Supplementary section 1

To estimate parameters given the thresholds, we used expectation-maximization (EM) algorithm, which is based on the maximum likelihood method. The similar algorithm can be found in (Godoy, et al., 2011; Chen, et al., 2012). In order to simplify the formula, the following notations are introduced:

$$\Theta = (\theta, \gamma^{\mathrm{T}}, \alpha_0)^{\mathrm{T}}, \qquad \varphi_i = \left(G_i, X_i^{\mathrm{T}}, 1\right)^{\mathrm{T}}$$
$$f(i,j) = f\left(c_j - \varphi_i^T \cdot \hat{\theta}^k\right), \qquad F(i,j) = \Phi\left(c_j - \varphi_i^T \cdot \hat{\theta}^k\right)$$

where$\Phi(t)$ and *f(t)* are the value of cumulative distribution function and probability density function of a normal distribution with mean 0 and variance σ^2 at point *t*, respectively. $c = \left(c_1, c_2, ..., c_r\right)^{\mathrm{T}}$ is the threshold vector. The likelihood function is as following:

$$L(\theta, \gamma_1, \dots, \gamma_m, \alpha_0) = \prod_{\{i:s_i=0\}} F(i,1) \cdot \prod_{\{i:s_i=1\}} [F(i,2) - F(i,1)] \cdot \cdots \cdot \prod_{\{i:s_i=r-1\}} [F(i,r) - F(i,r-1)] \cdot \prod_{\{i:s_i=r-1\}} [1 - F(i,r)]$$

Based on the EM algorithm, the parameters which maximize the likelihood function can be obtained by an iterative process. Assume the estimate value of parameter $\hat{\theta}$at k^{th} iteration is denoted by $\hat{\theta}^k = \left(\hat{\theta}^k, \hat{\gamma}_1^k, \hat{\gamma}_2^k, \dots, \hat{\gamma}_m^k, \hat{\alpha}_0^k\right)^{\mathrm{T}}$, then the detailed iterative algorithm is as following:

$$\hat{\Theta}^{k+1} = \hat{\Theta}^k - \left(\sum_{i=1}^{N} \varphi_i \cdot \varphi_i^T\right)^{-1} \left[\sum_{i=1}^{N} \sigma^2 \varphi_i \left(\sum_{j=0}^{r} I_{\{s_i=j\}} \cdot \frac{f(i,j+1) - f(i,j)}{F(i,j+1) - F(i,j)}\right)\right]$$

The iteration initial value can be selected arbitrarily. Here we use zero vector as the initial value. As the iteration goes on, the parameters converge to the value $\hat{\theta} = \left(\hat{\theta}, \hat{\gamma}_1, \hat{\gamma}_2, \dots, \hat{\gamma}_m, \hat{\alpha}_0\right)^{\mathrm{T}}$ that maximizes the likelihood function. Furthermore, the corresponding fisher information matrix $i(\hat{\Theta})$ can be obtained:

$$\mathrm{i}(\hat{\Theta}) = -E\left[\frac{\partial^2}{\partial\Theta^2}\log L(\Theta)\,|\,\hat{\Theta}\right]$$
$$= \sum_{i=1}^{N}\left[\left(\frac{\hat{f}(i,1)^2}{\hat{F}(i,1)} + \sum_{j=1}^{r-1}\frac{\left(\hat{f}(i,j+1)-\hat{f}(i,j)\right)^2}{\hat{F}(i,j+1)-\hat{F}(i,j)} + \frac{\hat{f}(i,r)^2}{1-\hat{F}(i,r)}\right)\cdot\varphi_i\cdot\varphi_i^{\mathrm{T}}\right].$$

In the iteration process, the denominator which is close to 0 sometimes induces some elements of information matrix to not be a number or infinity. Here we assign the value 10^7 to continue the algorithm procedure. In the simulations we fix σ=1 because we are interested in association testing not estimating association effect size. Actually likelihood function is homogeneous for parameters and variance so that the different variance would change the estimate of parameter and fisher matrix simultaneously, which result in no effect on the association test.

Chen T, Zhao Y, Ljung L. 2012.Impulse Response Estimation with Binary Measurements: A Regularized FIR Model Approach System Identification. 16th IFAC Symposium on System Identification.16(1): 113-118.

Godoy B, Goodwin G, Aguero J, Marelli D, Wigren T. 2011. On identification of FIR systems having quantized output data. Automatica 47: 1905-1915.

Supplementary section 2

Algorithm implementation

Below is the detailed algorithm implementation of the proposed SVSI method.

If the threshold is known, the iteration (3) is repeated starting from the default initial value origin until the termination condition is satisfied or the default maximum number of iteration steps (200) is reached. The difference between adjacent estimate parameters (which is a vector) is calculated for each iteration step, and the termination condition is that every component of the difference is less than the product of the maximum value of the corresponding component of the estimate parameters and a default value of 10^{-6}.

If the threshold is unknown, then the default initial value of threshold c is set to be a sequence of 1:r, i.e., where r is the total number of categories of the observed outcome minus 1, and the default of the initial value of parameters estimated is set to be origin. The iterations (3) and (6) are alternatively repeated to estimate the parameter and threshold simultaneously. The algorithm stops when the L_2 norm of the difference between adjacent estimate parameters and the difference between adjacent estimate thresholds are smaller than a pre-specified prevision threshold; the default value is 10^{-8}, or the default maximum number of iteration steps (100) is reached.

Once MLE $\hat{\theta}$ is obtained, the observed Fisher information matrix of $\hat{\theta}$ and Wald statistics can be constructed by equation (4) and (5), respectively. Under the null hypothesis, Wald statistics asymptotically follows the chi-square distribution with df of 1.

Supplementary section 3

Computational time

The main difference between **SV** and **oPRB** lies in their respective core algorithm. The **SV** method adopts EM algorithm and **oPRB** mostly uses IRWLS algorithm. These two algorithms have their own advantages. EM algorithm has outstanding robustness so that no matter what the initial value is, the likelihood is non-decreasing in the number of the iterations (McLachlan and Krishnan, 1997). And due to no need of calculating the complicated weighted matrix and its corresponding inverse matrix, EM algorithm takes much less time (about dozens of times less) than IRWLS algorithm per iteration. IRWLS algorithm, however, has local superlinear convergence property that guarantees very fast convergence ability near the optimal position. Hence, given fixed convergence accuracy, the IRWLS algorithm sometimes needs less iteration steps, especially when initial value is near the optimal value. These are consistent with the traditional view and reported in Bi et al. (2014).

To compare the computational costs of the four methods above, we applied the four methods above to the simulated datasets and calculated the elapse time under different circumstances. For **SVSI**, we used the maximum number of iteration steps 200 not 100 for this cost comparison of the computational cost. For commercial software Matlab, the built-in *mnrfit* function for **oPRB** and **oLG** was used. For freely available software R, *polr* function in MASS package (http://stat.ethz.ch/R-manual/R-patched/library/MASS/html/polr.html) was used for **oLG** and **oPRB**.

Supplementary Table S3 presents the comparison of computational costs of the proposed **SV** and the three other methods in Matlab. As displayed in supplementary Table 3S, 49% ~ 121% of the computational cost could be saved by using the proposed new **SV** method instead of **oLG** and **oPRB**. The computational cost of **SV** was similar to that of **LG** but the power of **LG** was significantly smaller than that of **SV** in most situations. As expected, similar to Matlab, **oLG** and **oPRB** were also slower than **LG** in R. **SV** was identical to or slightly slower than **oLG** and **oPRB** but had better performance in terms of type I error and power as shown in the main text.

McLachlan G.J & Krishnan T. 1997. The EM algorithm and extensions, Wiley Series in Probability and Statistics.

Bi W, Zhao Y, 2014. Iterative parameter estimate with batched binary-valued observations: convergence with an exponential rate. Proceedings of the 19th International Federation of Automatic Control World Congress **19**, 3220-3225.

Supplementary matrix 1

The population 3×3 tables for θ = 1:

	Genotype			
Phenotype	**0**	**1**	**2**	Subtotal
0	1448195	93995	896	1543086
1	293398	63564	1694	358656
2	63045	32857	2356	98258
Subtotal	1804638	190416	4946	2000000

The population 3×3 tables for θ = 2

	Genotype			
Phenotype	**0**	**1**	**2**	Subtotal
0	904255	7659	0	911914
1	568847	29971	24	598842
2	332655	151581	5008	489244
Subtotal	1805757	189211	5032	2000000

Supplementary Table S1: True and estimated thresholds by SVSI approach for the parameter combination used in Table 2. Data is generated using SVsimu.

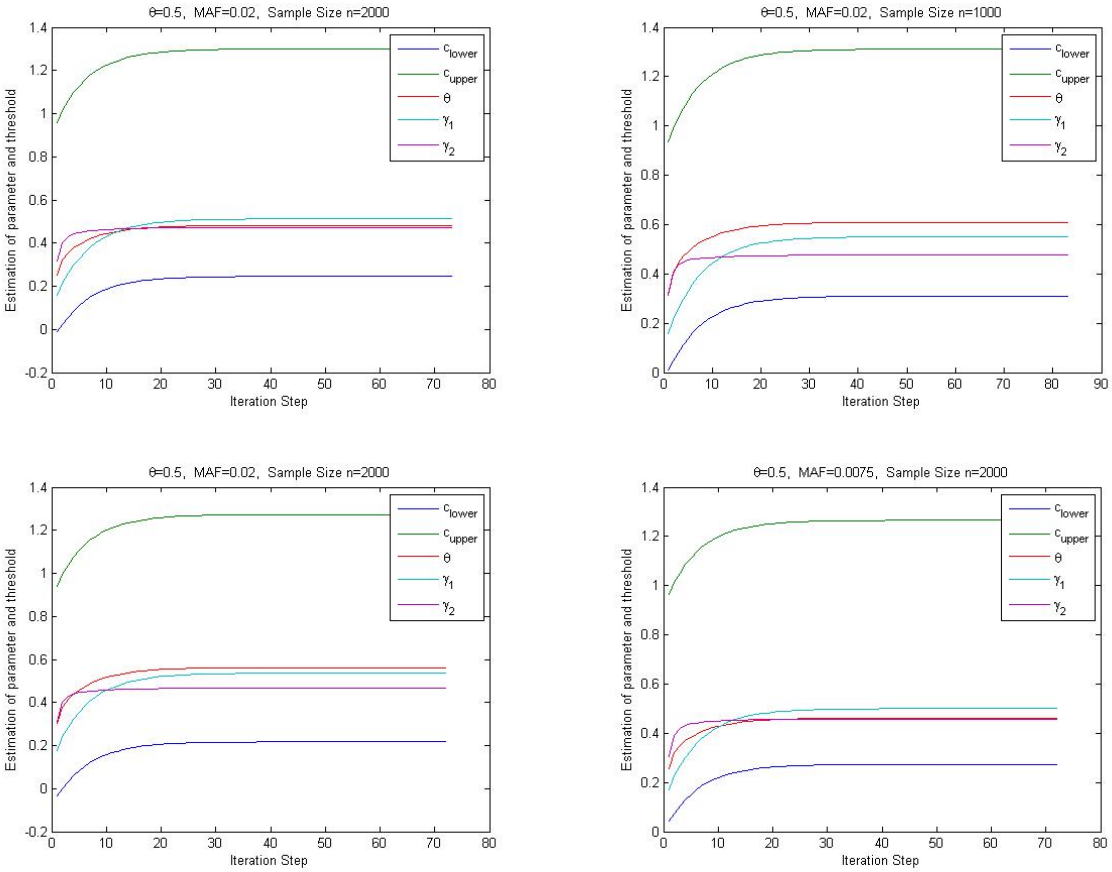
n	p_A	θ	Disease Model	c_1	c_2	$\overline{\widehat{c_1}}$	$\overline{\widehat{c_2}}$
Rand							
1000	0.0075	0	H_0	0.2534	1.2816	0.2526	1.2876
		0.5	ADD	0.2534	1.2816	0.2595	1.2845
		2	ADD	0.2534	1.2816	0.2448	1.2761
	0.2	0	H_0	0.2534	1.2816	0.2548	1.2755
		0.5	ADD	0.2534	1.2816	0.2557	1.2951
		2	ADD	0.2534	1.2816	0.2476	1.2844
2000	0.0075	0	H_0	0.2534	1.2816	0.2478	1.2877
		0.5	ADD	0.2534	1.2816	0.2545	1.2817
		2	ADD	0.2534	1.2816	0.2474	1.2763
	0.2	0	H_0	0.2534	1.2816	0.2480	1.2801
		0.5	ADD	0.2534	1.2816	0.2550	1.2840
		2	ADD	0.2534	1.2816	0.2619	1.2915
Same							
999	0.0075	0	H_0	0.2534	1.2816	-0.1423	0.8550
		0.5	ADD	0.2534	1.2816	-0.1336	0.8667
		2	ADD	0.2534	1.2816	-0.1195	0.8936
	0.2	0	H_0	0.2534	1.2816	-0.1387	0.8604
		0.5	ADD	0.2534	1.2816	0.0087	1.0311
		2	ADD	0.2534	1.2816	0.1579	1.4587
1998	0.0075	0	H_0	0.2534	1.2816	-0.1378	0.8592
		0.5	ADD	0.2534	1.2816	-0.1315	0.8665
		2	ADD	0.2534	1.2816	-0.1243	0.8883
	0.2	0	H_0	0.2534	1.2816	-0.1430	0.8539
		0.5	ADD	0.2534	1.2816	0.0080	1.0290
		2	ADD	0.2534	1.2816	0.1495	1.4426

n is the number of individuals sampled from the population; p_A is minor allele frequency of SNP; θ is the association coefficient of SNP with outcome; Disease model is representing the

underlying genetic disease model. c_1 and c_2 are two thresholds used to define ordered categorical phenotype with three categories.

Supplementary Figure Legend

Supplementary Figure S1.Performance of the estimation of the threshold and parameter as the iteration proceeds. The data is simulated by SVsimu model and randomly sampling process.



Supplementary Table S2. The mean of $\hat{\theta}$, mean of estimated standard error of $\hat{\theta}$, and standard deviation of $\hat{\theta}$ across simulation repetitions for the set-valued (SV) and logistic regression (LG and oLG) methods based on 1000 simulations*.

					LG					oLG					SV					oPRB				
n	p_A	θ	**SM**	**DM**	$\bar{\hat{\theta}}$	$\overline{\widehat{se}(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\bar{\hat{\theta}}}{\overline{\widehat{se}(\hat{\theta})}}$	$\frac{\bar{\hat{\theta}}}{sd(\hat{\theta})}$	$\bar{\hat{\theta}}$	$\overline{\widehat{se}(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\bar{\hat{\theta}}}{\overline{\widehat{se}(\hat{\theta})}}$	$\frac{\bar{\hat{\theta}}}{sd(\hat{\theta})}$	$\bar{\hat{\theta}}$	$\overline{\widehat{se}(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\bar{\hat{\theta}}}{\overline{\widehat{se}(\hat{\theta})}}$	$\frac{\bar{\hat{\theta}}}{sd(\hat{\theta})}$	$\bar{\hat{\theta}}$	$\overline{\widehat{se}(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\bar{\hat{\theta}}}{\overline{\widehat{se}(\hat{\theta})}}$	$\frac{\bar{\hat{\theta}}}{sd(\hat{\theta})}$
rand																								
1000	0.0075	0	LGsimu	ADD	-0.013	0.573	0.586	-0.022	-0.022	-0.039	0.536	0.541	-0.073	-0.073	-0.024	0.32	0.32	-0.077	-0.076	-0.024	0.317	0.321	-0.077	-0.076
1000	0.0075	0.5	LGsimu	ADD	0.573	1.021	0.772	0.5613	0.742	0.5084	0.512	0.515	0.993	0.9862	0.3	0.31	0.31	0.976	0.977	0.3	0.307	0.308	0.9783	0.9766
1000	0.0075	2	LGsimu	ADD	3.33	**38.53**	**4.103**	0.0864	0.812	2.0356	0.535	0.575	3.805	3.5423	1.208	0.32	0.34	3.819	3.556	1.208	0.317	0.34	3.816	3.5556
1000	0.0075	0	SVsimu	ADD	-0.036	0.94	0.756	-0.039	-0.048	-0.058	0.538	0.711	-0.108	-0.081	-0.031	0.32	0.34	-0.095	-0.09	-0.033	0.479	0.367	-0.069	-0.09
1000	0.0075	0.5	SVsimu	ADD	0.94	2.742	1.198	0.3427	0.785	0.8312	0.522	0.531	1.593	1.5661	0.495	0.31	0.32	1.586	1.56	0.495	0.312	0.318	1.5856	1.5605
1000	0.0075	2	SVsimu	ADD	9.988	**224.5**	**6.583**	0.0445	1.517	3.9777	**12.06**	**2.556**	0.33	1.5563	2.139	0.5	0.57	4.269	3.733	2.194	**1.178**	**0.788**	1.8618	2.7843
1000	0.2	0	LGsimu	ADD	-0.002	0.117	0.117	-0.019	-0.019	-0.005	0.111	0.11	-0.049	-0.05	-0.004	0.07	0.07	-0.058	-0.059	-0.004	0.066	0.065	-0.058	-0.059
1000	0.2	0.5	LGsimu	ADD	0.502	0.12	0.119	4.1869	4.222	0.4968	0.108	0.106	4.609	4.6781	0.296	0.06	0.06	4.601	4.686	0.296	0.064	0.063	4.5923	4.6861
1000	0.2	2	LGsimu	ADD	2.013	0.176	0.173	11.452	11.65	2.0085	0.132	0.135	15.18	14.843	1.186	0.07	0.08	16.52	15.43	1.186	0.074	0.077	16.015	15.434
1000	0.2	0	SVsimu	ADD	-0.008	0.122	0.122	-0.067	-0.067	-0.008	0.112	0.11	-0.068	-0.069	-0.004	0.07	0.07	-0.067	-0.068	-0.004	0.067	0.065	-0.067	-0.068
1000	0.2	0.5	SVsimu	ADD	0.842	0.135	0.133	6.2487	6.342	0.8444	0.112	0.109	7.517	7.7608	0.505	0.07	0.06	7.659	7.869	0.505	0.066	0.064	7.6079	7.8688
1000	0.2	2	SVsimu	ADD	3.615	0.318	0.322	11.357	11.23	3.4337	0.188	0.18	18.29	19.069	2.012	0.1	0.1	20.95	20.06	2.012	0.101	0.1	19.957	20.057
same																								
999	0.0075	0	LGsimu	ADD	0.218	5.072	1.644	0.0429	0.132	0.0127	0.498	0.526	0.026	0.0242	0.007	0.3	0.32	0.022	0.021	0.007	0.302	0.318	0.0222	0.021
999	0.0075	0.5	LGsimu	ADD	0.899	8.848	2.113	0.1016	0.426	0.4908	0.473	0.48	1.038	1.0225	0.299	0.29	0.29	1.037	1.023	0.299	0.287	0.292	1.04	1.0229
999	0.0075	2	LGsimu	ADD	5.569	**115.3**	**5.989**	0.0483	0.93	1.9515	0.487	0.491	4.011	3.9759	1.182	0.28	0.28	4.235	4.182	1.182	0.278	0.283	4.2457	4.1816
999	0.0075	0	SVsimu	ADD	0.097	3.234	1.313	0.03	0.074	-0.011	0.523	0.541	-0.021	-0.021	-0.007	0.31	0.32	-0.023	-0.022	-0.007	0.313	0.323	-0.023	-0.022
999	0.0075	0.5	SVsimu	ADD	1.284	11.54	2.3	0.1112	0.558	0.8273	0.498	0.507	1.662	1.6309	0.499	0.3	0.3	1.681	1.645	0.499	0.296	0.303	1.6827	1.6449
999	0.0075	2	SVsimu	ADD	11.2	**285.5**	**6.432**	0.0392	1.742	3.9372	3.015	2.771	1.306	1.421	2.044	0.45	0.54	4.567	3.804	2.114	**1.249**	**0.8**	1.6923	2.6411
999	0.2	0	LGsimu	ADD	-0.005	0.124	0.128	-0.039	-0.038	-0.008	0.105	0.108	-0.073	-0.071	-0.005	0.06	0.07	-0.074	-0.072	-0.005	0.064	0.066	-0.074	-0.072
999	0.2	0.5	LGsimu	ADD	0.551	0.129	0.133	4.2645	4.137	0.5049	0.103	0.104	4.879	4.866	0.306	0.06	0.06	4.914	4.887	0.306	0.062	0.063	4.9069	4.8867
999	0.2	2	LGsimu	ADD	2.058	0.186	0.187	11.059	11.02	1.9861	0.131	0.132	15.18	15.051	1.173	0.07	0.07	16.54	15.71	1.173	0.073	0.075	16.033	15.715

					LG					oLG					SV					oPRB				
n	p_A	θ	**SM**	**DM**	$\hat{\theta}$	$\widehat{se(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\hat{\theta}}{\widehat{se(\hat{\theta})}}$	$\frac{\hat{\theta}}{sd(\hat{\theta})}$	$\hat{\theta}$	$\widehat{se(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\hat{\theta}}{\widehat{se(\hat{\theta})}}$	$\frac{\hat{\theta}}{sd(\hat{\theta})}$	$\hat{\theta}$	$\widehat{se(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\hat{\theta}}{\widehat{se(\hat{\theta})}}$	$\frac{\hat{\theta}}{sd(\hat{\theta})}$	$\hat{\theta}$	$\widehat{se(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\hat{\theta}}{\widehat{se(\hat{\theta})}}$	$\frac{\hat{\theta}}{sd(\hat{\theta})}$
999	0.2	0	SVsimu	ADD	0.006	0.131	0.128	0.0447	0.046	0.0027	0.108	0.106	0.025	0.0253	0.002	0.06	0.06	0.026	0.026	0.002	0.065	0.064	0.026	0.0264
999	0.2	0.5	SVsimu	ADD	0.881	0.144	0.143	6.1381	6.164	0.8446	0.11	0.113	7.647	7.4535	0.507	0.06	0.07	7.81	7.523	0.507	0.065	0.067	7.7637	7.5231
999	0.2	2	SVsimu	ADD	3.351	0.315	0.304	10.649	11.02	3.2821	0.181	0.17	18.18	19.288	1.919	0.09	0.1	21	20.17	1.919	0.098	0.095	19.625	20.172
rand																								
1000	0.0075	0	LGsimu	DOM	-0.013	0.573	0.586	-0.022	-0.022	-0.039	0.536	0.541	-0.073	-0.073	-0.024	0.32	0.32	-0.077	-0.076	-0.024	0.317	0.321	-0.077	-0.076
1000	0.0075	0.5	LGsimu	DOM	0.573	1.021	0.772	0.5613	0.742	0.5084	0.512	0.515	0.993	0.9862	0.3	0.31	0.31	0.976	0.977	0.3	0.307	0.308	0.9783	0.9766
1000	0.0075	2	LGsimu	DOM	3.33	**38.53**	**4.103**	0.0864	0.812	2.0356	0.535	0.575	3.805	3.5423	1.208	0.32	0.34	3.819	3.556	1.208	0.317	0.34	3.816	3.5556
1000	0.0075	0	SVsimu	DOM	-0.036	0.94	0.756	-0.039	-0.048	-0.058	0.538	0.711	-0.108	-0.081	-0.031	0.32	0.34	-0.095	-0.09	-0.033	0.479	0.367	-0.069	-0.09
1000	0.0075	0.5	SVsimu	DOM	0.94	2.742	1.198	0.3427	0.785	0.8312	0.522	0.531	1.593	1.5661	0.495	0.31	0.32	1.586	1.56	0.495	0.312	0.318	1.5856	1.5605
1000	0.0075	2	SVsimu	DOM	9.988	**224.5**	**6.583**	0.0445	1.517	3.9777	**12.06**	**2.556**	0.33	1.5563	2.139	0.5	0.57	4.269	3.733	2.194	**1.178**	**0.788**	1.8618	2.7843
1000	0.2	0	LGsimu	DOM	-0.002	0.117	0.117	-0.019	-0.019	-0.005	0.111	0.11	-0.049	-0.05	-0.004	0.07	0.07	-0.058	-0.059	-0.004	0.066	0.065	-0.058	-0.059
1000	0.2	0.5	LGsimu	DOM	0.502	0.12	0.119	4.1869	4.222	0.4968	0.108	0.106	4.609	4.6781	0.296	0.06	0.06	4.601	4.686	0.296	0.064	0.063	4.5923	4.6861
1000	0.2	2	LGsimu	DOM	2.013	0.176	0.173	11.452	11.65	2.0085	0.132	0.135	15.18	14.843	1.186	0.07	0.08	16.52	15.43	1.186	0.074	0.077	16.015	15.434
1000	0.2	0	SVsimu	DOM	-0.008	0.122	0.122	-0.067	-0.067	-0.008	0.112	0.11	-0.068	-0.069	-0.004	0.07	0.07	-0.067	-0.068	-0.004	0.067	0.065	-0.067	-0.068
1000	0.2	0.5	SVsimu	DOM	0.842	0.135	0.133	6.2487	6.342	0.8444	0.112	0.109	7.517	7.7608	0.505	0.07	0.06	7.659	7.869	0.505	0.066	0.064	7.6079	7.8688
1000	0.2	2	SVsimu	DOM	3.615	0.318	0.322	11.357	11.23	3.4337	0.188	0.18	18.29	19.069	2.012	0.1	0.1	20.95	20.06	2.012	0.101	0.1	19.957	20.057
2000	0.0075	0	LGsimu	DOM	0.02	0.391	0.409	0.0508	0.049	0.0047	0.367	0.387	0.013	0.0122	0.002	0.22	0.23	0.008	0.007	0.002	0.219	0.231	0.0078	0.0074
2000	0.0075	0.5	LGsimu	DOM	0.535	0.398	0.402	1.3444	1.329	0.5143	0.353	0.354	1.456	1.4522	0.304	0.21	0.21	1.43	1.426	0.304	0.212	0.213	1.4324	1.4264
2000	0.0075	2	LGsimu	DOM	2.229	**2.711**	**1.336**	0.8222	1.669	2.0093	0.366	0.374	5.49	5.3738	1.194	0.22	0.22	5.478	5.373	1.194	0.218	0.222	5.4744	5.373
2000	0.0075	0	SVsimu	DOM	-0.02	0.409	0.408	-0.049	-0.049	-0.026	0.372	0.386	-0.07	-0.067	-0.016	0.22	0.23	-0.073	-0.071	-0.016	0.222	0.229	-0.073	-0.071
2000	0.0075	0.5	SVsimu	DOM	0.84	0.44	0.47	1.9101	1.789	0.8075	0.361	0.377	2.236	2.1392	0.483	0.22	0.23	2.231	2.139	0.483	0.217	0.226	2.2309	2.1386
2000	0.0075	2	SVsimu	DOM	6.953	**83.61**	**5.785**	0.0832	1.202	3.4423	0.711	0.648	4.841	5.3102	2.033	0.28	0.3	7.15	6.709	2.035	0.298	0.314	6.838	6.4713
2000	0.2	0	LGsimu	DOM	-0.003	0.082	0.082	-0.036	-0.037	-0.004	0.078	0.078	-0.058	-0.057	-0.003	0.05	0.05	-0.064	-0.064	-0.003	0.047	0.047	-0.064	-0.064
2000	0.2	0.5	LGsimu	DOM	0.502	0.085	0.087	5.9412	5.798	0.5	0.076	0.077	6.566	6.4765	0.298	0.05	0.05	6.555	6.475	0.298	0.045	0.046	6.5417	6.4748
2000	0.2	2	LGsimu	DOM	2.003	0.124	0.128	16.201	15.66	1.9953	0.093	0.093	21.43	21.351	1.179	0.05	0.05	23.3	22.13	1.179	0.052	0.053	22.592	22.128
2000	0.2	0	SVsimu	DOM	1E-03	0.086	0.086	0.0114	0.011	######	0.079	0.078	1E-04	0.0001	-2E-04	0.05	0.05	-0.004	-0.004	-2E-04	0.047	0.046	-0.004	-0.004

					LG					oLG					SV					oPRB				
n	p_A	θ	**SM**	**DM**	$\bar{\hat{\theta}}$	$\overline{\widehat{se}(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\bar{\hat{\theta}}}{\overline{\widehat{se}(\hat{\theta})}}$	$\frac{\bar{\hat{\theta}}}{sd(\hat{\theta})}$	$\bar{\hat{\theta}}$	$\overline{\widehat{se}(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\bar{\hat{\theta}}}{\overline{\widehat{se}(\hat{\theta})}}$	$\frac{\bar{\hat{\theta}}}{sd(\hat{\theta})}$	$\bar{\hat{\theta}}$	$\overline{\widehat{se}(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\bar{\hat{\theta}}}{\overline{\widehat{se}(\hat{\theta})}}$	$\frac{\bar{\hat{\theta}}}{sd(\hat{\theta})}$	$\bar{\hat{\theta}}$	$\overline{\widehat{se}(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\bar{\hat{\theta}}}{\overline{\widehat{se}(\hat{\theta})}}$	$\frac{\bar{\hat{\theta}}}{sd(\hat{\theta})}$
2000	0.2	0.5	SVsimu	DOM	0.833	0.095	0.095	8.7808	8.779	0.8368	0.079	0.08	10.57	10.524	0.501	0.05	0.05	10.76	10.57	0.501	0.047	0.047	10.689	10.572
2000	0.2	2	SVsimu	DOM	3.581	0.221	0.214	16.194	16.73	3.4179	0.132	0.129	25.84	26.594	2.003	0.07	0.07	29.58	27.97	2.003	0.071	0.072	28.186	27.974
same																								
999	0.0075	0	LGsimu	DOM	0.218	5.072	1.644	0.0429	0.132	0.0127	0.498	0.526	0.026	0.0242	0.007	0.3	0.32	0.022	0.021	0.007	0.302	0.318	0.0222	0.021
999	0.0075	0.5	LGsimu	DOM	0.899	**8.848**	**2.113**	0.1016	0.426	0.4908	0.473	0.48	1.038	1.0225	0.299	0.29	0.29	1.037	1.023	0.299	0.287	0.292	1.04	1.0229
999	0.0075	2	LGsimu	DOM	5.569	**115.3**	**5.989**	0.0483	0.93	1.9515	0.487	0.491	4.011	3.9759	1.182	0.28	0.28	4.235	4.182	1.182	0.278	0.283	4.2457	4.1816
999	0.0075	0	SVsimu	DOM	0.097	3.234	1.313	0.03	0.074	-0.011	0.523	0.541	-0.021	-0.021	-0.007	0.31	0.32	-0.023	-0.022	-0.007	0.313	0.323	-0.023	-0.022
999	0.0075	0.5	SVsimu	DOM	1.284	**11.54**	**2.3**	0.1112	0.558	0.8273	0.498	0.507	1.662	1.6309	0.499	0.3	0.3	1.681	1.645	0.499	0.296	0.303	1.6827	1.6449
999	0.0075	2	SVsimu	DOM	11.2	**285.5**	**6.432**	0.0392	1.742	3.9372	3.015	2.771	1.306	1.421	2.044	0.45	0.54	4.567	3.804	2.114	**1.249**	**0.8**	1.6923	2.6411
999	0.2	0	LGsimu	DOM	-0.005	0.124	0.128	-0.039	-0.038	-0.008	0.105	0.108	-0.073	-0.071	-0.005	0.06	0.07	-0.074	-0.072	-0.005	0.064	0.066	-0.074	-0.072
999	0.2	0.5	LGsimu	DOM	0.551	0.129	0.133	4.2645	4.137	0.5049	0.103	0.104	4.879	4.866	0.306	0.06	0.06	4.914	4.887	0.306	0.062	0.063	4.9069	4.8867
999	0.2	2	LGsimu	DOM	2.058	0.186	0.187	11.059	11.02	1.9861	0.131	0.132	15.18	15.051	1.173	0.07	0.07	16.54	15.71	1.173	0.073	0.075	16.033	15.715
999	0.2	0	SVsimu	DOM	0.006	0.131	0.128	0.0447	0.046	0.0027	0.108	0.106	0.025	0.0253	0.002	0.06	0.06	0.026	0.026	0.002	0.065	0.064	0.026	0.0264
999	0.2	0.5	SVsimu	DOM	0.881	0.144	0.143	6.1381	6.164	0.8446	0.11	0.113	7.647	7.4535	0.507	0.06	0.07	7.81	7.523	0.507	0.065	0.067	7.7637	7.5231
999	0.2	2	SVsimu	DOM	3.351	0.315	0.304	10.649	11.02	3.2821	0.181	0.17	18.18	19.288	1.919	0.09	0.1	21	20.17	1.919	0.098	0.095	19.625	20.172
1998	0.0075	0	LGsimu	DOM	0.05	0.42	0.433	0.1198	0.116	0.0175	0.348	0.355	0.05	0.0491	0.01	0.21	0.22	0.049	0.048	0.01	0.211	0.216	0.0489	0.0479
1998	0.0075	0.5	LGsimu	DOM	0.632	0.717	0.631	0.8816	1.002	0.5061	0.329	0.332	1.538	1.5232	0.309	0.2	0.2	1.543	1.532	0.309	0.2	0.202	1.5472	1.5319
1998	0.0075	2	LGsimu	DOM	3.042	**17.62**	**3.062**	0.1727	0.993	1.9388	0.338	0.333	5.74	5.8158	1.174	0.19	0.19	6.035	6.091	1.174	0.194	0.193	6.0467	6.0909
1998	0.0075	0	SVsimu	DOM	0.004	0.442	0.46	0.0087	0.008	-0.002	0.362	0.369	-0.005	-0.004	-8E-04	0.22	0.22	-0.004	-0.004	-8E-04	0.217	0.221	-0.004	-0.004
1998	0.0075	0.5	SVsimu	DOM	0.946	1.116	0.806	0.8477	1.174	0.8187	0.345	0.353	2.371	2.317	0.494	0.21	0.21	2.392	2.348	0.494	0.206	0.21	2.3952	2.3485
1998	0.0075	2	SVsimu	DOM	8.881	**140.8**	**6.375**	0.0631	1.393	3.3684	**0.493**	**0.86**	6.836	3.9173	1.96	0.26	0.29	7.454	6.759	1.965	0.3	0.332	6.5603	5.926
1998	0.2	0	LGsimu	DOM	-0.003	0.087	0.089	-0.04	-0.039	-0.005	0.074	0.074	-0.068	-0.067	-0.003	0.04	0.05	-0.071	-0.071	-0.003	0.045	0.045	-0.071	-0.071
1998	0.2	0.5	LGsimu	DOM	0.548	0.091	0.089	6.0108	6.169	0.5003	0.073	0.072	6.853	6.9199	0.303	0.04	0.04	6.895	6.963	0.303	0.044	0.044	6.8861	6.963
1998	0.2	2	LGsimu	DOM	2.042	0.131	0.128	15.649	15.93	1.9792	0.092	0.092	21.46	21.563	1.17	0.05	0.05	23.37	22.39	1.17	0.052	0.052	22.667	22.385
1998	0.2	0	SVsimu	DOM	-0.002	0.092	0.093	-0.027	-0.026	-1E-03	0.076	0.077	-0.013	-0.012	-6E-04	0.05	0.05	-0.014	-0.014	-6E-04	0.046	0.046	-0.014	-0.014
1998	0.2	0.5	SVsimu	DOM	0.871	0.101	0.098	8.6146	8.866	0.8325	0.078	0.075	10.68	11.041	0.5	0.05	0.04	10.91	11.2	0.5	0.046	0.045	10.849	11.203

					LG					oLG					SV					oPRB				
n	p_A	θ	**SM**	**DM**	$\bar{\hat{\theta}}$	$\overline{\widehat{se}(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\bar{\hat{\theta}}}{\overline{\widehat{se}(\hat{\theta})}}$	$\frac{\bar{\hat{\theta}}}{sd(\hat{\theta})}$	$\bar{\hat{\theta}}$	$\overline{\widehat{se}(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\bar{\hat{\theta}}}{\overline{\widehat{se}(\hat{\theta})}}$	$\frac{\bar{\hat{\theta}}}{sd(\hat{\theta})}$	$\bar{\hat{\theta}}$	$\overline{\widehat{se}(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\bar{\hat{\theta}}}{\overline{\widehat{se}(\hat{\theta})}}$	$\frac{\bar{\hat{\theta}}}{sd(\hat{\theta})}$	$\bar{\hat{\theta}}$	$\overline{\widehat{se}(\hat{\theta})}$	$sd(\hat{\theta})$	$\frac{\bar{\hat{\theta}}}{\overline{\widehat{se}(\hat{\theta})}}$	$\frac{\bar{\hat{\theta}}}{sd(\hat{\theta})}$
1998	0.2	2	SVsimu	DOM	3.337	0.22	0.205	15.144	16.31	3.2729	0.127	0.119	25.73	27.401	1.915	0.06	0.07	29.7	28.93	1.915	0.069	0.066	27.769	28.931

$\bar{\hat{\theta}}$: The mean of $\hat{\theta}$ for 1000 replicates; $\overline{\widehat{se}(\hat{\theta})}$: The mean of estimated standard error of $\hat{\theta}$ for 1000 replicates; $sd(\hat{\theta})$: The empirical standard deviation of $\hat{\theta}$ for 1000 replicates; p_A is minor allele frequency of SNP; θ is the association coefficient of SNP with outcome; SM is simulation model; DM is disease model representing the underlying genetic disease model. LG stands for logistic regression model on the regrouped binary outcome (recoding as 0 or greater than 0); SV stands for the set-valued method; oLG stands for ordered logistic regression method.

Supplementary Table S3. Computational costs of the four methods of logistic regression (LG and oLG), oPRB, and SV on simulated datasets under different situations.

n	p_A	θ	DM	SV	oPRB	oLG	LG	Ratio of times taken		
								oPRB/SV	oLG/SV	LG/SV
1000	0.0075	0	H0	0.19	0.37	0.37	0.16	1.9	1.9	0.82
2000	0.0075	0	H0	0.3	0.59	0.62	0.18	1.95	2.03	0.61
1000	0.2	0	H0	0.19	0.3	0.3	0.14	1.62	1.62	0.73
2000	0.2	0	H0	0.27	0.49	0.49	0.15	1.78	1.8	0.56
1000	0.0075	0.5	ADD	0.14	0.29	0.3	0.13	2.02	2.06	0.92
2000	0.0075	0.5	ADD	0.21	0.44	0.46	0.14	2.12	2.21	0.68
1000	0.2	0.5	ADD	0.16	0.28	0.28	0.13	1.71	1.69	0.8
2000	0.2	0.5	ADD	0.23	0.43	0.43	0.14	1.86	1.86	0.62
1000	0.0075	1	ADD	0.14	0.28	0.28	0.13	1.98	2	0.91
2000	0.0075	1	ADD	0.21	0.45	0.46	0.15	2.13	2.18	0.7
1000	0.2	1	ADD	0.17	0.28	0.27	0.13	1.63	1.6	0.76
2000	0.2	1	ADD	0.25	0.44	0.43	0.14	1.75	1.72	0.58
1000	0.0075	1.5	ADD	0.15	0.29	0.29	0.14	1.96	1.97	0.93
2000	0.0075	1.5	ADD	0.22	0.46	0.46	0.15	2.14	2.14	0.71
1000	0.2	1.5	ADD	0.19	0.29	0.29	0.13	1.55	1.51	0.71
2000	0.2	1.5	ADD	0.28	0.47	0.46	0.15	1.65	1.62	0.54
1000	0.0075	2	ADD	0.16	0.34	0.34	0.16	2.16	2.18	1.02
2000	0.0075	2	ADD	0.22	0.48	0.48	0.17	2.2	2.21	0.8
1000	0.2	2	ADD	0.21	0.31	0.3	0.14	1.49	1.45	0.66
2000	0.2	2	ADD	0.33	0.52	0.51	0.16	1.57	1.53	0.48
Mean(range)								1.86(1.49,2.20)	1.86(1.45,2.21)	0.73(0.48,1.02)

n is the number of individuals sampled from the population; p_A is minor allele frequency of SNP; θ is the association coefficient of SNP with outcome; DM (Disease model) is representing the underlying genetic disease model. **LG** stands for logistic regression model on the regrouped

binary outcome (recoding as 0 or greater than 0); **SV** stands for the set-valued method; **oLG** stands for ordered logistic regression method; **oPRB** is the usual probit model with IRWLS estimation algorithm.