Reviewer's report

Title: Testing the treatment effect on competing causes of death in oncology clinical trials

Version: 2 Date: 29 December 2013

Reviewer: Arthur Allignol

Reviewer's report:

This paper presents a simulation study that compares three tests for assessing treatment effect in the competing risks setting, while varying the degree of misclassification of the cause of failure and the censoring rate. I feel the authors are comparing apples and oranges in this article. That might be due to a deep misunderstanding of the competing risks framework.

* Major Compulsory Revisions

1. I don't understand why one would like to compare head to head a test for the cause-specific hazards (CSH) and the cumulative incidence function (CIF). These are two different quantities and their use should be driven by the subject matter and not through power considerations. The former is the instantaneous probability of experiencing an event while the latter is the cumulative probability of experiencing an event as time progresses.

2. Some statement in the paper are downright wrong:

- "the cause-specific hazard approach, based on the concept of latent failure times" (p. 1, l. -2).

The cause-specific hazards are not based on the concept of latent failure times. There is two ways to model competing risks data, either the latent failure time approach, or a stochastic process point of view (Andersen, Borgan, Gill and Keiding, 1993; Aalen, Borgan and Gjessing, 2008; Beyersmann, Allignol, and Schumacher, 2011). In the latter, the "independence of competing risks" problem is rendered irrelevant. In both cases, the cumulative cause-specific hazards can be estimated from the observed data using the Nelson-Aalen estimator.
"this cause-specific strategy has been criticized because it
requires the assumption of independence between the times to
different events" (p.3, l. 12)

Only computing a "cause-specific survival probability" using
the Kaplan-Meier estimator requires the independence
assumption. In this case, the Kaplan-Meier estimates have the
interpretation of being the probability of experiencing an
event if all the other possible causes of failure are
removed. That's usually not of interest. If you are interest in
the probability of experiencing an event in the presence of the
competing causes of failure, look at the CIF.

- "Gray proposed an alternative approach [...] prone to
misinterpretation" (p. 3)

Again, the independence assumption does not belong here. I
don't see why the CIF is prone to misinterpretation. The CIF is
the probability of experiencing an event of a specific type,
and it reduces to the number of events of type j / number of
individuals at risk in the absence of censoring. Of course,
if you kill everybody, less relapse will be
observed. Understanding the shape of the CIF is done through
the analysis of *ALL* cause-specific hazards (e.g., your
reference 4).

- "Moreover, the definition of the hazard of the CIF implies that
patients who experience an event remain in the risk sets of the
other types of event, which is counterintuitive" (p.3)

That is just a mathematical trick so that (1 -
exp(subdistribution hazard)) = CIF like in the usual survival
situation. If you are not comfortable with that, there is other
tests or regression models for the CIF (e.g., Gerds, Scheike
and Andersen, 2012, "Absolute risk regression for competing
risks: interpretation, link functions, and prediction" and
references therein.)

3. Section "Survival plots"
- "cause-specific survival": It has no meaning (see point 2). If you really need to plot something to illustrate the cause-specific analysis, I suggest a plot of the Nelson-Aalen estimates of the cumulative cause-specific hazards. Also recall that in the competing risks setting, there is *2* cause-specific hazards and *both* should be analysed for a complete understanding of the data to be had (e.g., Putter at al's tutorial.)

Also note that to estimate the cause-specific hazards, "only deaths declared “due to cancer” (CD) are considered as events of interest while all other deaths are censored" is a programming trick in order to only count the number of CD events. See the definition of the Nelson-Aalen estimator for example. Death form other causes can be considered as independent censoring following Aalen's definition (see Andersen, Borgan, Gill and Keiding, 1993; Aalen, Borgan and Gjessing, 2008). Independent censoring is defined wrt the intensity process and has nothing to do with stochastic independence.

- "For death by cause, the curves correspond to the CIF, as described by Gray, with patients dying of competing causes remaining at risk."

That's not true. See e.g., your reference (3) for a definition of the estimator of the CIF. It is the integral of the overall survival probability times the increments of the cause-specific hazard for the event of interest. For all these ingredients, patients that experience a competing event are removed from the risk set.

4. "We simulated data with negative, null, and positive correlations, thereby covering an exhaustive range of dependence assumptions" (p. 2, l -9).

In terms of the process point of view, that is useless. See also Beyersmann et al. (2009) "Simulating competing risks data in
survival analysis”.

**Level of interest:** An article of limited interest

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interest