Peter McIntyre

Tipping the balance

Almost four in ten serious adverse drug reactions now listed in the labels of 12 targeted cancer therapies were not mentioned in the studies that led to their approval. Half the serious reactions that were missed are potentially fatal. How can we improve the way we investigate and report the side effects of new drugs?

Questions are being raised about the accuracy and integrity of reports from pivotal clinical trials that provide the evidence for licensing cancer drugs. There is increasing concern that reports overstate the effectiveness of innovative drugs in a real world setting, because patients on trials are healthier and fitter than most of the people it will be used in, and understate side effects. This distorts the information used by clinicians to define the recommended dose, by regulators to assess the risk–benefit profile, and by patients to choose between treatment options.

Researchers and patient groups are calling for changes in the way that trials are designed and reported, with fewer exclusions and a much more rigorous approach to reporting side effects.

A team at the Princess Margaret Hospital in Toronto has turned a spotlight on this issue in a series of papers which highlights the gap between adverse events reported from 'pivotal' trials (which form the basis for marketing approval) and the warnings eventually added to drug labels – often years later.

The first of these, published in 2011, showed that 39% of serious adverse drug reactions – half of them potentially fatal – were not described in any of the randomised clinical trial (RCT) reports associated with 12 targeted anti-cancer agents (*JCO* 2011, 29:174–185) They had to be added to drug labels at a later date.

The same team analysed anti-cancer drugs approved by the US Food and Drug Administration (FDA) between 2000 and 2010 and found that most were associated with increased odds of toxic death, treatment discontinuation or severe adverse events (*JCO* 2012, 30: 3012–19).

In 2014 the team demonstrated that adverse effects also led to increased costs of treatment (*JCO* 2014, 32: 3634–43)

Saroj Niraula, lead author on the 2012 and 2014 papers and now a medical oncologist at Cancer Care Manitoba, in Winnipeg, Canada, stresses that new therapies have saved tens of thousands of lives and that criticism of trial reports should be seen in that context. "My point is to do what we can to improve the reporting of the trials so we can make the best judgement about efficacy and toxicity, rather than pointing out flaws in reporting research."

However, he says that RCTs are focused on demonstrating clinical efficacy rather than testing toxicity. "Frequently when we read pivotal RCT published papers we see statements like 'no differences in toxicities', but most trials are not powered to support such statements."

"Pivotal RCT papers often contain statements like 'no differences in toxicities', but there is no real data to support that"

His 2012 paper noted that treatment-related mortality associated with bevacizumab (Avastin), the cardiovascular effects of aromatase inhibitors, and the increased risk of cardiopulmonary arrest with cetuximab (Erbitux) all went unreported in the original trials.

Bevacizumab was approved in the EU for treating metastatic breast cancer in 2007 and in the US in 2008, on the basis of trial reports that showed tumour shrinkage and an increase in progression-free survival. Further evidence on both safety and efficacy that emerged in the two years following the trial, however, prompted the FDA to withdraw that approval, on the grounds that patients would "risk potentially life-threatening side effects without proof that the use of Avastin will provide a benefit, in terms of delay in tumour growth, that would justify those risks." (The European regulators, the EMA made a different judgement call after deciding that the benefits of tumour shrinkage did outweigh the risks.)

The aromatase inhibitor Arimidex was approved by the FDA in 2002 as an adjuvant treatment for early breast cancer in postmenopausal women, on the basis of the ATAC trial, which showed improved disease-free survival compared to tamoxifen and a lower incidence of certain side effects associated with tamoxifen.

A secondary analysis of the ATAC data by the FDA later led to a warning being added to the drug label to indicate that "anastrozole may be linked to an increased risk for ischemic cardiovascular events in women with pre-existing ischemic heart disease."

Yet the report of a ten-year update on the trial, published in 2011, made no reference to the new evidence, or the additional warning.

The 2006 trial comparing cetuximab and radiotherapy with radiotherapy alone for people with squamous carcinoma of the head and neck reported 'similar' incidence rates of severe reactions for the two treatment arms. The 2% of patients who died on the cetuximab arm as a result of cardiopulmonary arrest went unreported because the trial only reported acute adverse events that affected at least 10% of patients.

Lapatinib (Tykerb), is another striking example, which was flagged up by Bostjan Seruga, one of the collaborators in the 'Toronto papers', at a presentation he made at the European Cancer

Congress in Vienna in September 2015. He pointed out that Tykerb's drug label has been revised 12 times since it was approved in 2007 to treat women with metastatic HER2-positive breast cancer. Added warnings include notice of potential damage to lungs, severe skin reactions, and a 'boxed' warning on hepatotoxicity – the strongest warning that the FDA can mandate.

Seruga, who is based at the Institute of Oncology in Ljubljana, Slovenia, pointed out that new evidence from post-marketing surveillance can significantly change the risk–benefit balance, but it is the early impression about lack of harm that sticks. "Patients do not know what symptoms to expect based on prior experience, drug developers may have a false impression as to how a drug is tolerated, regulators may not have confidence in the fidelity of information about balancing risks and benefits and payers cannot accurately predict the utilisation of health care services."

"New evidence can significantly change the risk–benefit balance, but it is the early impression about lack of harm that sticks"

A distorted picture

There are a number of ways in which trial reports paint a distorted picture: patient selection for trials, a failure to detect or report side effects, and the way data are presented are all implicated.

Patient selection

Patients who are fit enough to join clinical trials are not representative of the substantial proportion of patients with the condition in the wider public. Trials usually exclude those with heart or kidney disease or a previous history of cancer. It has been half-seriously suggested that to enter a clinical trial you need to be "a marathon runner who happens to have cancer".

This means that when drugs are used in clinical practice, results very often don't live up to expectations. Niraula says, "Drug companies put a lot of investment into clinical trials, and mostly with good intentions want the drug to work for the benefit of the patient and understandably, want a return on their investment. When it enters the real population, the result is a higher likelihood of toxicities and a lower likelihood of benefits."

A study at the Princess Margaret Hospital in Toronto provides some confirmatory evidence. It compared outcomes for patients with metastatic castration-resistant prostate cancer, treated at the same hospital, to identical standards of care, according to whether or not they were on a trial. They found that the trial patients were younger, had less comorbidity and better performance status. Patients treated in routine practice had shorter survival and experienced more toxicity, notably fever and infection (*Ann Oncol* 2013, 24:2972–77). This difference between outcomes inside and outside clinical trials even has its own label: "the efficacy–effectiveness gap".

The likelihood is that differences in outcomes will be even greater for patients treated away from major centres, since patients are likely to have poorer access to supportive care to address side effects. As quality of life worsens, patients may suspend treatment or reduce the recommended dose.

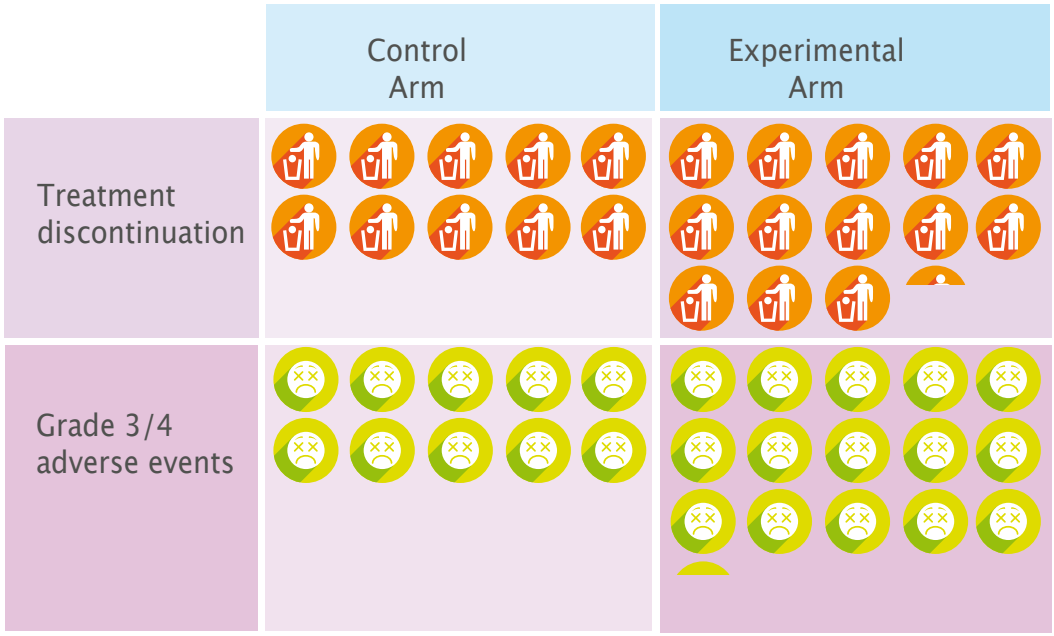
Age discrepancy is widespread within clinical trials, as a by-product of excluding patients with comorbidities. The CML Advocates Network found that the average age of CML patients on phase III trials was 47, while the average age of real world patients in Europe is nearer 65, meaning that side effects in the older population with comorbidities are not discovered in trials.

There is, however, evidence that exclusions do not necessarily invalidate trial results. In a study with some similarities to that conducted in Toronto, Joseph Unger and a team at the Fred Hutchinson Memorial Hospital in Seattle studied 21 RCTs supported by the National Cancer Institute. By comparing the survival of patients on the control arm – who were receiving standard care – to similar patients treated outside trials, they were able to gain insight into differences relevant to being in a trial (*JNCI* 2014, 106:dju002 doi:10.1093/jnci/dju002).

Unger and his colleagues found that, while being on a trial was associated with better survival, the difference lasted for only one year after diagnosis. They believe the difference is simply due to patients in the trial being younger and fitter with fewer comorbidities.

Survival curves for standard treatment patients in trials and non-trial patients were very similar in the longer term. Of course, looking at control arm patients does not say anything about the efficacy of treatments, but Unger says it suggests that any benefits found for new treatments should translate to a real-world setting. "The fact that over the long term patients had very similar outcomes suggests that trials are not picking off qualitatively different cancer patients, they are just excluding those with comorbid conditions that affect survival in the short term." However, this conclusion would not be valid, he says, if

The price we pay for progress



A 2012 study of 12 widely used targeted cancer drugs approved since 2000 (*JCO* 30: 3012–19) showed that most are associated with higher rates of toxic death (odds ratio1.4), treatment discontinuation (OR 1.33) or severe adverse events (OR 1.52)

the treatments have too much toxicity or poor compliance.

Unger is also concerned that some of the exclusions of patients due to concern over safety are outdated, and that it is increasingly unrealistic to exclude patients from trials simply because they have had a previous cancer.

However, the tendency to exclude patients appears to be increasing. A study of 86 practice-changing RCTs showed that the proportion of patients excluded from trials had doubled to 18% after 2010, compared with the 9% before 2000 (*Cancer Treat Rev* 2016, 43:67–73). There were increases in exclusions of patients with cerebrovascular events, gastrointestinal bleeding or cardiac conditions. There was also a decrease in the average upper age limit.

Trials in denial

Deciding on which side effects to look for can influence what is found. Ian Tannock, a leading member of the Toronto group, believes the ATAC trial led to a distorted view of the relative safety of Arimidex to tamoxifen, because it was left to the doctors involved in the trial to make a judgement on which events could be connected with the treatment. In a letter to *The Lancet* (March 2011), he argues that this created a bias "because side-effects due to tamoxifen were recognised better at the start of the ATAC trial than were those due to anastrozole," and he suggests it would be better to have a prespecified checklist.

Patients, however, point out that prespecified checklists can also lead to under/non-reporting of important side effects that have been omitted from the list. This is a particular problem for side effects such as exhaustion and diarrhoea, which are not life-threatening but can make life almost unbearable.

Gilly Spurrier-Bernard, president administrator of MelanomaFrance, describes how difficult it was for her husband to record side effects on a trial of vemurafenib (Zelboraf), despite being under the care of the Gustave Roussy Institute, one of Europe's best cancer centres.

"Clinicians only want to report the effects that the trial pharmaceutical companies have identified as a high risk. My husband had a number of skin reactions which we knew were to do with the drugs, because he had never had them before, and they look down the list and say, that is nothing to do with the trial.

"We were treated at a very good centre but it used to drive me up the wall that what you were reporting as potential side effects did not even get recorded."

Several studies show that clinicians under-report adverse events that are very significant for patients.

In 2015 a study from the Italian National Cancer Institute in Milan found extensive under-reporting by doctors of six symptoms that blight the lives of

patients in three randomised trials, including nausea, diarrhoea and anorexia (*JCO* 2015, 33:910–915). Six years previously, in 2009, a survey by Myeloma Patients Europe had shown fundamental differences in perceptions between patients, nurses and doctors in assessing the impact on quality of life of various side effects, including hair loss, fatigue, reduced body function, neuropathy and thrombotic events.

Eric Low, chief executive of Myeloma UK and the chief author of that report, says it shows why patients must have more of an input into reporting side effects to ensure that trial reports paint an accurate picture. He points out, however, that it is only when drugs come into everyday use that clinicians learn how to deal with side effects. He gives the example of bortezomib (Velcade), the first significant proteasome inhibitor, which was given accelerated approval in 2003 as a treatment for relapsing myeloma.

"Initially bortezomib had many side effects, particularly neuropathy, but over time we got a subcutaneous version and doctors moved to giving it once a week and that made a dramatic difference. Now peripheral neuropathy is quite rare.

"Almost half of patients who suffered severe side effects had their first episode after the treatment cycle used to define dosage"

"The real benefit of a new drug comes as clinical experience accrues and patient management and patient selection improves. At the point where a new drug is approved we don't have in depth data, and with a move towards accelerated approval we are going to have even less."

Data from general clinical practice is, however, only used to update clinical trial reports in a small minority of cases. Bostjan Seruga reported that his team had looked at 311 RCTs of prostate, breast and lung cases published over a 30 year period and found that only one in five had published updated reports. Where publications were updated they predominantly showed a smaller magnitude of effect and a greater number of side effects, than the original reports.

There is increasing support, by EMA in Europe and the NCI and the FDA in the US, for moving towards patient-reported outcomes to mitigate the inaccurate reporting of side effects. The EMA completed a public consultation on this issue in 2015 and is expected to report back early this year.

The issue is complicated by the fact that, in the context of certain clinical trials, patients themselves may feel they have an incentive to downplay the seriousness of side effects. Gilly Spurrier-Bernard knows this from her own family experience, when her husband was on a trial for ipilimumab, and in her advocacy role hosting online forums for melanoma patients.

"I spent four years filling in patient questionnaires and as far as I am concerned they are totally useless. Patients lie through their teeth because they know that patients get kicked off the trial if they show any slightly scary signs of side effects. With the ipilimumab trial the slightest sign of colitis or diarrhoea of significant amount you were pretty much kicked off. This is all discussed on patient forums."

She fears for what will happen when the trial treatments come into general use. "People with brain mets, or comorbidities or lupus are excluded from most of these trials. How will side effects affect people who already have autoimmune problems? None of this has been recorded properly. They need to get it sorted."

Misreporting data

Whether by accident or by design, the process of writing up clinical trials offers further opportunities to downplay the negatives and talk up the positives.

In 2004 An-Wen Chan and colleagues reported on 122 journal articles from 102 clinical trials and found that 50% of efficacy outcomes and 65% of harm outcomes were incompletely reported (*JAMA* 2004, 291:2457–65). In 62% of trials, at least one primary outcome from the trial protocol was changed or omitted. The authors concluded that "reporting of trial outcomes is not only frequently incomplete but also biased," and that "published articles may overestimate the benefits of an intervention."

Another of the landmark studies from the Princess Margaret Hospital, Toronto, found that a third of clinical trials for women with breast cancer showed "bias in reporting" in primary endpoints, and two thirds showed bias in reporting toxicity (*Ann Oncol* 2013, 24:1238–44). Positive trials were particularly associated with under-reporting toxicity.

Peter Jüni, Founding Director of the Clinical Trials Unit of Bern University Hospital, outlined at the 2015 European Cancer Congress how the reported results of clinical trials are often distorted. Common practices include 'fishing' through data for spurious positive outcomes, swapping primary and secondary outcomes because the primary outcomes are not very good, and excluding outliers to make results statistically significant.

Perhaps the most pernicious practice

Incomplete information

Laws governing the marketing of medicinal products in European Union member states require that all medicinal products must "be accompanied by labelling and package leaflet which provide a set of comprehensible information enabling the use of the medicinal product safely and appropriately".

But research into the reporting of side effects for some of the most widely used targeted anti-cancer drugs shows the majority are not reported in the pivotal trial and are added to the label, sometimes many years later.

The TKI HER2-blocker lapatinib (Tykerb) has had 12 amendments to its label since it received marketing approval in March 2007, even though safety had already been evaluated in clinical trials in more than 3,500 patients with advanced or metastatic breast cancer. The most common adverse reactions (i.e. in more than 20% of patients) initially recorded for Tykerb plus capecitabine were diarrhoea, hand-foot syndrome, nausea, rash, vomiting, and fatigue.

Warnings given on the label included:

- Reports of decreases in left ventricular ejection fraction
- Foetal harm if administered during pregnancy
- Dose reduction to be considered for patients with severe hepatic impairment.
- Prolonged QT interval in the heart's electrical cycle in some patients.

In **August 2007** further warnings were added about:

- Interstitial lung disease and
- Pneumonitis

In **2008** a boxed warning (highest grade of warning) was added about:

- Reports of severe and sometimes fatal hepatotoxicity – "If changes in liver function are severe, therapy … should be discontinued"

Various notices were added about drug-drug and drug-food interactions in the intervening period.

In **June 2013** the label was amended to warn about:

- Grade 3/4 diarrhoea. "The diarrhea may be severe, and deaths have been reported," says the label. (Most cases of diarrhoea are less severe, occur early in treatment and last 4 to 5 days.)

is selectively omitting inconvenient results, such as the 2% of patients on the cetuximab arm who died as a result of cardiopulmonary arrest. A bigger problem may be the non-publication of entire trials that generate inconvenient results. It is such practices that sparked the launch of the AllTrials campaign in January 2013, which calls for "all trials past and present [to be] registered, and the full methods and the results reported" (alltrials.net).

Wrong dosage, worse effects

While many of the biases listed above may be nothing new, it seems that reporting of side effects from targeted drugs may be a particular problem. One reason is that cytotoxics are prescribed for fixed protocols, whereas targeted drugs are often continued until resistance develops, and adverse effects that are not immediately apparent often occur later.

The big problem here is not just that, as Seruga remarked, it is the early impression about lack of clinical harm that sticks, but that early toxicity results set the basis for defining dosing, and as a result recommended dosage levels may be set too high.

Research led by Sophie Postel-Vinay from the Gustave Roussy Institute found that more than half of the most serious toxicities in phase I trials occurred after the end of the 'dose-limiting-toxicity' period used to determine tolerability (*JCO*

2011, 29:1728–35). Although the severity of toxicities decreased during the trials, the proportion of unresolved toxicities increased, more medication had to be prescribed to deal with side effects, and dose reduction became more frequent, suggesting that "benign late toxicities may not be bearable over time and might require specific management."

This was confirmed in a much larger study led by Postel-Vinay and coordinated by the EORTC, which gained unprecedented access to raw patient data from institutions and pharmaceutical companies covering more than 2,000 patients in 54 phase I trials (*EJC* 2014, 50:2040–49). Almost half of patients who suffered severe side effects (grade 3 or worse) had their first episode after the cycle of treatment that was used to define dosage. One in 11 patients experienced dose-limiting side effects (i.e. the medication had to be paused or reduced), of which the most common were fatigue, nausea, vomiting, gastro-intestinal disorders and hypertension.

The way forward

Most experts agree on a number of steps to improve reporting on data from clinical trials and assess the value of new drugs.

Saroj Niraula, in Winnipeg, says that good-quality population-based studies are required from real-world use after a drug receives full approval, along with stricter regulations about reporting. "We as physicians should be able to provide our patients with the most comprehensive information possible on efficacy and toxicity before they come to a decision about the amount of toxicity that is acceptable to them for a given benefit.

"Journals have to be more stringent. There should be academic incentives to report toxicities well. We want honest and exhaustive information from pivotal drug trials."

Joseph Unger at the Fred Hutchinson in Seattle believes that trials should have fewer exclusions. "From a patient perspective access to trials is a huge issue. But also from a researcher's perspective we want to be able to do these trials as quickly as possible. If we are excluding patients for reasons that are unnecessary, that is hindering our efforts."

"The key recommendations are that everything about late toxicities is reported, which is not the case at the moment"

At the Gustave Roussy, Sophie Postel-Vinay is calling for data on adverse effects to be collected more comprehensively and for longer periods. "The key recommendations are that everything about late toxicities is reported, which is not the case at the moment, and that the recommended phase II dosage is based on everything that is seen over the whole trial."

These recommendations are already being adopted in protocols or written into guidelines for some phase I trials, although there is as yet no settled methodology for deciding on the dose limiting toxicity definition and duration, or the phase II dose recommendation.

Gilly Spurrier-Bernard from MelanomaFrance is campaigning for a patient-driven reporting system filled in on laptops or phones whenever there is a significant event, as some patients already do with pain diaries. "Patient issues change over time and according to how healthy you are feeling. Researchers need to be asking how it impacts on daily life. Then you need some clever algorithms for data mining."

Bettina Ryll, who founded Melanoma Patient Network Europe after her husband Peter developed malignant melanoma, agrees. "We see more and more selected trial populations and it automatically becomes less representative of the entire patient population," she says. "RCTs are the wrong way to tackle safety. We need a much better pharmacovigilance system where we capture data much more systematically and then act upon it.

"We need new drugs, as every patient with a life-threatening condition will tell you. We also need a way to study them meaningfully and in a way that does not prevent access for patients, does not drive up cost and captures reality."

Melanoma Patient Network Europe is preparing a project with the Uppsala Monitoring Centre to harvest direct patient reports of symptoms and side effects. The Centre runs the WHO international drug monitoring programme, which was set up after the thalidomide disaster, and has the world's largest dataset of adverse events, publishing data from 120 national health authorities worldwide on an open website at vigiaccess.com.

Bostjan Seruga from Ljubljana would like to see the American NCI initiative on patient-reported adverse events (PRO-CTCAE) fully incorporated into clinical trials, along with updated reports to capture data not originally reported by RCTs, and specific trials to address the needs of patients who were ineligible.

"Oncologists, journal editors and societies like ESMO and ASCO need to introduce measures to ensure complete reporting of toxicity to serve our patients better."