Heritability of non-HLA genetics in Celiac Disease: a population based study in 107,000 twins - Appendix A, method for estimation of heritability when accounting for necessary cause allele

Ralf Kuja-Halkola[a], *Benjamin Lebwohl*[a,b], *Jonas Halfvarson*[c], *Cisca Wijmenga*[d], *Patrik K.E Magnusson*[a], *Jonas F Ludvigsson*[a,e,f]

From the

[a] Department Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

[b]Celiac Disease Center, Department of Medicine, Columbia University Medical Center, Columbia University, New York, USA

[c] Department of Gastroenterology, Faculty of Medicine and Health, Örebro University, Örebro, Sweden

[d] University of Groningen, University Medical Center, Department of Genetics, Groningen, the Netherlands

[e] Department of Pediatrics, Örebro University Hospital, Örebro, Sweden

[f] Division of Epidemiology and Public Health, School of Medicine, University of Nottingham, Clinical Sciences Building 2, City Hospital, Nottingham, UK

Introduction

Almost 100% of all celiac disease cases carry either the DQ2 or DQ8 human leukocyte antigen (HLA) allele, making the presence of DQ2/DQ8 essentially a necessary cause for the disease. However, not all with DQ2/DQ8 develops the disease, i.e. the penetrance is not 100%, making the presence of DQ2 or DQ8 alleles a necessary, but not sufficient, cause. Further, DQ2 has a carrier frequency of about 25% in the (Swedish) population, while the carrier frequency of DQ8 is much smaller. In the derivations below, no distinction between homo and heterozygous carrier status of DQ2/DQ8 is made (i.e., only one necessary-cause allele is assumed). About 4% of the DQ2/DQ8-carriers develop the disease; meaning that the penetrance is about 4%. Therefore, in the Swedish population we would expect approximately 1% to have celiac disease, however if data is gathered from registers, as in the current study, this number is expected to be lower due to non-detected cases.

Here we derive a way to account for HLA in analyses of twin data where only disease status is observed.

Classical twin model for binary traits

The standard approach to estimating association between twins in pairs in the classical twin design for a binary phenotype is called the liability-threshold model. An underlying standard normally distributed liability (mean=0, variance=1) of the disease is assumed, and a threshold is estimated; where, if an individual have the disease, the liability is assumed to be above the threshold, and if the individual does not have the disease the liability is assumed to be below the threshold. The correlation between two underlying liability distributions, one for each twin in a pair, is the measure of association, and is referred to as the tetrachoric correlation. This correlation – with separate estimates for for monozygotic (MZ) and dizygotic (DZ) twin pairs – is the basis for heritability calculations.

To find the likelihood of the correlation and threshold; let $Y_{ij} = 1$ be the event that individual j in twin pair i has the disease, and $Y_{ij} = 0$ the event that he/she do not have the disease. Further, let ρ be the tetrachoric correlation, τ the threshold, and let y_{ij} be the observed event for twin j in pair i (=0 or 1). The likelihood of the observed data is

$$\boldsymbol{L}(\rho, \tau) = \prod_{i}^{N} \Pr(Y_{i1} = y_{i1}, Y_{i2} = y_{i2} | \rho, \tau), \qquad \text{Eq. 1}$$

where

$$
\begin{aligned}
\Pr(Y_{i1}=0, Y_{i2}=0|\rho,\tau) &= \int_{-\infty}^{\tau}\int_{-\infty}^{\tau}\frac{1}{2\pi}\left|\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}\right|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\begin{bmatrix}z_1 & z_2\end{bmatrix}\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}^{-1}\begin{bmatrix}z_1\\ z_2\end{bmatrix}\right)dz_2dz_1,\\
\Pr(Y_{i1}=0, Y_{i2}=1|\rho,\tau) &= \int_{-\infty}^{\tau}\int_{\tau}^{\infty}\frac{1}{2\pi}\left|\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}\right|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\begin{bmatrix}z_1 & z_2\end{bmatrix}\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}^{-1}\begin{bmatrix}z_1\\ z_2\end{bmatrix}\right)dz_2dz_1,\\
\Pr(Y_{i1}=1, Y_{i2}=0|\rho,\tau) &= \int_{\tau}^{\infty}\int_{-\infty}^{\tau}\frac{1}{2\pi}\left|\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}\right|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\begin{bmatrix}z_1 & z_2\end{bmatrix}\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}^{-1}\begin{bmatrix}z_1\\ z_2\end{bmatrix}\right)dz_2dz_1,\\
\Pr(Y_{i1}=1, Y_{i2}=1|\rho,\tau) &= \int_{\tau}^{\infty}\int_{\tau}^{\infty}\frac{1}{2\pi}\left|\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}\right|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\begin{bmatrix}z_1 & z_2\end{bmatrix}\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}^{-1}\begin{bmatrix}z_1\\ z_2\end{bmatrix}\right)dz_2dz_1.
\end{aligned}
\qquad \text{Eq. 2}
$$

The maximum likelihood estimates of ρ and τ are found by maximizing $\boldsymbol{L}(\rho,\tau)$ with regards to the parameters.

Heritability estimation is conducted by including both the group of MZ and DZ twins in a model and, under the assumption that MZ twins are genetically identical and DZ twins share 50% of their segregating alleles on average, and MZ and DZ twin pairs share environment to an equal extent, incorporating the constraints

$$
\begin{aligned}
\rho_{MZ} &= h^2 + c^2,\\
\rho_{DZ} &= \frac{1}{2}h^2 + c^2,\\
e^2 &= 1 - h^2 - c^2.
\end{aligned}
\qquad \text{Eq. 3}
$$

Here h^2 is the (narrow sense) heritability, c^2 is the shared environment, and e^2 is the non-shared environment.

Alternative likelihood given necessary alleles

The problem with Eq. 1 and 2 is that we are assuming that all individuals can have the disease when, in fact, only those having the necessary cause are at risk for developing the disease. To alleviate this problem we may introduce information about whether the twins in pairs have the necessary cause or not. The preferred way to include the information would be to measure the presence of DQ2/DQ8-alleles in the population; however this is not feasible in many studies, including our study with over 100,000 twins.

We want to estimate the likelihood given the necessary causes, but do not have them observed. Luckily we may calculate probabilities for the distribution of necessary causes, and thus calculate the likelihood of the data. Let r be the new (non-HLA) correlation between the liabilities of disease, and t be the new threshold, both defined similar as above. Note that r and t represent slightly different entities than ρ and τ; they refer to the correlation and threshold *after the necessary cause has been accounted for*. Let $X_{ij}=1$ be the event that individual j in twin pair i have the necessary cause, and $X_{ij}=0$ be the event that he/she does not have the necessary cause. The likelihood may be re-written as

$$\boldsymbol{L}(r,t) = \prod_i \Pr(Y_{i1} = y_{i1}, Y_{i2} = y_{i2} | r, t)$$
$$= \prod_i \sum_{k=0}^{1} \sum_{l=0}^{1} \Pr(Y_{i1} = y_{i1}, Y_{i2} = y_{i2} | X_{i1} = k, X_{i2} = l, r, t) \Pr(X_{i1} = k, X_{i2} = l | r, t) \quad \text{Eq. 4}$$
$$= \prod_i \sum_{k=0}^{1} \sum_{l=0}^{1} \Pr(Y_{i1} = y_{i1}, Y_{i2} = y_{i2} | X_{i1} = k, X_{i2} = l, r, t) \Pr(X_{i1} = k, X_{i2} = l),$$

where the last step is due to that the probability of having the allele is independent of the penetrance of the disease, as well as the correlation, not due to HLA, of disease between the co-twins. Here $\Pr(Y_{i1} = y_{i1}, Y_{i2} = y_{i2} | X_{i1} = k, X_{i2} = l, r, t)$ is known, see **eTable 1**.

eTable 1: Probabilities of observing the data given the distribution of necessary cause

$\boldsymbol{y_{i1}}$	$\boldsymbol{y_{i2}}$	$\boldsymbol{k}$	$\boldsymbol{l}$	$\boldsymbol{\Pr(Y_{i1} = y_{i1}, Y_{i2} = y_{i2} \mid X_{i1} = k, X_{i2} = l, r, t)}$
0	0	0	0	1
0	0	0	1	$\int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_2^2\right) dz_2$
0	0	1	0	$\int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_1^2\right) dz_1$
0	0	1	1	$\int_{-\infty}^{t} \int_{-\infty}^{t} \frac{1}{2\pi} \left\lvert \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix} \right\rvert^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \begin{bmatrix} z_1 & z_2 \end{bmatrix} \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}^{-1} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) dz_2 dz_1$
0	1	0	1	$\int_{t}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_2^2\right) dz_2$
1	0	1	0	$\int_{t}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_1^2\right) dz_1$
0	1	1	1	$\int_{-\infty}^{t} \int_{t}^{\infty} \frac{1}{2\pi} \left\lvert \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix} \right\rvert^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \begin{bmatrix} z_1 & z_2 \end{bmatrix} \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}^{-1} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) dz_2 dz_1$
1	0	1	1	$\int_{t}^{\infty} \int_{-\infty}^{t} \frac{1}{2\pi} \left\lvert \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix} \right\rvert^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \begin{bmatrix} z_1 & z_2 \end{bmatrix} \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}^{-1} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) dz_2 dz_1$
1	1	1	1	$\int_{t}^{\infty} \int_{t}^{\infty} \frac{1}{2\pi} \left\lvert \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix} \right\rvert^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \begin{bmatrix} z_1 & z_2 \end{bmatrix} \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}^{-1} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) dz_2 dz_1$
1	0	0	0	0
0	1	0	0	0
1	0	0	1	0
0	1	1	0	0
1	1	0	0	0
1	1	1	0	0
1	1	0	1	0

Thus, if we derive $\Pr(X_{i1} = k, X_{i2} = l)$, which depends on the distributions of alleles in the population, we may find the likelihood.

Prevalences of alleles and necessary cause

Let p be the probability of an individual having the necessary cause (i.e., having at least one DQ2/DQ8 allele), and let p^* be the probability that an allele is DQ2/DQ8. Assume the probability of allele 1 and allele 2 being DQ2/DQ8 to be independent (corresponding to no assortative mating and no strong selection), then

$$p = \Pr(X_{i1} = 1 \cup X_{i2} = 1) = \Pr(X_{i1} = 1) + \Pr(X_{i2} = 1) - \Pr(X_{i1} = 1 \cap X_{i2} = 1)$$
$$= p^* + p^* - {p^*}^2 = p^*(2 - p^*) \Leftrightarrow p^* = 1 - \sqrt{1 - p}. \quad \text{Eq. 5}$$

For example, if $p = \frac{1}{4}$ then $p^* = 1 - \sqrt{1 - \frac{1}{4}} \approx 0.134$.

Inheritance of alleles

Given known allele frequency in a population, the probability of offspring having at least one DQ2/DQ8-allele may be calculated (i.e., the probability of having the necessary cause). One can use this to calculate the probability of two offspring (DZ twins or full sibling pair) of a parental pair having the necessary cause. There are 16 possible combinations of parental alleles, see **eTable 2**. An offspring will inherit one of the two alleles from each parent, thus four different allele combinations are possible per offspring in this specific locus when considering both chromosomes. When the probability of having the DQ2/DQ8-allele can be assumed independent for each parent and chromosome (i.e., no assortative mating) we can tabulate the different scenarios as in **eTable 2**.

eTable 2: Distribution of alleles in parents and offspring

Parents	○○ ○○	●○ ○○	●● ○○	●○ ●○	●○ ○●	●● ●○	●● ○●	●○ ●●
Children	- - - -	x x - -	x x x x	x x x -	x x x -	x x x x	x x x x	x x x x
Probability of parents	$(1-p^*)^4$	$p^*(1-p^*)^3$	$p^{*2}(1-p^*)^2$	$p^{*2}(1-p^*)^2$	$p^{*2}(1-p^*)^2$	$p^{*3}(1-p^*)$	$p^{*3}(1-p^*)$	$p^{*3}(1-p^*)$
Pairs of offspring where none have necessary cause	$\left(\frac{4}{4}\right)^2 = 1$	$\left(\frac{2}{4}\right)^2 = \frac{1}{4}$	$\left(\frac{0}{4}\right)^2 = 0$	$\left(\frac{1}{4}\right)^2 = \frac{1}{16}$	$\left(\frac{1}{4}\right)^2 = \frac{1}{16}$	$\left(\frac{0}{4}\right)^2 = 0$	$\left(\frac{0}{4}\right)^2 = 0$	$\left(\frac{0}{4}\right)^2 = 0$
Pairs of offspring where only one have necessary cause	$2\frac{0}{4}\frac{0}{4} = 0$	$2\frac{2}{4}\frac{2}{4} = \frac{2}{4}$	$2\frac{0}{4}\frac{0}{4} = 0$	$2\frac{3}{4}\frac{1}{4} = \frac{6}{16}$	$2\frac{3}{4}\frac{1}{4} = \frac{6}{16}$	$2\frac{0}{4}\frac{0}{4} = 0$	$2\frac{0}{4}\frac{0}{4} = 0$	$2\frac{0}{4}\frac{0}{4} = 0$
Pairs of offspring where both have necessary cause	$\left(\frac{0}{4}\right)^2 = 0$	$\left(\frac{2}{4}\right)^2 = \frac{1}{4}$	$\left(\frac{4}{4}\right)^2 = 1$	$\left(\frac{3}{4}\right)^2 = \frac{9}{16}$	$\left(\frac{3}{4}\right)^2 = \frac{9}{16}$	$\left(\frac{4}{4}\right)^2 = 1$	$\left(\frac{4}{4}\right)^2 = 1$	$\left(\frac{4}{4}\right)^2 = 1$

Parents	●● ●●	○● ○○	○○ ●○	○○ ○●	○● ●○	○● ○●	○○ ●●	○● ●●
Children	x x x x	x x - -	x x - -	x x - -	x x x -	x x x -	x x x x	x x x x
Probability of parents	p^{*4}	$p^*(1-p^*)^3$	$p^*(1-p^*)^3$	$p^*(1-p^*)^3$	$p^{*2}(1-p^*)^2$	$p^{*2}(1-p^*)^2$	$p^{*2}(1-p^*)^2$	$p^{*3}(1-p^*)$
Pairs of offspring where none have necessary cause	$\left(\frac{0}{4}\right)^2 = 0$	$\left(\frac{2}{4}\right)^2 = \frac{1}{4}$	$\left(\frac{2}{4}\right)^2 = \frac{1}{4}$	$\left(\frac{2}{4}\right)^2 = \frac{1}{4}$	$\left(\frac{1}{4}\right)^2 = \frac{1}{16}$	$\left(\frac{1}{4}\right)^2 = \frac{1}{16}$	$\left(\frac{0}{4}\right)^2 = 0$	$\left(\frac{0}{4}\right)^2 = 0$
Pairs of offspring where only one have necessary cause	$2\frac{0}{4}\frac{0}{4} = 0$	$2\frac{2}{4}\frac{2}{4} = \frac{2}{4}$	$2\frac{2}{4}\frac{2}{4} = \frac{2}{4}$	$2\frac{2}{4}\frac{2}{4} = \frac{2}{4}$	$2\frac{3}{4}\frac{1}{4} = \frac{6}{16}$	$2\frac{3}{4}\frac{1}{4} = \frac{6}{16}$	$2\frac{0}{4}\frac{0}{4} = 0$	$2\frac{0}{4}\frac{0}{4} = 0$
Pairs of offspring where both have necessary cause	$\left(\frac{4}{4}\right)^2 = 1$	$\left(\frac{2}{4}\right)^2 = \frac{1}{4}$	$\left(\frac{2}{4}\right)^2 = \frac{1}{4}$	$\left(\frac{2}{4}\right)^2 = \frac{1}{4}$	$\left(\frac{3}{4}\right)^2 = \frac{9}{16}$	$\left(\frac{3}{4}\right)^2 = \frac{9}{16}$	$\left(\frac{4}{4}\right)^2 = 1$	$\left(\frac{4}{4}\right)^2 = 1$

Note: "●", the necessary DQ2/DQ8 allele. "○", not the necessary DQ2/DQ8 allele. "x", necessary cause. "-", not necessary cause. "Parents", distribution of alleles in parents. "Children", possible combinations of children having the necessary cause. "Probability of parents", probability of specific distribution of alleles in parents.

From this a total of 6 different combinations of parental probabilities and offspring frequencies may be found, see **eTable 3**.

eTable 3: Probabilities of pairs having necessary cause as a function of distribution of alleles in parents

Parental combination of alleles	Number of instances parental combination arises	Probability for parental combination of alleles	Proportion of offspring pairs where neither have the necessary cause	Proportion of offspring pairs where only one have necessary cause	Proportion of offspring pairs where both have necessary cause
No alleles	1	$(1-p^*)^4$	1	0	0
One allele in one parent	4	$p^*(1-p^*)^3$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
Two alleles in one parent	2	$p^{*2}(1-p^*)^2$	0	0	1
One allele in each parent	4	$p^{*2}(1-p^*)^2$	$\frac{1}{16}$	$\frac{6}{16}$	$\frac{9}{16}$
Two alleles in one parent, one allele in other	4	$p^{*3}(1-p^*)$	0	0	1
Two alleles in both parents	1	p^{*4}	0	0	1

The probabilities of observing concordant no necessary cause, discordant necessary cause, and concordant necessary cause in MZ twin pairs and DZ twin pairs (or any full sibling pairs) may be calculated. Note that for MZ twins, who are genetically identical, this is simple; they follow the distribution in the population and will never be discordant for the necessary cause. The distributions of necessary causes are presented in **eTable 4**, and are based on the probabilities and frequencies presented in **eTable 3**.

eTable 4: Probabilities for concordance and discordance in having necessary cause

	MZ twin pairs	DZ twin pairs
Concordant no necessary cause	$\Pr(X_{i1}=0, X_{i2}=0\|MZ) = C00_{MZ} = 1-p = (1-p^*)^2$	$\Pr(X_{i1}=0, X_{i2}=0\|DZ) = C00_{DZ} = 1\cdot(1-p^*)^4\cdot 1 + 4\cdot p^*(1-p^*)^3\cdot\frac{1}{4} + 4\cdot p^{*2}(1-p^*)^2\cdot\frac{1}{16} = (1-p^*)^4 + p^*(1-p^*)^3 + \frac{1}{4}\cdot p^{*2}(1-p^*)^2$
Discordant necessary cause	$\Pr(X_{i1}=0, X_{i2}=1\|MZ) = \Pr(X_{i1}=1, X_{i2}=0\|MZ) = C01_{MZ} = 0$	$\Pr(X_{i1}=0, X_{i2}=1\|DZ) = \Pr(X_{i1}=1, X_{i2}=0\|DZ) = C01_{DZ} = \left(4\cdot p^*(1-p^*)^3\cdot\frac{1}{2} + 4\cdot p^{*2}(1-p^*)^2\cdot\frac{6}{16}\right)/2 = p^*(1-p^*)^3 + \frac{3}{4}\cdot p^{*2}(1-p^*)^2$
Concordant necessary cause	$\Pr(X_{i1}=1, X_{i2}=1\|MZ) = C11_{MZ} = p = p^*(2-p^*)$	$\Pr(X_{i1}=1, X_{i2}=1\|DZ) = C11_{DZ} = 4\cdot p^*(1-p^*)^3\cdot\frac{1}{4} + 2\cdot p^{*2}(1-p^*)^2\cdot 1 + 4\cdot p^{*2}(1-p^*)^2\cdot\frac{9}{16} + 4\cdot p^{*3}(1-p^*)\cdot 1 + 1\cdot p^{*4}\cdot 1 = p^*(1-p^*)^3 + \frac{17}{4}\cdot p^{*2}(1-p^*)^2 + 4\cdot p^{*3}(1-p^*) + p^{*4}$

Note: $C00_{MZ}$, $C01_{MZ}$, $C11_{MZ}$, $C00_{DZ}$, $C01_{DZ}$, and $C11_{DZ}$ are the indicated probabilities for each scenario.

We may combine the information in **eTable 1** and **eTable 4** to find the likelihood of the data, presented in **eTable 5**. The likelihood will depend on the observed disease status and be summed over the distribution of necessary causes in twins in a pair.

eTable 5: Likelihood contribution from MZ and DZ pairs with different observed disease status

MZ	y_{i1}	y_{i2}	$L_{i,MZ}(r_{MZ},t)=\sum_{k=0}^{1}\sum_{l=0}^{1}\Pr(Y_{i1}=y_{i1},Y_{i2}=y_{i2}\mid X_{i1}=k,X_{i2}=l,r_{MZ},t)\Pr(X_{i1}=k,X_{i2}=l)$
	0	0	$1\cdot C00_{MZ}+\int_{-\infty}^{t}\int_{-\infty}^{t}\frac{1}{2\pi}\left\|\begin{bmatrix}1 & r_{MZ}\\ r_{MZ} & 1\end{bmatrix}\right\|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\begin{bmatrix}z_1 & z_2\end{bmatrix}\begin{bmatrix}1 & r_{MZ}\\ r_{MZ} & 1\end{bmatrix}^{-1}\begin{bmatrix}z_1\\ z_2\end{bmatrix}\right)dz_2dz_1\cdot C11_{MZ}$
	0	1	$\int_{-\infty}^{t}\int_{t}^{\infty}\frac{1}{2\pi}\left\|\begin{bmatrix}1 & r_{MZ}\\ r_{MZ} & 1\end{bmatrix}\right\|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\begin{bmatrix}z_1 & z_2\end{bmatrix}\begin{bmatrix}1 & r_{MZ}\\ r_{MZ} & 1\end{bmatrix}^{-1}\begin{bmatrix}z_1\\ z_2\end{bmatrix}\right)dz_2dz_1\cdot C11_{MZ}$
	1	0	$\int_{-t}^{\infty}\int_{-\infty}^{t}\frac{1}{2\pi}\left\|\begin{bmatrix}1 & r_{MZ}\\ r_{MZ} & 1\end{bmatrix}\right\|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\begin{bmatrix}z_1 & z_2\end{bmatrix}\begin{bmatrix}1 & r_{MZ}\\ r_{MZ} & 1\end{bmatrix}^{-1}\begin{bmatrix}z_1\\ z_2\end{bmatrix}\right)dz_2dz_1\cdot C11_{MZ}$
	1	1	$\int_{t}^{\infty}\int_{t}^{\infty}\frac{1}{2\pi}\left\|\begin{bmatrix}1 & r_{MZ}\\ r_{MZ} & 1\end{bmatrix}\right\|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\begin{bmatrix}z_1 & z_2\end{bmatrix}\begin{bmatrix}1 & r_{MZ}\\ r_{MZ} & 1\end{bmatrix}^{-1}\begin{bmatrix}z_1\\ z_2\end{bmatrix}\right)dz_2dz_1\cdot C11_{MZ}$

DZ	y_{i1}	y_{i2}	$L_{i,DZ}(r_{DZ},t)=\sum_{k=0}^{1}\sum_{l=0}^{1}\Pr(Y_{i1}=y_{i1},Y_{i2}=y_{i2}\mid X_{i1}=k,X_{i2}=l,r_{DZ},t)\Pr(X_{i1}=k,X_{i2}=l)$
	0	0	$1\cdot C00_{DZ}+2\cdot\int_{-\infty}^{t}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)dz\cdot C01_{DZ}$ $+\int_{-\infty}^{t}\int_{-\infty}^{t}\frac{1}{2\pi}\left\|\begin{bmatrix}1 & r_{DZ}\\ r_{DZ} & 1\end{bmatrix}\right\|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\begin{bmatrix}z_1 & z_2\end{bmatrix}\begin{bmatrix}1 & r_{DZ}\\ r_{DZ} & 1\end{bmatrix}^{-1}\begin{bmatrix}z_1\\ z_2\end{bmatrix}\right)dz_2dz_1\cdot C11_{DZ}$
	0	1	$\int_{t}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)dz\cdot C01_{DZ}$ $+\int_{-\infty}^{t}\int_{t}^{\infty}\frac{1}{2\pi}\left\|\begin{bmatrix}1 & r_{DZ}\\ r_{DZ} & 1\end{bmatrix}\right\|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\begin{bmatrix}z_1 & z_2\end{bmatrix}\begin{bmatrix}1 & r_{DZ}\\ r_{DZ} & 1\end{bmatrix}^{-1}\begin{bmatrix}z_1\\ z_2\end{bmatrix}\right)dz_2dz_1\cdot C11_{DZ}$
	1	0	$\int_{t}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)dz\cdot C01_{DZ}$ $+\int_{t}^{\infty}\int_{-\infty}^{t}\frac{1}{2\pi}\left\|\begin{bmatrix}1 & r_{DZ}\\ r_{DZ} & 1\end{bmatrix}\right\|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\begin{bmatrix}z_1 & z_2\end{bmatrix}\begin{bmatrix}1 & r_{DZ}\\ r_{DZ} & 1\end{bmatrix}^{-1}\begin{bmatrix}z_1\\ z_2\end{bmatrix}\right)dz_2dz_1\cdot C11_{DZ}$
	1	1	$\int_{-\infty}^{t}\int_{-\infty}^{t}\frac{1}{2\pi}\left\|\begin{bmatrix}1 & r_{DZ}\\ r_{DZ} & 1\end{bmatrix}\right\|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\begin{bmatrix}z_1 & z_2\end{bmatrix}\begin{bmatrix}1 & r_{DZ}\\ r_{DZ} & 1\end{bmatrix}^{-1}\begin{bmatrix}z_1\\ z_2\end{bmatrix}\right)dz_2dz_1\cdot C11_{DZ}$

Note: Observed disease status is indicated by y_{i1} and y_{i2}. Prevalence assumed equal across zygosities (t equal for MZ and DZ). r_{MZ} and r_{DZ} represent the correlation in MZ and DZ twin pairs, respectively.

It is now possible to use r_{MZ} and r_{DZ} in standard twin modelling to estimate heritability.