Supplementary Materials for

Early Termination of a Two-Stage Study to Develop and Validate a Panel of Biomarkers

Joseph S. Koopmeiners[1,†] and Rachel Isaksson Vogel[2]

[1]Division of Biostatistics, University of Minnesota, Minneapolis, Minnesota

[2]Biostatistics and Bioinformatics Core, Masonic Cancer Center, University of Minnesota, Minneapolis, Minnesota

[†]*Corresponding author's address: koopm007@umn.edu*

1 Setting the Simulation Parameters

Let X be a 15 dimensional vector of biomarkers, which is composed of n_s signal markers, X_s, (i.e. biomarkers that are truly associated with PC recurrence) and $15 - n_s$ noisy markers, X_n (i.e. biomarkers that are not associated with PC recurrence). That is, X is composed of two sub-vectors,

$$X = (X_s, X_n),$$

where X_s is a n_s dimensional sub-vector of signal markers and X_n is a $(15 - n_s)$ dimensional sub-vector of noisy markers. The signal markers and noisy markers are randomly sampled from multivariate normal distributions, $X_s \sim N(0, \Sigma_s)$ and $X_n \sim N(0, I)$, respectively. Biochemical failure times, Y, are randomly sampled from an exponential distribution with rate parameter, λ, where λ is a function of X_s,

$$\log \lambda = \beta_0 + X_s^t \beta_{1,s},$$

with $\beta_{1,s} = \beta_s W_s$, where β_s is a scaler and W_s is a vector of weights indicating the relative importance of each marker.

The following parameters must be specified: n_s, W_s, Σ_s, β_0 and $\beta_{1,s}$. We varied n_s, W_s and Σ_s to determine their effect on the operating characteristics of our study and we set β_0 and β_1 by solving a system of equations to achieve the desired prevalence and true value of $ROC(0.1)$ (either $ROC(0.1)_0$ or $ROC(0.1)_1$). We note that the following equations only consider the signal markers as the recurrence times are not associated with the noisy markers. Our desired prevalence, a 5-year PC recurrence rate of 0.15, defines the following equation,

$$0.15 = P\left(Y \leq 5\right) = \int_{-\infty}^{5} \int_{x_s \in \Re^{n_s}} f\left(y|x_s, \beta_0, \beta_{1,s}\right) f\left(x_s|\Sigma_s\right) dx_s dy, \tag{1}$$

and the desired true value for $ROC(0.1)$ defines the following equations for the specificity (equal to 0.90),

$$0.90 = P\left(X_s^t \beta_{1,s} \leq c|Y > 5\right) = \frac{\int_5^{\infty} \int_{x_s^t \beta_{1,s} \leq c} f\left(y|x_s, \beta_0, \beta_{1,s}\right) f\left(x_s|\Sigma_s\right) dx_s dy}{\int_5^{\infty} \int_{x_s \in \Re^{n_s}} f\left(y|x_s, \beta_0, \beta_{1,s}\right) f\left(x_s|\Sigma_s\right) dx_s dy}, \tag{2}$$

and the sensitivity (equal to $ROC(0.1)$),

$$ROC(0.1) = P\left(X_s^t \beta_{1,s} > c|Y \leq 5\right) = \frac{\int_{-\infty}^{5} \int_{x_s^t \beta_{1,s} > c} f\left(y|x_s, \beta_0, \beta_{1,s}\right) f\left(x_s|\Sigma_s\right) dx_s dy}{\int_{-\infty}^{5} \int_{x_s \in \Re^{n_s}} f\left(y|x_s, \beta_0, \beta_{1,s}\right) f\left(x_s|\Sigma_s\right) dx_s dy}. \tag{3}$$

We note that Equations 2 and 3 introduce a third parameter, c, leaving us with three equations and three unknowns. These equations can be simplified by defining,

$$Z = X_s^t \beta_{1,s},$$

where $Z \sim N\left(0, \beta_{1,s}^t \Sigma_s \beta_{1,s}\right)$, and noting that,

$$\int_0^5 f\left(y|x_s, \beta_0, \beta_{1,s}\right) dy = 1 - e^{-5e^{\beta_0 + z}}.$$

This allows us to redefine Equations 1, 2 and 3 as

$$0.15 = P\left(Y \le 5\right) = \int_{-\infty}^{\infty} \left(1 - e^{-5e^{\beta_0 + z}}\right) f\left(z|\beta_{1,s}, \Sigma_s\right) dz, \tag{4}$$

$$0.90 = P\left(X_s^t \beta_{1,s} \le c|Y > 5\right) = \frac{\int_{-\infty}^{c} e^{-5e^{\beta_0 + z}} f\left(z|\beta_{1,s}, \Sigma_s\right) dz}{\int_{-\infty}^{\infty} e^{-5e^{\beta_0 + z}} f\left(z|\beta_{1,s}, \Sigma_s\right) dz}, \tag{5}$$

and,

$$ROC(0.1) = P\left(X_s^t \beta_{1,s} > c|Y \le 5\right) = \frac{\int_{c}^{\infty} \left(1 - e^{-5e^{\beta_0 + z}}\right) f\left(z|\beta_{1,s}, \Sigma_s\right) dz}{\int_{-\infty}^{\infty} \left(1 - e^{-5e^{\beta_0 + z}}\right) f\left(z|\beta_{1,s}, \Sigma_s\right) dz}. \tag{6}$$

Solving Equations 4, 5 and 6 provides the correct values of β_0 and β_1 for simulating with the desired prevalence and true value of $ROC(0.1)$.

2 Additional Simulation Results

Tables 1, 2 and 3 presents simulation results evaluating the impact of marker parameters n_s, W_s, V_s and ρ on the operating characteristics of our study when two, three and four stopping times are included in stage 2, respectively. The results are similar to what was observed for the design with only one, final analysis at the conclusion of stage. We see that increasing the number of signal markers, n_s, results in in a slight increase in the expected sample size under the null and a decrease in the power, while increasing the correlation between signal markers, ρ, results in a decreased expected sample size under the null and an increase in power. Overall, though, we see that the operating characteristics of our study are reasonably robust to varying the underlying marker distribution and that differences due to varying the marker parameters are small compared to differences due to varying the design parameters.

Table 1: Simulation results evaluating the impact of marker parameters n_s, W_s, V_s and ρ on the operating characteristics of our study when two stopping times are included in stage 2. Simulations were completed using $P = 0.3$ and $ROC(0.1)_{co} = 0.45$. 10000 simulations were completed for each scenario.

	$n_s = 3$				$n_s = 5$				$n_s = 7$			
	Null hypothesis[1]		Alt hypothesis[2]		Null hypothesis[1]		Alt hypothesis[2]		Null hypothesis[1]		Alt hypothesis[2]	
	E(SS)	Unc α	E(SS)	Unc $1-\beta$	E(SS)	Unc α	E(SS)	Unc $1-\beta$	E(SS)	Unc α	E(SS)	Unc $1-\beta$
$W_s = (1,\ldots,1)$, $V_s = (1,\ldots,1)$												
$\rho = 0.0$	323	0.003	604	0.909	325	0.002	609	0.887	325	0.002	614	0.867
$\rho = 0.2$	317	0.004	602	0.902	322	0.002	609	0.898	325	0.002	614	0.879
$\rho = 0.4$	320	0.003	599	0.912	320	0.002	606	0.893	319	0.002	610	0.887
$\rho = 0.6$	315	0.003	599	0.903	319	0.002	606	0.899	319	0.003	607	0.892
$W_s = (1,\ldots,1)$, $V_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$												
$\rho = 0.0$	323	0.003	604	0.903	326	0.003	613	0.885	325	0.001	617	0.869
$\rho = 0.2$	319	0.004	603	0.903	320	0.002	608	0.890	323	0.002	614	0.876
$\rho = 0.4$	320	0.003	601	0.904	320	0.002	606	0.897	321	0.002	612	0.875
$\rho = 0.6$	319	0.004	600	0.905	320	0.004	608	0.894	321	0.002	613	0.883
$W_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$, $V_s = (1,\ldots,1)$												
$\rho = 0.0$	322	0.003	604	0.895	323	0.002	611	0.884	321	0.002	613	0.874
$\rho = 0.2$	319	0.003	604	0.904	320	0.003	608	0.898	325	0.003	612	0.887
$\rho = 0.4$	320	0.003	601	0.911	320	0.002	604	0.895	319	0.002	609	0.892
$\rho = 0.6$	316	0.004	599	0.905	318	0.002	606	0.906	321	0.003	608	0.895
$W_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$, $V_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$												
$\rho = 0.0$	318	0.003	604	0.901	321	0.002	612	0.887	322	0.001	613	0.873
$\rho = 0.2$	319	0.003	600	0.899	320	0.003	608	0.888	322	0.002	612	0.882
$\rho = 0.4$	320	0.002	600	0.906	319	0.003	609	0.899	319	0.002	612	0.881
$\rho = 0.6$	316	0.003	600	0.906	319	0.003	604	0.902	320	0.002	607	0.884
$W_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$, $V_s = \left(\frac{1}{n_s}, \frac{2}{n_s}, \ldots, \frac{n_s}{n_s}\right)$												
$\rho = 0.0$	315	0.003	605	0.905	321	0.002	611	0.889	324	0.002	616	0.872
$\rho = 0.2$	317	0.002	600	0.897	321	0.003	608	0.888	323	0.002	613	0.876
$\rho = 0.4$	319	0.004	602	0.907	325	0.002	608	0.896	322	0.002	610	0.879
$\rho = 0.6$	316	0.003	603	0.906	319	0.002	607	0.896	324	0.003	609	0.882

[1]: $ROC\,(0.1)_0 = 0.40$
[2]: $ROC\,(0.1)_a = 0.65$

Table 2: Simulation results evaluating the impact of marker parameters n_s, W_s, V_s and ρ on the operating characteristics of our study when three stopping times are included in stage 2. Simulations were completed using $P = 0.3$ and $ROC(0.1)_{co} = 0.45$. 10000 simulations were completed for each scenario.

	$n_s = 3$				$n_s = 5$				$n_s = 7$			
	Null hypothesis[1]		Alt hypothesis[2]		Null hypothesis[1]		Alt hypothesis[2]		Null hypothesis[1]		Alt hypothesis[2]	
	E(SS)	Unc α	E(SS)	Unc $1-\beta$	E(SS)	Unc α	E(SS)	Unc $1-\beta$	E(SS)	Unc α	E(SS)	Unc $1-\beta$
$\rho = 0.0$	308	0.002	564	0.899	308	0.002	572	0.885	310	0.002	577	0.862
$\rho = 0.2$	308	0.004	560	0.898	308	0.003	568	0.886	307	0.002	573	0.871
$\rho = 0.4$	305	0.004	560	0.900	306	0.003	565	0.886	306	0.003	574	0.882
$\rho = 0.6$	303	0.004	561	0.906	306	0.003	563	0.894	307	0.003	570	0.892
$W_s = (1, \ldots, 1)$, $V_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$												
$\rho = 0.0$	309	0.003	561	0.907	307	0.002	570	0.878	310	0.002	577	0.869
$\rho = 0.2$	304	0.003	560	0.899	307	0.002	568	0.886	308	0.001	574	0.871
$\rho = 0.4$	306	0.003	561	0.903	307	0.003	569	0.892	304	0.002	574	0.874
$\rho = 0.6$	301	0.004	562	0.903	305	0.003	567	0.894	307	0.003	573	0.872
$W_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$, $V_s = (1, \ldots, 1)$												
$\rho = 0.0$	302	0.002	560	0.903	305	0.002	572	0.883	308	0.002	577	0.867
$\rho = 0.2$	303	0.002	561	0.895	304	0.003	567	0.890	306	0.002	570	0.882
$\rho = 0.4$	305	0.004	561	0.900	307	0.003	565	0.898	305	0.003	569	0.888
$\rho = 0.6$	304	0.003	557	0.906	304	0.002	560	0.900	306	0.002	566	0.895
$W_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$, $V_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$												
$\rho = 0.0$	304	0.003	563	0.898	305	0.002	572	0.881	306	0.001	575	0.866
$\rho = 0.2$	303	0.003	564	0.899	307	0.003	568	0.884	308	0.002	573	0.876
$\rho = 0.4$	300	0.004	561	0.898	305	0.003	568	0.892	306	0.003	569	0.888
$\rho = 0.6$	302	0.003	560	0.900	305	0.004	564	0.899	302	0.003	573	0.885
$W_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$, $V_s = \left(\frac{1}{n_s}, \frac{2}{n_s}, \ldots, \frac{n_s}{n_s}\right)$												
$\rho = 0.0$	305	0.003	563	0.898	308	0.003	570	0.886	304	0.001	575	0.871
$\rho = 0.2$	305	0.003	561	0.895	308	0.004	567	0.882	307	0.002	572	0.874
$\rho = 0.4$	304	0.003	561	0.899	308	0.003	567	0.887	308	0.002	571	0.880
$\rho = 0.6$	304	0.004	561	0.902	304	0.003	566	0.889	307	0.002	569	0.886

[1]: $ROC\,(0.1)_0 = 0.40$

[2]: $ROC\,(0.1)_a = 0.65$

Table 3: Simulation results evaluating the impact of marker parameters n_s, W_s, V_s and ρ on the operating characteristics of our study when four stopping times are included in stage 2. Simulations were completed using $P = 0.3$ and $ROC(0.1)_{co} = 0.45$. 10000 simulations were completed for each scenario.

	$n_s = 3$				$n_s = 5$				$n_s = 7$			
	Null hypothesis[1]		Alt hypothesis[2]		Null hypothesis[1]		Alt hypothesis[2]		Null hypothesis[1]		Alt hypothesis[2]	
	E(SS)	Unc α	E(SS)	Unc $1-\beta$	E(SS)	Unc α	E(SS)	Unc $1-\beta$	E(SS)	Unc α	E(SS)	Unc $1-\beta$
$W_s = (1,\ldots,1)$, $V_s = (1,\ldots,1)$												
$\rho = 0.0$	299	0.004	541	0.898	302	0.002	551	0.878	302	0.002	558	0.862
$\rho = 0.2$	298	0.002	539	0.897	298	0.002	547	0.890	300	0.002	553	0.870
$\rho = 0.4$	297	0.002	540	0.895	298	0.003	545	0.889	301	0.002	553	0.878
$\rho = 0.6$	296	0.003	538	0.906	298	0.003	545	0.893	300	0.004	545	0.890
$W_s = (1,\ldots,1)$, $V_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$												
$\rho = 0.0$	300	0.003	546	0.893	300	0.003	550	0.880	300	0.002	556	0.859
$\rho = 0.2$	299	0.004	540	0.898	300	0.003	548	0.880	297	0.002	556	0.868
$\rho = 0.4$	297	0.005	540	0.897	300	0.003	545	0.886	300	0.002	551	0.868
$\rho = 0.6$	297	0.004	539	0.904	299	0.003	547	0.883	298	0.004	552	0.873
$W_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$, $V_s = (1,\ldots,1)$												
$\rho = 0.0$	298	0.003	541	0.894	299	0.003	550	0.880	299	0.002	556	0.862
$\rho = 0.2$	298	0.004	541	0.899	299	0.003	541	0.884	299	0.002	550	0.871
$\rho = 0.4$	294	0.003	539	0.904	300	0.004	545	0.892	301	0.003	548	0.885
$\rho = 0.6$	297	0.004	538	0.901	296	0.003	541	0.898	300	0.004	546	0.892
$W_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$, $V_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$												
$\rho = 0.0$	299	0.004	543	0.893	298	0.003	550	0.876	300	0.003	555	0.862
$\rho = 0.2$	297	0.002	540	0.894	296	0.002	549	0.887	297	0.003	553	0.876
$\rho = 0.4$	297	0.003	540	0.897	295	0.003	545	0.887	297	0.003	551	0.876
$\rho = 0.6$	298	0.002	539	0.896	297	0.002	544	0.890	296	0.003	548	0.884
$W_s = \left(\frac{n_s}{n_s}, \frac{n_s-1}{n_s}, \ldots, \frac{1}{n_s}\right)$, $V_s = \left(\frac{1}{n_s}, \frac{2}{n_s}, \ldots, \frac{n_s}{n_s}\right)$												
$\rho = 0.0$	294	0.002	545	0.891	298	0.003	550	0.877	299	0.002	555	0.865
$\rho = 0.2$	296	0.004	542	0.895	300	0.001	547	0.876	299	0.002	552	0.870
$\rho = 0.4$	297	0.005	542	0.894	299	0.003	547	0.888	299	0.002	551	0.878
$\rho = 0.6$	296	0.004	540	0.900	299	0.003	546	0.887	299	0.002	548	0.882

[1]: $ROC\,(0.1)_0 = 0.40$

[2]: $ROC\,(0.1)_a = 0.65$