Supplementary Section A Illustration of Stratified Trend Tests.

We use case-control data examining the relationship between endometrial cancer and two estrogen-related variants in *CYP19A1* [1] to illustrate the comparison of stratified trend tests. Using case-control genotype counts published in the paper, we conducted trend tests stratified on age group ~~and~~ , body mass index (BMI) group, and race/ethnicity, without adjustment for other variables. Supplementary Table 1 contains results of these analyses. As in the original analyses [1], we found a strong relationship between endometrial cancer and the genetic variants of interest, except when the genetic model was assumed to be heterozygous. Statistics based on the monotone genetic models are generally consistent with each other, though the additive model statistics are larger than dominant and recessive in this example, perhaps because the additive model is closer to the truth for these data. ~~In the analyses stratified by age group, the values of the statistics are similar whether the statistic makes the case-control model assumption or the population model assumption for H. In the analyses stratified by BMI group and race/ethnicity, the values of the statistics are larger when the statistic makes the case-control model assumption than when it makes the population model assumption.~~

The values of the statistics are lower under the case-control model assumption than they are under the population assumption in the analyses stratified by age group, and higher under the case-control model assumption than they are under the population assumption in the analyses stratified by BMI and race/ethnicity. These differences are due to differences in the distribution of cases and controls over the strata of the adjustment variable. As age increases, the proportion of cases R_i/N_i in each age group decreases slightly, causing the weight function $J'_\pi(J_\pi^{-1}(R_i/N_i))$ under the population model $H = J_\pi$ in equation (2) to increase with increasing age. Under the case-control model, where $H = L$, the weight function $L'(L^{-1}(R_i/N_i))$ is constant or decreases as age increases. Because the association between genotype and disease is stronger in the higher age groups, the statistic is higher under the population model assumption than it is under the case-control model assumption.

The reverse is true for the proportion of cases in each BMI group. The genotype-disease relationship is stronger in the higher BMI groups, which also have a higher proportion of cases. This means that the high-association BMI strata are weighted more highly by the case-control model weights, and less highly by the population model weights, causing the comparison between case-control model and population model statistics to be reversed from what it was under age stratification. The distribution of cases and controls over race/ethnicity strata and its relationship to the strata with the strongest association is most similar to the distribution pattern in BMI groups. Blacks have the smallest proportion of cases and the weakest gene-disease association. The proportion of cases and the gene-disease association are stronger in whites and Asians, creating a similar pattern to the one in BMI groups.

Supplementary Section B Derivation of ARE for stratified case-control data using population model.

To calculate ARE for case-control stratified data under a population model, we first calculated unstratified components of Noether's theorem [2], then used sums over strata to deduce the ARE formula. We used the notation in Randles' and Wolfe's text [3], letting $S_{nm} = U_m = \frac{H'_m(\hat{\delta}_m)}{H_m(\hat{\delta}_m)[1-H_m(\hat{\delta}_m)]}\left[\frac{S}{N}\sum_j d_{jm}r_j - \frac{R}{N}\sum_j d_{jm}s_j\right]$, where $\hat{\delta}_m = H_m^{-1}(R/N)$ and m refers to the statistic being used. This statistic was compared to the statistic $T_{nm} = S_{n\tilde{m}}$. Both evaluated the null hypothesis $\beta = \beta_0$. Given sequences $\mu_{S_{nm}}(\beta)$, $\mu_{S_{n\tilde{m}}}(\beta)$, $\sigma^2_{S_{nm}}(\beta)$, and $\sigma^2_{S_{n\tilde{m}}}(\beta)$, the conditions necesssary for use of Noether's theorem that ARE(S_{nm},$S_{n\tilde{m}}$)= $\frac{e_{mL}}{e_{\tilde{m}L}}$ [3] are met.

Note that, for case-control data, the r_j and s_j have multinomial distributions whose probabilities are

$$P(\text{j alleles}|\text{disease, sampled}) = \frac{J_\pi(\alpha+\beta d_{jL})g_j^{(cc)}}{\sum_l J_\pi(\alpha+\beta d_{lL})g_l^{(cc)}}$$

and

$$P(\text{j alleles}|\text{no disease, sampled}) = \frac{[1-J_\pi(\alpha+\beta d_{jL})]g_j^{(cc)}}{\sum_l [1-J_\pi(\alpha+\beta d_{lL})]g_l^{(cc)}}, \text{ respectively.}$$

ARE(S_{nm},$S_{n\tilde{m}}$)= $\frac{e_{mL}}{e_{\tilde{m}L}}$ can be calculated as follows:

1. Rewrite S_{nm}:

$$S_{nm} = \frac{H'_m(\hat{\delta}_m)}{H_m(\hat{\delta}_m)[1-H_m(\hat{\delta}_m)]}\left[N\sum_j d_{jm}\left\{\frac{r_j}{n_j}-\frac{R}{N}\right\}\frac{n_j}{N}\right]$$

2. Find a sequence of numbers $\mu_{S_{nm}}$, which can be the limiting expected value with respect to the truth:

$$\mu_{S_{nm}} = \frac{H'_m(H_m^{-1}(J_\pi(\alpha)))}{J_\pi(\alpha)[1-J_\pi(\alpha)]}\left[N\sum_j d_{jm}\{Ep_j - Ep\}g_j^{(cc)}\right]$$

   where $Ep_j = P[\text{disease}|j \text{ alleles, sampled}]$ and $Ep = P[\text{disease}|\text{sampled}]$.

3. Take the derivative with respect to β, and evaluate at $\beta = 0$:

$$\left.\frac{\partial}{\partial\beta}\mu_{S_{nm}}\right|_{\beta=0} = \frac{H'_m(H_m^{-1}(J_\pi(\alpha)))}{J_\pi(\alpha)[1-J_\pi(\alpha)]}\left[N\sum_j d_{jm}\left(\left.\frac{\partial}{\partial\beta}\{Ep_j-Ep\}\right|_{\beta=0} g_j^{(cc)}\right)\right] \text{ because } \{Ep_j-Ep\}=0 \text{ when } \beta=0$$

4. Calculate $\left.\frac{\partial}{\partial\beta}\{Ep_j-Ep\}\right|_{\beta=0}$:

$$\left.\frac{\partial}{\partial\beta}\{Ep_j-Ep\}\right|_{\beta=0} = \frac{\pi L'(\alpha)\{d_{jL}-\sum_l d_{lL}g_l^{(cc)}\}}{\{\pi L(\alpha)+[1-L(\alpha)]\}^2}$$

5. Substitute expression for $\left.\frac{\partial}{\partial\beta}\{Ep_j-Ep\}\right|_{\beta=0}$ to get $\left.\frac{\partial}{\partial\beta}\mu_{S_{nm}}\right|_{\beta=0}$:

$$\left.\frac{\partial}{\partial\beta}\mu_{S_{nm}}\right|_{\beta=0} = \frac{H'_m(H_m^{-1}(J_\pi(\alpha)))}{J_\pi(\alpha)[1-J_\pi(\alpha)]}N\frac{\pi L'(\alpha)}{\{\pi L(\alpha)+[1-L(\alpha)]\}^2}\left[\sum_j d_{jm}g_j^{(cc)}\{d_{jL}-\sum_l d_{lL}g_l^{(cc)}\}\right]$$

6. For the sequence $\sigma^2_{S_{nm}}$ use variance under null:

$$\sigma^2_{S_{nm}} = \frac{[H'_m(H_m^{-1}(J_\pi(\alpha)))]^2}{J_\pi(\alpha)[1-J_\pi(\alpha)]}N\left[\sum_j d_{jm}^2 g_j^{(cc)} - (\sum_j d_{jm}g_j^{(cc)})^2\right]$$

7. Calculate two efficacies, one using statistic m and the other using $\tilde{m}$, and summing over strata:

$$e_{mL} = \frac{\left[\sum_i \left(N_i \frac{H'_m(H_m^{-1}(J_\pi(\alpha_i)))L'(\alpha_i)}{L(\alpha_i)[1-L(\alpha_i)]} \sum_j d_{jm} g_{ij}^{(cc)} \{d_{jL} - \sum_l d_{lL} g_{il}^{(cc)}\}\right)\right]^2}{\sum_i \left(N_i \left[H'_m(H_m^{-1}(J_\pi(\alpha_i)))\right]^2 \frac{\{\pi L(\alpha_i)+1-L(\alpha_i)\}^2}{\pi L(\alpha_i)[1-L(\alpha_i)]} \left[\sum_j d_{jm}^2 g_{ij}^{(cc)} - (\sum_j d_{jm} g_{ij}^{(cc)})^2\right]\right)}$$

$$e_{\tilde{m}L} = \frac{\left[\sum_i \left(N_i \frac{H'_{\tilde{m}}(H_{\tilde{m}}^{-1}(J_\pi(\alpha_i)))L'(\alpha_i)}{L(\alpha_i)[1-L(\alpha_i)]} \sum_j d_{j\tilde{m}} g_{ij}^{(cc)} \{d_{jL} - \sum_l d_{lL} g_{il}^{(cc)}\}\right)\right]^2}{\sum_i \left(N_i \left[H'_{\tilde{m}}(H_{\tilde{m}}^{-1}(J_\pi(\alpha_i)))\right]^2 \frac{\{\pi L(\alpha_i)+1-L(\alpha_i)\}^2}{\pi L(\alpha_i)[1-L(\alpha_i)]} \left[\sum_j d_{j\tilde{m}}^2 g_{ij}^{(cc)} - (\sum_j d_{j\tilde{m}} g_{ij}^{(cc)})^2\right]\right)}$$

8. Use the efficacies to calculate the ARE of statistic m versus $\tilde{m}$:

$$ARE(m,\tilde{m}) = \frac{\left[\sum_i v_i \frac{H'_m(H_m^{-1}(J_\pi(\alpha_i)))L'(\alpha_i)}{L(\alpha_i)[1-L(\alpha_i)]} \left[\sum_j d_{jm} d_{jL} g_{ij}^{(cc)} - \sum_j d_{jm} g_{ij}^{(cc)} \sum_j d_{jL} g_{ij}^{(cc)}\right]\right]^2}{\left[\sum_i v_i \frac{H'_{\tilde{m}}(H_{\tilde{m}}^{-1}(J_\pi(\alpha_i)))L'(\alpha_i)}{L(\alpha_i)[1-L(\alpha_i)]} \left[\sum_j d_{j\tilde{m}} d_{jL} g_{ij}^{(cc)} - \sum_j d_{j\tilde{m}} g_{ij}^{(cc)} \sum_j d_{jL} g_{ij}^{(cc)}\right]\right]^2}$$
$$\times \frac{\sum_i v_i \left[H'_{\tilde{m}}(H_{\tilde{m}}^{-1}(J_\pi(\alpha_i)))\right]^2 \frac{1}{J_\pi(\alpha_i)[1-J_\pi(\alpha_i)]} \left[\sum_j d_{j\tilde{m}}^2 g_{ij}^{(cc)} - (\sum_j d_{j\tilde{m}} g_{ij}^{(cc)})^2\right]}{\sum_i v_i \left[H'_m(H_m^{-1}(J_\pi(\alpha_i)))\right]^2 \frac{1}{J_\pi(\alpha_i)[1-J_\pi(\alpha_i)]} \left[\sum_j d_{jm}^2 g_{ij}^{(cc)} - (\sum_j d_{jm} g_{ij}^{(cc)})^2\right]}$$

References

[1] Setiawan VW, Doherty JA, Shu X, Akbari MR, Chen C, DeVivo I, DeMichele A, Garcia-Closas M, Goodman MT, Haiman CA, Hankinson SE, Henderson BE, Horn-Ross PL, Lacey JV, LeMarchand L, Levine DA, Liang X, Lissowska J, Lurie G, McGrath M, Narod SA, Rebbeck TR, Ursin G, Weiss NS, Xiang Y-B, Yang HP, Zheng W, Olson SH: Two Estrogen-Related Variants in CYP19A1 and Endometrial Cancer Risk: A Pooled Analysis in the Epidemiology of Endometrial Cancer Consortium. *Cancer Epidemiol Biomarkers Prev* 2009; 18:242–247.

[2] GE Noether. On a Theorem of Pitman. *Ann Math Stat*, 26(1):64–68, 1955.

[3] RH Randles and DA Wolfe. *Introduction to the Theory of Nonparametric Statistics*. John Wiley, New York, 1979.

Supplementary Section C Tables.

Supplementary Section D Figures.

Supplementary Figure 1: Asymptotic Relative Efficiency (ARE) for stratified tests using cohort data in the presence of confounding, comparing the statistic using the additive model to one using the true model underlying the data (low disease frequency, low allele frequency variance). ARE is calculated using equation (4). There are six strata with equal distribution of subjects across strata, stratum-specific population disease frequencies of 0.01, 0.025, 0.0375, 0.0625, 0.075, and 0.1, and stratum-specific high-risk allele frequencies as follows. Subfigure (a) uses constant high-risk allele frequencies across strata, while all other subfigures use stratum-specific high-risk allele frequencies with mean value displayed on the x-axis and variance across strata of 0.525. High-risk allele frequencies have arbitrary association with disease frequencies in subfigure (b); monotone positive association in subfigure (c); and monotone negative association in subfigure (d).

Supplementary Figure 2: Asymptotic Relative Efficiency (ARE) for stratified tests using case-control data in the presence of confounding, comparing the statistic using the additive model to one using the true model underlying the data (low disease frequency, low allele frequency variance). ARE is calculated using equation (5). The sampling ratio π is 100, meaning that cases are 100 times more likely to be sampled than controls. There are six strata with equal distribution of subjects across strata, stratum-specific population disease frequencies of 0.01, 0.025, 0.0375, 0.0625, 0.075, and 0.1, and stratum-specific high-risk allele frequencies as follows. Subfigure (a) uses constant high-risk allele frequencies across strata, while all other subfigures use stratum-specific high-risk allele frequencies with mean value displayed on the x-axis and variance across strata of 0.525. High-risk allele frequencies have arbitrary association with disease frequencies in subfigure (b); monotone positive association in subfigure (c); and monotone negative association in subfigure (d).

Supplementary Figure 3: Asymptotic Relative Efficiency (ARE) for stratified tests using cohort data in the presence of confounding, comparing the statistic using the additive model to one using the true model underlying the data (high disease frequency, high allele frequency variance). ARE is calculated using equation (4). There are six strata with equal distribution of subjects across strata, stratum-specific population disease frequencies of 0.25, 0.375, 0.5, 0.75, 0.875, and 0.95, and stratum-specific high-risk allele frequencies as follows. Subfigure (a) uses constant high-risk allele frequencies across strata, while all other subfigures use stratum-specific high-risk allele frequencies with mean value displayed on the x-axis and variance across strata of 2.1. High-risk allele frequencies have arbitrary association with disease frequencies in subfigure (b); monotone positive association in subfigure (c); and monotone negative association in subfigure (d).

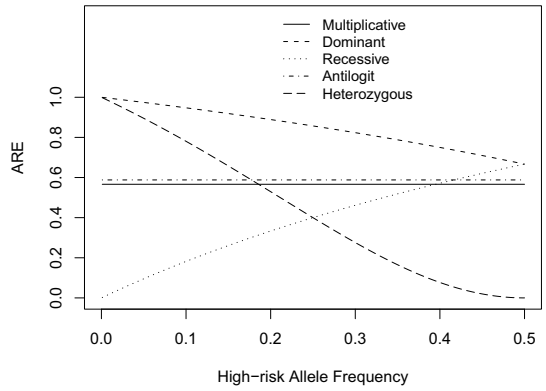
Supplementary Figure 4: Asymptotic Relative Efficiency (ARE) for stratified tests using case-control data in the presence of confounding, comparing the statistic using the additive model to one using the true model underlying the data (high disease frequency, high allele frequency variance). ARE is calculated using equation (5). The sampling ratio π is 100, meaning that cases are 100 times more likely to be sampled than controls. There are six strata with equal distribution of subjects across strata, stratum-specific population disease frequencies of 0.25, 0.375, 0.5, 0.75, 0.875, and 0.95, and stratum-specific high-risk allele frequencies as follows. Subfigure (a) uses constant high-risk allele frequencies across strata, while all other subfigures use stratum-specific high-risk allele frequencies with mean value displayed on the x-axis and variance across strata of 2.1. High-risk allele frequencies have arbitrary association with disease frequencies in subfigure (b); monotone positive association in subfigure (c); and monotone negative association in subfigure (d).

Supplementary Figure 5: Asymptotic Relative Efficiency (ARE) for stratified tests using cohort data in the presence of confounding, comparing the statistic using the additive model to one using the true model underlying the data (high disease frequency, low allele frequency variance). ARE is calculated using equation (4). There are six strata with equal distribution of subjects across strata, stratum-specific population disease frequencies of 0.25, 0.375, 0.5, 0.75, 0.875, and 0.95, and stratum-specific high-risk allele frequencies as follows. Subfigure (a) uses constant high-risk allele frequencies across strata, while all other subfigures use stratum-specific high-risk allele frequencies with mean value displayed on the x-axis and variance across strata of 0.525. High-risk allele frequencies have arbitrary association with disease frequencies in subfigure (b); monotone positive association in subfigure (c); and monotone negative association in subfigure (d).

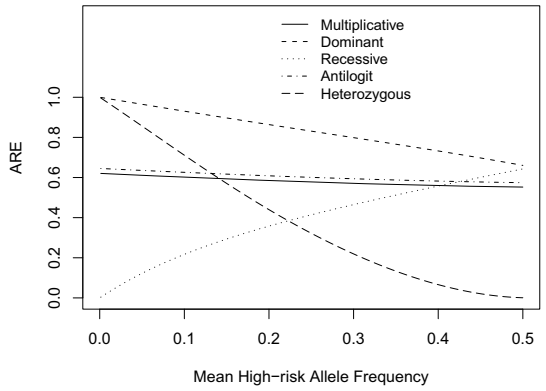
Supplementary Figure 6: Asymptotic Relative Efficiency (ARE) for stratified tests using case-control data in the presence of confounding, comparing the statistic using the additive model to one using the true model underlying the data (high disease frequency, low allele frequency variance). ARE is calculated using equation (5). The sampling ratio π is 100, meaning that cases are 100 times more likely to be sampled than controls. There are six strata with equal distribution of subjects across strata, stratum-specific population disease frequencies of 0.25, 0.375, 0.5, 0.75, 0.875, and 0.95, and stratum-specific high-risk allele frequencies as follows. Subfigure (a) uses constant high-risk allele frequencies across strata, while all other subfigures use stratum-specific high-risk allele frequencies with mean value displayed on the x-axis and variance across strata of 0.525. High-risk allele frequencies have arbitrary association with disease frequencies in subfigure (b); monotone positive association in subfigure (c); and monotone negative association in subfigure (d).

Supplementary Figure 7: Asymptotic Relative Efficiency (ARE) for stratified tests, comparing two statistics both different from the true model underlying the data. ARE is calculated using equation (4) for cohort data with stratum-specific population disease frequencies of 0.01, 0.025, 0.0375, 0.0625, 0.075, and 0.1. There are six strata with equal distribution of subjects across strata and constant high-risk allele frequencies across strata.

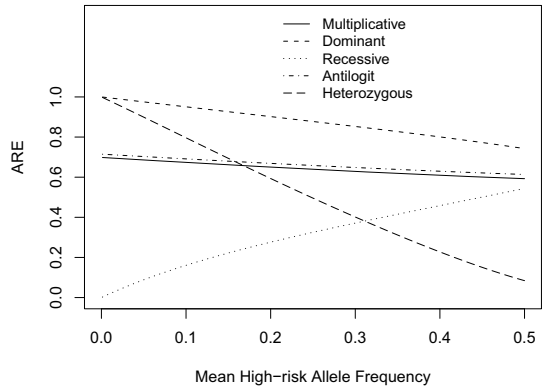
Supplementary Figure 8: Asymptotic Relative Efficiency (ARE) for stratified tests, comparing two statistics both different from the true model underlying the data. ARE is calculated using equation (5) for case-control data with stratum-specific population disease frequencies of 0.01, 0.025, 0.0375, 0.0625, 0.075, and 0.1. There are six strata with equal distribution of subjects across strata and constant high-risk allele frequencies across strata. Truth is assumed at the population level, and the sampling ratio π is 100, meaning that cases are 100 times more likely to be sampled than controls.



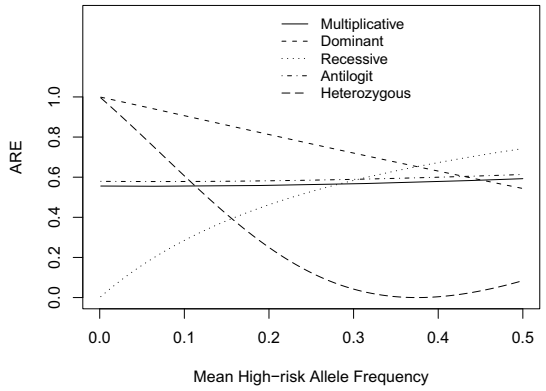
(a)



(b)

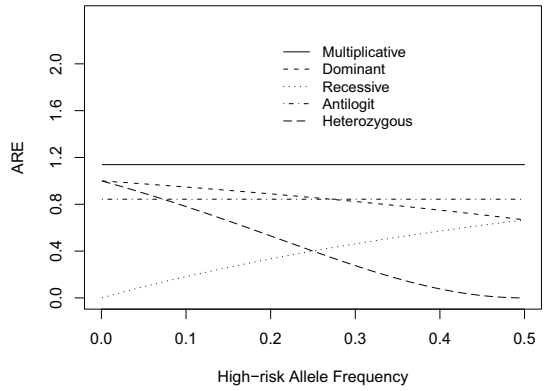


(c)

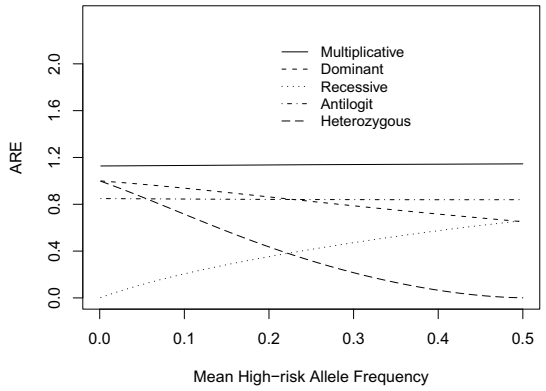


(d)

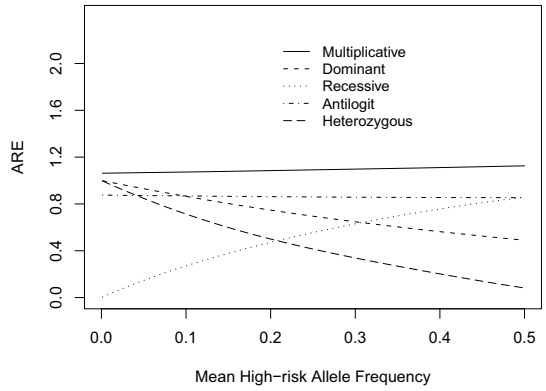
Supplementary Figure 1



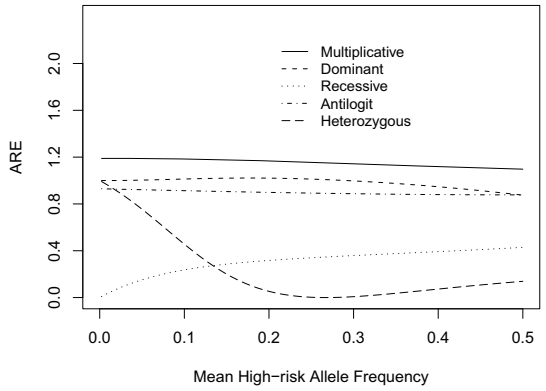
(a)



(b)

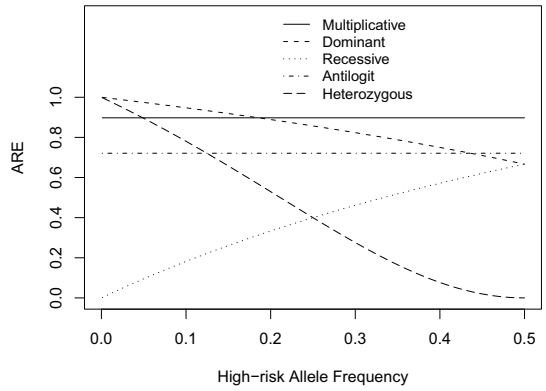


(c)

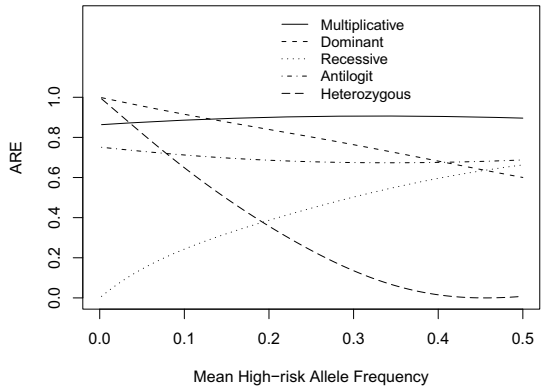


(d)

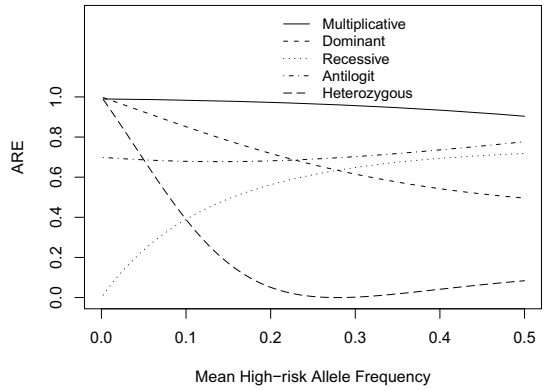
Supplementary Figure 2



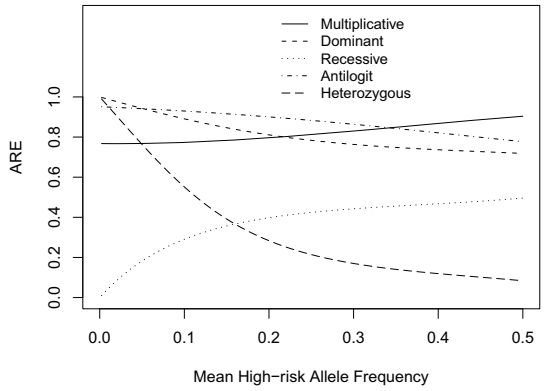
(a)



(b)

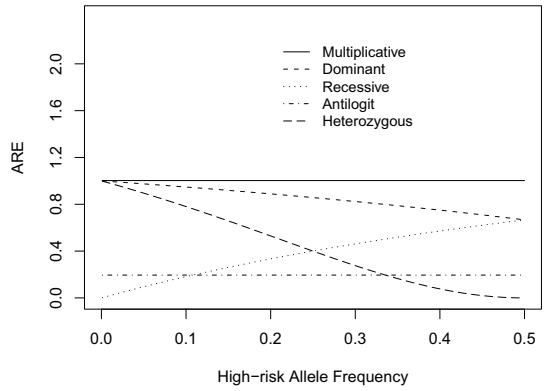


(c)

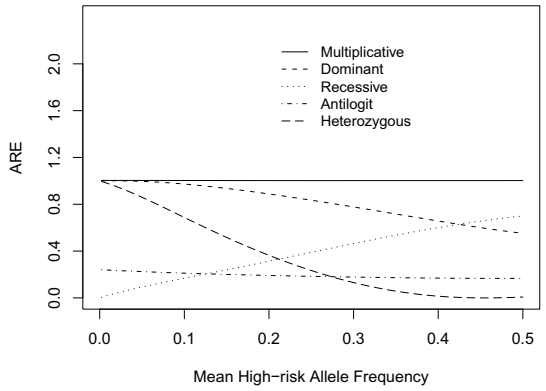


(d)

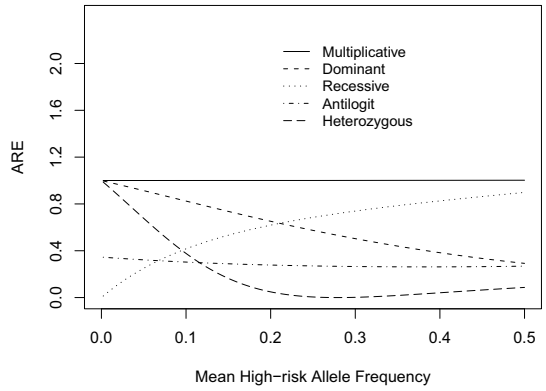
Supplementary Figure 3



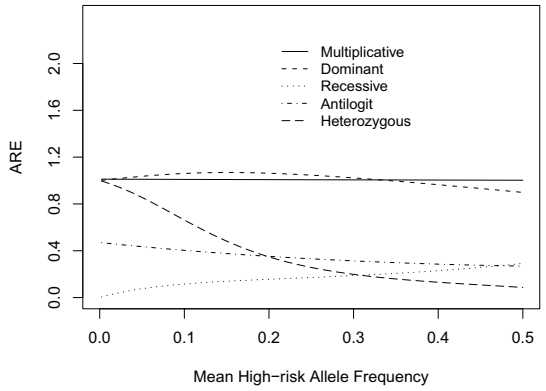
(a)



(b)

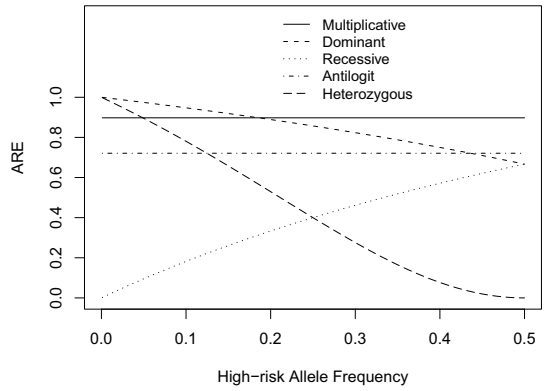


(c)

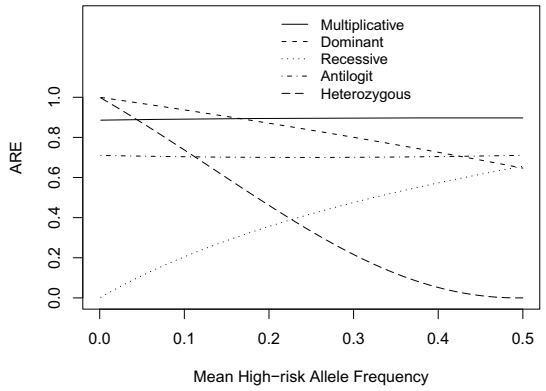


(d)

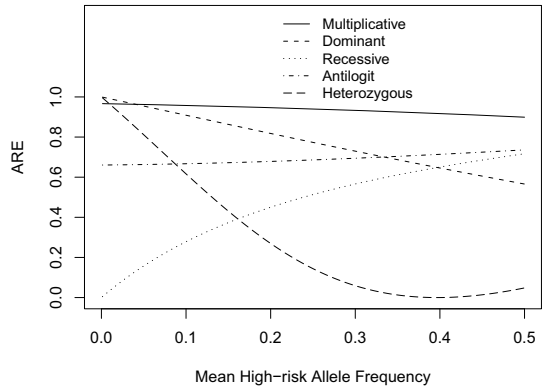
Supplementary Figure 4



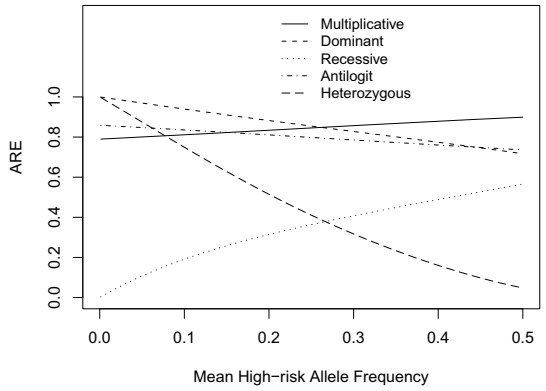
(a)



(b)

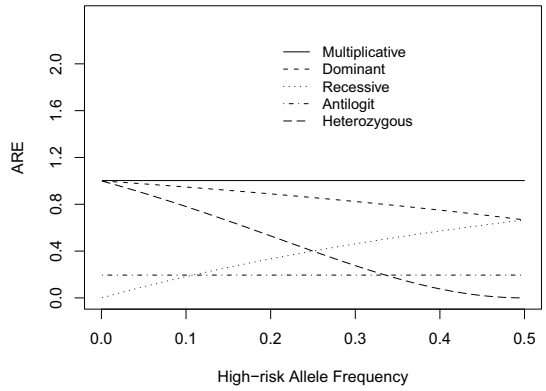


(c)

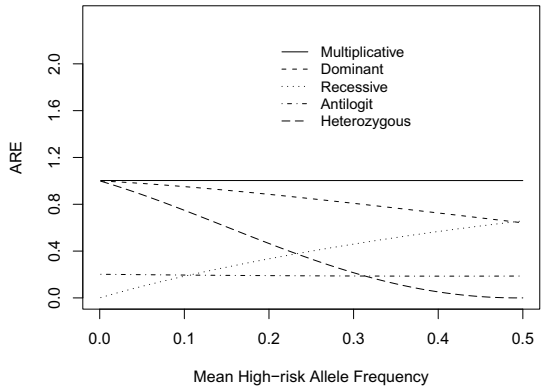


(d)

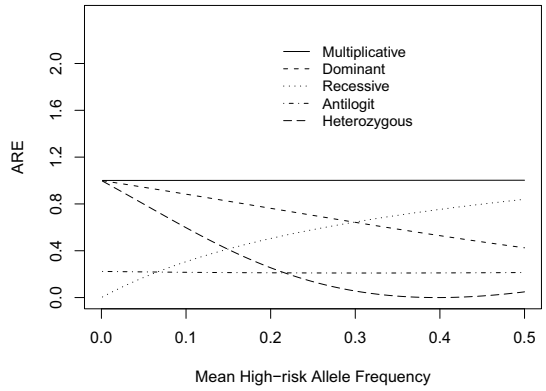
Supplementary Figure 5



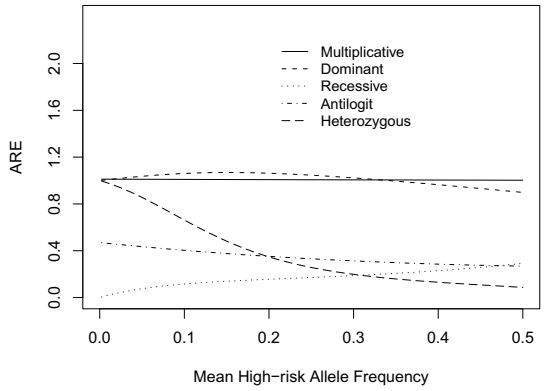
(a)



(b)

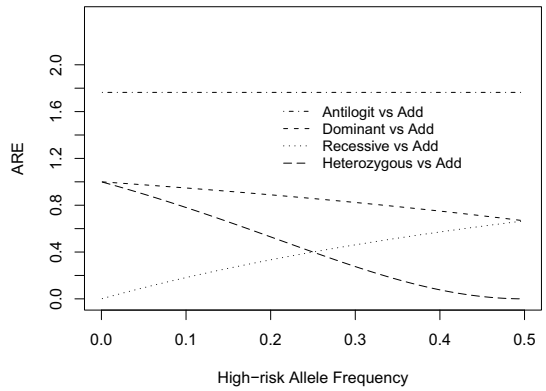


(c)

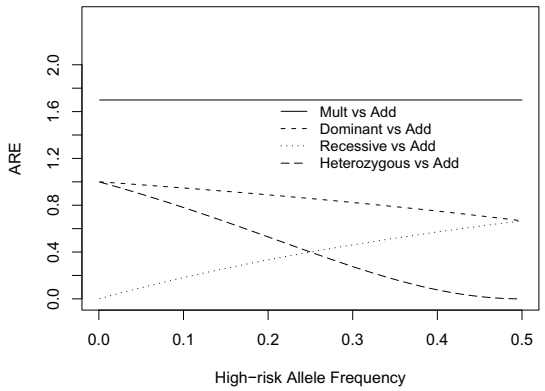


(d)

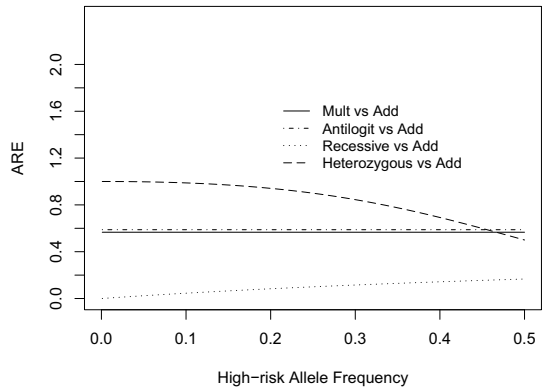
Supplementary Figure 6



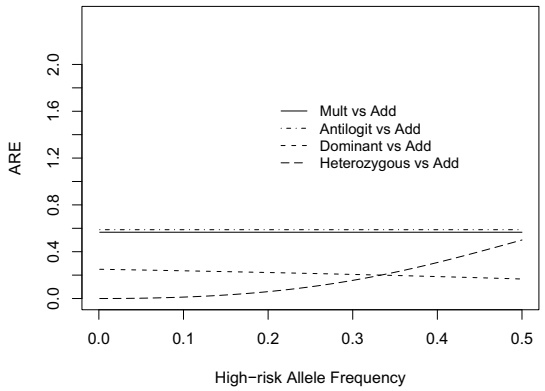
(a) Multiplicative Truth



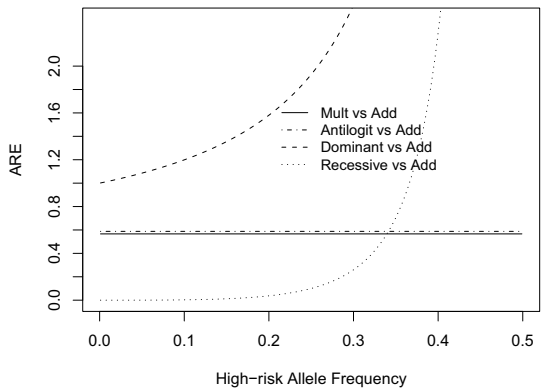
(b) Antilogit Truth



(c) Dominant Truth

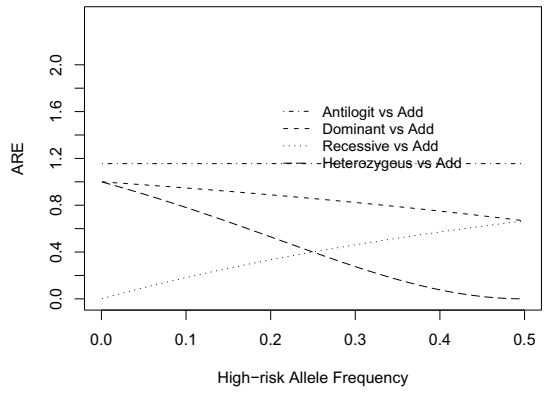


(d) Recessive Truth

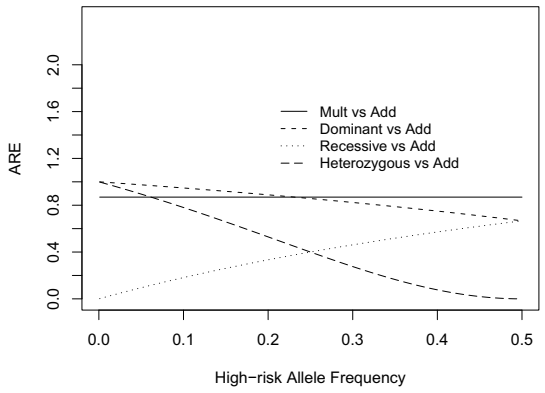


(e) Heterozygous Truth

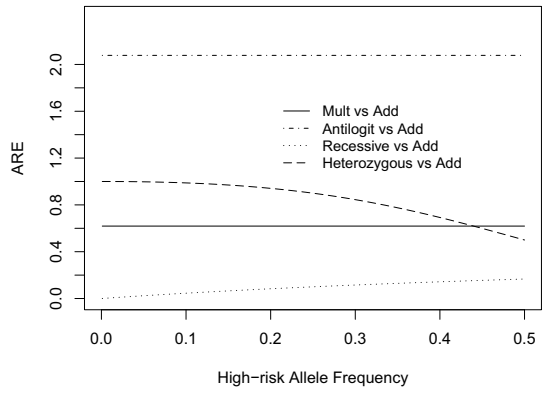
Supplementary Figure 7



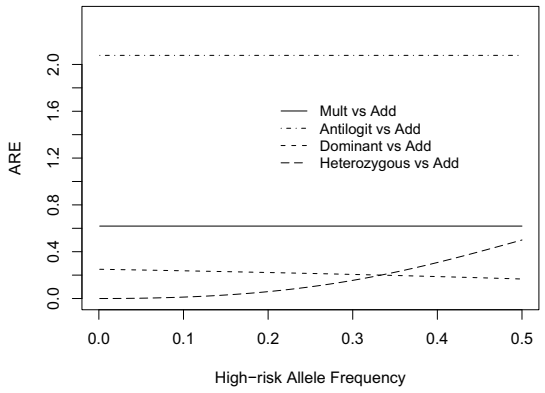
(a) Multiplicative Truth



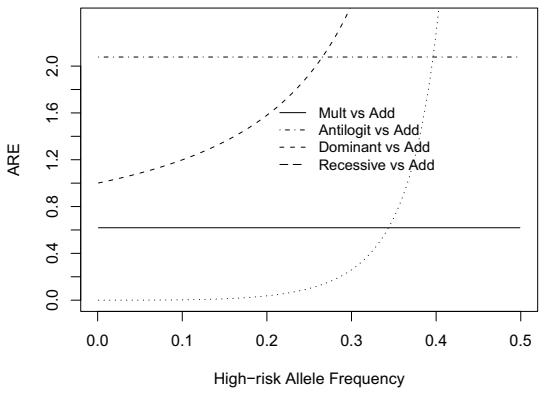
(b) Antilogit Truth



(c) Dominant Truth



(d) Recessive Truth



(e) Heterozygous Truth

Supplementary Figure 8