APPENDIX

Here we generalized the results of Ghosh and Lin (2000) to obtain the decomposition of $\sqrt{m}\{\hat{F}_1(t) - F_1(t)\}$ for clustered competing risks data. We define the following random processes:

$$dM_{1ik}(t) = dN_{1ik}(t) - Y_{ik}(t)dH_1(t)\text{ , and } dM_{ik}(t) = dN_{ik}(t) - Y_{ik}(t)dH(t)\,.$$

Let $M_{1ik}(t) = \int_0^t dM_{1ik}(u)$ and $M_{ik}(t) = \int_0^t dM_{ik}(u)$. Since event times for individuals within the same cluster are correlated, both $M_{1ik}(t)$ and $M_{ik}(t)$ are no longer martingales with respect to the overall joint filtration $\mathcal{F}_t = \vee_{i=1}^{n} \vee_{k=1}^{n_i} \sigma\{Y_{ik}(s), N_{ik}(s) : 0 \le s \le t\}$. Therefore, the martingale central limit theory cannot be applied. Following the procedure of Ghosh and Lin (2000), we have the following equation

$$\sqrt{m}\{\hat{F}_1(t) - F_1(t)\} = \sqrt{m}\left[\int_0^t S(u)d\{\hat{H}_1(u) - H_1(u)\} - \int_0^t \{\hat{H}(u) - H(u)\}dF_1(u)\right] + o_p(1)\,.$$

Using results from empirical processes (Pollard, 1984; Spiekerman and Lin, 1998), both $\left\{\hat{H}_1(u) - H_1(u)\right\}$ and $\left\{\hat{H}(u) - H(u)\right\}$ can be decomposed into a sum of mean zero random variables. Therefore, we have

$$\sqrt{m}\{\hat{F}_1(t) - F_1(t)\} = \sum_{i=1}^{n}\sum_{k=1}^{n_i} Z_{ik}(t) + o_p(1)\,, \qquad \text{(A.1)}$$

where

$$Z_{ik}(t) = m^{-1/2}\int_0^t \frac{S(u)}{y(u)} dM_{1ik}(u) + m^{-1/2}\int_0^t \frac{F_1(u) - F_1(t)}{y(u)} dM_{ik}(u)\,,$$

and $y(t) = P(t \le X_{ik})$. For any fixed time t, by the law of large numbers for a weakly dependent series (Feller, 1957), $\hat{F}_1(t) - F_1(t)$ converges to zero in probability. Therefore, $\hat{F}_1(t)$ is a consistent estimator for the cumulative incidence function $F_1(t)$.

Since the decomposed terms $Z_{ik}(t)$ only depend on observation *k* in cluster *i*, they are correlated within the same cluster, but independent between different clusters. By using the methods for the clustered linear statistic (Williams, 2000), the between-cluster variance estimator of the linear statistic $Z(t)=\sum_{i=1}^{n}\sum_{k=1}^{n_i} Z_{ik}(t)$ is given by $\hat{V}(t)=\frac{n}{n-1}\sum_{i=1}^{n}\left\{\widehat{Z}_{i\cdot}(t)-\overline{Z}_{\cdot\cdot}(t)\right\}^2$,

where $\widehat{Z}_{i\cdot}(t)=\sum_{k=1}^{n_i}\widehat{Z}_{ik}(t)$, and $\overline{Z}_{\cdot\cdot}(t)=\frac{1}{n}\sum_{i=1}^{n}\widehat{Z}_{i\cdot}(t)$. Since,

$$
\begin{aligned}
E\{\hat{V}(t)\} &= \frac{n}{n-1}\left(\sum_{i=1}^{n} E\left\{\sum_{k}^{n_i}\widehat{Z}_{ik}(t)\right\}^2 - nE\left\{\frac{1}{n}\sum_{i=1}^{n}\sum_{k}^{n_i}\widehat{Z}_{ik}(t)\right\}^2\right) \\
&= \frac{n}{n-1}\left(\sum_{i=1}^{n}\sum_{k_1}^{n_i}\sum_{k_2}^{n_i} E\widehat{Z}_{ik_1}(t)\widehat{Z}_{ik_2}(t) - \frac{1}{n}\sum_{i=1}^{n}\sum_{k_1}^{n_i}\sum_{k_2}^{n_i} E\widehat{Z}_{ik_1}(t)\widehat{Z}_{ik_2}(t)\right) \\
&= \sum_{i=1}^{n}\sum_{k_1}^{n_i}\sum_{k_2}^{n_i} E\widehat{Z}_{ik_1}(t)\widehat{Z}_{ik_2}(t) \\
&= \sum_{i=1}^{n}\operatorname{var}\left\{\sum_{k}^{n_i} Z_{ik}(t)\right\},
\end{aligned}
$$

we conclude that $\hat{V}(t)$ is a consistent estimator for the variance of $Z(t)$.

For the variance of the Gray test under the null, assuming that process $\widehat{W}(t)$ converges to $W(t)$ in probability, and $\frac{m_0}{m}\to\rho_0$ as $m\to\infty$, we show that $\hat{Q}_G$ can be rewritten as sum of some mean zero random variables:

$$
\begin{aligned}
\hat{Q}_G &= \sqrt{\frac{m_1 m_0}{m}}\int \widehat{W}(t)\,d\left\{\widehat{F}_1^{(1)}(t)-\widehat{F}_1^{(0)}(t)\right\} \\
&= \sqrt{\frac{m_1 m_0}{m}}\left[\int W(t)\,d\left\{\widehat{F}_1^{(1)}(t)-F_1^{(1)}(t)\right\}-\int W(t)\,d\left\{\widehat{F}_1^{(0)}(t)-F_1^{(0)}(t)\right\}\right]+o_p(1) \quad \text{(A.2)} \\
&= \frac{1}{\sqrt{m}}\sum_{i=1}^{n}\sum_{k=1}^{n_i} g_{ik}\int W(t)\,dZ_{ik}(t)+o_p(1).
\end{aligned}
$$

By applying the robust variance estimator method of Williams (2000) to equation (A.2), one can consistently estimate the variance for test statistics $\hat{Q}_G$ by

$$\widehat{V}_G = \frac{1}{m} \times \frac{n}{n-1} \sum_{i=1}^{n} \left\{ \sum_{k=1}^{n_i} g_{ik} \int \widehat{W}(t) d\widehat{Z}_{ik}(t) - \int \widehat{W}(t) d\widehat{Z}^{*}_{\cdot\cdot}(t) \right\}^2 ,$$

When individuals in groups 0 and 1 are from different clusters, above formula reduce to

$$\widehat{V}_G = \frac{1}{m} \sum_{g=0}^{1} \frac{m_{1-g} n^{(g)}}{m_g (n^{(g)} - 1)} \sum_{i=1}^{n^{(g)}} \left[\int \widehat{W}(t) \left\{ d\widehat{Z}^{(g)}_{i\cdot}(t) - d\widehat{Z}^{(g)}_{\cdot\cdot}(t) \right\} \right]^2 .$$

Similarly, the robust variance for test statistics $\hat{Q}_{PM}$ can be consistently estimated by

$$\widehat{V}_{PM} = \frac{1}{m} \times \frac{n}{n-1} \sum_{i=1}^{n} \left\{ \sum_{k=1}^{n_i} g_{ik} \int \widehat{W}(t) \widehat{Z}_{ik}(t) dt - \int \widehat{W}(t) \widehat{Z}^{*}_{\cdot\cdot}(t) dt \right\}^2 ,$$

and

$$\widehat{V}_{PM} = \frac{1}{m} \sum_{g=0}^{1} \frac{m_{1-g} n^{(g)}}{m_g (n^{(g)} - 1)} \sum_{i=1}^{n^{(g)}} \left[\int \widehat{K}(t) \left\{ \widehat{Z}^{(g)}_{i\cdot}(t) - \widehat{Z}^{(g)}_{\cdot\cdot}(t) \right\} dt \right]^2 ,$$

when individuals in groups 0 and 1 are from different clusters.