

Web Appendix for: Combining Multiple Biomarkers Linearly to Maximize the Partial Area under the ROC Curve

Qingxiang Yan^{1*}, Leonidas E. Bantis¹, Janet L. Stanford^{2,3}, and Ziding Feng¹

Derivation of proposed pAUC objective function under the normality assumption:

The normality assumption implies that $Y_D \sim \text{MVN}(\mu_D, \Sigma_D)$ and $Y_{\bar{D}} \sim \text{MVN}(\mu_{\bar{D}}, \Sigma_{\bar{D}})$. Then the combined scores follow univariate normal distributions:

$$W_D = \beta^T Y_D \sim N(\beta^T \mu_D, \beta^T \Sigma_D \beta), \quad W_{\bar{D}} = \beta^T Y_{\bar{D}} \sim N(\beta^T \mu_{\bar{D}}, \beta^T \Sigma_{\bar{D}} \beta).$$

The pAUC of this composite marker can be written as:

$$\text{pAUC}(\beta, t_0) = \Pr \left(W_D > W_{\bar{D}}, W_{\bar{D}} > S_{W_{\bar{D}}}^{-1}(t_0) \right),$$

where $S_{W_{\bar{D}}}^{-1}(\cdot)$ is the inverse survival function of $W_{\bar{D}}$. If we let $S_{W_{\bar{D}}}^{-1}(t_0) = \xi$, then $\Pr(W_{\bar{D}} > \xi) = t_0$; and we can rewrite the above equation as:

$$\begin{aligned} \text{pAUC}(\beta, t_0) &= \Pr \left(W_D > W_{\bar{D}}, W_{\bar{D}} > S_{W_{\bar{D}}}^{-1}(t_0) \right) \\ &= \Pr \left(W_D - W_{\bar{D}} > 0, W_{\bar{D}} > \xi \right) \\ &= \Pr \left(\frac{W_D - W_{\bar{D}} - (\beta^T \mu_D - \beta^T \mu_{\bar{D}})}{\sqrt{\beta^T (\Sigma_D + \Sigma_{\bar{D}}) \beta}} > \frac{-(\beta^T \mu_D - \beta^T \mu_{\bar{D}})}{\sqrt{\beta^T (\Sigma_D + \Sigma_{\bar{D}}) \beta}}, \frac{W_{\bar{D}} - \beta^T \mu_{\bar{D}}}{\sqrt{\beta^T \Sigma_{\bar{D}} \beta}} > \frac{\xi - \beta^T \mu_{\bar{D}}}{\sqrt{\beta^T \Sigma_{\bar{D}} \beta}} \right). \end{aligned}$$

¹Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, U.S.A.

²Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98195, U.S.A.

³Department of Epidemiology, University of Washington School of Public Health, Seattle, WA 98109, U.S.A.

*Correspondence to: Qingxiang Yan, 1400 Pressler St. Pickens Tower, Dept. of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, U.S.A. E-mail: qyan@mdanderson.org

Now let

$$Z = \frac{W_D - W_{\bar{D}} - (\beta^T \mu_D - \beta^T \mu_{\bar{D}})}{\sqrt{\beta^T (\Sigma_D + \Sigma_{\bar{D}}) \beta}}$$

and

$$X = \frac{W_{\bar{D}} - \beta^T \mu_{\bar{D}}}{\sqrt{\beta^T \Sigma_{\bar{D}} \beta}},$$

then Z and X both follow $N(0, 1)$. The pAUC function becomes:

$$\begin{aligned} \text{pAUC}(\beta, t_0) &= \Pr \left(Z > \frac{-(\beta^T \mu_D - \beta^T \mu_{\bar{D}})}{\sqrt{\beta^T (\Sigma_D + \Sigma_{\bar{D}}) \beta}}, X > \frac{\xi - \beta^T \mu_{\bar{D}}}{\sqrt{\beta^T \Sigma_{\bar{D}} \beta}} \right) \\ &= \Pr \left(Z < \frac{(\beta^T \mu_D - \beta^T \mu_{\bar{D}})}{\sqrt{\beta^T (\Sigma_D + \Sigma_{\bar{D}}) \beta}}, X < \frac{-\xi + \beta^T \mu_{\bar{D}}}{\sqrt{\beta^T \Sigma_{\bar{D}} \beta}} \right) \end{aligned}$$

Based on our previous definition $\Pr(W_{\bar{D}} > \xi) = t_0$, we have

$$\begin{aligned} t_0 &= \Pr(W_{\bar{D}} > \xi) = 1 - \Pr(W_{\bar{D}} < \xi) \\ &= 1 - \Pr \left(X < \frac{\xi - \beta^T \mu_{\bar{D}}}{\sqrt{\beta^T \Sigma_{\bar{D}} \beta}} \right) = 1 - \Phi \left(\frac{\xi - \beta^T \mu_{\bar{D}}}{\sqrt{\beta^T \Sigma_{\bar{D}} \beta}} \right) = \Phi \left(\frac{-\xi + \beta^T \mu_{\bar{D}}}{\sqrt{\beta^T \Sigma_{\bar{D}} \beta}} \right). \end{aligned}$$

Therefore,

$$\Phi^{-1}(t_0) = \frac{-\xi + \beta^T \mu_{\bar{D}}}{\sqrt{\beta^T \Sigma_{\bar{D}} \beta}}.$$

The correlation between the two standard normal variables Z and X is:

$$\begin{aligned} \text{corr}(Z, X) &= \text{cov}(Z, X) \\ &= \text{cov} \left(\frac{W_D - W_{\bar{D}} - (\beta^T \mu_D - \beta^T \mu_{\bar{D}})}{\sqrt{\beta^T (\Sigma_D + \Sigma_{\bar{D}}) \beta}}, \frac{W_{\bar{D}} - \beta^T \mu_{\bar{D}}}{\sqrt{\beta^T \Sigma_{\bar{D}} \beta}} \right) \\ &= \text{cov}(W_D - W_{\bar{D}}, W_{\bar{D}}) \cdot \frac{1}{\sqrt{\beta^T (\Sigma_D + \Sigma_{\bar{D}}) \beta} \sqrt{\beta^T \Sigma_{\bar{D}} \beta}} \\ &= -\text{cov}(W_{\bar{D}}, W_{\bar{D}}) \cdot \frac{1}{\sqrt{\beta^T (\Sigma_D + \Sigma_{\bar{D}}) \beta} \sqrt{\beta^T \Sigma_{\bar{D}} \beta}} \quad (W_D \text{ and } W_{\bar{D}} \text{ are independent}) \\ &= \frac{-\beta^T \Sigma_{\bar{D}} \beta}{\sqrt{\beta^T (\Sigma_D + \Sigma_{\bar{D}}) \beta} \sqrt{\beta^T \Sigma_{\bar{D}} \beta}} \\ &= \frac{-\sqrt{\beta^T \Sigma_{\bar{D}} \beta}}{\sqrt{\beta^T (\Sigma_D + \Sigma_{\bar{D}}) \beta}}. \end{aligned}$$

To summarize, the pAUC function can be written as the joint distribution function of two correlated standard normal random variables:

$$\text{pAUC}(\boldsymbol{\beta}, t_0) = F_{BVN} \left(\frac{\boldsymbol{\beta}^T (\boldsymbol{\mu}_D - \boldsymbol{\mu}_{\bar{D}})}{\sqrt{\boldsymbol{\beta}^T (\boldsymbol{\Sigma}_{\bar{D}} + \boldsymbol{\Sigma}_D) \boldsymbol{\beta}}}, \Phi^{-1}(t_0); -\frac{\sqrt{\boldsymbol{\beta}^T \boldsymbol{\Sigma}_{\bar{D}} \boldsymbol{\beta}}}{\sqrt{\boldsymbol{\beta}^T (\boldsymbol{\Sigma}_{\bar{D}} + \boldsymbol{\Sigma}_D) \boldsymbol{\beta}}} \right)$$

Web Table 1. Partial-AUC comparison between the proposed parametric method and the truth. The marginal distribution of biomarker values of the non-disease group is always $N(0,1)$; for the disease group, the marginal means are randomly selected from interval $[0,1]$ and the marginal standard deviations are all 2s. Correlation between any two biomarkers is 0.3. Under this construction, the covariance matrices of the biomarker distribution of the disease and non-disease groups are proportional; hence the truth can be obtained by applying Su and Liu’s method. This experiment does not require data generation. Both methods are directly applied to the true population parameters. The FPR range of interest is $[0, 0.1]$. The number of markers (m) and the absolute difference in pAUC between the PMuN method and the Su & Liu’s method are listed below.

m	Absolute Difference
2	9.71E-17
3	<1.00E-17
4	<1.00E-17
5	<1.00E-17
6	9.71E-17
7	<1.00E-17
8	1.04E-16
9	<1.00E-17
10	1.30E-15
11	2.98E-16
12	1.04E-16
13	2.01E-16
14	9.71E-17
15	2.01E-16
16	4.02E-16
17	2.01E-16
18	1.20E-15
19	2.30E-15
20	5.97E-16

Web Table 2. Parameters and settings used for simulation study In Section 4.1: known distributions

Parameters	Settings
Number of biomarkers	3
Number of Monte Carlo samples	1000
Distribution	normal, lognormal, gamma
μ_D	$(1, 1, 1)^T$
$\mu_{\bar{D}}$	$(1.3, 1.5, 1.8)^T$
σ_D	$(0.5, 0.5, 0.5)$
$\sigma_{\bar{D}}$	Proportional to σ_D : $(0.7, 0.7, 0.7)$ Disproportional to σ_D : $(0.5, 0.8, 1)$
Correlation	Exchangeable, unstructured
FPR range	$[0, 0.3]$
$(n_D, n_{\bar{D}})$	$(25, 25), (100, 100)$

Web Table 3. Simulation results for Section 4.1: Multivariate distributions. Number of markers is three; FPR range is $[0, 0.3]$; and sample size is $(25, 25)$. The correlatoin structure is exchangeable which means all biomarkers from both disease and non-disease groups share the same correlation.

$(n_D, n_{\bar{D}})$	Σ	ρ	Method	Normal Dist.			Gamma Dist.		Lognormal Dist.	
				Mean pAUC	Coef. MSE	pAUC Rank	Mean pAUC	pAUC Rank	Mean pAUC	pAUC Rank
(25,25)	Proportional	0.3	Truth	0.193	-	-	-	-	-	-
			PMuN	0.185	0.281	3.516	0.170	3.685	0.169	3.924
			Liu2005	0.165	1.031	6.551	0.157	5.724	0.157	5.604
			Su1993	0.185	0.275	3.599	0.171	3.574	0.169	3.879
			KS	0.182	0.372	4.342	0.167	4.528	0.167	4.355
			Stepdown	0.180	0.407	4.689	0.166	4.873	0.165	4.801
			SW-Pepe	0.180	0.401	4.729	0.165	4.982	0.165	4.907
			Logistic	0.185	0.274	3.656	0.171	3.448	0.170	3.438
			GS	0.179	0.516	4.918	0.165	5.186	0.164	5.091
(25,25)	Disproportional	0.3	Truth	0.182	-	-	-	-	-	-
			PMuN	0.174	0.384	3.452	0.162	3.519	0.161	3.631
			Liu2005	0.172	0.461	3.954	0.160	3.878	0.158	4.328
			Su1993	0.169	0.625	5.048	0.158	5.071	0.155	5.414
			KS	0.171	0.513	4.177	0.160	4.292	0.160	4.130
			Stepdown	0.169	0.530	4.695	0.157	4.870	0.157	4.811
			SW-Pepe	0.169	0.541	4.846	0.156	5.043	0.157	4.859
			Logistic	0.169	0.595	4.787	0.160	4.169	0.160	3.974
			GS	0.168	0.675	5.041	0.156	5.158	0.157	4.853
(25,25)	Proportional	0.7	Truth	0.198	-	-	-	-	-	-
			PMuN	0.190	0.614	3.417	0.170	3.175	0.166	3.563
			Liu2005	0.149	2.089	7.643	0.145	6.927	0.146	6.555
			Su1993	0.191	0.587	3.329	0.170	3.574	0.165	3.982
			KS	0.187	0.854	4.247	0.166	4.307	0.163	4.341
			Stepdown	0.182	0.540	5.162	0.161	5.354	0.159	5.146
			SW-Pepe	0.186	0.392	4.061	0.164	4.449	0.162	4.391
			Logistic	0.190	0.547	3.497	0.170	3.287	0.167	3.250
			GS	0.185	1.064	4.644	0.163	4.926	0.161	4.771
(25,25)	Disproportional	0.7	Truth	0.181	-	-	-	-	-	-
			PMuN	0.174	0.754	2.880	0.162	2.982	0.159	3.181
			Liu2005	0.171	0.681	4.425	0.159	4.112	0.157	3.821
			Su1993	0.165	1.376	5.514	0.153	5.579	0.146	6.101
			KS	0.172	0.920	3.775	0.159	3.879	0.157	3.815
			Stepdown	0.167	0.733	5.054	0.153	5.508	0.151	5.295
			SW-Pepe	0.169	0.703	4.496	0.156	4.705	0.153	4.624
			Logistic	0.166	1.297	5.037	0.157	4.313	0.155	4.347
			GS	0.168	1.385	4.819	0.155	4.921	0.153	4.816

Web Table 4. Simulation results for Section 4.1: Multivariate distributions. Number of markers is three; FPR range is $[0, 0.3]$; and sample size is $(100, 100)$. The correlatoin structure is exchangeable which means all biomarkers from both disease and non-disease groups share the same correlation.

$(n_D, n_{\bar{D}})$	Σ	ρ	Method	Normal Dist.			Gamma Dist.		Lognormal Dist.	
				Mean pAUC	Coef. MSE	pAUC Rank	Mean pAUC	pAUC Rank	Mean pAUC	pAUC Rank
(100,100)	Proportional	0.3	Truth	0.193	-	-	-	-	-	-
			PMuN	0.191	0.070	3.273	0.177	3.684	0.176	3.934
			Liu2005	0.173	0.953	7.864	0.169	6.847	0.169	6.362
			Su1993	0.191	0.068	3.156	0.178	3.451	0.176	4.026
			KS	0.190	0.110	4.221	0.177	4.292	0.176	4.358
			Stepdown	0.189	0.129	4.615	0.176	4.710	0.175	4.619
			SW-Pepe	0.189	0.132	4.662	0.176	4.762	0.175	4.665
			Logistic	0.191	0.070	3.327	0.178	3.207	0.177	3.210
			GS	0.189	0.137	4.881	0.176	5.048	0.175	4.827
(100,100)	Disproportional	0.3	Truth	0.182	-	-	-	-	-	-
			PMuN	0.180	0.112	3.075	0.169	3.541	0.169	3.404
			Liu2005	0.179	0.177	4.105	0.169	3.874	0.167	4.770
			Su1993	0.177	0.357	5.886	0.166	6.001	0.161	6.775
			KS	0.179	0.150	3.868	0.168	4.011	0.169	3.771
			Stepdown	0.179	0.185	4.516	0.168	4.647	0.168	4.242
			SW-Pepe	0.179	0.189	4.612	0.167	4.723	0.168	4.330
			Logistic	0.177	0.325	5.240	0.168	4.236	0.168	4.047
			GS	0.179	0.197	4.697	0.167	4.966	0.168	4.660
(100,100)	Proportional	0.7	Truth	0.198	-	-	-	-	-	-
			PMuN	0.196	0.049	3.218	0.176	3.118	0.173	3.353
			Liu2005	0.154	2.142	8.000	0.152	7.872	0.153	7.687
			Su1993	0.196	0.050	3.260	0.176	3.306	0.172	4.150
			KS	0.195	0.100	3.944	0.175	4.105	0.172	3.922
			Stepdown	0.191	0.173	5.654	0.171	5.601	0.169	5.273
			SW-Pepe	0.195	0.058	3.954	0.175	4.257	0.172	4.118
			Logistic	0.196	0.052	3.428	0.176	2.930	0.173	2.952
			GS	0.194	0.209	4.543	0.174	4.811	0.171	4.545
(100,100)	Disproportional	0.7	Truth	0.181	-	-	-	-	-	-
			PMuN	0.180	0.115	2.497	0.167	2.525	0.165	3.049
			Liu2005	0.175	0.339	6.425	0.164	5.458	0.164	4.096
			Su1993	0.176	0.394	5.538	0.163	5.934	0.154	7.066
			KS	0.179	0.209	3.280	0.166	3.420	0.164	3.479
			Stepdown	0.177	0.247	5.181	0.163	5.606	0.162	5.233
			SW-Pepe	0.178	0.143	3.979	0.165	4.431	0.163	4.395
			Logistic	0.177	0.335	4.731	0.166	4.144	0.163	4.293
			GS	0.178	0.446	4.369	0.165	4.482	0.163	4.389

Web Table 5. Simulation results for Section 4.1: Multivariate distributions. Number of markers is three; FPR range is $[0, 0.3]$; and sample size is $(100, 100)$. The correlatoin structure is unstructured so different pairs of biomarkers may have different correlations.

$(n_D, n_{\bar{D}})$	Σ	ρ	Method	Normal Dist.			Gamma Dist.		Lognormal Dist.	
				Mean pAUC	Coef. MSE	pAUC Rank	Mean pAUC	pAUC Rank	Mean pAUC	pAUC Rank
(100, 100)	Proportional	$\rho_D : (0.2, 0.3, 0.4)$ $\rho_{\bar{D}} : (0.2, 0.3, 0.4)$	Truth	0.190	-	-	-	-	-	-
			PMuN	0.188	0.072	2.955	0.174	2.403	0.173	2.637
			Liu2005	0.172	0.943	6.815	0.167	4.838	0.167	4.447
			Su1993	0.188	0.070	2.936	0.174	2.092	0.173	2.716
			KS	0.187	0.115	3.901	0.173	2.971	0.173	2.842
			Stepdown	0.186	0.133	4.171	0.173	3.375	0.172	3.162
			SW-Pepe	0.186	0.132	4.195	0.173	3.426	0.172	3.239
			Logistic	0.188	0.071	3.027	0.174	1.896	0.174	1.957
(100, 100)	Disproportional	$\rho_D : (0.2, 0.3, 0.4)$ $\rho_{\bar{D}} : (0.8, 0.7, 0.6)$	Truth	0.178	-	-	-	-	-	-
			PMuN	0.177	0.089	3.321	0.164	2.897	0.164	3.088
			Liu2005	0.172	0.602	6.860	0.162	4.843	0.163	4.013
			Su1993	0.173	0.271	5.873	0.161	4.985	0.157	5.779
			KS	0.176	0.140	4.124	0.164	3.343	0.163	3.405
			Stepdown	0.175	0.181	4.785	0.163	3.812	0.163	3.759
			SW.Pepe	0.175	0.180	4.829	0.163	3.924	0.163	3.864
			Logistic	0.174	0.243	5.208	0.163	4.197	0.162	4.093

Web Table 6. Simulation results from Secton 4.2: Data are generated using logistic model. Number of markers is five. FPR rang is $[0, 0.3]$; and sample size is $(100, 100)$. "Mean" stands for "Mean pAUC" and "Rank" stands for "pAUC rank".

$(n_D, n_{\bar{D}})$	Method	Linear									Nonlinear	
		Normal Dist.			Gamma Dist.			Lognormal Dist.			Normal Dist.	
		Mean	Coef. MSE	Rank	Mean	Coef. MSE	Rank	Mean	Coef. MSE	Rank	Mean	Rank
(100, 100)	Risk Score	0.154	-	-	0.153	-	-	0.142	-	-	0.146	-
	pAUC	0.146	0.403	3.013	0.144	0.208	3.941	0.132	0.277	4.096	0.087	2.801
	Liu2005	0.079	2.736	6.999	0.125	0.702	6.867	0.119	0.664	6.657	0.083	4.402
	Su	0.147	0.297	2.370	0.147	0.403	2.135	0.136	0.886	2.492	0.082	5.342
	KS	0.144	0.234	3.810	0.144	0.234	3.586	0.134	0.282	3.386	0.087	2.998
	Stepdown	0.140	0.863	4.814	0.141	1.107	4.620	0.130	1.714	4.681	0.086	3.632
	Step.Pepe	0.141	0.670	4.674	0.141	1.077	4.646	0.130	1.565	4.567	0.086	3.593
	Logistic	0.147	0.315	2.320	0.147	0.334	2.205	0.136	0.600	2.120	0.082	5.232

Web Table 7. A summary of the marker distributions generated by the logistic model (Section 4.2). The following table was summarized from the corresponding large validation sets with $n = (100000, 100000)$. We can see that when the initial set was drawn from a normal distribution, the generated markers all have symmetric distributions (mean = median) with similar standard deviation across the cases and controls. When the initial set was drawn from positively skewed distributions such as gamma and lognormal, the resulting markers are also positively skewed (mean > median). The column 'P(KS)' contains the p-values calculated from the Kolmogorov-Smirnov test for testing the null hypothesis of normal distribution.

Sampled From:		Controls				Cases			
Dist.(mean,sd)	Marker	Median	Mean	Std.	P (KS)	Median	Mean	Std.	P (KS)
Normal(1,1)	M1	0.63	0.63	0.93	0.8340	1.36	1.36	0.93	0.1811
	M2	0.78	0.78	0.98	0.8145	1.22	1.22	0.97	0.5909
	M3	1.10	1.11	0.99	0.9117	0.89	0.89	0.99	0.5871
	M4	0.85	0.85	0.99	0.9445	1.15	1.15	0.99	0.9140
	M5	1.18	1.18	0.98	0.8252	0.82	0.82	0.98	0.6435
Gamma(1,1)	M1	0.49	0.68	0.64	< 0.0001	1.03	1.33	1.18	< 0.0001
	M2	0.56	0.79	0.76	< 0.0001	0.89	1.23	1.16	< 0.0001
	M3	0.77	1.11	1.09	< 0.0001	0.62	0.89	0.88	< 0.0001
	M4	0.59	0.85	0.84	< 0.0001	0.82	1.16	1.13	< 0.0001
	M5	0.83	1.16	1.12	< 0.0001	0.58	0.82	0.80	< 0.0001
Lognormal(1,1)	M1	0.58	0.72	0.54	< 0.0001	0.92	1.30	1.27	< 0.0001
	M2	0.62	0.81	0.65	< 0.0001	0.83	1.20	1.21	< 0.0001
	M3	0.76	1.09	1.12	< 0.0001	0.66	0.89	0.78	< 0.0001
	M4	0.65	0.86	0.75	< 0.0001	0.78	1.15	1.20	< 0.0001
	M5	0.80	1.15	1.19	< 0.0001	0.63	0.83	0.71	< 0.0001

Web Table 8. Compare RB-LOO with LOPO using logistic regression and AUC. We use the data generated in Section 4.1: sample size is (100, 100), disproportional covariance matrices, and correlation between any two markers is 0.3. Multivariate normal, lognormal and gamma distributions are examined. Under each setting, for each of the 1000 repetitions, we use only the training set and calculate the RB-LOO and LOPO cross-validated AUC for logistic regression. The mean(standard deviation) of the 1000 repetitions for each method under each setting is then summarized below.

Distribution	RB-LOO	LOPO
Normal	0.7790(0.050)	0.7821(0.050)
Lognormal	0.8037(0.046)	0.8067(0.046)
Gamma	0.7865(0.048)	0.7896(0.048)

Web Table 9. Average running time in seconds for the different methods. Sample sizes are 100 cases and 100 controls. Specifications of the computer: Inter Core i7-4720HQ CPU 2.6GHz, 16GB RAM.

# of Biomarkers	PMuN	Liu2005	Su1993	KS	Stepdown	SW-pepe	logistic
3	4.17	< 0.01	< 0.01	16.36	2.41	5.01	< 0.01
5	7.16	< 0.01	< 0.01	87.05	4.98	19.72	< 0.01
10	15.15	0.01	< 0.01	1027.1	10.99	99.32	0.01
15	23.72	< 0.01	< 0.01	2380.7	17.00	237.05	0.01
20	34.09	< 0.01	< 0.01	3981.1	22.92	442.01	< 0.01

Web Table 10. Simulation studies for comparing the proposed partition-based algorithm (PA) to direct-optimization algorithm (DA). The correlations between the biomarkers from the disease group are assumed to be highly negative and follow uniform(-0.6,-0.9) (to emphasize the complementary property of multiple biomarkers), while the correlations between the biomarkers from the non-disease group are assumed to be low and follow uniform(0.1,0.3). Other population parameters are drawn from different uniform distributions and should satisfy the following assumption: for each biomarker, the disease group has higher standard deviation than the non-disease group; the direction of some biomarker is allowed to be the opposite; and all markers should have an AUC around 0.6. For 5, 10, 15, and 20 biomarkers, we draw 1000 sets of population parameters and use the two algorithms to search for the optimal pAUC on each set of population parameters. The *relative differences* defined as $(pAUC_{PA} - pAUC_{DA}) / pAUC_{DA} \times 100\%$ are summarized below. The targeted FPR region is [0,0.3]. Note that the initial value for the direct-optimization algorithm is chosen as Fishers discriminant coefficients, which should be a robust initial guess. Four different commonly used optimization procedures are used.

The summary of the relative difference				
# of Markers	Min	Median	Mean	Max
<i>fminbnd()</i> from <i>neldermead</i> .				
5	-2.1%	0.4%	3.7%	39.7%
10	-1.9%	0.5%	5.3%	27.4%
15	-0.4%	0.7%	4.2%	27.7%
20	-0.6%	1.3%	6.1%	24.0%
<i>nloptr</i> (with <i>LNAIgrthm</i> ='NLOPT_LN_BOBYQA') from <i>nloptr</i> .				
5	0%	> 100%	> 100%	> 100%
10	0%	91%	> 100%	> 100%
15	0%	67%	> 100%	> 100%
20	0%	73%	82%	92%
<i>optim</i> (with <i>method</i> ='L-BFGS-B') from <i>stats</i> .				
5	100%	> 100%	> 100%	> 100%
10	0%	94%	> 100%	> 100%
15	0%	78%	> 100%	> 100%
20	0%	69%	> 100%	> 100%
<i>Generalized simulated annealing GenSA()</i> from <i>GenSA</i> .				
5	-1e-12%	-1e-15%	-5e-15%	1e-15%
10	-7e-11%	-6e-12%	-1e-11%	2e-12%
15	-6e-10%	-7e-11%	-6e-11%	3e-12%
20	-1e-4%	-7e-11%	-1e-6%	1e-11%

Web Table 11. Comparing time-to-complete for generalized simulated annealing between the proposed algorithm and direct optimization. Times summarized below are *system times* calculated by *proc.time()* from base R.

# of Markers	using Proposed Algorithm				Using Direct Optimization			
	Min	Median	Mean	Max	Min	Median	Mean	Max
5	0	0.01	0.017	0.08	0	0.03	0.034	0.11
10	0	0.03	0.057	0.17	0.03	0.08	0.086	0.22
15	0	0.06	0.069	0.16	0.03	0.13	0.128	0.23
20	0.04	0.14	0.137	0.25	0.08	0.18	0.189	0.32