
**eAppendix 1.** Generalized Pairwise Comparisons

**eAppendix 2.** Details of the Simulation Parameters

This supplementary material has been provided by the authors to give readers additional information about their work.
eAppendix 1. Generalized Pairwise Comparisons

A full description of generalized pairwise comparisons has been previously published.\textsuperscript{12} We restrict our discussion to the analysis of survival data from randomized trials comparing an experimental to a control group. Pairwise comparisons are carried out on all possible pairs of patients, one from the experimental group (group T) and the other from the control group (group C). Let $x_i$ and $y_j$ be the survival times of a patient $i$ from group T ($i = 1, \ldots, n_T$) and of a patient $j$ from group C ($j = 1, \ldots, n_C$). A pair is classified as ‘favorable’ if the survival of patient $i$ is longer than that of patient $j$ by at least $m$ months, i.e., $x_i - y_j > m$; ‘unfavorable’ in the opposite situation, i.e., $x_i - y_j < -m$; and ‘neutral’ in all other cases, i.e. $-m \leq x_i - y_j \leq m$, or one or both of the survival times is censored such that the pair cannot be classified as favorable or unfavorable. A pairwise score, noted $p_{ij}(m)$, takes the value 1, 0, or -1, in each of these respective cases. The net chance of a longer survival, called “proportion in favor of treatment” in the original publication, is calculated as the sum of the pairwise scores over all the pairs that can be formed between one patient from the treatment group and one patient from the control group:

$$\Delta(m) = \frac{\sum_{i=1}^{n_T} \sum_{j=1}^{n_C} p_{ij}(m)}{n_T \cdot n_C}$$

This formula can be rewritten as

$$\Delta(m) = \mathbb{P}[x_i > y_j + m] - \mathbb{P}[y_j > x_i + m]$$

where $\mathbb{P}[x_i > y_j + m]$ and $\mathbb{P}[y_j > x_i + m]$ are, respectively, the probabilities for a random pair to be favorable or unfavorable to the treatment. A confidence interval for $\Delta(m)$, and a test of statistical significance, can be computed using a randomization test. When $m = 0$, this test is exactly equivalent to the Gehan-Wilcoxon test statistic. When
$m > 0$, this test can be viewed as a generalization of the Gehan-Wilcoxon test statistic.\textsuperscript{13}

Further generalizations of the test statistic $\Delta(m)$ assign weights to the pairwise scores $p_{ij}(m)$ to achieve better power under proportional hazards.
eAppendix 2. Details of the Simulation Parameters

In the simulations, survival times have been assumed to follow a Weibull distribution, which is more general than the negative exponential distribution and as such, fits a wide variety of survival curves actually observed in many types of cancer. The Weibull distribution has a shape parameter denoted $k$ (if $k = 1$ the Weibull distribution reduces to a negative exponential distribution) and a scale parameter denoted $\lambda$. In all scenarios, survival times in the control group followed a Weibull distribution with $k = 2$ and $\lambda_c = 11.5$.

Scenario 1: proportional hazards

Survival times in the experimental group followed a Weibull distribution with $k = 2$ and $\lambda_T = 13.4$. The hazard ratio was constant and equal to 0.75 ($= (11.5)^2 / (13.4)^2$).

Scenario 2: early survival differences

Survival times in the experimental group followed a Weibull distribution with $k = 2$ and $\lambda_T$ such that the hazard ratio was set at 0.4 between 0 and 4 months, to 0.55 between 4 and 8 months, to 0.7 between 8 and 12 months, to 0.85 between 12 and 16 months, and to 1 (no effect) after 16 months.

Scenario 3: delayed survival differences

Survival times in the experimental group followed a Weibull distribution with $k = 2$ and $\lambda_T$ such that the hazard ratio was set at 1 (no effect) between 0 and 4 months. It decreased to 0.875 between 4 and 8 months, to 0.7 between 8 and 12 months, to 0.5 between 12 and 16 months and to 0.3 thereafter.

Scenario 4: curable disease

© 2016 American Medical Association. All rights reserved.
Survival times in the experimental group followed the same Weibull distribution as in the control group, except for 10 percent of the patients selected at random, who were assumed cured (very long survival times).

**Scenario 5: crossing hazards**

The survival times in the experimental group were generated from two distinct distributions. The survival times of half of the patients were generated using a constant hazard ratio equal to 2.5 (detrimental effect of the experimental treatment), and the survival times of the other half of the patients were generated using a constant hazard ratio equal to 0.5 (benefit of the experimental treatment).