HAL
open science

The Use of Model-Based Tumor-Size Metrics to Predict Survival

Benjamin Ribba, Nick Holford, France Mentré

To cite this version:

Benjamin Ribba, Nick Holford, France Mentré. The Use of Model-Based Tumor-Size Metrics to Predict Survival. Clinical Pharmacology and Therapeutics, 2014, 96 (2), pp.133 - 135. 10.1038/clpt.2014.111 . inserm-01084276

HAL Id: inserm-01084276

https://inserm.hal.science/inserm-01084276v1

Submitted on 18 Nov 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On The Use of Model-Based Tumor-Size Metrics To Predict Survival

Benjamin Ribba[1], Nick Holford[2], France Mentré[3]

[1] Inria, project-team NuMed, Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69007 Lyon Cedex 07, France.

[2] Department of Pharmacology and Clinical Pharmacology, University of Auckland, Auckland, New Zealand.

[3] IAME, UMR 1137, INSERM & Université Paris Diderot, Sorbonne Paris Cité, Paris, F-75018, France.

Acknowledgements of research support: None

Corresponding author: benjamin.ribba@inria.fr. Inria Sophia Antipolis – Méditerranée 2004, route des Lucioles – BP 93, 06902 Sophia Antipolis Cedex, France. Tel: +33 4 89 73 24 51

Running title: Analysis of survival using model-based tumor-size metrics

Previous presentation of the work: not applicable

Disclaimers: None

Mixed-effect models are increasingly being used to analyze the time-course of tumor size and to identify tumor size metrics as predictors of overall survival in cancer patients. However, caution is needed interpreting the results of such analyses and applying them to clinical research. Sparse individual tumor data as a result of patient mortality can lead to falsely conclusions about tumor size metrics as predictors of overall survival.

Body text

Using early tumor size response to predict long-term clinical outcome, such as overall survival, may shorten the time needed to evaluate the effectiveness of treatments [1]. A recent review has summarized efforts to model continuous longitudinal tumor size data to describe the observed time course of tumor size and the extension to prediction of survival [2]. By using time-to-event approaches, metrics derived from the time course of individual tumor size changes have been tested as predictors of overall survival. *Clinical Pharmacology and Therapeutics* was the first journal to publish such an analysis [1]. Similar studies have also been reported for colorectal cancer, thyroid cancer, breast cancer and gastrointestinal stromal tumor (see [2]).

In all these studies, the authors used a mixed-effects model to analyze the observed tumor size data. They then estimated individual patient parameters using all available tumor size measurements (so called individual empirical Bayes estimates). These individual parameter estimates were used to calculate certain tumor size response metrics, which were subsequently tested as predictors of overall survival.

In the study published in the journal in 2009, Wang et al. identified the percentage tumor size reduction from baseline at week 8 as a tumor metric which appeared to be a significant predictor of overall survival in non-small cell lung cancer patients [1]. The authors concluded that this metric might serve as an early biomarker for survival prediction to help early drug development decisions. The same year, Claret et al. identified the change in tumor size from baseline at week 7 as a predictor of overall survival in colorectal cancer patients [3]. In a similar study published in 2013 [4], Claret et al. compared two tumor metrics, the tumor size reduction at week 6 and the time to

tumor growth, as predictors of overall survival in patients with colorectal cancer and concluded the time to tumor growth metric was a superior predictor.

It has been known for some time, however, that analysis of survival by tumor response may be misleading about the effect of treatment on survival [5, 6]. Herein, we focus on a methodological problem related to the use of mixed-effects models to predict individual tumor size metrics with limited numbers of observations per patient. When the individual data (here the tumor size measurements) are not informative enough, the individual parameter estimates will shrink towards the typical value in the population with similar fixed effects (e.g. treatment exposure) without accounting for actual individual differences [7]. Early occurrence of death is the primary reason why some individuals may have a reduced number of tumor size measurements – and consequently greater shrinkage of parameter estimates. Thus, it appeared important to address how shrinkage of individual parameter estimates might influence the Type 1 error of falsely detecting or failing to detect specific tumor size metrics as predictors of overall survival.

To investigate the problem, we simulated tumor size data by using a simplified version of an empirical tumor size model reported by Claret et al. [4]. This model describes tumor size *TS* as a function of time *t* since treatment initiation:

$$TS(t) = TS\,(0) \cdot exp\left[KL \cdot t - \frac{KDE}{\lambda} \cdot \left(1 - exp(-\lambda \cdot t)\right)\right] \qquad (1)$$

and involves the tumor growth rate constant (*KL*, week^{-1}) and tumor decrease rate constant (*KDE*, week^{-1}) plus an exponential change with time of decrease according to the parameter λ (week^{-1}). λ may be attributed either to washout of drug treatment or development of resistance to continuing treatment. This model is similar to the one used by Wang et al. in 2009 [1].

We used this model (equation 1) to fit simulated tumor size data and to estimate tumor size response metrics. As in [4], we investigated two tumor size metrics: the first one is

the tumor size ratio, *TSR6*, defined by the predicted tumor size at week 6, *TS(6)*, relative to the predicted baseline, *TS(0)*:

$$TSR6 = \frac{TS(6)}{TS\ (0)} \tag{2}$$

This metric is similar to the metric used in [1].

The second is the time to tumor growth in weeks, *TTG*, calculated from:

$$TTG = \frac{\log(KDE) - \log\ (KL)}{\lambda} \tag{3}$$

To simulate the tumor size data, we used the parameter values and distributions as reported in [4]. We assumed random effects on all four parameters. Overall 1,000 replicates of datasets each with 500 patients were created with a simulated tumor size measurement every 6 weeks from baseline (0) until 102 weeks.

Survival times were simulated independently from tumor size by sampling from an exponential distribution. The influence of parameter shrinkage on Type 1 error was explored through 9 different simulation scenarios with mean survival times of 6, 12, 24, 36, 48, 60, 80, 120 and 180 weeks. As in the studies mentioned earlier, the simulated tumor size data was used to obtain empirical Bayes estimates (EBE) of individual tumor size parameters using a mixed-effect approach (NONMEM, version 7.2; Icon Development Solutions, Ellicot City, MD). The EBE of individual parameters were used to calculate individual derived metrics for TSR6 and TTG.

In all the different scenarios, the estimation was performed using individual simulated observations until the time of death or 102 weeks (end of study). The median (2.5th and 97.5th percentile) simulated TSR6 metric was 0.84 (0.49 – 1.08). For the TTG metric, the median simulated TTG was 22.7 weeks (-1.6 – 102.2). Reference values of the two metrics were also calculated from the same individual "true" parameters used to simulate tumor size course in each patient. These reference values for the two metrics are not affected by shrinkage. Shrinkage on a metric was calculated from 1 – (variance of EBE metric)/(variance of "true" metric) [7]. The potential influence of each of the tumor

metrics on survival was analyzed using a Cox model assuming a linear relationship of the covariate to the log hazard. All patients were included in the survival analysis including those patients with only baseline measurements (approximately 3% of patients in the scenario with mean survival time 180 weeks and 60% in the scenario with mean survival time 6 weeks).

The results of the analysis are summarized in Figure 1. The type 1 error of TSR6 was outside the prediction interval in the scenario with a mean survival time of 6 weeks. At longer survival times (36 to 180 weeks) the Type 1 error was at or below the prediction interval. The type 1 error was increased for TTG for all scenarios, and approached a peak of 43.4% in the scenario with a mean survival time of 48 weeks. The smaller inflation of Type 1 error observed with very short mean survival times of 6 and 12 weeks is linked to a very high shrinkage for TTG reducing the variability of the metrics across patients. At the extreme, all patients would have the same TTG and thus, there would be no link with the distribution of survival times.

In contrast TSR6 only had an increased Type 1 error in the scenario with a mean survival of 6 weeks. With longer mean survival times TSR6 is expected to have little shrinkage because later observations have little influence on the predicted tumor size at 6 weeks. However, at even longer survival times (36 to 180 weeks) the Type 1 error was at or below the prediction interval.

When the "true" simulated individual parameters were used to calculate reference values of TTG (ITTG) and TSR6 (ITSR6), the Type 1 error rates were all within the prediction interval (see Figure 1). These results demonstrate that the increased Type 1 error is a consequence of different degrees of shrinkage of the individual estimates. It is not due to differences in survival distribution in relation to the distribution of TTG because the Type 1 error is not inflated when the "true" parameter values are used to calculate tumor metrics for each survival time scenario.

On the basis of our findings we want to highlight that the use of a metric similar to TSR6 as used by Wang et al. [1] could be problematic in making appropriate drug development decisions because the Type 1 error rate is too high when the mean survival

time is similar to the time of tumor size ratio evaluation and too low when mean survival times are longer. The use of the metric TTG is even more problematic with substantial inflation of Type 1 error, especially, when mean survival time is close to TTG (Figure 1).

In conclusion, mixed-effect models of tumor size dynamics may be useful for leveraging the available tumor size data in order to improve assessment of drug effectiveness in the Phase I and Phase II of drug development. However, when using these approaches to derive individual model-based metrics used in survival analysis, investigators should be aware of the potential for misleading results such as biased population parameters [8] and inaccurate Type 1 errors arising from limited individual tumor size data. We assumed survival as the only event determining shrinkage and did not include dropout events due to other causes, and in particular, disease progression. Note that accounting for informative dropout due to disease or tumor size progression might lead to different results. Herein, a simple exponential distribution (constant hazard) was assumed for survival times. Even in this straightforward case, we show that the use of substantially shrunken individual parameter estimates can lead to false acceptance or rejection of the null hypothesis when evaluating model-based tumor size metrics as predictors of survival. Further methodological work has to be performed to evaluate the relevance of more advanced metrics (e.g. using the full time course of tumour shrinkage and growth [9] and/or using joint modeling) linking mixed-effect models of tumor size to survival times.

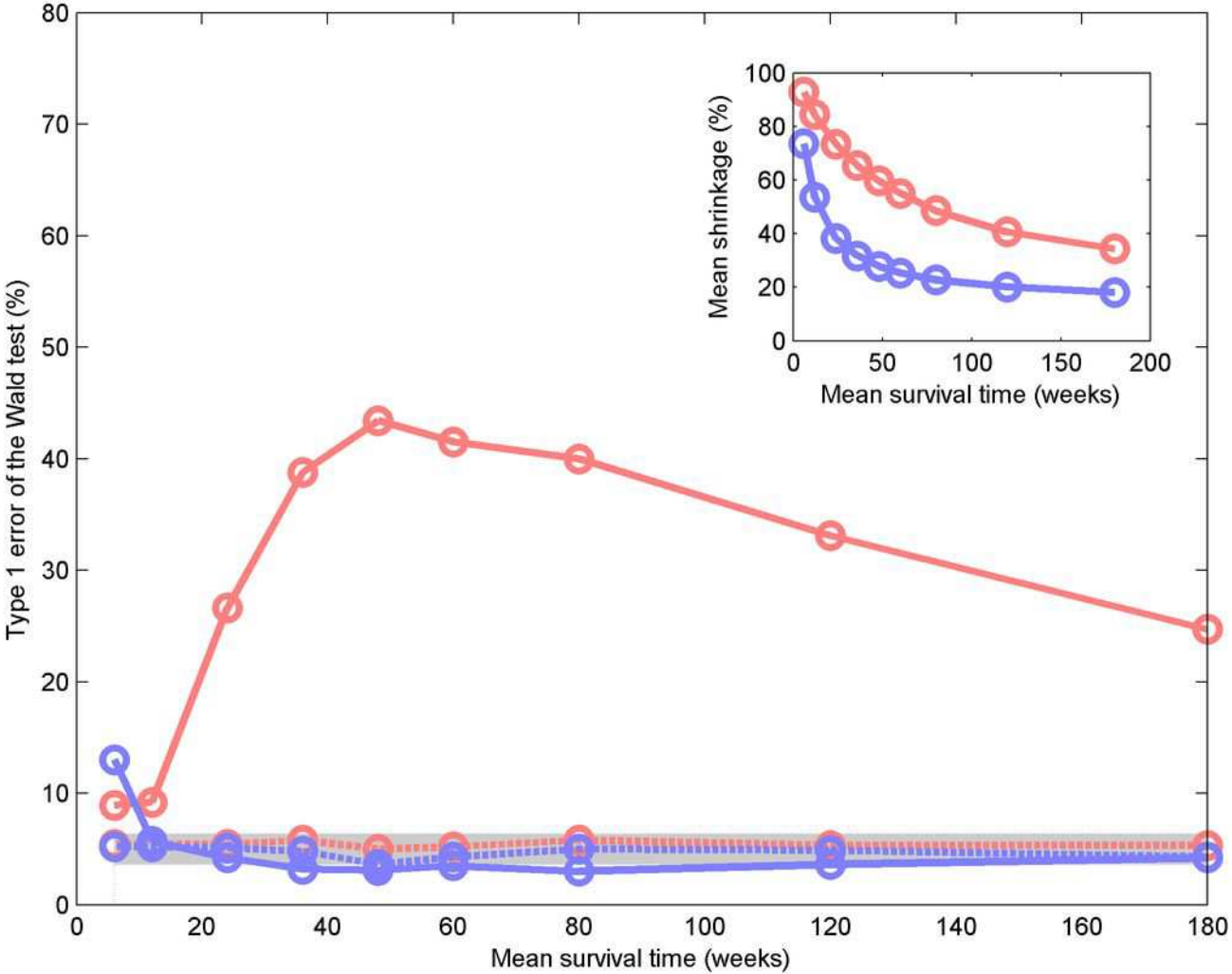
There is no conflict of interest

References

1. Wang, Y., et al., *Elucidation of relationship between tumor size and survival in non-small-cell lung cancer patients can aid early decision making in clinical drug development.* Clinical pharmacology and therapeutics, 2009. **86**(2): p. 167-74.
2. Bender, B.C., E. Schindler, and L.E. Friberg, *Population pharmacokinetic pharmacodynamic modelling in oncology: a tool for predicting clinical response.* Br J Clin Pharmacol, 2013.
3. Claret, L., et al., *Model-based prediction of phase III overall survival in colorectal cancer on the basis of phase II tumor dynamics.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2009. **27**(25): p. 4103-8.
4. Claret, L., et al., *Evaluation of tumor-size response metrics to predict overall survival in Western and Chinese patients with first-line metastatic colorectal cancer.* J Clin Oncol, 2013. **31**(17): p. 2110-4.
5. Buyse, M. and P. Piedbois, *On the relationship between response to treatment and survival time.* Stat Med, 1996. **15**: p. 2797-812.
6. Anderson, J.R., K.C. Cain, and R.D. Gelber, *Analysis of survival by tumor response and other comparisons of time-to-event by outcome variables.* J Clin Oncol, 2008. **26**(24): p. 3913-5.
7. Karlsson, M.O. and R.M. Savic, *Diagnosing model diagnostics.* Clin Pharmacol Ther, 2007. **82**(1): p. 17-20.
8. Bonate, P.L. and B. Suttle, *Effect of censoring due to progressive disease on tumor size kinetic parameter estimates.* AAPS J, 2013. **15**(3): p. 832-9.
9. Hansson, E.K., et al., *PKPD Modeling of VEGF, sVEGFR-2, sVEGFR-3, and sKIT as Predictors of Tumor Dynamics and Overall Survival Following Sunitinib Treatment in GIST.* CPT Pharmacometrics Syst Pharmacol, 2013. **2**: p. e84.

Legend

Figure 1: Relationship of the Type 1 error of the Wald test with mean survival times for tumor size reduction (TSR6 blue) and time to tumor growth (TTG red continuous line) metrics. The Type 1 error of falsely concluding an association between each tumor metric and hazard was estimated from the frequency of rejecting the null hypothesis using a Wald test of size 5%. 1,000 data sets were created for each of the 9 survival simulation scenarios. By design, the survival distribution was independent of tumor size, so that the empirical Type 1 error rates are all expected within the 95% prediction interval. The grey band represents the 95% prediction interval around 5% Type 1 error (for 1,000 replicates: 3.65%-6.35%). The dashed lines stand for the values of the metrics calculated using the reference "true" individual parameter values. Inset: Shrinkage of TSR6 and TTG with mean survival time.


Type 1 error of the Wald test (%)
Mean survival time (weeks)
0
10
20
30
40
50
60
70
80
0
20
40
60
80
100
120
140
160
180
Mean shrinkage (%)
Mean survival time (weeks)
0
20
40
60
80
100
0
50
100
150
200