

## Introduction

A **composite endpoint** is generally defined as an outcome combining several endpoints of interest within a single variable in order to achieve a higher overall event rate [5].

We propose a **composite** endpoint of either death, or progression (**only if death is not observed**), measured from randomization, to be used as a surrogate for OS in clinical trials. The main advantage is that a higher proportion of events will be related to the true endpoint of interest (OS), as compared to other commonly used surrogates (e.g., DFS).

## Methodology

Although meaningful, the proposed composite endpoint has to be a **validated surrogate** in the disease setting of interest. In the current project, we conduct a survival analysis to examine the clinical benefit captured by the true endpoint of interest (OS) and the proposed composite endpoint, in a dataset with 5099 breast cancer patients randomized either to trastuzumab or observation (8 years median follow-up). We have also compared DFS with the proposed composite endpoint in terms of clinical benefit and surrogacy to OS.

Kaplan-Meier survival curves are produced for each of the three endpoints and a comparison among treatment arms is performed using the log-rank test with the following hypotheses and test statistic:

- $H_0$ : No difference in survival between treatment arms.
- $H_1$ : The survival difference between treatment arms is statistically significant.

$$\chi^2 = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2}, \quad (1)$$

where  $O_1$ ,  $O_2$  and  $E_1$ ,  $E_2$  are the total number of observed and expected events in each treatment arm, respectively.

Multivariate Cox proportional-hazards (PH) models were deployed to investigate the impact of other independent variables to the survival function. The hazard  $h(t)$  is defined as the probability of dying at a given time, assuming that the patients have survived until that time [4].

$$h(t, Z) = h_0(t) \exp^{\beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_p Z_p}, \quad (2)$$

where  $Z_1, \dots, Z_p$  are the covariates of interest and  $h_0(t)$  is the baseline hazard function.

We then assess the surrogacy at trial-level ( $R_{trial}^2$ ) and individual-level ( $R_{ind}^2$ , *data not presented*), using the information theory [2] [3].

At the trial-level, a two-stage approach will be used [3]. In stage 1, the following trial-specific Cox proportional hazard models are fitted:

$$S_{ij}(t) = S_{i0}(t) \exp^{\alpha_i Z_{ij}}, \quad (3)$$

$$T_{ij}(t) = T_{i0}(t) \exp^{\beta_i Z_{ij}}, \quad (4)$$

where  $S_{i0}(t)$  and  $T_{i0}(t)$  are the trial-specific base-line hazard functions,  $Z_{ij}$  is the treatment indicator for subject  $j$  in trial  $i$ , and  $\alpha_i$ ,  $\beta_i$  are the trial-specific treatment effects on  $S$  and  $T$ , respectively. Next, the second stage of the analysis is conducted:

$$\hat{\beta}_i = \lambda_0 + \lambda_1 \hat{\alpha}_i + \varepsilon_i, \quad (5)$$

where the parameter estimates for  $\beta_i$  and  $\alpha_i$  are based on the full model that was fitted in stage 1. The classical coefficient of determination of the fitted stage 2 model provides an estimate of  $R_{trial}^2$ .

At the individual-level, the surrogate endpoint  $S$  is redefined as a time-dependent covariate  $S(t)$ , taking value 0 until the surrogate endpoint occurs and 1 thereafter [1].

## Results

The composite endpoint encompasses as events deaths and relapses (only when death is not observed). Here, 40% in each arm are relapses.

Results

Trastuzumab has a clear survival benefit over observation for OS, DFS and the composite endpoint (log-rank for all  $p < 0.0001$ ).

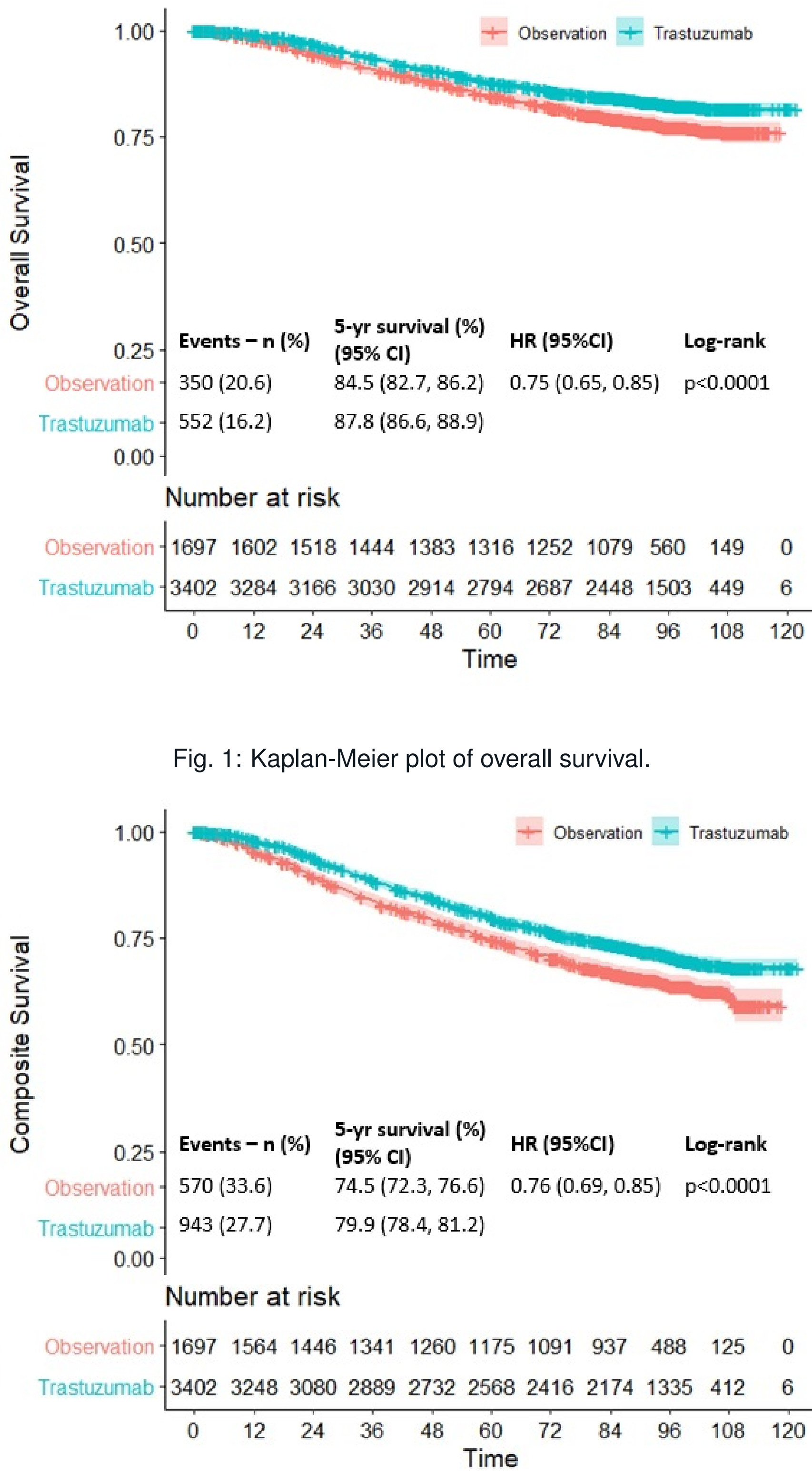
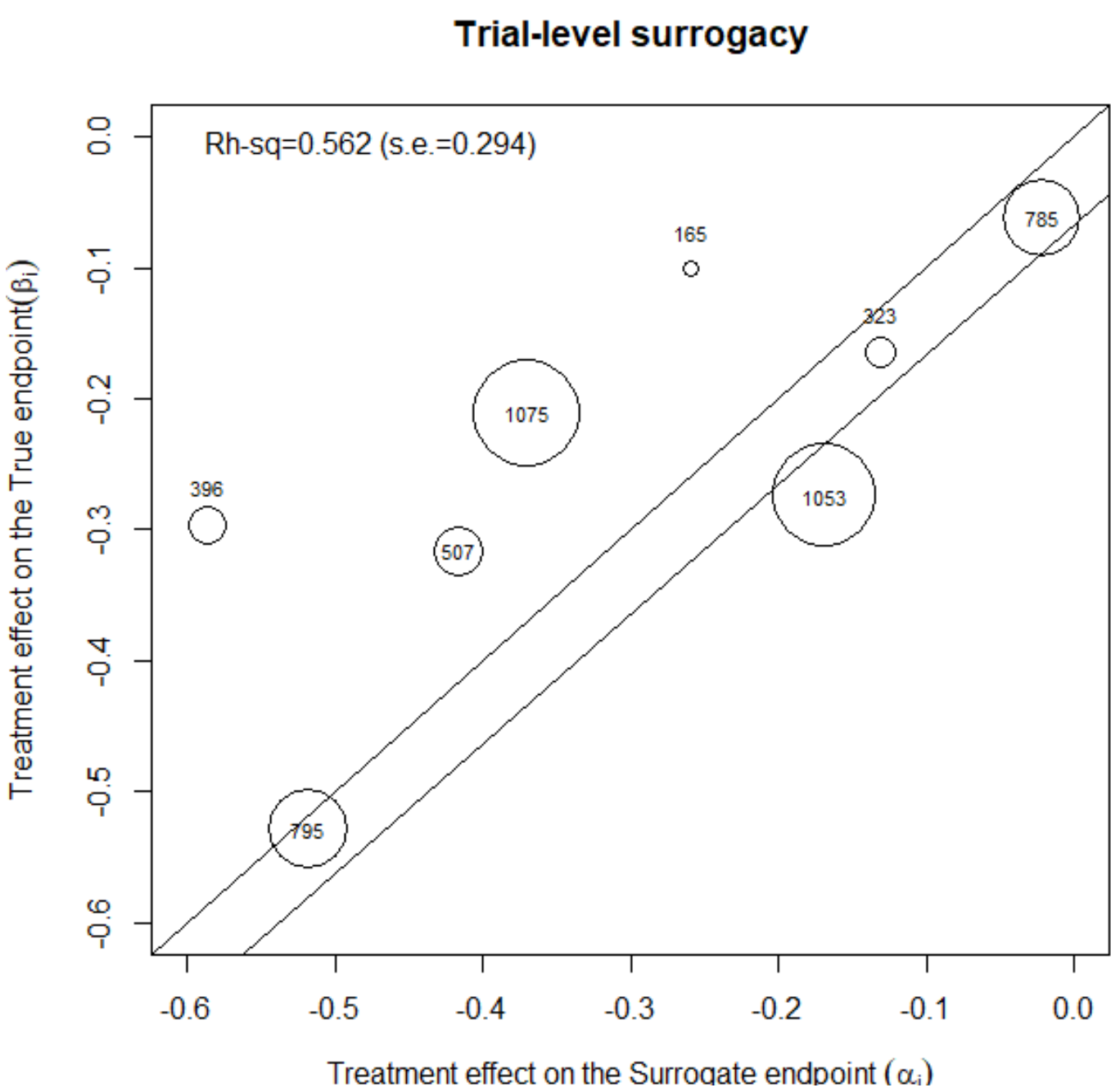


Fig. 2: Kaplan-Meier plot of the composite endpoint.

In multivariate Cox PH models, except from treatment arm, the clinically relevant ER status and nodal status were found to be statistical significant for OS, while for both surrogate endpoints, menopausal status was also found significant.

| Model                                     | OS    |              |         | Composite |              |         | DFS   |              |         |
|---|-------|--------------|---------|-----------|--------------|---------|-------|--------------|---------|
|   | HR    | 95% CI       | p-value | HR        | 95% CI       | p-value | HR    | 95% CI       | p-value |
| <b>Treatment arm</b>                      |       |              |         |           |              |         |       |              |         |
| Trastuzumab vs. Observation               | 0.734 | (0.64, 0.84) | <.0001  | 0.757     | (0.68, 0.84) | <.0001  | 0.742 | (0.67, 0.82) | <.0001  |
| <b>Estrogen-receptor status</b>           |       |              |         |           |              |         |       |              |         |
| Positive vs. Negative                     | 0.687 | (0.60, 0.79) | <.0001  | 0.764     | (0.69, 0.85) | <.0001  | 0.756 | (0.68, 0.84) | <.0001  |
| <b>Nodal status</b>                       |       |              |         |           |              |         |       |              |         |
| 1-3 vs. >=4                               | 0.398 | (0.34, 0.47) | <.0001  | 0.549     | (0.48, 0.62) | <.0001  | 0.530 | (0.47, 0.60) | <.0001  |
| Negative vs. >=4                          | 0.228 | (0.19, 0.28) | <.0001  | 0.357     | (0.31, 0.41) | <.0001  | 0.345 | (0.30, 0.40) | <.0001  |
| Not assessed vs. >=4                      | 0.918 | (0.76, 1.10) | 0.3637  | 0.820     | (0.70, 0.96) | 0.0137  | 0.826 | (0.71, 0.97) | 0.0177  |
| <b>Menopausal status at randomisation</b> |       |              |         |           |              |         |       |              |         |
| Post- vs. Pre-menopausal                  | -     | -            | -       | 0.844     | (0.73, 0.98) | 0.0242  | 0.828 | (0.71, 0.96) | 0.0122  |
| Uncertain vs. Pre-menopausal              | -     | -            | -       | 0.801     | (0.69, 0.93) | 0.0042  | 0.791 | (0.68, 0.92) | 0.0025  |

For the surrogacy assessment, the recruiting centers were used for the trial-level surrogacy. The resulting treatment effects on the composite was only moderately correlated with treatments effects on OS but the variance of the estimate was very large  $R^2_{trial} = 0.56$  (s.e.=0.29). When excluding the three smaller centers,  $R^2_{trial}$  is increased to 0.66 (s.e.=0.39).



Future steps

- Assess the impact of crossover to the treatment benefit of the composite endpoint and OS.
- Explore individual-level surrogacy measures and adjust for significant predictors, to possibly increase the trial-level surrogacy.
- Evaluate the applicability and usefulness of the proposed composite endpoint to different treatment settings or various simulated scenarios (e.g. different percentage of similarity between OS and the composite).

References

[1] A. Alonso and G. Molenberghs. Evaluating time-to-cancer recurrence as a surrogate marker for survival from an information theory perspective. *Statistical Methods in Medical Research*, 17:497–504, 2008.

[2] T. Burzykowski, G. Molenberghs, and M. B. et al. Validation of surrogate endpoints in multiple randomized clinical trials with failure-time endpoints. *Applied Statistics*, 50:405–22, 2001.

[3] M. Buyse, S. Michiels, and K. L. et al. Leukemia-free survival as a surrogate end point for overall survival in the evaluation of maintenance therapy for patients with acute myeloid leukemia in complete remission. *Haematologica*, 96:1106–12, 2011.

[4] D. Collett. *Modelling Survival Data in Medical Research*. Chapman and Hall/CRC, 2015.

[5] G. Rauch, S. Schöler, and M. Kieser. *Planning and Analyzing Clinical Trials with Composite Endpoints*. Springer International Publishing, Switzerland, 2017.