

When is an endophenotype useful to detect association to a disease? Exploring the relationships between disease status, endophenotype and genetic polymorphisms

Alexandre Bureau^{1,2}, Jordie Croteau²

Received _____; accepted _____

¹Département de médecine sociale et préventive, Université Laval

²Institut universitaire en santé mentale de Québec du Centre intégré de santé et de services sociaux de la Capitale-Nationale

Supplementary tables

Table S1: Values of the coefficients of linear models of the population risk allele frequency at locus 2 for polytomous phenotype scenario P2.

	pRAF ^a	Model conditional on X_1				Model ignoring X_1		
		Intercept	Y_2	X_1Y_1	$X_1Y_1Y_2$	Intercept	Y_2	Y_1Y_2
Unrelated ^b	0.1	0.100	0.081	-0.026	0.149	0.099	0.082	0.013
	0.3	0.299	0.161	-0.063	0.241	0.298	0.162	0.023
	0.5	0.498	0.166	-0.079	0.225	0.497	0.168	0.022
Ascertained families ^c	0.3	0.389	0.094	-0.077	0.207	0.387	0.104	0.007

^a Population risk allele frequency

^b Least square fit to expected allele frequency at locus 2 for each combination of explanatory variables

^c Least square fit to empirical allele frequency at locus 2 for each combination of explanatory variables in data simulated under the familial dependence simulation model with $\sigma_2 = 1.0$, $\alpha = \log(4) = 1.386$ and $\nu = 0.01$ in 3-generation 16-members family depicted in Figure 2 ascertained for at least a cousin pair affected by the disease

Table S2: Values of the coefficients of linear models of the population risk allele frequency at locus 2 for polytomous phenotype scenario P3.

	pRAF ^a	Model conditional on X_1				Model ignoring X_1		
		Intercept	X_1	X_1Y_1	$X_1Y_1Y_2$	Intercept	Y_1	Y_1Y_2
Unrelated ^b	0.1	0.100	0.004	-0.029	0.106	0.100	-0.004	0.018
	0.3	0.300	0.009	-0.070	0.228	0.298	-0.008	0.043
	0.5	0.500	0.010	-0.086	0.254	0.501	-0.010	0.052
Ascertained families ^c	0.3	0.306	0.025	-0.066	0.181	0.311	-0.010	0.035

^a Population risk allele frequency

^b Least square fit to expected allele frequency at locus 2 for each combination of explanatory variables

^c Least square fit to empirical allele frequency at locus 2 for each combination of explanatory variables in data simulated under the familial dependence simulation model with $\sigma_2 = 1.0$, $\alpha = \log(4) = 1.386$ and $\nu = 0.01$ in 3-generation 16-members family depicted in Figure 2 ascertained for at least a cousin pair affected by the disease

Table S3: Values of the coefficients of linear models of the population risk allele frequency at locus 2 for polytomous phenotype scenario P4.

	pRAF ^a	Model conditional on X_1				Model ignoring X_1		
		Intercept	Y_1	X_1Y_1	$X_1Y_1Y_2$	Intercept	Y_1	Y_1Y_2
Unrelated ^b	0.1	0.096	0.034	-0.033	0.129	0.096	0.030	0.014
	0.3	0.290	0.075	-0.069	0.246	0.290	0.067	0.030
	0.5	0.487	0.085	-0.075	0.246	0.487	0.076	0.033
Ascertained families ^c	0.3	0.342	0.063	-0.044	0.193	0.342	0.061	0.017

^a Population risk allele frequency

^b Least square fit to expected allele frequency at locus 2 for each combination of explanatory variables

^c Least square fit to empirical allele frequency at locus 2 for each combination of explanatory variables in data simulated under the familial dependence simulation model with $\sigma_2 = 1.0$, $\alpha = \log(4) = 1.386$ and $\nu = 0.01$ in 3-generation 16-members family depicted in Figure 2 ascertained for at least a cousin pair affected by the disease

Table S4: Values of the coefficients of linear models of the population risk allele frequency at locus 2 for polytomous phenotype scenario P5.

	pRAF ^a	Model conditional on X_1				Model ignoring X_1	
		Intercept	X_1	$X_1Y_2^*$ ^b	$X_1Y_1Y_2^*$	Intercept	Y_2
Unrelated ^c	0.1	0.100	0.003	0.175	-0.029	0.100	0.020
	0.3	0.300	0.006	0.324	-0.070	0.300	0.053
	0.5	0.500	0.006	0.318	-0.086	0.499	0.067
Ascertained families ^d	0.3	0.315	0.072	0.217	-0.091	0.324	0.037

^a Population risk allele frequency

^b $Y_2^* = 1 - Y_2$

^c Least square fit to expected allele frequency at locus 2 for each combination of explanatory variables

^d Least square fit to empirical allele frequency at locus 2 for each combination of explanatory variables in data simulated under the familial dependence simulation model with $\sigma_2 = 1.0$, $\alpha = \log(4) = 1.386$ and $\nu = 0.01$ in 3-generation 16-members family depicted in Figure 2 ascertained for at least a cousin pair affected by the disease

Table S5: Values of the coefficients of linear models of the population risk allele frequency at locus 2 for transition scenario T2.

	pRAF ^a	Model conditional on X_1			Model ignoring X_1	
		Intercept	X_1	X_1Y_2	Intercept	Y_2
Unrelated ^b	0.1	0.100	-0.001	0.074	0.100	0.008
	0.3	0.300	-0.003	0.160	0.300	0.019
	0.5	0.500	-0.003	0.170	0.501	0.023
Ascertained families ^c	0.3	0.307	0.027	0.109	0.311	0.015

^a Population risk allele frequency

^b Least square fit to expected allele frequency at locus 2 for each combination of explanatory variables

^c Least square fit to empirical allele frequency at locus 2 for each combination of explanatory variables in data simulated under the familial dependence simulation model with $\sigma_2 = 1.0$, $\alpha = \log(4) = 1.386$ and $\nu = 0.01$ in 3-generation 16-members family depicted in Figure 2 ascertained for at least a cousin pair affected by the disease

Table S6: Values of the coefficients of linear models of the population risk allele frequency at locus 2 for transition scenario T3.

	pRAF ^a	Model conditional on X_1				Model ignoring X_1		
		Intercept	Y_1Y_2	X_1Y_1	$X_1Y_1Y_2$	Intercept	Y_1	Y_1Y_2
Unrelated ^b	0.1	0.100	0.036	-0.006	0.100	0.100	-0.002	0.051
	0.3	0.300	0.080	-0.016	0.180	0.300	-0.004	0.111
	0.5	0.500	0.090	-0.017	0.180	0.500	-0.006	0.123
Ascertained families ^c	0.3	0.334	0.053	-0.060	0.223	0.336	-0.021	0.107

^a Population risk allele frequency

^b Least square fit to expected allele frequency at locus 2 for each combination of explanatory variables

^c Least square fit to empirical allele frequency at locus 2 for each combination of explanatory variables in data simulated under the familial dependence simulation model with $\sigma_2 = 1.0$, $\alpha = \log(4) = 1.386$ and $\nu = 0.01$ in 3-generation 16-members family depicted in Figure 2 ascertained for at least a cousin pair affected by the disease