LEAPING TO CONCLUSIONS

Clinical trials are sometimes stopped prematurely if they show early benefit.

But this might pose a risk to patients, says **Margaret McCartney**

“Major trial stopped early because results so good”—such urgent, feel good headlines about the latest medical breakthrough make for an eye catching health story. Cancer and HIV drug trials seem especially susceptible to a high media profile when there is talk of a newly discovered advantage.

The impression that this is happening more often is borne out by a recent study published in *Annals of Oncology* by Trotta and colleagues.¹ The research team looked for all published clinical trials of cancer drugs in the past 11 years that were stopped early after an interim analysis found benefits. Over half of the 25 trials they identified were published in the past three years (2005–7).

Is this something to be concerned about? Surely an effective drug needs to be licensed and made available to as many people as possible, as quickly as possible. People who are seriously ill cannot wait for tedious committees to wait for long clinical trials to end before deciding about effective drugs—at least, that is how the argument goes. Indeed, Trotta and colleagues found that most of the recent trials that were stopped early were used in support of an application for marketing authorisation at the European Medicines Agency or the US Food and Drugs Administration.

Yet this may be exactly the reason to be concerned. Is the trend for stopping trials

early in order to quickly license apparently effective drugs good for patients?

Trials are often stopped by a data and safety monitoring committee, which is responsible for assessment of the results obtained during the live running of a clinical trial. Part of the role of these committees is to ensure that participants in the trial are getting ethically acceptable care. For example, if an intervention being tested is found, early on, to cause unacceptable harm or have serious side effects, there may be ample justification for stopping the trial on the basis of such interim results. Nor would it be ethical to continue with a trial where it becomes known far earlier than expected which group is being disadvantaged by the treatment they receive. Yet counterintuitively, stopping trials early may actually contribute to both greater uncertainty about the intervention and thus more risk to patients.

Dangers of not finishing

Iain Chalmers, an editor of the James Lind Library and a previous director of the UK Cochrane Centre, says: “The issue is basically about the play of chance. You can get big swings by chance which may be statistically significant in both directions.” He explains how preliminary research is vitally important in interim analyses. “ISIS I was a trial of atenolol in acute myocardial infarction. One of the reasons this trial was done was that a systematic review had suggested that it would decrease risk of early death by 15%.

On the basis of this promising information, it was decided to do a very large trial to get a confident handle on the size of effect.

“The data monitoring committee found that early results suggested that atenolol was harmful, and statistically significantly so. But in assessing these interim results, it took into account that the evidence from the systematic review of all previous trials suggested that they might well be a reflection of the play of chance. So it allowed recruitment to the trial to continue.

“By the time the trial had ended it had shown that atenolol reduced mortality by 15%, exactly as predicted by the systematic review of all previous trials. In other words, the systematic review protected the committee from making an inference which, in retrospect, would have been dangerous. The lesson is, ‘Don’t embark on a new trial before you know what happened in all previous trials.’ And this means all existing trials—not just those in print. If people are not publishing trials—for example, because of stopping them early due to adverse effects—then you will have a biased sample.”

However, interim results may also overestimate the extent of harm of an intervention. Referring to the CRASH trial, which assessed the effects of steroids given to people with acute traumatic brain injury, Sir Iain says, “A systematic review had shown that this treatment might be helpful but that it was equally likely that it was harmful. This was the basis for doing the trial. The trial was stopped when

an interim analysis showed a statistically significant harmful effect of the drug. Although the study provided strong evidence to justify abandoning this treatment, the actual size of the harmful effect was probably an overestimate, as the authors noted. In other words, both false and true harmful effects may be overestimated."

Indisputable evidence

So how convincing should the interim results be for a trial to stop early? Stuart Pocock is professor of medical statistics at the London School of Hygiene and Tropical Medicine. "If you are going to stop a trial claiming superiority for a new treatment you need proof beyond reasonable doubt. This means that you have to have very convincing evidence and a very high level of statistical significance." He suggests that the significance should be more stringent than the conventional P value of 0.05 used in medical journals as an acceptable level of significance. "Some people have recommend P values of 0.001 or less—the Peto rule—that is the sort of strength of evidence that you need on your primary end point to stop a trial early for benefit of a new treatment. But there are lots of other judgments needed: what is the evidence from other studies, what are the side effects, what are the secondary outcomes? That's why you need a data monitoring committee with good judgment and experience."

Clearly the role of the data monitoring committee, which is responsible for the interpretation of interim analyses and the action that this should prompt, is centrally important. However, there is no absolute requirement to even have one—or even for its members to be independent of the trial and of any commercial interests. Professor Pocock says: "Any major pivotal trial will have a data monitoring committee that is primarily there to ensure patient safety is ethically looked after. Certainly the industry sponsored trials I've collaborated in do have independent committees. But there are trials, perhaps in fields with less of a track record of large scale quality trials where data monitoring committees still don't exist." However, even this does not guarantee easy decision making. "The first step is to have an independent committee. The second step to have a wise committee," he adds.

There are numerous examples of trials that have looked, at an early stage, to be producing apparently compelling positive evidence

which the end result does not go on to confirm. For example, interim analysis of the results of the OPTIMIST trial (optimised phase III tifacogin in multicenter international sepsis trial) seemed to show a survival benefit of both clinical and statistical significance compared to placebo, with a P value of 0.006. But this was not enough to stop the trial by predetermined standards. The study continued and the apparent benefit of the intervention disappeared.² This is what Professor Pocock calls regression to the truth.

Healthy scepticism

Victor Montori and colleagues published a systematic review of trials stopping because of early benefits in 2005.³ The trials were typically industry funded and evaluated drugs in cardiology, cancer, and HIV. In almost all (94%) of the trials they identified, the resulting paper did not report at least one of the following pieces of information: the planned sample size, the interim analysis after the trial was stopped, whether a predetermined stopping rule had been adhered to, or if there had been an adjusted analysis that accounted for interim monitoring and truncation of the trial. They concluded: "RCTs [randomised controlled trials] stopped early for benefit are becoming more common, often fail to adequately report relevant information about the decision to stop early, and show

implausibly large treatment effects, particularly when the number of events is small. These findings suggest clinicians should view the results of such trials with skepticism."

The scepticism they recommend would affect the validity of such results which are used for licensing purposes, as the study in the *Annals of Oncology* found. Lead researcher Giovanni Tafuri, a drug assessor at the Italian Medicines Agency, says: "As reported in our paper, more than 78% of the RCTs published in the past three years with an interim analysis ending the trial were used for registration purposes. Obviously this can guarantee quicker access to the market, and we are aware that trials stopped early for benefit may result in the quick identification of promising new treatments for patients. However, findings obtained with this strategy should be considered preliminary results and thus should be subsequently confirmed with post-marketing confirmative studies."

So how much confidence can prescribers

have when trials of interventions have been stopped prematurely for benefit? Dr Tafuri is cautious: "Prescribers should be aware that the interim results, leading to an immature benefit-risk balance of a drug, cannot be directly translated into clinical practice without further confirmative studies. The use of interim analyses to stop trials early for benefit should always be clearly stated in study publications to allow prescribers to be aware of the trial methodology used."

So what should patients be told about the trial they are taking part in when it comes to the assessment of interim results? "I would hope that the patient would come to the trial in the knowledge of how much the play of chance has played us false in the past," Sir Iain says. "Right now, there are five trials trying to work out what level of oxygen in the blood is best for premature babies. People are going to have to hold their nerve. Early results may show an increased level of blindness in higher oxygenation and an increased death rate in lower oxygenation. What's the best in terms of death or disability as measured two years after birth? We have the parent of a premature baby on our data monitoring committee—we look to her for help with these issues. If we are worried that we may be looking at a chance finding in an interim analysis, we may well contact the committees of the other four trials and ask them if they have any interim results that have made them concerned about the same things that we are concerned about. Cooperation of that kind is another way of coping with the dilemmas that result from the play of chance."

Dr Tafuri agrees. "Patients involved in clinical studies should always be informed about the progress of the trial. It should be pointed out that patients are always fully informed about the objectives of the study before being enrolled. In fact, at the moment they decide to join the trial, they must sign an informed consent. Interim analyses are usually pre-planned and, before the trial starts, investigators should inform patients that interim results should be considered as partial and that only completing the trial allows the achievement of study objectives."

Competing interests: None declared.

- 1 Trotta F, Apolone G, Garattini S, Tafuri G. Stopping a trial early in oncology: for patients or for industry? *Ann Oncol* 2008 Apr 9 [Epub ahead of print].
- 2 Angus DC, Crowther MA. Unraveling severe sepsis: why did OPTIMIST fail and what's next? *JAMA* 2003;290:256-8.
- 3 Montori V, Devereaux PJ, Adhikari NK, Burns KE, Eggert CH, Briel M, et al. Randomised trials stopped early for benefit. *JAMA* 2005;294:2203-9.

Margaret McCartney is a general practitioner in Glasgow Margaret@margaretmccartney.com