Taking a Measured Approach to Toxicity Data in Phase I Oncology Clinical Trials

Authors: Manish R. Sharma¹, Mark J. Ratain¹

Affiliations of authors:
¹Department of Medicine and Committee on Clinical Pharmacology & Pharmacogenomics, University of Chicago, Chicago, IL

Corresponding author:
Mark J. Ratain, MD
mratain@medicine.bsd.uchicago.edu
5841 S. Maryland Avenue, MC 2115
Chicago, IL 60637-1470
Phone: (773) 702-4400
Fax: (773) 702-3163

Running Title: Measuring Toxicity in Phase I Clinical Trials

Conflicts of interest: None

Summary word count (max 50): 50

Word count (max 1200): 1,190

References (max 12): 9

Grant Support

This work was supported by UM1CA186705 from the National Cancer Institute (to M.J. Ratain) and K23GM112128 from the National Institute of General Medical Sciences (to M.R. Sharma).
Summary

The standard categorical system for assessing attribution of toxicity to study drug(s) in phase I trials is cumbersome and uninformative. While a binary system (‘related’ versus ‘unrelated’) would be sufficient to define maximum tolerated dose, a probability estimation would better support dose selection for randomized dose ranging phase II trials.
In this issue of *Clinical Cancer Research*, Eaton and colleagues (1) analyzed data from 11,909 toxicities on 38 phase I trials sponsored by the National Cancer Institute (NCI) and report that the rate of drug related toxicity increased with dose, while the rate of unrelated toxicity did not. They found that these relationships to dose were similar when ‘unrelated’ and ‘unlikely’ related toxicities were considered separately or grouped together as ‘unrelated’, and when ‘possibly’, ‘probably’ and ‘definitely’ related toxicities were considered separately or grouped together as ‘related’. They propose a simplified binary system of ‘related’ versus ‘unrelated’ when assessing the attribution of adverse events (AEs) to study drug(s) in phase I trials.

The Common Toxicity Criteria (now the Common Terminology Criteria for Adverse Events) was first developed in 1983 by the Cancer Treatment and Evaluation Program of NCI to standardize the language of AEs reported on NCI-sponsored trials. With respect to attribution, the International Conference on Harmonisation (ICH) in 1994 stated in its E2A guideline: “the expression ‘reasonable causal relationship’ is meant to convey in general that there are facts (evidence) or arguments to suggest a causal relationship” between the AE and the investigational agent(s)/intervention (2). Citing the ICH E2A guideline, CTEP requires that AEs be reported with one of five attributions to study drug: ‘definite’, ‘probable’, ‘possible’, ‘unlikely’ and ‘unrelated’, where these mean the AE ‘is clearly related’, ‘is likely related’, ‘may be related’, ‘is doubtfully related’, or ‘is clearly not related’ to the intervention, respectively (3).

The distinctions between adjacent categories (e.g., ‘possibly’ versus ‘unlikely’ related) are inherently ambiguous and subjective, although studies to formally assess interrater variability are lacking. Misattribution is common, with nearly 50% of AEs in the placebo arms of two randomized phase III trials attributed to study drug (4). Disease-related symptoms and drug toxicities can look similar, especially for AEs that are common, not previously known to be
associated with the drug (or its class), and not causally related to its mechanism of action. Fatigue is an AE that can easily be misattributed to study drug, while the acneiform rash associated with anti-EGFR therapy would almost never be misattributed.

The study by Eaton and colleagues has sound methodology and should inform the conduct of traditional first-in-human studies that aim to define a maximum tolerated dose (MTD). Most such studies are being conducted by industry sponsors (rather than NCI) and are already using a binary attribution system. However, several limitations of the study minimize the generalizability of the findings to modern oncology drug development. First, their analysis was restricted to monotherapy trials, whereas many current phase I trials are combination studies with the potential for pharmacokinetic and/or pharmacodynamic interactions between the drugs, and the issue of attribution to the investigational drug versus the standard drug (5). Second, they only considered toxicities recorded during the first cycle of therapy as defined in each protocol. Long-term and often less severe toxicities are important in assessing the tolerability of oral targeted therapies (6), and we are now in the era of investigational immunotherapy where immune-related adverse events may be delayed or insidious (7).

An underemphasized finding of this study is the high rate of misattribution of toxicities as being related to study drug when they are almost certainly not related. The y-intercepts on Figure 1A, from Eaton and colleagues’ article, demonstrate that ~60% of patients have Grade ≥ 1 toxicities and ~12% of patients have Grade ≥ 3 toxicities attributed to study drug at doses that are approaching zero as a percentage of the maximum administered dose. There are major differences between ‘possibly’, ‘probably’ and ‘definitely’ related attributions in this regard, as Figures 2C-E, from Eaton and colleagues’ article, show that ~60%, ~20%, and ~10% of patients have Grade ≥ 1 toxicities, respectively, that are attributed to study drug at these lowest doses.
The authors hypothesize that toxicities observed at the lowest doses may be idiosyncratic (i.e., dose independent), but it is not plausible for this to account for toxicities in ~60% of patients across a variety of drugs and therapeutic classes. Instead, it is much more plausible that these toxicities were documented by investigators as being related to study drug when they were actually due to disease or other factors, as in the placebo arms of the study by Hillman and colleagues (4).

A binary classification of ‘related’ versus ‘unrelated’ with respect to toxicity attribution is analogous to a classification of ‘responders’ versus ‘non-responders’ with respect to efficacy. Just as response rates do not tell the whole story regarding efficacy and are supplemented by waterfall plots and spider plots that depict change in tumor size as a continuous variable (8), a ‘related’ versus ‘unrelated’ classification would not tell the whole story regarding toxicity. A ‘possibly’ related toxicity is like ‘stable disease’ in that it can be due to either drug or disease. In modern oncology drug development, the primary objective of phase I trials should be to select doses for randomized dose ranging phase II trials, rather than to identify a MTD (9). The optimal selection of doses for a dose ranging phase II trial requires a sophisticated understanding of the relationship between dose-toxicity and dose-efficacy. The understanding of the relationship between dose and toxicity would be enhanced by assigning more weight to toxicities that are more likely to be related to study drug(s), as this would minimize the impact of misattribution discussed above. Weighting by attribution would also support quantitative comparisons between randomized arms in a phase II trial with respect to toxicity.

Figure 1 illustrates a proposed system for assessing attribution of AEs to study drugs using probability estimation. Investigators would electronically click on an arrow and move it along a probability distribution that ranges between 0 and 1. For the purpose of defining a dose-
limiting toxicity in a traditional phase I study, a probability of \( \geq 0.5 \) would be considered ‘related’ and a probability of \(< 0.5\) would be considered ‘unrelated’. The probability of attribution to study drug would change over time in a Bayesian manner, with revision in the setting of new information. For example, Grade 2 fatigue might be assigned a probability of 75% on Day 8, but this might be revised to 10% on Day 15 if the study drug was held for 7 days and the toxicity did not improve. While there would be variability in the probabilities assigned by different investigators evaluating the same patient, a continuous measure would minimize the impact of this variability compared to a categorical measure. Toxicities with a higher probability of attribution to study drug would be assigned greater weight. For example, if 10 patients in both Arms A and B had Grade 3 diarrhea but the average probability of attribution to drug was 60% in Arm A and 90% in Arm B, then Arm A would be considered more tolerable with respect to diarrhea. Quantitative measures of toxicity (based on a combination of frequency, severity, duration, and attribution), along with quantitative measures of efficacy, would provide maximum information to guide go/no-go decisions and dose selection for future studies. It is also important to avoid oversimplification of data in an era where there is great value in having rich data available for secondary analyses.
References

Figure 1. A proposed system for assessing attribution of adverse events (AEs) to study drugs using probability estimation. The investigator would move the arrow to provide a quantitative (and visual) assessment of the probability that an AE is related to study drug(s) at a point in time. In the hypothetical example below, the probability decreases between Day 8 and Day 15 because the toxicity did not resolve after the study drug was discontinued.
Figure 1:

“Related” = more likely than not to be related

Probability (Day 8):
0 0.5 1

“Unrelated” = more likely than not to be unrelated

Probability (Day 15):
0 0.5 1
Clinical Cancer Research

Taking a Measured Approach to Toxicity Data in Phase I Oncology Clinical Trials

Manish R. Sharma and Mark J. Ratain

Clin Cancer Res Published OnlineFirst October 14, 2015.

Updated version
Access the most recent version of this article at:

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.