308
9 Giladi H, Choinière M, Fitzcharles MA, Ware MA, Tan X, Shir Y. Pregabalin for chronic pain: does one medication fit all?*Curr Med Res Opin* 2015;31:1403-11. 10.1185/03007995.2015.1040750 25868712
10 Ratner M. Pfizer settles largest ever fraud suit for off-label promotion. *Nat Biotechnol* 2009;27:961-2. 10.1038/nbt1109-961 19898429
11 US Department of Justice. Press release. 2009. https://www.justice.gov/opa/pr/justice-department-announces-largest-health-care-fraud-settlement-its-history
12 Goodman CW, Brett AS. Gabapentin and pregabalin for pain—is increased prescribing a cause for concern?*N Engl J Med* 2017;377:411-4. 10.1056/NEJMp1704633 28767350
13 Shanthanna H, Gilron I, Rajarathinam M, etal . Benefits and safety of gabapentinoids in chronic low back pain: A systematic review and meta-analysis of randomized controlled trials. *PLoS Med* 2017;14:e1002369. 10.1371/journal.pmed.1002369 28809936
15 Radley DC, Finkelstein SN, Stafford RS. Off-label prescribing among office-based physicians. *Arch Intern Med* 2006;166:1021-6. 10.1001/archinte.166.9.1021 16682577
16 Poonacha TK, Go RS. Level of scientific evidence underlying recommendations arising from the National Comprehensive Cancer Network clinical practice guidelines. *J Clin Oncol* 2011;29:186-91. 10.1200/JCO.2010.31.6414 21149653
17 Kesselheim AS, Myers JA, Avorn J. Characteristics of clinical trials to support approval of orphan vs nonorphan drugs for cancer. *JAMA* 2011;305:2320-6. 10.1001/jama.2011.769 21642684
18 Mullard A. Learning from exceptional drug responders. *Nat Rev Drug Discov* 2014;13:401-2. 10.1038/nrd4338 24875081
19 Halpern SD, Karlawish JHT, Berlin JA. The continuing unethical conduct of underpowered clinical trials. *JAMA* 2002;288:358-62. 10.1001/jama.288.3.358 12117401
20 Vera-Badillo FE, Shapiro R, Ocana A, Amir E, Tannock IF. Bias in reporting of end points of efficacy and toxicity in randomized, clinical trials for women with breast cancer. *Ann Oncol* 2013;24:1238-44. 10.1093/annonc/mds636 23303339
21 Boutron I, Altman DG, Hopewell S, Vera-Badillo F, Tannock I, Ravaud P. Impact of spin in the abstracts of articles reporting results of randomized controlled trials in the field of cancer: the SPIIN randomized controlled trial. *J Clin Oncol* 2014;32:4120-6. 10.1200/JCO.2014.56.7503 25403215

Table

Table 1| Indications recommended by the NCCN for sorafenib and sunitinib monotherapy that are not approved by the FDA

Indication	Citation provided by NCCN	Randomised?	Primary endpoint	Sample size*	Trial result on primary endpoint	Confirmatory RCT registered
Sorafenib						
Angiosarcoma	Maki et al, 2009	No	ORR	(37)	–	None
Chordoma	Bompas et al, 2015	No	9 month PFS	(27)	–	None
Desmoid tumours	Gounder et al, 2011	Retrospective case series				NCT02066181 (ongoing, PFS primary outcome)
GIST	Montemurro et al, 2009	Retrospective analysis				None
	Kindler et al, 2011	No	ORR	38	+	
	Park et al, 2012	No	DCR	31	+	
Hemangiopericytoma	Valentin et al, 2013	No	9 month PFS	(5)	–	None
Osteosarcoma	Grignani et al, 2012	No	4 month PFS	35	+	None
Sunitinib						
Alveolar soft part sarcoma	Stacchiotti et al, 2009	Retrospective case series				None
	Stacchiotti et al, 2011	Retrospective case series				
Angiosarcoma	George et al, 2009	No	ORR	(2)	–	None
Chordoma	George et al, 2009	No	ORR	(9)	–	None
Hemangiopericytoma	George et al, 2009	No	ORR	(3)	–	None
	Stacchiotti et al, 2012	Retrospective case series				
Meningioma	Kaley et al, 2015	No	6 month PFS	36	+	None
Thymic carcinoma	Thomas et al, 2015	No	ORR	(25)	–	None
Thyroid	Carr et al, 2010	No	ORR	35	+	None
	Cabanillas et al, 2010	Retrospective analysis				

* Brackets indicate subgroup size. Confirmatory status was defined based on whether trials used randomised design and a survival or progression free endpoint. Confirmatory trials were sought by a search (12 December 2017) of clinicaltrials.gov for trials using the same drug in the same indication. Recommendations based on subgroup analysis are indicated by parenthetical sample size. ORR=objective response rate; PFS=progression-free survival; DCR=disease control rate; NCCN= National Comprehensive Cancer Network