

# COMMENT

**PHYSICS** Behind the scenes with the people who found gravity waves at LIGO **p.28**

**PHILOSOPHY** Daniel Dennett returns to the evolution of consciousness **p.30**



**CONFERENCES** Keynote speakers must engage audiences **p.31**

**OBITUARY** Vera Rubin, confirmer of dark matter, remembered **p.32**

ILLUSTRATION BY RICHARD WILKINSON



## Consider drug efficacy before first-in-human trials

Ethical review boards must focus on clinical promise as well as safety to hold the first tests of drugs in humans to a higher standard, say **Jonathan Kimmelman** and **Carole Federico**.

**O**n 17 January 2016, a healthy man was declared brain-dead after receiving an experimental drug in a first-in-human trial in France. Four of five other subjects receiving the same dose have serious, ongoing neurological complications. Investigations into the trial described many troubling safety practices, such as steep increases in dose levels delivered to sequential subjects

without sufficient delays to check for safety.

The year since has brought intense scrutiny about how the debacle could have been anticipated and prevented. However, another issue is still largely overlooked: the duty to evaluate whether an experimental treatment is promising enough to warrant testing on people.

In the wake of the tragedies, the French

medicines safety agency (ANSM) ordered an examination of the information that the drug developer, Bial, based in Trofa, Portugal, had supplied to ethics committees and potential researchers before the trial (see [go.nature.com/2j88gqy](http://go.nature.com/2j88gqy)). The report notes that the 63-page Investigator Brochure describing the trial included fewer than two pages of evidence that the drug had the ►

► desired pharmacological activity. It identified only two studies presented as evidence for efficacy, both problematic. In one, Bial had data for a different marketed drug showing it was more effective than Bial's drug at relieving pain in animals, but did not include that information in a summary figure. Both preclinical studies showed only "moderate" positive effects. Moreover, Bial's drug had been tested at a range of doses in mice that made it impossible to estimate the most likely effective dose in humans.

Press coverage following the tragedy quoted independent experts concluding that there was little evidence to support a trial, and that at least five other drugs designed to act in a similar way had been tested in people without success. (Bial maintains that toxicities were not predictable and that it has followed all human-testing norms. We approached the company for more information about the event for the purposes of this Comment but received no response.)

As bioethicists, we have studied the ethics of first-in-human (FIH) and early-phase research for more than a decade. We discuss ethics review with regulators, ethics oversight committee members, investigators and others. We also have personal experiences serving as reviewers on dozens of early phase trials.

We contend that a lack of emphasis on evidence for the efficacy of drug candidates is all too common in decisions about whether an experimental medicine can be tested in humans. We call for infrastructure, resources and better methods to rigorously evaluate the clinical promise of new interventions before testing them on humans for the first time.

### EFFICACY NOT CONSIDERED

More-thorough assessments of clinical potential before trials begin could lower failure rates and drug-development costs. Currently, more than half of drugs that reach later-stage (phase II and III trials) human testing fail because they do not demonstrate efficacy. Today, the evaluation of preclinical evidence is especially important. Favoured picks for the next commissioner of the US Food and Drug Administration (FDA) are likely to lower the current requirements that a drug must demonstrate efficacy in humans before entering the market. If so, low standards for launching clinical trials in the United States could result in ineffective drugs being approved, while also decreasing incentives.

Regulators in Europe and North America evaluate safety before human trials can proceed, but they do not currently demand evidence for potential efficacy. At a workshop of the US National Academy of Sciences in September, Robert Temple, a veteran at the FDA's Center for Drug Evaluation and Research, said that the agency largely left it to drug sponsors to evaluate their rationale

## JUDGING DRUG EFFICACY

### Three questions to assess clinical promise

Ethics requires clear-eyed evaluation of a drug's potential. These questions can help provide clarity.

#### What is the likelihood that the drug will prove clinically useful?

- How have other drugs in the same class or against the same target performed in human trials?
- How have other drugs addressing the same disease process fared?

#### Assume the drug works in humans. What is the likelihood of observing the preclinical results?

- Are the treatment effects seen in animals large and consistent enough to suggest a tangible benefit to patients?

- How well do animal models reflect human disease?

#### Assume the drug does not work in humans. What is the likelihood of observing the preclinical results?

- Have effects of random variation and bias been minimized (for example by sample sizes, randomization, blinding, dose-response curves and proper controls)?
- Do the conditions of the experiment (for instance age of animal models, timing of treatments and outcomes) match clinical scenarios?
- Have effects been reproduced in different models and/or in independent laboratories?

that an experimental drug was likely to work. "I can't think of any cases where [FDA has] said you can't do this [phase I] study because we're just too sceptical." The European Medicines Agency (EMA) — Europe's drug regulator — is similarly silent about the evaluation of clinical promise, even in proposed revisions to guidelines prompted by the Bial affair.

Commercial interests cannot be trusted to ensure that human trials are launched only when the case for clinical potential is robust. We believe that many FIH studies are launched on the basis of flimsy, under-scrutinized evidence. The ALS Therapy Development Institute, which studies the motor-neuron disease amyotrophic lateral sclerosis, has concluded — through its own animal studies — that several compounds that failed clinical trials entered human testing on the basis of poorly conducted or poorly designed preclinical experiments<sup>1</sup>.

Across medical science, preclinical studies are plagued by poor design, implementation and reporting<sup>2</sup>. Several investigations suggest that the magnitude of effects seen in many preclinical studies are not reproducible<sup>3</sup>, or do not reflect intended clinical scenarios<sup>4</sup>. For example, scores of compounds aimed at protecting the brain after stroke have been put to trials on the basis of preclinical studies with very modest effects, or run under clinically unrealistic conditions, such as administering a stroke drug to animals without reflecting the typical delay between a person having a stroke and receiving care<sup>5</sup>.

Early human trials rarely have dire outcomes<sup>6</sup>. According to the EMA<sup>7</sup>, severe accidents have occurred in only 2 of 3,100 FIH trials overseen by the agency since 2005. But, even if individual participants

are not harmed, trials of ineffective therapies place burdens on society. Drug development is costly, in terms of money and people. Patients, healthy volunteers and experts involved in testing a dud treatment are not available for more promising ones. Expenses wasted on ineffective therapies and uninformative trials result in higher drug prices. Investigators, host institutions and sponsors have a responsibility to consider all this before embarking on new research programmes.

Moreover, researchers have ethical obligations to "assure that the risks to subjects are reasonable in relation to the anticipated benefits", according to FDA guidance. Such regulators explicitly delegate these appraisals to ethics review committees. By definition, blood draws, administration of foreign substances and inconvenience are justified only insofar as the research in which they are embedded is likely to advance medical knowledge and potential treatments. The battery of animal-toxicity and dosage tests that regulators require before allowing human trials do not provide this evidence. Nonetheless, ethics boards often take regulatory approval as a signal of clinical promise.

### ETHICS ENABLED

What is to be done? First, the documents that drug sponsors submit to investigators and ethics committees should include negative and unfavourable results from animal studies, if they exist. They should also summarize outcomes from clinical tests of

**"We must abandon the fiction that current oversight systems are adequate."**

benefits", according to FDA guidance. Such regulators explicitly delegate these appraisals to ethics review committees. By definition, blood draws, administration of foreign substances



other products in the same drug class. One small way to discourage data cherry-picking would be to have drug sponsors sign a statement testifying that the clinical and preclinical evidence presented on clinical promise is complete and unbiased. Potential investigators should also, like manuscript editors and peer reviewers, be encouraged to request further information after reading company materials.

Second, FIH trials should proceed only after careful vetting of the preclinical evidence by people with the appropriate expertise who are independent of the drug sponsor. In our own experience, institutional review boards (IRBs) and clinical investigators often claim they lack the resources and background to conduct such assessments.

Instead, we suggest the creation of a centralized FIH advisory system that combines ethical and scientific review. Several precedents exist. The Recombinant DNA Advisory Committee (which reviews new gene-transfer protocols) has assessed evidence of both risk and efficacy since it began reviewing human gene-transfer studies in 1989. Further examples of centralized, expert review of clinical trials in the United States include the SMART IRB Reliance Platform at the National Center for Advancing Translational Sciences; the National Cancer Institute's Central IRB; and the Office for Human Research Protections' '407 review process' for certain paediatric trials.

The FIH advisory mechanism we envision would consist of subcommittees that specialize in clinical areas (for example, neurodegenerative disease, cancer and cardiovascular disease). Advisory-committee assessments would, like most of the above examples, be included in materials presented to physician-investigators and local ethics committees.

Although an FIH advisory panel might be maintained by scientific funders, the presentations compiled by drug companies often contain commercially sensitive information, which regulatory authorities have greater capacities for protecting. The logical home for such a review mechanism would thus be within an authority such as the FDA (in the United States), the EMA (in the European Union) or the Pharmaceuticals and Medical Devices Agency (in Japan).

Third, the appraisal of clinical promise should be rigorous and structured (see 'Three questions to assess clinical promise'). It should encourage reviewers to consider a broader evidence base, as well as whether positive effects in preclinical studies might reflect chance or bias<sup>8</sup>. The International Society of Stem Cell Research (whose ethics committee one of us, J.K., serves on) has articulated a similar set of structured recommendations for cell-based interventions<sup>9</sup>.

Critics of our proposal might raise several objections. First, it requires investments in

new regulatory infrastructures. However, central review systems might actually diminish costs and investigator burden<sup>10</sup>. Another objection is that it could increase the costs and time for drug development. We believe that these would be partly offset by a more sound basis for late-stage trials, for which the expense of clinical failures is greater.

Critics also object that FIH advisory panels could stop truly promising drug candidates from being tested. However, we are not arguing that the preclinical evidence must be strong, rather that it be examined critically to inform ethical judgement. For diseases in which robust preclinical evidence is impossible — for instance, where animal models are clearly inadequate as in many neurodegenerative disorders — a limited suggestion of clinical promise might be enough to justify trials for a relatively benign candidate aimed at a great unmet medical need.

Several steps can be taken quickly. One would be to convene a National Academy of Sciences panel to advise on how best to harness preclinical evidence to evaluate clinical potential. This could set priorities on what evaluations are most needed and examine how more-rigorous review of preclinical evidence could best be absorbed into existing oversight structures.

Another step would be to encourage university ethics boards to appoint ad hoc members with relevant medical expertise to summarize and evaluate proposed early phase trials. More ambitiously, funders could create boards to offer counsel on tricky issues at the intersection of animal research and human trials. These would function much like the Recombinant DNA Advisory Committee, but focus on treatments that are risky or that have mechanisms of action that have never been tested.

We must abandon the fiction that current oversight systems are adequate to protect volunteers in first-in-human trials or to steward scientific efforts. ■

**Jonathan Kimmelman** is associate professor of biomedical ethics, and **Carole Federico** is a PhD candidate, at McGill University, Montreal, Canada.  
e-mail: jonathan.kimmelman@mcgill.ca

1. Perrin, S. *Nature* **507**, 423–425 (2014).
2. Rice, A. S. C. *et al. Pain* **139**, 243–247 (2008).
3. Steward, O. *et al. Exp. Neurol.* **233**, 597–605 (2012).
4. Henderson, V. C. *et al. eLife* **4**, e08351 (2015).
5. Ginsberg, M. D. *Stroke* **40**, S111–S114 (2009).
6. Johnson, R. A., Rid, A., Emanuel, E. & Wendler, D. *Clin. Trials* **13**, 149–160 (2016).
7. Bonini, S. & Rasi, G. *N. Engl. J. Med.* **375**, 1788–1789 (2016).
8. Kimmelman, J. & Henderson, V. J. *Med. Ethics* **42**, 50–53 (2016).
9. Kimmelman, J. *et al. Nature* **533**, 311–313 (2016).
10. Wagner, T. H., Murray, C., Goldberg, J., Adler, J. M. & Abrams, J. J. *Clin. Oncol.* **28**, 662–666 (2010).

J.K. declares competing financial interests: see [go.nature.com/2jtmepq](http://go.nature.com/2jtmepq).