Published in final edited form as:

Int J Radiat Oncol Biol Phys. 2014 June 1; 89(2): 232–234. doi:10.1016/j.ijrobp.2014.01.030.

Comparing long-term treatment-associated toxicities in cancer patients: Approaches, caveats, and recommendations

Matthew J. Schipper, Ph.D.,

Department of Radiation Oncology and Biostatistics, University of Michigan, M2531 SPHII, 1415 Washington Heights, Ann Arbor, MI 48109. Phone: (734) 232-1076, Fax: 734-936-4540

Jeremy MG. Taylor, Ph.D.,

Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109

Grace L. Smith, M.D., Ph.D., M.P.H., and

Department of Radiation Oncology, The University of Texas M. D. Anderson Cancer Center

Reshma Jagsi, M.D., Ph.D

Department of Radiation Oncology, University of Michigan

Matthew J. Schipper: mjschipp@umich.edu

A common research goal in radiation oncology is to compare the toxicity associated with two different treatment techniques. For example, one might compare toxicity that develops after intensity modulated radiotherapy (IMRT) versus three-dimensional conformal radiotherapy (3D-CRT) for a particular condition. Interest in such studies is growing as the field of radiation oncology attempts to demonstrate that the utilization of new technology is supported by meaningful clinical benefits. The gold standard for comparing two treatments remains the randomized controlled trial. However, in many circumstances, randomized trials are not practical or feasible. When no randomized data is available, observational data may be used. In this context, the choice of analytic methods is particularly important. Therefore, we wish to call attention to an important methodological issue in the analysis of toxicity after radiation treatment.

Specifically, several recent studies have used a metric known as the “rate ratio” to quantify the relative risk of toxicity between treatments. Often, the ‘rate’ of toxicity for a given treatment group is calculated as the number of events per total person-years of follow-up time. The rate ratio is simply the ratio of these rates for one treatment versus another. For example, in a recent study published in *Gynecologic Oncology*, Wright et al. used SEER-Medicare data to compare the incidence of toxicity between IMRT and conformal radiation among women with uterine cancer. They concluded that ‘women who received IMRT had a higher rate of bowel obstruction (rate ratio=1.41)’ [1]. Similarly, in a study published in *JAMA* in 2013, Sheets et al. also used SEER-Medicare data to compare various toxicities between IMRT and proton therapy for the treatment of men with prostate cancer [2]. Darby et al present a population based case-control study of the effect of radiation dose to the heart

in breast cancer patients on the subsequent risk of ischemic heart disease [3]. Other examples include a study of the effect of type of boost in breast cancer patients [4] and the effect of IMRT in HN cancer [5].

Although the rate ratio can be straightforward to calculate and may appear to be simple to interpret, its use is vulnerable to certain biases that are particularly important when applied to comparing toxicities with standard versus emerging radiation treatments. More specifically, rate ratios are likely to lead to biased conclusions if both of the following two conditions hold:

1. The toxicity does not occur at a constant rate over time.
2. The average length of follow-up differs between the two treatment groups.

Other than in randomized trials (with fixed randomization probabilities and contemporaneous treatment), these two conditions will often be true. For example, acute radiation toxicity is generally more common in the year immediately following treatment than in later years. Late effects are, by their very definition, not expected in early years of follow-up. Also, since analyses of interest in radiation oncology often compare a newer treatment modality (e.g. IMRT) with an older treatment (e.g. 3D-CRT) the average follow-up will often differ, as the newer treatment gains in popularity and use [6], [7].

Of note, the bias in the rate ratio could be in either direction (i.e. either over or under-estimating the relative difference). A simple example helps to illustrate the point. Imagine treatment A is the old standard treatment, whose popularity has waned over time, such that 90% of patients were treated with A in 2000, but by 2010, only 10% of patients are treated with A. Meanwhile, treatment B, a newer technique, grows in popularity over time, such that in 2000 when it has been newly introduced, only 10% of patients receive it, but by 2010, 90% of patients receive B. In addition, assume that the probability of toxicity is identical over time for treatments A and B. In this circumstance, regardless of the true rate of toxicity, one would expect the rate ratio to equal 1, signifying equal toxicity for both treatments. However, this is not necessarily the case. The curves in Figure 1 depict three hypothetical scenarios for the risk of toxicity. In scenario 1, risk (or as statisticians would call it, the “hazard”) of toxicity is high shortly after treatment but declines with time, as can be seen from the ‘flattening’ of the curve. This type of curve is typical of situations in which acute toxicity is common but late toxicity is less common. In scenario 2 a patient’s risk of toxicity is constant over the 10 year period. In scenario 3, the risk is near zero initially but then increases, as would be the case if acute toxicity were uncommon but late toxicity was more common. For each of these 3 scenarios, we used a simulation study to calculate the expected rates of toxicity along with the corresponding rate ratios (Table 1).

As noted above, because the true probability of toxicity is the same for both treatments at each timepoint, the rate ratios, if they were valid measures, should be very close to 1. However, for scenarios 1 and 3, the rate ratios are not, thus incorrectly implying that the risk of toxicity differs between the treatment groups. Depending on the exact shape of the toxicity curve over time, the estimated rate ratios could be greater than or less than 1. The intuitive explanation for the result under scenario 1, is that treatment A patients have longer follow-up times on average than do treatment B patients and toxicity is less common in later

years. Thus in treatment A, a greater proportion of the ‘person-years’ are in the earlier years, where the risk of toxicity is lower, resulting in a rate ratio (B vs A) greater than 1. In Scenario 3 where late toxicity is more common and treatment A patients tend to have longer follow-up, the rate ratio (B vs A) is less than 1, incorrectly implying less toxicity in treatment B patients. Only in scenario 2, where the risk of toxicity is constant over time, is the rate ratio estimate, as desired, equal to 1. Otherwise, the differential follow-up leads to bias.

The five papers cited above each use the rate ratio to capture relative risk, but do so using a variety of analyses and two types of study design (cohort or case-control). In the analysis by Wright et al, the estimated rate ratios for bowel obstruction and other GI toxicity were 1.41 and 1.49, respectively, seemingly implying an increased risk associated with IMRT. However the Kaplan-Meier (KM) curves for GI toxicity, which are completely overlapping, clearly show no difference between treatments. Rate ratios are calculated in a similar manner in [4]. The two treatment groups were matched with respect to several factors but not by treatment year and median follow-up was 16.7 years for the implant group versus 12.6 years for the electron group. Thus these rate ratios may also be subject to the bias described above. Each of the remaining 3 papers appear to estimate the rate ratio in a manner that appropriately deals with differential follow-up. The adjusted analysis in [2] weights patients according to their propensity to be treated with IMRT. This propensity score is calculated in part on the year of treatment and thus in this analysis, the differential follow-up mostly disappears (Table 2). However, this will not always occur with a propensity score-based analysis but only when the propensity score is based in part on treatment year. In [3] and [5] the authors appropriately use conditional logistic regression to analyze matched pairs data, and based on the rare disease assumption, utilize odds ratios to estimate rate ratios. Importantly, the matching and the subsequent strata for analysis were based on year of treatment ensuring similar duration of follow-up.

Therefore, we caution radiation oncologists interpreting analyses conducted using the rate ratio to consider whether there has been adequate attention to this potential for bias. Moreover, we suggest that those who are conducting future studies of toxicity outcomes with two different radiation technologies give consideration to alternative approaches. For example, rather than calculating rate ratios as defined above, we recommend use of hazard ratios from a survival model (e.g. Cox model). KM plots or cumulative incidence plots are useful graphical displays of the data. These techniques naturally account for differing follow-up times as seen in the example, where the expected hazard ratio was 1.00 in all 3 of the scenarios.

Finally, we note that while we have focused on toxicity outcomes, the same issues apply for any time to event type outcome, including progression or death. These approaches to quantifying risk help clinicians translate population-based study results into clinically meaningful estimates for assessing an individual patient’s risks and benefits with various treatment options.

References

1. Wright JD, Deutsch I, Wilde ET, et al. Uptake and outcomes of IMRT for uterine cancer. *Gynecologic Oncology*. 2013; 130(1):43–48. [PubMed: 23500087]
2. Sheets NC, Goldin GH, Meyer AM, et al. Intensity-modulated radiation therapy, proton therapy, or conformal radiation therapy and morbidity and disease control in localized prostate cancer. *Journal American Med Association*. 2012; 307(15):1611–20.10.1001/jama.2012.460
3. Darby SC, Ewertz M, McGale P, et al. Risk of Ischemic Heart Disease in Women after Radiotherapy for Breast Cancer. *New England Journal of Medicine*. 2013; 368:987–998.10.1056/NEJMoa120982 [PubMed: 23484825]
4. Hill-Kayser CE, Chacko D, Hwang W, et al. Long-term clinical and cosmetic outcomes after breast conservation treatment for women with early-stage breast carcinoma according to the type of breast boost. *Int J Radiation Oncology Biol Phys*. 2011; 79(4):1048–1054.
5. Graff P, Lapeyre M, Desandes E, et al. Impact of intensity-modulated radiotherapy on health-related quality of life for head and neck cancer patients: matched-pair comparison with conventional radiotherapy. *Int J Radiat Oncol Biol Phys*. 2013; 67(5):1309–1317. 2007. [PubMed: 17289292]
6. Bortfeld T, Marks L. Hype cycle in Radiation Oncology. *Int J Radiation Oncology Biol Phys*. 2013 Aug 1; 86(5):819–21.10.1016/j.ijrobp.2013.03.027
7. Goldin GH, Sheets NC, Meyer AM, Kuo TM, Wu Y, Stürmer T, Godley PA, Carpenter WR, Chen RC. Comparative effectiveness of intensity-modulated radiotherapy and conventional conformal radiotherapy in the treatment of prostate cancer after radical prostatectomy. *JAMA Intern Med*. 2013 Jun 24; 173(12):1136–43.10.1001/jamainternmed.2013.1020 [PubMed: 23689844]

Figure 1.
Hypothetical Example: Three possible scenarios of cumulative proportion of patients experiencing toxicity over time.

Table 1

Average rates of toxicity under each of 3 toxicity scenarios.

Scenario	Rate* for Treatment Group A	Rate* for Treatment Group B	Rate Ratio (B vs A)	Hazard Ratio **
1	10.1	13.2	1.31	1.00
2	7.1	7.1	1.00	1.00
3	3.6	2.2	.61	1.00

* Per 100 person-years

** Estimated from Cox model with treatment as covariate