Joint modeling of linkage and association for binary traits

J Concepción Loredo-Osti

joint work with

Haiyan Yang and Michael Woods

Memorial University
St. John's, Newfoundland and Labrador, Canada

jcloredoosti@mun.ca

August 7, 2018

MEMORIAL UNIVERSITY

Linkage and association

- Linkage and association methods are widely used in genetic analysis.
- Linkage describes a relationship between phenotype and loci while association describes a relationship between phenotype and alleles.
- Linkage is a consequence of co-segregation, a fundamental genetic process.
- Association is plainly a statistical statement about co-occurrence of alleles and phenotype, embellished with some population genetics arguments that most of times are of no use in the analysis.

MEMORIAL UNIVERSITY

- Linkage is a phenomenon to be studies within families.
- However, when two or more supposedly unrelated individuals from a particular population share the same phenotype, it is natural to think that the population genetic processes have affected all the members of the population and not only unrelated singletons.
- On the other hand, association can be studied not only in families but is unrelated individuals as well.

- This indicates that both linkage and association can be modelled together and, in fact, there are some methods that use both linkage and association (TDT, FBAT).
- Of course, joint modelling of linkage and association does not consist of performing a linkage study followed by an association study with case-controls (or any other design) and a denser set of markers on the regions where 'a linkage peak' was found.
- A 'good' methodology takes int consideration both linkage and association in the development of estimation and testing procedures and provides some flexibility to deal with general family configurations and 'unrelated' individuals.

MEMORIAL UNIVERSITY

Logistic model with over-dispersion

- Suppose that we have a random sample $y_1, y_2, \ldots, y_T$, with $y_t \in \{0, 1\},\ t = 1, 2, \ldots, N$ and each y_t has an associated vector of covariates $\{\boldsymbol{z}_t \in \mathbb{R}^q,\ t = 1, 2, \ldots, T\}$. Additionally, assume that for each t there is an independent random variable $\xi_t \in \mathbb{R},\ t = 1, 2, \ldots, T$ such that given ξ_t and $\boldsymbol{z}_t$, the random variable y_t has a Bernoulli distribution with parameter $\pi_t(\xi_t)$ where

$$\begin{aligned}\pi_t(\xi_t) &= \Pr\left(y_t = 1 \mid \boldsymbol{z}_t, \xi_t\right) \\ &= \frac{\mathrm{e}^{\boldsymbol{\theta}'\boldsymbol{z}_t + \xi_t}}{1 + \mathrm{e}^{\boldsymbol{\theta}'\boldsymbol{z}_t + \xi_t}} \qquad \text{for } t = 1, 2, \ldots, T.\end{aligned}$$

- Define $\zeta_t = e^{\boldsymbol{\theta}' \boldsymbol{z}_t}$, $\zeta_t(\xi) = \mathrm{e}^z \, \zeta_t$ and $\lambda_t(\xi) = (1 + \zeta_t(\xi))^{-1}$ so that $\pi_t(\xi) = \zeta_t(\xi) \, \lambda_t(\xi)$. Thus, if every ξ_t has a density parameterized by $\boldsymbol{\vartheta}$, say $\mathrm{f}_{\boldsymbol{\vartheta}}(\cdot)$, for $t = 1, 2, \ldots, T$

$$\Pr(y_t \mid \boldsymbol{z}_t, \xi) = \zeta_t^{y_t}(\xi) \, \lambda_t(z) \tag{1}$$

$$\begin{aligned} \Pr(y_t \mid \boldsymbol{z}_t) &= \int_{-\infty}^{\infty} \Pr(y_t \mid \boldsymbol{z}_t, \xi) \, \mathrm{f}_{\boldsymbol{\vartheta}}(\xi) \, d\xi \qquad (2) \\ &= y_t + (-1)^{y_t} \int_{-\infty}^{\infty} \lambda_t(\xi) \, \mathrm{f}_{\boldsymbol{\vartheta}}(\xi) \, d\xi \\ &= \pi_t, \end{aligned}$$

i.e., π_t is a function of y_t and $\boldsymbol{z}_t$ only.

Consequently, the conditional on $\{\boldsymbol{z}_t\}$,

$$\Pr(\boldsymbol{y} \mid \{\boldsymbol{z}_t\}) = \prod_{t=1}^{T} \pi_t \tag{3}$$

and the log-likelihood can be written as

$$\ell(\boldsymbol{\theta}, \boldsymbol{\vartheta}) = \sum_{t=1}^{T} \log \pi_t$$

with $\boldsymbol{\theta}$ and $\boldsymbol{\vartheta}$ being the parameters of the logistic model and over-dispersion distribution, respectively.

- Under mild assumptions regarding $\mathrm{f}_{\boldsymbol{\vartheta}}()$,

$$\frac{\partial \ell(\boldsymbol{\theta}, \boldsymbol{\vartheta})}{\partial \boldsymbol{\theta}} = \sum_{t=1}^{T} \boldsymbol{z}_t \left(y_t - \mathrm{E}_{\xi_t}\left(\pi_t(\xi_t) \,\middle|\, \boldsymbol{z}_t, y_t \right) \right) \tag{4}$$

$$= \sum_{t=1}^{T} \boldsymbol{z}_t \left(\frac{y_t - \mu_t}{\mu_t(1-\mu_t)} \right) \mathrm{E}_{\xi_t}\left(\mathrm{Var}(y_t \mid \boldsymbol{z}_t, \xi_t) \right)$$

where $\mu_t = \mathrm{E}\left(y_t \mid \boldsymbol{z}_t\right)$ and $\mathrm{Var}(y_t \mid \boldsymbol{z}_t, \xi_t) = \pi_t(\xi_t)\big(1 - \pi_t(\xi_t)\big)$, and

$$\frac{\partial \ell(\boldsymbol{\theta}, \boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}} = \sum_{t=1}^{T} \mathrm{E}_{\xi_t} \left(\frac{\partial \mathrm{f}_{\boldsymbol{\vartheta}}(\xi_t)}{\partial \boldsymbol{\vartheta}} \,\middle|\, \boldsymbol{z}_t, y_t \right) \tag{5}$$

- To evaluate the expectations involved in the previous expressions we can use the following relationship. For $r = -1, 0, 1, 2, \ldots$, define

$$P_t^{(r)} = \int_{-\infty}^{\infty} \pi_t^{r+1}(\xi)\, \mathrm{f}_{\vartheta}(\xi)\, dz,$$

so that for $r = 0, 1, 2, \ldots$,

$$\pi_t\, \mathrm{E}_{\xi_t}\left(\pi_t^r(\xi_t) \,\middle|\, \boldsymbol{z}_t, y_t\right) = (1 - y_t)\, P_t^{(r-1)} + (-1)^{1-y_t}\, P_t^{(r)}$$

and

$$\mu_t = P_t^{(0)}$$

as well as

$$\mathrm{E}_{\xi_t}\left(\mathrm{Var}(y_t \,|\, \boldsymbol{z}_t, \xi_t)\right) = P_t^{(0)} - P_t^{(1)}$$

Normal random effects

- If we assume that $\mathrm{f}_{\boldsymbol{\vartheta}}(\cdot)$ is the normal distribution with null mean and variance $\sigma^2 > 0$ variance. Then

$$\frac{\partial \ell(\boldsymbol{\theta}, \sigma^2)}{\partial \sigma^2} = -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{t=1}^{T} \mathrm{E}_{\xi_t} \left(\xi_t^2 \mid \boldsymbol{z}_t, y_t \right)$$

$$= -\frac{1}{2\sigma^2} \sum_{t=1}^{T} \frac{y_t - \mu_t}{1 - \mu_t} + \frac{1}{2\sigma^4} \sum_{t=1}^{T} \frac{y_t - \mu_t}{\mu_t(1 - \mu_t)} \mathrm{E}_{\xi_t} \left(\xi_t^2 \, \pi_t(\xi_t) \right)$$

and the remaining components of the Fisher information matrix can be obtained in a similar way.

A model for association

- To accommodate for more than one family, lets replace the sub-index 't' in y_t, $\boldsymbol{z}_t$, ξ_t and π_t by 'ij', to represent the jth individual from the ith family, $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, n_i$.
- Furthermore, assume that $\boldsymbol{z}_{ij} = (1, g_{ij}, \boldsymbol{x}'_{ij})' \in \mathbb{R}^{S+2}$ with g_{ij} being the marker phenotype and $\boldsymbol{x}_{ij}$ a set of covariates and, as before $y_{ij} \in \{0, 1\}$.
- Thus

$$\ell(\boldsymbol{\theta}, \boldsymbol{\vartheta}) = \sum_{i=1}^{N} \sum_{j}^{n_i} \log \pi_{ij}$$

  etc...
- When the ξ_{ij} represent non-genetic over-dispersion, this is a simple model for association.

Mixed model with genetic random effects

- In this case, we assume that the vector of ranom effects for the ith family, $\boldsymbol{\xi}_i$ has null mean and variance $\sigma^2 \boldsymbol{\Sigma}_i$, then

$$\Pr(\boldsymbol{y}_i \mid \{\boldsymbol{z}_{ij}\}, \boldsymbol{\xi}_i) = \prod_{j=1}^{n_i} \eta_{ij}^{y_{ij}}(\xi_{ij})\, \lambda(\xi_{ij})$$

and

$$\Pr(\boldsymbol{y}_i \mid \{\boldsymbol{x}_{ij}\}) = \int_{\mathbb{R}^{n_i}} \Pr(\boldsymbol{y}_i \mid \{\boldsymbol{z}_{ij}\}, \boldsymbol{\xi}_i)\, \mathrm{f}(\boldsymbol{\xi}_i)\, d\boldsymbol{\xi}_i$$

MEMORIAL UNIVERSITY

- Thus, under some regularity conditions,

$$\frac{\partial \Pr(\boldsymbol{y}_i \mid \{\boldsymbol{z}_{ij}\}, \boldsymbol{\xi}_i)}{\partial \boldsymbol{\theta}} = \Pr(\boldsymbol{y} \mid \{\boldsymbol{z}_{ij}\}, \boldsymbol{\xi}_i) \sum_{j=1}^{n_i} \boldsymbol{z}_{ij} \left(y_{ij} - \pi_{ij}(\xi_{ij})\right)$$

and, consequently,

$$\frac{\partial \ell_i(\boldsymbol{\theta}, \sigma^2)}{\partial \boldsymbol{\theta}} = \mathrm{E}_{\boldsymbol{\xi}_i} \left(\sum_{j=1}^{n_i} \boldsymbol{z}_{ij} \left(y_{ij} - \pi_{ij}(\xi_{ij})\right) \,\middle|\, \{\boldsymbol{z}_{ij}\}, \boldsymbol{y}_i \right)$$

- When $\boldsymbol{\xi}_i \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 \boldsymbol{\Sigma}_i\right)$ with $\sigma^2 > 0$, then

$$\frac{\partial \ell_i(\boldsymbol{\theta}, \sigma^2)}{\partial \sigma^2} = -\frac{n_i}{2\sigma^2} + \frac{1}{2\sigma^4} \mathrm{E}_{\boldsymbol{\xi}_i}\left(\boldsymbol{\xi}_i' \boldsymbol{\Sigma}^{-1} \boldsymbol{\xi}_i \,|\, \{\boldsymbol{z_{ij}}\}, \boldsymbol{y}_i\right)$$

- Furthermore, if $\boldsymbol{\Sigma}_i$ can be written as $\boldsymbol{\Sigma}_i = \boldsymbol{I} + \beta \boldsymbol{\Phi}_i$, for some $\beta \geq 0$ and a n.n.d. $\boldsymbol{\Phi}_i$, then

$$\frac{\partial \ell_i(\boldsymbol{\theta}, \sigma^2, \beta)}{\partial \beta} = \frac{1}{2\sigma^2} \mathrm{E}_{\boldsymbol{\xi}_i}\left(\boldsymbol{\xi}_i' \boldsymbol{\Sigma}^{-1} \boldsymbol{\Phi}_i \boldsymbol{\Sigma}^{-1} \boldsymbol{\xi}_i - \sigma^2 \operatorname{tr}\left(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Phi}_i\right) \,|\, \{\boldsymbol{z_{ij}}\}, \boldsymbol{y}_i\right.$$

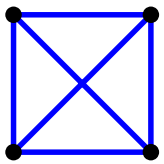
- We can think of $\boldsymbol{\Phi}_i$ as the Malécot kinship matrix at any given locus for the ith family, say $\boldsymbol{\Phi}_i^{\langle m \rangle}$, and β as the signal-to-noise ratio at such a locus.
- Computation of $\varphi_{ij}^{\langle m \rangle}$ involves the estimation of IBD at the mth locus.

IBD state $g_i = (a, b)$	IBD state $g_j = (c, d)$	State description: Partition	State description: Ewens'	Condensed states $\mathcal{S}$
● ■	● ■	(a, b, c, d)	(0,0,0,1)	$\mathcal{S}_1 = (1, 1, 1)$
● ■	● ■	$(a, b)(c, d)$	(0,2,0,0)	$\mathcal{S}_2 = (1, 1, 0)$
● ■	● ■	$(a, b, c)(d)$	(1,0,1,0)	$\mathcal{S}_3 = (1, 0, 1)$
● ■	● ■	$(a, b, d)(c)$	(1,0,1,0)	
● ■	● ■	$(a, b)(c)(d)$	(2,1,0,0)	$\mathcal{S}_4 = (1, 0, 0)$
● ■	● ■	$(a, c, d)(b)$	(1,0,1,0)	$\mathcal{S}_5 = (0, 1, 1)$
● ■	● ■	$(a)(b, c, d)$	(1,0,1,0)	
● ■	● ■	$(a)(b)(c, d)$	(2,1,0,0)	$\mathcal{S}_6 = (0, 1, 0)$
● ■	● ■	$(a, c)(b, d)$	(0,2,0,0)	$\mathcal{S}_7 = (0, 0, 2)$
● ■	● ■	$(a, d)(b, c)$	(0,2,0,0)	
● ■	● ■	$(a, c)(b)(d)$	(2,1,0,0)	$\mathcal{S}_8 = (0, 0, 1)$
● ■	● ■	$(a, d)(b)(c)$	(2,1,0,0)	
● ■	● ■	$(a)(b, c)(d)$	(2,1,0,0)	
● ■	● ■	$(a)(b, d)(c)$	(2,1,0,0)	
● ■	● ■	$(a)(b)(c)(d)$	(4,0,0,0)	$\mathcal{S}_9 = (0, 0, 0)$

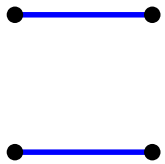
Table. IBD states for a pair of individuals at a given locus.

The kinship coefficient φ_{ij} between the ith and jth individuals can be written as

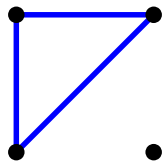
$$\varphi_{ij} = \Delta_{1,ij} + \frac{1}{2}\left(\Delta_{3,ij} + \Delta_{5,ij} + \Delta_{7,ij}\right) + \frac{1}{4}\Delta_{8,ij}$$



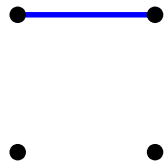
$\mathcal{J}_1$



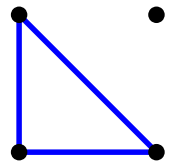
$\mathcal{J}_2$



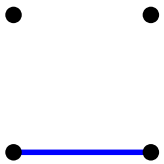
$\mathcal{J}_3$



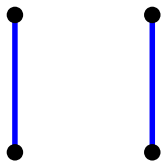
$\mathcal{J}_4$



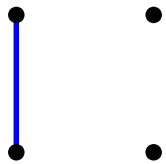
$\mathcal{J}_5$



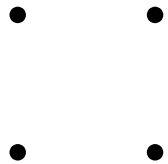
$\mathcal{J}_6$



$\mathcal{J}_7$



$\mathcal{J}_8$



$\mathcal{J}_9$

$$\Delta_r = \Pr(\mathcal{S} = \mathcal{J}_r)$$

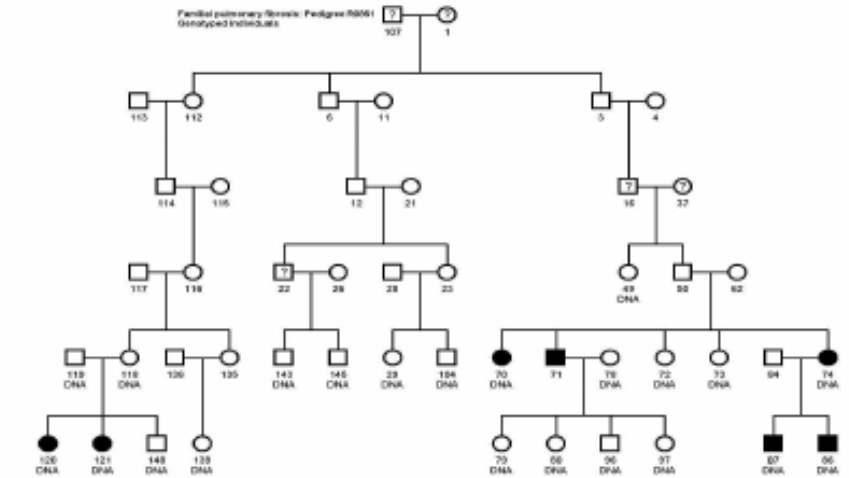
IBD at a particular locus

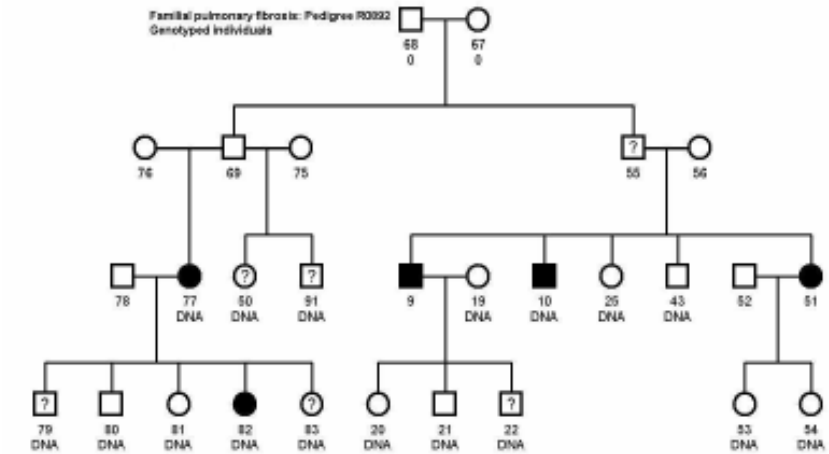
- The random process underlying changes in the patterns of *IBD* across the genome is recombination, so the natural context for estimation of *IBD* in founders and singletons is the ancestral recombination graph, which specifies the complete ancestry of a collection of chromosomes.
- For families, the natural frameworks are *inheritance vectors*, or *descent graphs* and/or some variations of these methodologies some times written in FFT terms.

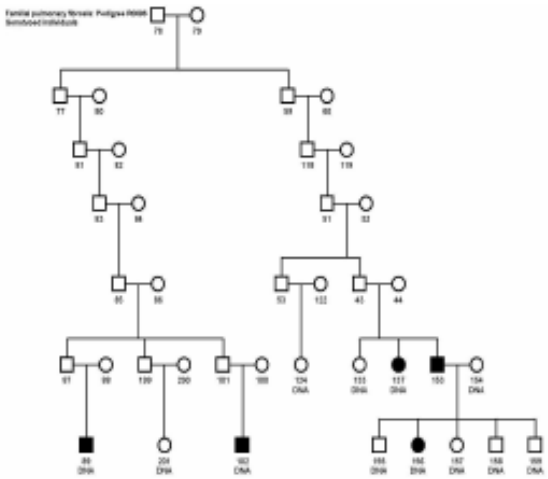
MEMORIAL UNIVERSITY

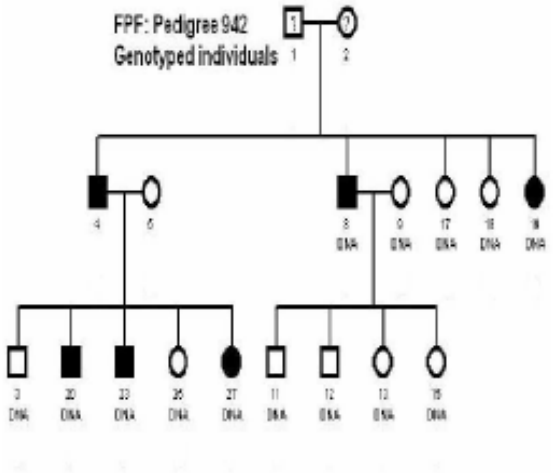
- Abecais & Wigginton proposed a generalization of the Green & Lander algorithm to account for LD. However, this method does not scale well even for pedigrees of modest size.
- Han & Abney have developed a method for estimating the IBD at any position, given a dense genotype data and a pedigree of any size and complexity. Their method is based in a linear approximation of a HMM simplified version of the continuous-chromosome IBD process.
- Browning provides a population-based framework for the inference of *IBD*.
- Thompson and Glazner & Thompson address the issue for families and individuals within a population also through a HMM approximation of the *IBD* process.

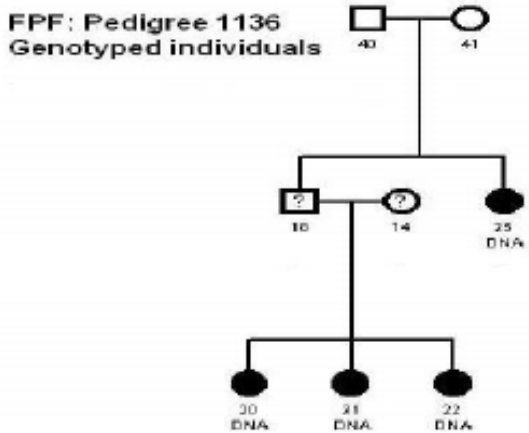
MEMORIAL UNIVERSITY




Familial pulmonary fibrosis: Pedigree R0892
Genotyped individuals




FPF: Pedigree 942
Genotyped individuals


FPF: Pedigree 1136
Genotyped individuals

A crude approximation

- In practice, much simpler approximations may provide satisfactory results. For example, we have estimated the set $\left\{\Delta_{r,ij}^{\langle m\rangle}\right\}$ for the ijth pair of individuals at the mth marker with a simple application of the Bayes Theorem

$$\Delta_{r,ij}^{\langle m\rangle} = \Pr\left(\mathcal{S}_{ij}^{\langle m\rangle} = \mathcal{s}_r \,\middle|\, g_i^{\langle m\rangle}, g_j^{\langle m\rangle}\right)$$

$$= \frac{\Delta_{r,ij}^{o} \Pr\left(g_i^{\langle m\rangle}, g_j^{\langle m\rangle} \,\middle|\, \mathcal{S}_{ij}^{\langle m\rangle} = \mathcal{s}_r\right)}{\sum_{s=1}^{9} \Delta_{s,ij}^{o} \Pr\left(g_i^{\langle m\rangle}, g_j^{\langle m\rangle} \,\middle|\, \mathcal{S}_{ij}^{\langle m\rangle} = \mathcal{s}_s\right)}$$

where the conditional *AIS* probabilities come from the next *Tables* and $\Delta_{r,ij}^{o}$ can be either the pedigree based estimate of the IBD coefficients or an estimate obtained for the ijth pair from properly chosen unlinked markers.

With a bi-allelic marker there are six distinguishable marker genotype pairs and their conditional probabilities given the condensed identity states, $\Pr(g_i, g_j \mid \mathcal{S} = \mathcal{J}_r)$, are

	Condensed identity states								
g_i, g_j	$\mathcal{J}_1$	$\mathcal{J}_2$	$\mathcal{J}_3$	$\mathcal{J}_4$	$\mathcal{J}_5$	$\mathcal{J}_6$	$\mathcal{J}_7$	$\mathcal{J}_8$	$\mathcal{J}_9$
aa, aa	p_a	p_a^2	p_a^2	p_a^3	p_a^2	p_a^3	p_a^2	p_a^3	p_a^4
aa, ab	0	0	$p_a p_b$	$2p_a^2 p_b$	0	0	0	$p_a^2 p_b$	$2p_a^3 p_b$
ab, aa	0	0	0	0	$p_a p_b$	$2p_a^2 p_b$	0	$p_a^2 p_b$	$2p_a^3 p_b$
aa, bb	0	$p_a p_b$	0	$p_a p_b^2$	0	$p_a^2 p_b$	0	0	$p_a^2 p_b^2$
ab, ab	0	0	0	0	0	0	$2p_a p_b$	$p_a p_b (p_a + p_b)$	$4p_a^2 p_b^2$
bb, aa	0	$p_a p_b$	0	$p_a^2 p_b$	0	$p_a p_b^2$	0	0	$p_a^2 p_b^2$
ab, bb	0	0	0	0	$p_a p_b$	$2p_a p_b^2$	0	$p_a p_b^2$	$2p_a p_b^3$
bb, ab	0	0	$p_a p_b$	$2p_a p_b^2$	0	0	0	$p_a p_b^2$	$2p_a p_b^3$
bb, bb	p_b	p_b^2	p_b^2	p_b^3	p_b^2	p_b^3	p_b^2	p_b^3	p_b^4

MEMORIAL UNIVERSITY

With a polymorphic marker, the relevant conditional probabilities $\Pr(g_i, g_j \mid \mathcal{S} = \mathcal{J}_r)$ are

	Condensed identity states								
g_i, g_j	$\mathcal{J}_1$	$\mathcal{J}_2$	$\mathcal{J}_3$	$\mathcal{J}_4$	$\mathcal{J}_5$	$\mathcal{J}_6$	$\mathcal{J}_7$	$\mathcal{J}_8$	$\mathcal{J}_9$
aa, aa	p_a	p_a^2	p_a^2	p_a^3	p_a^2	p_a^3	p_a^2	p_a^3	p_a^4
aa, ab	0	0	$p_a p_b$	$2p_a^2 p_b$	0	0	0	$p_a^2 p_b$	$2p_a^3 p_b$
ab, aa	0	0	0	0	$p_a p_b$	$2p_a^2 p_b$	0	$p_a^2 p_b$	$2p_a^3 p_b$
aa, bb	0	$p_a p_b$	0	$p_a p_b^2$	0	$p_a^2 p_b$	0	0	$p_a^2 p_b^2$
ab, ab	0	0	0	0	0	0	$2p_a p_b$	$p_a p_b (p_a + p_b)$	$4p_a^2 p_b^2$
aa, bc	0	0	0	$2p_a p_b p_c$	0	0	0	0	$2p_a^2 p_b p_c$
ab, ac	0	0	0	0	0	0	0	$p_a p_b p_c$	$4p_a^2 p_b p_c$
ab, cc	0	0	0	0	0	$2p_a p_b p_c$	0	0	$2p_a p_b p_c^2$
ab, cd	0	0	0	0	0	0	0	0	$4p_a p_b p_c p_d$

MEMORIAL UNIVERSITY

- Of course, $\Delta_r^{\langle m \rangle}$s obtained this way do not possess the beautiful properties of the estimates obtained through any of the HHM approaches mentioned above. In fact, they may not even be compatible with the legal states of Markov chain. However, for all practical purposes, they produce very much the same numerical results for this problem.
- Be aware that estimating $\Delta^o_{s,ij}$ by method of moments from SNP data is not possible because in such a case the matrix of conditional probabilities given by the bi-allelic *Table* is not of full rank. This also implies that the log-likelihood is not strictly convex so one needs to take extra precaution to find the MLE estimates. Nonetheless, kinship and inbreeding are identifiable.

MEMORIAL UNIVERSITY

Bottom line is once $\Delta_{r,ij}^{\langle m \rangle}$ for a fixed locus m has been obtained, we can compute

$$\varphi_{ij}^{\langle m \rangle} = \Delta_{1,ij}^{\langle m \rangle} + \frac{1}{2}\left(\Delta_{3,ij}^{\langle m \rangle} + \Delta_{5,ij}^{\langle m \rangle} + \Delta_{7,ij}^{\langle m \rangle}\right) + \frac{1}{4}\Delta_{8,ij}^{\langle m \rangle}$$

where

$$\Delta_{r,ij}^{\langle m \rangle} = \Pr\left(\mathcal{S}^{\langle m \rangle} = \mathcal{s}_r \,\middle|\, \text{`marker data'}\right)$$

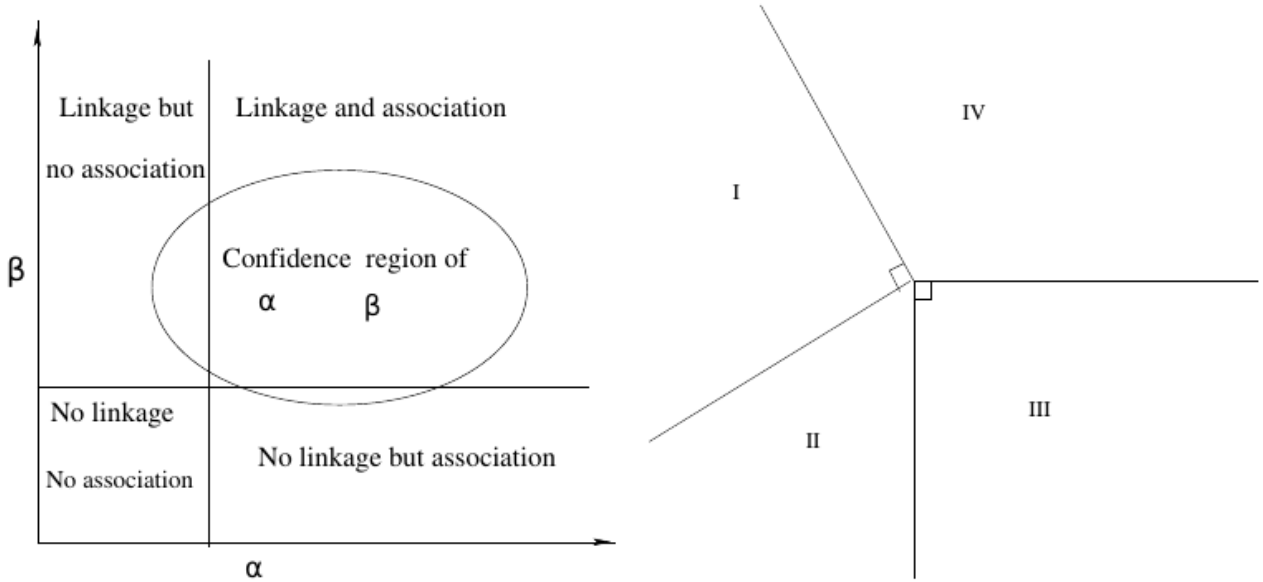
so that $\boldsymbol{\Sigma} = \boldsymbol{I} + \beta\,\boldsymbol{\Phi}^{\langle m \rangle >}$ in our model, β captures the linkage signal from the mth locus, while α does association.

- The joint linkage and association test is

$$H_0 : \alpha = 0 \text{ and } \beta = 0$$
$$\textit{vs} \quad H_A : \alpha \neq 0 \text{ and/or } \beta > 0$$

- The alternative has three cases:

$$H_{A_1} : \alpha \neq 0 \text{ and } \beta = 0$$
$$H_{A_2} : \alpha = 0 \text{ and } \beta > 0$$
$$H_{A_1} : \alpha \neq 0 \text{ and } \beta > 0$$



- Region I represents the LRT for H_{A_1} with angle $\pi/2$ and dist. χ^2_1; region II represents H_{A_2} with algle $\pi/2$ and dist. $\frac{1}{2}\chi^2_0 + \frac{1}{2}\chi^2_1$; region IV has an angle ρ and dist. $\frac{1}{2}\chi^2_1 + \frac{1}{2}\chi^2_2$; and Region II with angle $\pi - \rho$ and dist. χ^2_0,
$$\left(\rho = \arccos \frac{i_{\alpha\beta}}{\sqrt{i_{\alpha\alpha} i_{\beta\beta}}} \text{ under } H_0\right).$$
- So, the LRT has dist. $\frac{3}{8}\chi^2_0 + \frac{1}{2}\chi^2_1 + \frac{1}{8}\chi^2_2$.

α	β	μ = -3			μ = -5		
		Joint	Association	Linkage	Joint	Association	Linkage
.2	.2	.394	.142	.129	.350	.117	.147
	.4	.463	.140	.192	.416	.109	.180
	.6	.562	.150	.235	.498	.106	.278
	.8	.604	.118	.298	.615	.095	.378
.4	.2	.632	.382	.115	.461	.216	.109
	.4	.709	.362	.146	.545	.213	.181
	.6	.758	.353	.221	.628	.251	.271
	.8	.790	.372	.268	.722	.232	.371
.6	.2	.884	.746	.112	.623	.385	.124
	.4	.883	.692	.142	.701	.398	.190
	.6	.930	.695	.201	.749	.392	.275
	.8	.938	.660	.260	.836	.434	.380
.8	.2	.978	.920	.106	.773	.586	.117
	.4	.969	.901	.127	.840	.603	.196
	.6	.980	.898	.164	.882	.636	.273
	.8	.984	.887	.213	.928	.619	.368

Table: Empirical power of score tests for binary mixed models

- What about 'joint families and case/controls'?

- We started with

$$\pi_{ij}(\xi_{ij}) = \Pr\left(y_{ij} = 1 \mid \boldsymbol{z}_{ij}, \xi_{ij}\right)$$

$$= \frac{\mathrm{e}^{\boldsymbol{\theta}'\boldsymbol{z}_{ij}+\xi_{ij}}}{1+\mathrm{e}^{\boldsymbol{\theta}'\boldsymbol{z}_{ij}+\xi_{ij}}} \qquad \text{for } i = 1, 2, \ldots, N,\ j = 1, 2, \ldots, n_i.$$

and $\boldsymbol{\theta} = (\mu, \alpha, \boldsymbol{\gamma}')'$, $\boldsymbol{\xi}_i \sim \mathcal{N}\left(0,\, \sigma^2 \boldsymbol{\Sigma}_i\right)$ with $\boldsymbol{\Sigma}_i = \boldsymbol{I} + \beta \boldsymbol{\Phi}_i$, etc...

- This model can be easily adapted to other situations.

Thank you

CIHR IRSC

Canadian Institutes of Health Research

Instituts de recherche en santé du Canada

MEMORIAL UNIVERSITY