APPENDIX

A DERIVATION OF TEST STATISTIC FOR $K > 2$ GROUPS

Here, we extend our weighted log rank test procedure to compare $K > 2$ groups. We will assume an incomplete within-cluster group structure, as the complete within-cluster group structure is a special case. Recall that the weighted countinga nd at risk processes for group k were defined as

$$\begin{aligned}\hat{N}_k(t) &= \sum_{i=1} M \sum_{j=1}^{n_i} \omega_{ik} I\left[X_{ij} \le t, \delta_{ij}, G_{ij} = k\right] \\ \hat{Y}_k(t) &= \sum_{i=1} M \sum_{j=1}^{n_i} \omega_{ik} I\left[X_{ij} \ge t, G_{ij} = k\right],\end{aligned}$$

where the weights ω_{ik} are as defined in the main manuscript. When more than two groups are present, the counting processes $\hat{N}_k(t)$ and $\hat{Y}_k(t)$ follow the same formula, with weights defined as

$$\omega_{ik} = \begin{cases} (K_i^c n_{ik})^{-1} & \text{if} n_{ik} > 0 \\ 0 & \text{if} n_{ik} = 0 \end{cases}$$

where K_i^c is the number of groups having observations in cluster i, i.e., $K_i^c = \sum_{k=1}^{K} I\left[n_{ik} > 0\right]$. The rationale for this particular weight comes from consideration of the modified WCR procedure discussed in the main manuscript. In cluster i,a group is selected with equal probability from the groups available in the given cluster, i.e., with probability $(K_i^c)^{-1}$. Then, an observation is selected with equal probability from all observations from the selected group, i.e., with probability n_{ik}^{-1} for each group $k = 1, \ldots, K$. The two-step WCR procedure then suggests the weight $(K_i^c n_{ik})^{-1}$ for the weighted counting processes. If group k is not represented in cluster i, then, clearly, this group is selected with probability 0, and thus the weight for cluster i is also 0.The test statistic for group k is then defined as before

$$\hat{Z}_k(t) = \int_0^t \mathrm{d}\hat{N}_k(s) - \frac{\hat{Y}_k(s)}{\hat{Y}(s)} \mathrm{d}\hat{N}(s),$$

where $\hat{N}(t) = \sum_{i=1}^{K} \hat{N}_k(t)$ and $\hat{Y}(t) = \sum_{i=1}^{K} \hat{Y}_k(t)$. Let $\hat{\mathbf{Z}}(t) = \left(\hat{Z}_1(t), \ldots, \hat{Z}_K(t)\right)$ be the vector of test statistics. The estimated covariance matrix of $\hat{\mathbf{Z}}(t)$, denoted $\hat{\mathbf{\Sigma}}(t)$, can be estimated via jackknife as described in the main manuscript. The covariance matrix has rank $K - 1$, so we construct the test statistic as the quadratic form

$$X^2(t) = \hat{\mathbf{Z}}^T_{-k}(t) \hat{\mathbf{\Sigma}}_{-k}(t) \hat{\mathbf{Z}}_{-k}(t),$$

where the subscript $_{-k}$ denotes removal of the k^{th} element of $\hat{\mathbf{Z}}(t)$ and the k_{th} row and column of $\hat{\mathbf{\Sigma}}(t)$, where k can be arbitrarily selected. Under typical regularity conditions, it is expected that $X^2(t)$ follows the χ^2_{K-1} distribution and we reject $H_0 : S_1(t) = \ldots = S_K(t)$ if $X^2(t) > \chi^2_{K-1,1-\alpha}$. This procedure, calculating the quadratic form with the removal of one group, corresponds to the traditional K sample test for i.i.d. data.[11]