

# Safety Reporting in Randomized Clinical Trials

## A Need for Improvement

Yusuf Yazici, M.D.

### Abstract

*The reporting of adverse events (AEs) in randomized clinical trials (RCTs) is often lacking in the publication of trials. Part of the problem is the way safety data are reported in RCTs. Reporting of “time to event,” use of standardized incidence ratios for comparison to normal population or disease controls, use of “patient years” when reporting AE, and adequate sample size and power calculations are some of the problems that need to be addressed and improved in RCTs.*

Most of the data leading to a new drug coming to the market comes from randomized clinical trials (RCTs). By design, RCTs are efficacy trials and are not powered to look at adverse events (AEs), especially rare AEs. Additionally, the reporting of AEs in RCTs, especially of new drugs, is frequently incomplete or inadequate.<sup>1</sup> True AE risk can be assessed more robustly in long-term follow-up databases or cohorts. Yet, safety data from RCTs are frequently used as evidence of a lack of difference between active treatment and control arms in AE risk. Even though RCTs are not designed as safety trials and are mostly conducted to determine efficacy, this should not impact reporting on the AEs observed during these investigations. The current limitations of RCTs and problems with the way safety data are reported need to be recognized and possibly corrected in future RCTs.

Time-to-event data are closely related to causality and can provide information for the practicing physician on when to expect certain AEs after initiating therapy. They are one of the most important components of AE reporting and should

be tracked and then reported in publication of the study. It is important to remember that cumulative risk, as is commonly reported in RCTs, may be similar between two treatments at the end of the study period, but instantaneous risk might be different for the same drugs. “Average risk” does not reflect the true risk faced by patients, especially when decisions are being made at the beginning of treatment.

We conducted a study looking at AE reporting across multiple TNF (tumor necrosis factor) inhibitor and COX-2 (cyclooxygenase-2) inhibitor RCTs.<sup>2</sup> There were 70 studies (44 with TNF inhibitors, 26 with COX-2 inhibitors), of which 91% were industry sponsored. Only one-third of the studies provided the time-to-event for AE data in their reports, as a table, in the text, or as a Kaplan-Meier curve. No better reporting was noted for serious adverse events (SAEs). When SAE details were reported, only select cases were mentioned, and there was no uniform method of discussing the data.

In the same review, 8/17 rheumatoid arthritis (RA) TNF trials reporting malignancy used the Surveillance, Epidemiology, and End-Results (SEER) database as the comparison group in the normal population. SEER data assumes an even distribution of events in any given year, and the data is commonly presented as occurring over a 12-month period. If the SEER database is to be used as a comparator, the distribution and timing of AEs need to be even in any 12-month period. For example, if the expected number of a malignancy in the population is 12 malignancies in 1 year, the assumption is that one malignancy is seen each month. When 12 malignancies are seen among TNF users, it is reported that “there were no significant differences” among TNF users and the SEER database. This may be true where the “average” 12-month risk is concerned, but if eight of the 12 malignancies occur in the first 2 months after starting treatment, for those 2 months, the risk is much higher than expected from the SEER data.

Yusuf Yazici, M.D., is an Assistant Professor of Medicine, New York University School of Medicine, and in the Division of Rheumatology, Department of Medicine, NYU Hospital for Joint Diseases, NYU Langone Medical Center, New York, New York. Correspondence: Yusuf Yazici, M.D., 246 East 20th Street, New York, New York 10003; yusuf.yazici@nyumc.org.

Reports of lymphomas in TNF users show that over 60% of the malignancies happen within the first 4 to 6 months.<sup>3,4</sup> This trend also has been demonstrated for serious bacterial infections.<sup>5</sup> Knowing the early increased risk for malignancy or any other SAE is very important to the practicing physician. The risk-benefit analysis and treatment decisions will be very different when these SAEs are clustered in time rather than when they are represented as an even distribution through the year.

One more concern related to the time element in AEs is the use of "patient years," utilized to define the time frame of AE incidence. However, there are problems with using this method in reporting AEs that do not happen randomly during a trial and may be clustered at certain time points.<sup>6</sup> Especially relatively rare idiosyncratic drug reactions usually occur early in the treatment course and in only a few individuals. Apart from the few patients with AEs, remaining patients who are prescribed the drug will likely never get these reactions, however long they use the drug. This leads to the unduly inflated denominator of the related incidence ratio when "patient-years" are used and leads to potential under-representation of AEs. Late onset AEs are also apt to be missed when "patient-years" are used. In short, "... when an event is (or is believed) likely to occur at any stage during continuous treatment with a drug, then an event rate with a time component (rate per person year, etc.) has a true mean."<sup>7</sup> In our survey, six manuscripts were found using "patient years" when reporting AEs in TNF trials.

In addition to the concerns above, RCTs are not good tools to assess the safety of a drug. RCTs are usually of short duration and of a limited number of select patients, usually with few comorbid conditions. This is understandable for studying efficacy outcomes; however, it decreases the external validity of these trials, especially for safety outcomes. Many RCTs are labeled efficacy and safety, while due consideration for power is provided only for efficacy outcomes. This, in turn, necessitates a discussion of the inadequacy of sample size (type II error) for identifying harm. This is particularly important in RCTs of TNF inhibitors, as harm related to these agents is still a matter of debate.

A recent survey<sup>8</sup> of all published RCTs examining TNF inhibitors in RA, psoriatic arthritis, and ankylosing spondylitis looked at whether: 1. trials were labeled as efficacy, safety, or efficacy and safety investigations; 2. power calculations were adequately explained; 3. the statistical tests of significance were given for harm; and, last, 4. any discussion of type II error for harm was present. Of the 34 articles surveyed, 24 (71%) were labeled as efficacy and safety. Among these, 23 (96%) did not include safety as a formal primary or secondary end point. In only 2/24 (8.3%), power calculations were given. Finally, in only 3/22 (14%) was there any discussion about the inadequate sample size (type II error) for detecting harm. Even for efficacy outcomes, the primary point of the RCT of adequate information about power calculation was only given in about two-thirds of the

published papers.

Along with this problem, is the image that is created that if there are no major signals in RCTs regarding safety of a new drug, it can be perhaps assumed as "safe" based on the use of the word "safety" in the title or the abstract. The real safety data may only be available in postmarketing reports and from the long-term registries and databases when the drug is used by many more patients than had been enrolled in the RCTs and when a sizable groups of patients with other comorbid conditions are exposed to the drug.

## Conclusions

RCTs are good tools for establishing efficacy of one drug over another or relative to a placebo control. However, because of their limitations, RCTs provide only limited safety data. This may seem obvious to many, yet data from these RCTs regarding safety are frequently presented as evidence of lack of harm. To better study the safety of new drugs, drug companies and regulatory authorities need to establish mandatory registries to follow any and all patients started on new medications, ideally, not only for safety but also for efficacy. This is the only way to ascertain whether a new medication is truly an improvement over already available treatments.

## Disclosure Statement

Yusuf Yazici, M.D., participates in the Speaker Bureau of BMS and is a Consultant to BMS, Celgene, Centocor, Genentech, Roche, and UCB.

## References

1. Ioannidis JP, Contopoulos-Ioannidis DG. Reporting of safety data from randomised trials. *Lancet*. 1998;352(9142):1752-3.
2. Yazici Y, Yazici H. A survey of inclusion of the time element when reporting adverse effects in randomised controlled trials of cyclo-oxygenase-2 and tumour necrosis factor alpha inhibitors. *Ann Rheum Dis*. 2007;66(1):124-7.
3. Yazici H. Lymphomas after tumor necrosis factor antagonist therapy: comment on the article by Brown et al. *Arthritis Rheum*. 2003;48(8):2389.
4. Bongartz T, Sutton AJ, Sweeting MJ, et al. Anti-TNF antibody therapy in rheumatoid arthritis and the risk of serious infections and malignancies: systematic review and meta-analysis of rare harmful effects in randomized controlled trials. *JAMA*. 2006;295(19):2275-85.
5. Curtis JR, Patkar N, Xie A, et al. Risk of serious bacterial infections among rheumatoid arthritis patients exposed to tumor necrosis factor alpha antagonists. *Arthritis Rheum*. 2007 Apr;56(4):1125-33.
6. Windeler J, Lange S. Events per person year--a dubious concept. *BMJ*. 1995 Feb 18;310(6977):454-6.
7. Paterson KR. Events per person year. Can obscure the true risk of certain adverse drug reactions. *BMJ*. 1995;310(6992):1470.
8. Yazici Y, Adler NM, Yazici H. Most tumour necrosis factor inhibitor trials in rheumatology are undeservedly called 'efficacy and safety' trials: a survey of power considerations. *Rheumatology (Oxford)*, 2008;47(7):1054-7.