

# Genotype-Based Score Test for Association Testing in Families

Hae-Won Uh · Marian Beekman · Ingrid Meulenbelt ·  
Jeanine J. Houwing-Duistermaat

Received: 28 November 2013 / Revised: 20 November 2014 / Accepted: 17 February 2015 /  
Published online: 17 March 2015  
© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** The multiplex-case and control design in which multiple cases are sampled from the same family is considered. In such studies phenotype information of the un-genotyped relatives might be available. We intend to use additional family information when performing genetic association tests. A score test is revisited to provide a flexible framework to accommodate various genetic models and to improve power of the association test by adding available family information. The proposed test accounts for correlations induced by multiple cases from the same pedigree, directly deals with X-linked SNPs in mixed-sex-related samples, and incorporates additional phenotypic information such as the number of (un-genotyped) siblings and parents with similar symptoms by assigning the weights to (genotyped) multiplex cases. In addition, the score test directly incorporates posterior probabilities of imputed genotypes, which leads to an efficiency measure that reflects imputation uncertainty on the test conducted. The proposed test is applied to real applications for illustration. Its efficiency is demonstrated via simulations.

**Keywords** Score test · Multiplex-case and control design · Ascertained familial cases · X-linked SNPs · imputation uncertainty · Incorporation of additional family information

---

H.-W. Uh (✉) · J. J. Houwing-Duistermaat  
Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, S-5-P,  
P.O. Box 9600, 2300 RC Leiden, The Netherlands  
e-mail: h.uh@lumc.nl

M. Beekman · I. Meulenbelt  
Department of Molecular Epidemiology, Leiden University Medical Center,  
S-5-P, P.O. Box 9600, 2300 RC Leiden, The Netherlands

## 1 Introduction

We investigate the use of family-based samples for conducting case–control association analysis. The family-based association tests such as the FBAT [33] are robust against population structure, but lack power because they utilize only the within-family component for the construction of the association test [19,34,40]. The population-based tests that incorporate between-family information for family data may be more powerful. A hybrid case–control family-based analysis has been proposed, in which family-based designs integrate unselected controls from other studies into the analysis [21]. In contrast, we incorporate multiplex cases into population-based case–control study in the framework of score test. Although cases as well as controls can be sampled from families, we focus on the multiplex-case and control design using familial cases. The primary advantage of this design would be that familial cases are enriched for genetic factors and therefore may be more informative for genetic research [3]. Such families may have higher frequencies of susceptibility alleles, and the expected difference in frequencies can be greater using multiplex cases and unrelated controls than using independent samples [19,34]. Since the test statistic to detect association typically depends on the difference in genotype frequencies between cases and controls this design may improve power to detect association, in particular using next-generation sequencing data [48].

When cases are ascertained via multiple affected individuals within pedigrees, the ascertainment issue should be addressed. It is argued that the ascertainment event depends on the phenotype but is conditionally independent of the genotype given disease outcome. The retrospective likelihood, therefore, is appropriate under selection [44,45]. Using a score test simplifies this matter, in which both prospective and retrospective model can be dealt with [28]. There are other advantages of using a score test based on genotypes. Firstly, it does not require the genotype frequencies to comply with Hardy–Weinberg proportions (HWP). Earlier work by Sasieni [35] demonstrated that statistical tests based on the comparison of allele frequencies rather than genotype frequencies between unrelated cases and controls can have an increased rate of false-positive conclusions when genotype frequencies do not satisfy HWP. Secondly, it provides a simple framework; different dosage scores of the high-risk allele can be used to test for multiplicative, dominant, or recessive effect. Lastly, the score test can be equally applicable for a quantitative phenotype, even when the sample is selected by extremes of phenotype [44].

To consider family-based samples for conducting case–control association analysis, we need to take into account within-family correlations to obtain the proper type I error rates. Generalized estimating equations (GEE) can be used to account for correlations between related individuals [25]. Slager et al. found that this method often fails due to the singularity in the working correlation matrix [37]. Slager and Schaid [36] proposed a statistical test based on the Cochran–Armitage test for trend in proportions [4,11] which included a variance appropriately accounted for family relationships. Bourgain et al. [5] constructed a quasi-likelihood score (QLS) test statistic that accounts for correlations between individuals by including kinship coefficients, thereby utilizing information from the known pedigree structure. In order to utilize additional phenotype information of un-genotyped family members, Thornton and McPeck [40]

proposed the *more powerful* quasi-likelihood score (MQLS) test. This extension of the QLS test incorporates additional phenotypic information of relatives who are not genotyped. That is, the phenotype data of un-genotyped family members are used to give corresponding weights to the (genotyped) multiplex cases. Note that both QLS and MQLS tests are based on the best linear unbiased estimator of allele frequency. These are more suited for related samples from a large complex pedigree for which maximum likelihood estimation is impractical [30]; data collected on the Hutterites and such are outside the scope of this work. Being an allelic test the MQLS test does not have flexibility to test for multiplicative, dominant, or recessive effect. Uh et al. [43] extended the MQLS to genotypic MQLS (gMQLS) test to accommodate different genetic models. Note that the weighting scheme using positive family history, which is similar to that of the MQLS test, has been proposed by Callegaro et al. in allele-sharing statistics for genetic linkage analysis [7]. In this work we incorporate this weighting scheme for the association analysis directly in the score statistic.

Working within the framework of the score test makes other extensions feasible. Firstly, we wish to test for association on the X chromosome in a related sample. For the X chromosome, females have two chromosomes but males have only one. As the X chromosome represents 2.5 % of the human genome for males and 5 % for females, information coming from the X chromosome cannot be ignored. Until recently little research has been reported on performance of such test statistics for association on the X chromosome. For unrelated samples, Loley et al. conducted a broadly conceived simulation study comparing different tests for association on the X chromosome [26]. One option is that after applying an allele-based test to males and females separately the two statistics of  $\chi^2_{(1)}$  distribution can be combined to a test statistic of  $\chi^2_{(2)}$  [46]. Although this approach is straightforward to apply, it often is not a valid test for family data. For example, when a sibling consists of a brother and sister pair, the two  $\chi^2_{(1)}$  tests of males and females are not independent; combining these two statistics becomes rather complicated. Alternatively, for construction of X-linked score test for the multiplex-case and control design, we follow the line of reasoning by Clayton [10]. While males carry 1 copy, in females most loci on the X chromosome are subject to X inactivation [9]; a female will have approximately half her cells with 1 copy active while the remainder of her cells have the other copy activated. In the absence of interaction with other loci or environmental factors, males should be equivalent to homozygous females. Therefore, X loci in males are coded 0 or 2. To account for relatedness in multiplex cases, an X-linked correlation matrix can be calculated either using the ITO matrices of Li and Sacks [22] or using MINX (MERLIN in X, [1]).

In the Genome-Wide Association Studies (GWAS) era, another important point to be considered is the imputation of the genotypes. By borrowing external information of reference haplotypes from the Haplotype Mapping Project (HapMap, <http://www.hapmap.org/>) and 1000 Genomes Project (<http://www.1000genomes.org/>), the number of SNPs to be tested increased from 2.5 to 6.7 million [38]. Considering the number of imputed SNPs will increase, and providing computationally efficient software is as important as guarding both accuracy and precision of the test using these imputed genotypes. Therefore, we propose a one-step approach to test and to deal with uncertainty of imputed genotypes. Based on the well-known results concerning the score function for incomplete data [13], we replace the genotype by its poste-

rior expectation given imputed data. The variance of the score statistic measures the statistical information contained in the data. As in Louis [27], Marchini et al. [29] incorporated in the variance term the loss of information from not observing the real genotype. In this manner, the score test provides an efficiency measure  $R_T^2$  that reflects the impact of imputed genotypes on the specific test conducted [41, 42]. In order to provide computationally feasible software for dealing with GWAS using imputed SNPs, C++ executable programs (CCassoc and QTassoc) are available at <http://www.lumc.nl/uh>.

## 2 Methods

### 2.1 Score Test for the Ascertained Cases

We first address the ascertainment of the independent subjects. Let  $Y = (Y_1, \dots, Y_n)$  be the phenotype,  $X = (X_1, \dots, X_n)$  denote genotype dosage 0, 1, or 2. Further,  $\bar{Y}$  is the mean of  $Y$  in the whole sample, or the proportion of cases in case–control studies. Since the ascertainment event  $\mathcal{S}$  depends on the phenotype but is conditionally independent of the genotype given  $Y$ ,  $P(X | Y, \mathcal{S}) = P(X | Y)$ . Therefore, the retrospective likelihood based on  $P(X | Y)$  is appropriate under selection [44, 45]. Based on the retrospective likelihood, the score statistic for testing for an additive effect of a genotyped locus on phenotype is

$$U_X = (Y - \bar{Y})^\top X. \quad (1)$$

Although the score statistic initially is to model for the effect of genotype on phenotype, the subscript  $X$  is to note that  $U_X$  is based on the distribution of  $P(X | Y)$  [12, 28]. The score statistic  $U_X$  is asymptotically normally distributed under  $H_0$  with the zero mean and variance

$$\text{Var}U_X = (Y - \bar{Y})^\top \text{Var}X(Y - \bar{Y}). \quad (2)$$

Under  $H_0$ ,  $U_X^2/\text{Var}U_X$  is asymptotically distributed as chi-squared with 1 degree of freedom. The variance of genotype  $X$  can be estimated by

$$\widehat{\text{Var}}X = \sigma_X^2 = \tilde{p}_1(1 - \tilde{p}_1) + 4\tilde{p}_2(1 - \tilde{p}_2) - 4\tilde{p}_1\tilde{p}_2,$$

where  $\tilde{p} = (\tilde{p}_0, \tilde{p}_1, \tilde{p}_2)^\top$  denotes genotype frequency estimate. Here the genotype frequencies do not need to comply with Hardy–Weinberg proportions (HWP). Note that this test can be applied to a binary trait as well as a quantitative phenotype, even when the sample is selected by extremes of phenotype [44]. The score statistic and its variance can be adapted to test other genetic models and for X-linked SNPs.

## 2.2 Score Test for Related Sample Using Genetic Correlation

When using multiplex cases from the same pedigree, we need to take into account the correlations. We define the correlation matrix  $\mathbf{K}$  for  $n$  subjects as follows:

$$\mathbf{K} = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1n} \\ \rho_{12} & 1 & \dots & \rho_{2n} \\ \vdots & \dots & \dots & \vdots \\ \rho_{1n} & \rho_{2n} & \dots & 1 \end{pmatrix}. \quad (3)$$

The off-diagonal entries,  $\rho_{ij}$ s, are twice the kinship coefficient between individuals  $i$  and  $j$  ( $i \neq j$ ). Here we use prior kinship coefficients [6]; these are based on expected identical by descent (IBD) sharing of randomly chosen alleles from each of the two individuals, and calculated from the pedigree structure.

Let  $\Sigma = \sigma_X^2 \mathbf{1}\mathbf{1}^\top$  be  $n \times n$  matrix where  $\mathbf{1}$  represents a vector of ones of length  $n$ . To take into account correlations induced using multiplex cases from the same pedigree, the expression of the variance of  $U_X$  in (2) can be replaced by

$$\text{Var}U_X = n^{-1}(Y - \bar{Y})^\top [\mathbf{K} \circ \Sigma](Y - \bar{Y}), \quad (4)$$

where  $\circ$  denotes the (Hadamard) termwise product. Because  $\Sigma$  is a covariance matrix, it must be symmetric and positive semidefinite and is invertible, and  $\mathbf{K}$  inherits these properties. Note that  $\mathbf{K}$  is invertible if and only if the sample does not include both members of any MZ twin pair [40]. Using this modified variance of  $U_X$ , under  $H_0$  the ratio  $U_X^2/\text{Var}U_X$  is chi-squared distributed with 1 degree of freedom [43].

We next describe methods to calculate correlation coefficients of  $\mathbf{K}$  in (4) for different genetic models and X-linked SNPs to modify the variance of the score test.

## 2.3 Correlation coefficients for Autosomal Loci

### 2.3.1 Additive Model

To calculate  $\rho_{ij}$  in (3), we consider the ITO method [22]. The ITO method uses three stochastic matrices: I, T, and O. Each row of each matrix corresponds to the conditional probability of genotype, given both genotype and IBD status, 2 IBD (I), 1 IBD (T), and 0 IBD (O), respectively. Let  $\rho_T$  be the correlation for the T-component, then the correlation between the  $i$ th and  $j$ th relatives for autosomal loci is

$$\rho_{ij} = \pi_2 + \pi_1 \rho_T \quad (5)$$

where  $\pi_k$  is the probability that the specified relatives should share  $k$  alleles IBD. Note that  $\rho_{ij}$  is twice the value of kinship coefficient, and  $\rho_{ij} = (1/2)^R$  where  $R$  is the degree of relationship [39]. For autosomal loci  $\rho_T$  equals to 1/2. Hence, for example, the correlations between a sib-pair are  $\rho_{\text{sib-pair}} = 1/4 + 1/2\rho_T = 1/2$ , and the correlations between double first cousins are  $\rho_{\text{double-first-cousins}} = 1/16 +$

$6/16\rho_T = 1/4$ . For computation of  $\mathbf{K}$  existing software such as MERLIN [1] or KinInbcoef [2] can be used. Because we also want to extend the multiplicative models to alternative ones and to test X-linked SNPs, we apply the approach of Li and Sacks. For an outbred population, the correlation matrix consists of sub-blocks of families.

### 2.3.2 Recessive and Dominant Model

For the recessive model the correlation coefficient is no longer independent of allele frequencies. Let  $\mathbf{K}_{rec}$  denote the correlation matrix for the recessive model containing  $\rho_{recij}$  off the diagonal. We denote  $\hat{p}$  the minor allele frequency estimate. The formula for the additive model (5) should be replaced by  $\rho_{recij} = \pi_2 + \pi_1\rho_{Trec}$ , where  $\rho_{Trec} = \hat{p}/(1 + \hat{p})$ , when random mating condition is satisfied. For example, the correlation of a sib-pair is

$$\rho_{recsib-pair} = \frac{1}{4} + \frac{1}{2} \left( \frac{\hat{p}}{1 + \hat{p}} \right) = \frac{1 + 3\hat{p}}{4(1 + \hat{p})}.$$

The expression in (4) is then  $\text{Var}U_X = n^{-1}(Y - \bar{Y})^\top [\mathbf{K}_{rec} \circ \Sigma](Y - \bar{Y})$  and the element of  $\Sigma$ ,  $\sigma_X^2$ , can be replaced by  $\tilde{p}_2(1 - \tilde{p}_2)$ , where  $\tilde{p}_2$  denotes the frequency of the 2/2 genotype. Analogously assuming complete dominance of one allele over the other allele the above can be applied by flipping the values 1 and 0 for the dominant model.

### 2.4 Correlation Coefficients for X-linked Loci

An X-linked analog of the autosomal correlation matrix (3) is denoted as  $\mathbf{K}_X$ . For an additive model  $\mathbf{K}_X$  can be calculated either by the algorithm used in MINX (“MERLIN in X”) [1] or by KinInbcoefX [2]. Deriving the correlation for an X-linked gene is exactly the same as for the autosomal correlation, except that there are four basic correlations,  $\rho_T$ , in (5). For an additive model we have

$$\rho_{T_{f,f}} = 1/2, \quad \rho_{T_{f,m}} = \rho_{T_{m,f}} = 1/\sqrt{2}, \quad \rho_{T_{m,m}} = 0,$$

where the subscripts indicate female pairs, mixed pairs, and male pairs, respectively. Thus, the correlation between full sisters with respect to sex-linked traits is

$$\rho_{X,sisters} = 1/2 + 1/2\rho_{T_{f,f}} = 3/4,$$

and the correlation for maternal uncle and niece is

$$\rho_{X,maternal-uncle-niece} = 1/4\rho_{T_{m,f}} = 1/4 * 1/\sqrt{2}.$$

For females we denote genotypes 0, 1, and 2, and genotypes of males are coded as 0 and 2 by treating males as homozygous females. The expectation of the genotype X is the same in both sexes, while the genotypic variance due to the marker on X

chromosome in males ( $\sigma_m^2$ ) is twice that in females ( $\sigma_f^2$ ). Expected covariances for X-linked are  $3/4\sigma_f^2$  for sister pairs, and

$$1/4\rho_{T_{m,f}} = 1/4 * 1/\sqrt{2} * \sigma_m * \sigma_f = 1/4\sigma_f^2$$

for maternal uncle–niece pair.

The score statistic and its variance can be extended to test recessive effects and for X-linked SNPs using the corresponding correlation matrices.

### 2.5 Testing for Association at Imputed SNPs

To deal with the uncertainty caused by imputation, we followed the same line of reasoning as in Uh et al. [42] for modifying the score test. Based on the statistical theory for missing data, the genotype data can be partitioned into two parts:  $X_{\text{comp}} = [X_{\text{obs}}, X_{\text{mis}}]$  [13,27]. The log likelihoods for the complete data ( $\ell_{\text{comp}}$ ) and observed (incomplete) data ( $\ell_{\text{obs}}$ ) are given by

$$\begin{aligned} \ell_{\text{comp}}(\theta) &= \log P(X_{\text{obs}}, X_{\text{mis}} | \theta), \\ \ell_{\text{obs}}(\theta) &= \log \int P(X_{\text{obs}}, X_{\text{mis}} | \theta) dX_{\text{mis}}. \end{aligned}$$

Let  $U(\theta)$  be the complete data score,  $\partial\ell_{\text{comp}}(\theta)/\partial\theta$ , and  $I(\theta)$  the complete data information,  $-\partial^2\ell_{\text{comp}}(\theta)/\partial^2\theta$ , respectively. Instead of observing  $X$ , for imputed genotypes the posterior probability,  $\varpi_i = (\varpi_{i0}, \varpi_{i1}, \varpi_{i2})$ , is given for subject  $i = 1, \dots, n$ . Let the expected dosage for the genotype counts of the  $i$ th individual be

$$\tilde{X}_i = EX_i = \varpi_{i1} + 2\varpi_{i2}. \tag{6}$$

Then we replace the genotype counts  $X$  by

$$U_{\tilde{X}} = (Y - \bar{Y})^\top \tilde{X} \tag{7}$$

in the score statistic (1). As before the score statistic  $U_{\tilde{X}}$  is asymptotically normally distributed under  $H_0$  with zero mean. Its variance can be determined as in Louis [27]:  $\text{Var}(U) = E(J) - [E(UU^\top) - UU^\top]$ , where  $J$  is the negative of the second derivative matrix for the complete data and where  $E(J)$  is the expected complete data information over the posterior distribution. The first term of  $\text{Var}(U)$  is the variance if there were no “missing data,” and the second term is the penalty for using imputed genotypes. We apply this variance computation to our score  $U_{\tilde{X}}$ .

Let  $\Sigma = \sigma_{\tilde{X}}^2 \mathbf{1}\mathbf{1}^\top$  be  $n \times n$  matrix with the genotypic variance  $\sigma_{\tilde{X}}^2$  where  $\mathbf{1}$  represents a vector of ones of length  $n$ . And, the  $n \times n$  matrix  $\Sigma_{\text{loss}}$  denotes the loss of information. Then, the score and information for the observed data likelihood are given by

$$\begin{aligned} U_{\text{obs}}(\theta) &= E_{X_{\text{mis}} | X_{\text{obs}}} U(\theta), \\ I_{\text{obs}}(\theta) &= E_{X_{\text{mis}} | X_{\text{obs}}} I(\theta) - \text{Var}_{X_{\text{mis}} | X_{\text{obs}}} U(\theta) = \Sigma - \Sigma_{\text{loss}}. \end{aligned}$$

Here, the term  $\text{Var}_{X_{\text{mis}} | X_{\text{obs}}}(\cdot)$  represents the loss of information due to imputation uncertainty.

The elements of  $\Sigma_{\text{loss}}$  are defined by the outer product of the square root of individual loss  $l_i$ ,

$$l_i = \varpi_{i1}(1 - \varpi_{i1}) + 4\varpi_{i2}(1 - \varpi_{i2}) - 4\varpi_{i1}\pi_{i2}.$$

Thus, on the diagonal we have  $\Sigma_{\text{loss};ii} = l_i$  and off the diagonal we have  $\Sigma_{\text{loss};ij} = \sqrt{l_i l_j}$  for  $i, j = 1, \dots, n$ . Then the variance of the score can be expressed as

$$\text{Var}_{X_{\text{obs}}} U_{\tilde{X}} = n^{-1}(Y - \bar{Y})^\top [\mathbf{K} \circ (\Sigma - \Sigma_{\text{loss}})](Y - \bar{Y}), \tag{8}$$

where  $\circ$  denotes the (Hadamard) termwise product. The relative efficiency measure for case control design of Uh et al. [42] can be used as a quality control measure for accuracy with respect to the association parameter:

$$R_T^2 = \frac{(Y - \bar{Y})^\top [\mathbf{K} \circ (\Sigma - \Sigma_{\text{loss}})](Y - \bar{Y})}{(Y - \bar{Y})^\top [\mathbf{K} \circ \Sigma](Y - \bar{Y})}. \tag{9}$$

Consequently with genotyped data  $\Sigma_{\text{loss}} = 0$ , hence  $R_T^2$  equals to 1. In contrast to the imputation accuracy such as  $r^2$  of MACH [24], which is a pre-analysis measure, this post-analysis information measure assigns more weights to associated SNPs [41].

### 2.6 Incorporation of Family History

Suppose that we have  $n + m$  individuals with  $n$  and  $m$  the number of individuals with non-missing and with missing genotype data at the given marker, respectively. We consider the case–control status of these individuals, and let  $Y_n$  and  $Y_m$  denote case–control status of genotyped and un-genotyped individuals, respectively. Observe that the MQLS test [40] also treats the genotype  $X$  as a random variable, and that the phenotype and family relationship are treated as the weight. The cases get a weight of 1 and population controls get weight of 0. Let  $\mathbf{K}_{n,m}$  denote  $n \times m$  matrix that give correlations between non-missing and missing individuals. Then we propose the new weight for the score test as

$$Y^* = Y_n + \mathbf{K}_{n,m} Y_m. \tag{10}$$

For example, when a case has two un-genotyped affected siblings, the weight is

$$Y^* = 1 + (1/2 \ 1/2) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2$$

Note that this weight is based only on phenotype and genetic relationship. Therefore, we can replace  $Y$  with the corresponding new weight  $Y^*$  in the score statistic

$$U_X^* = (Y^* - \bar{Y}^*)^\top X, \quad (11)$$

and in its variances

$$\text{Var}U_X^* = n^{-1}(Y^* - \bar{Y}^*)^\top [\mathbf{K} \circ \Sigma](Y^* - \bar{Y}^*).$$

### 3 Simulations

To evaluate the performance of the test statistics, the following ascertainment schemes were considered: one affected and two unaffected (ASP1), two affected and one unaffected (ASP2), and three affected siblings (ASP3). For testing we select 1 genotyped individual from each family as cases and use the score test, which is the Cochran-Armitage test for trend in proportions [4, 11]. In addition, we explore whether the power increases by including phenotypic information of un-genotyped relatives. When the type of families is uniform in all families—for example in ASP2 each family consists of 1 affected genotyped sibling and 2 additional phenotype of un-genotyped siblings, the both score tests with and without family information are equivalent. Therefore, we mix the three types of datasets (Mixed): 200 cases are selected from each of the 3 types of ascertainment schemes (ASP1, ASP2, and ASP3). In particular, we implemented the following simulation procedure:

- (a) For case population, parental genotypes are generated assuming random mating and Hardy–Weinberg equilibrium for minor allele frequencies of 0.01, 0.05, 0.07, 0.10, and 0.30.
- (b) Conditional on parental genotypes we generate genotypes of 3 offspring assuming Mendelian transmission.
- (c) Using the logistic model, disease indicators for offspring were generated. The model we considered was

$$\text{logit}(\mu_i) = \beta_0 + \beta_g x_{ig}, \quad (12)$$

where  $x_{ig}$  is the genotypic score for a diallelic gene. The parameter  $\beta_0$  denotes the intercept and was determined by  $\beta_0 = \text{logit}(K_b)$ , where  $K_b = 10\%$ , the baseline disease risk under  $H_0$ , is used. For evaluation of type 1 error rate, we simulate data under the null, so that the odds ratio,  $\exp(\beta_g)$ , was set to 1. For simulating data under alternative hypothesis the genotype effect in the logistic regression model in Eq. (12) was modeled as follows:  $\exp(\beta_g)$ , the odds ratio, was equal to 1.2 and 1.5.

- (d) Since we are dealing with a complex phenotype we also included some residual familial correlation due to the polygenic or environmental sources in our disease model. The broad sense heritability can be written as  $h^2 \sim \sigma_u^2 / (\sigma_u^2 + \sigma_E^2)$ , where the residual error term  $\sigma_E^2$  represents the non-genetic residual familial correlations. This  $E$  part quantifies components of variance on the logit scale rather than on the original scale of the (underlying) phenotype,  $\sigma_E^2$  can be approximated as

- $\exp(1) \approx 3$  assuming  $E \sim N(0, 1)$ . Consequently following the error distribution  $N(0, \sigma_u^2)$  and by setting  $\sigma_u^2$  equal to 1, we obtained the broad sense heritability  $h^2$  equal to 25 %. Similarly, by setting  $\sigma_u^2 = 0.5$ , we obtained  $h^2$  equal to 14 %.
- (e) For control population the steps (a) and (b) are performed with exception that in the step (b) only the genotypes of one individual are generated.
  - (f) Each sampling experiment under additive and recessive genetic model and various ascertainment schemes consisted of 1000 independent replicates: 500 cases 500 controls for the ascertainment schemes ASP1, ASP2, ASP3, and 600 cases and 600 controls for the Mixed ascertainment.
  - (g) For each replicate, we tested for an association using (i) the score test (ST) and (ii) where appropriate the score test that includes phenotypic information of ungenotyped relatives ( $ST_{fam}$ ).

For generating X-linked genotypes in the step 1, maternal genotypes are generated as above. For paternal genotypes, only one allele is generated, which follows the Bernoulli distribution; the other allele is fixed as Y, which contributes to determine the gender of offspring in the step 2.

### 3.1 Simulation Results

Type I error rates are reported for each statistic, where the type I error rate is the proportion of significant replicates at the nominal level of 5 % out of the total number of replicates. With 1000 replicates and a true type I error rate of 5 %, 95 % of the empirical estimates are expected to fall between 0.036 and 0.064. In Table 1 the results from the simulation are shown. The baseline prevalence rate has been assumed to be 10 %. For the additive model the type I error rate of the score test under the ascertainment schemes ASP1 and ASP2 was within the expected range of (0.036, 0.064), except the type I error rate of 0.35 under the ascertainment scheme of ASP3, 1 selected from 3 affected siblings. For the Mixed ascertainment, both tests showed appropriate type I error rate. For the recessive model, under each ascertainment scheme the type I error rate of the both tests was too conservative to test rare variants with MAF 5 %, or the corresponding frequency of the homozygotes 0.25 %.

Table 2 presents the results of power analysis: the score test (ST) and the score test including positive family history ( $ST_{fam}$ ). We simulated the same models as those found in Table 1, namely the same ascertainment schemes discussed previously. Since no statistic examined in Table 1 showed the inflated type I error rates, the comparisons of the estimated power for all simulated models are valid. The power was calculated as the proportion of significant replicates at the nominal level of 5 % out of the total number of replicates. From Table 2, in general selecting cases with positive family history is advantageous; power increased by selecting cases from multiple affected siblings. For the additive model, with weaker association ( $OR = 1.2$ ) the gain in power was greater when residual variance was small ( $h^2 = 14$  %). For the recessive model for rare variants (frequency of homozygote  $\leq 0.01$ ) there was little power to detect association, whereas the gain in power was remarkable under the ascertainment schemes that required a minimum of two affected individuals in each family. Table 2 also presents power results from the Mixed ascertainment scheme of 600 cases and

**Table 1** Empirical type I error rates: calculated by the proportion of significant replicates at the nominal level of 5 % out of 1000 simulations

Model	Ascertainment	Tests	Case/control	Under H0					
				Minor allele frequency (MAF)					
				0.01	0.05	0.07	0.1	0.3	
Additive	ASP1	ST <sup>a</sup>	500/500	0.050	0.050	0.046	0.046	0.053	
	ASP2			0.050	0.052	0.04	0.056	0.048	
	ASP3			0.049	0.035	0.049	0.036	0.035	
	Mixed	ST	600/600	0.050	0.046	0.052	0.049	0.049	
		ST <sub>fam</sub> <sup>b</sup>		0.044	0.044	0.047	0.049	0.045	
	Recessive				Frequency of homozygote <sup>c</sup>				
						0.0025	0.0049	0.010	0.090
		ASP1	ST	500/500	0.029	0.054	0.048	0.049	
		ASP2			0.020	0.045	0.052	0.050	
		ASP3			0.018	0.038	0.036	0.038	
Mixed		ST	600/600	0.027	0.051	0.049	0.048		
		ST <sub>fam</sub>		0.028	0.040	0.054	0.048		

<sup>a</sup> Score test<sup>b</sup> Score test including phenotypic information of un-genotyped relatives<sup>c</sup> Equal to square of MAF

600 controls, in which each 1/3 of cases were selected from ASP1, ASP2, and ASP3 families. For the rare homozygotes, type 1 error rate was too conservative to test the recessive effects. Overall the ST<sub>fam</sub> test was more powerful than the ST test.

Table 3 shows the results of power analysis for X-linked SNPs. Overall the results show the similar findings as for autosomal SNPs. While in case of rare variants (MAF = 0.01) our proposed tests (ST and ST<sub>fam</sub>) have little power; for more common variants the gain in power can be observed.

## 4 Applications to Real Data

### 4.1 The Leiden Longevity Study (LLS)

For association of X-linked SNPs and illustration of using imputed probabilities of the genotypes, we apply the proposed test to data from the Leiden Longevity Study (LLS) [17,41].

In the Leiden Longevity Study (LLS), long-lived families are investigated for parameters contributing to the longevity phenotype. Families were included if at least two long-lived siblings were alive and fulfilled the age criterion of 89 years or older for men and 91 years or older for women. In total, 944 long-lived proband siblings were included with a mean age of 94 years (range 89–104), 1671 offspring (mean age 61, range 39–81), and 744 partners (mean age 60, 36–79). Nonagenarian siblings were

**Table 2** Empirical power for autosomal SNPs: calculated as the proportion of significant replicates at the nominal level of 5 % out of the total number of replicates

Model	h <sup>2</sup>	Ascertainment	ST	Case/control	OR = 1.2	OR = 1.5	0.01	0.05	0.07	0.1	0.3	0.01	0.05	0.07	0.1	0.3	
Additive																	
Minor allele frequency (MAF)																	
25 %	ASP1	ST <sup>a</sup>	500/500	0.066	0.080	0.093	0.139	0.240	0.092	0.244	0.307	0.092	0.244	0.307	0.383	0.727	
				0.072	0.134	0.163	0.217	0.431	0.126	0.440	0.630	0.722	0.954				
				0.086	0.187	0.254	0.328	0.680	0.171	0.709	0.755	0.914	0.999				
	14 %	Mixed	ST	600/600	0.078	0.118	0.182	0.226	0.522	0.147	0.545	0.653	0.147	0.545	0.653	0.761	0.981
					0.082	0.149	0.212	0.287	0.558	0.136	0.579	0.695	0.837	0.992			
					0.049	0.115	0.117	0.157	0.301	0.039	0.127	0.218	0.325	0.624			
ST <sub>fam</sub> <sup>b</sup>	ASP1	ST	500/500	0.097	0.257	0.209	0.355	0.568	0.086	0.519	0.59	0.086	0.519	0.59	0.833	0.984	
				0.105	0.284	0.714	0.76	0.772	0.112	0.824	0.969	0.986	1				
				0.090	0.242	0.386	0.453	0.622	0.245	0.808	0.908	0.963	0.999				
Mixed	ASP2	ST	600/600	0.095	0.259	0.520	0.555	0.664	0.312	0.877	0.959	0.312	0.877	0.959	0.984	1	
				0.095	0.259	0.520	0.555	0.664	0.312	0.877	0.959	0.984	1				
				0.095	0.259	0.520	0.555	0.664	0.312	0.877	0.959	0.984	1				
Recessive																	
Frequency of homozygote <sup>c</sup>																	
25 %	ASP1	ST	500/500	0.0025	0.0049	0.01	0.09	0.0025	0.0049	0.01	0.09	0.0025	0.0049	0.01	0.09	0.185	
				0.032	0.043	0.063	0.139	0.001	0.006	0.025	0.185						
14 %	ASP2	ST	600/600	0.020	0.048	0.069	0.221	0.002	0.014	0.066	0.482	0.002	0.014	0.066	0.482		
				0.020	0.048	0.069	0.221	0.002	0.014	0.066	0.482						

Table 2 continued

Model	h <sup>2</sup>	Ascertainment	ST	Case/control	OR = 1.2	OR = 1.5						
		ASP3	ST		0.024	0.064	0.08	0.392	0.004	0.032	0.126	0.81
		Mixed	ST	600/600	0.045	0.060	0.082	0.264	0.088	0.107	0.193	0.795
			ST <sub>fam</sub>		.031	0.061	0.077	0.307	0.085	0.161	0.253	0.833
14 %		ASP1	ST	500/500	0.029	0.049	0.057	0.179	0.002	0.012	0.034	0.256
		ASP2			0.059	0.077	0.115	0.294	0.010	0.019	0.136	0.800
		ASP3			0.034	0.178	0.209	0.607	0.018	0.084	0.374	0.996
		Mixed	ST	600/600	0.046	0.101	0.131	0.419	0.129	0.178	0.44	0.974
			ST <sub>fam</sub>		0.051	0.153	0.159	0.452	0.167	0.238	0.559	0.995

<sup>a</sup> Score test

<sup>b</sup> Score test including phenotypic information of un-genotyped relatives

<sup>c</sup> Equal to square of MAF

**Table 3** Empirical power for X-linked SNPs: calculated by the proportion of significant replicates at the nominal level of 5 % out of 1000 simulations

Model	h <sup>2</sup>	Ascertainment	ST	Case/control	OR = 1.2			OR = 1.5					
					0.01	0.05	0.07	0.1	0.3	0.01	0.05	0.07	0.1
Additive													
Minor allele frequency (MAF)													
25 %	ASP1	ST <sup>a</sup>	500/500	0.021	0.039	0.073	0.108	0.280	0.064	0.342	0.418	0.598	0.901
	ASP2			0.023	0.068	0.163	0.217	0.431	0.126	0.440	0.630	0.722	0.954
	ASP3			0.037	0.163	0.222	0.198	0.576	0.166	0.750	0.851	0.940	1.000
	Mixed	ST	600/600	0.019	0.072	0.118	0.133	0.321	0.089	0.402	0.554	0.693	0.944
14 %	ASP1	ST <sup>b</sup>	500/500	0.045	0.086	0.180	0.158	0.412	0.141	0.559	0.710	0.821	0.986
	ASP2			0.011	0.034	0.042	0.052	0.107	0.032	0.107	0.164	0.195	0.440
	ASP3			0.033	0.110	0.164	0.237	0.544	0.127	0.707	0.830	0.907	0.997
	Mixed	ST	600/600	0.025	0.387	0.474	0.572	0.952	0.315	0.994	0.9979	1.000	1.000
		ST <sup>b</sup>		0.021	0.151	0.224	0.292	0.646	0.126	0.860	0.909	0.967	1.000
		ST <sup>b</sup>		0.026	0.212	0.325	0.403	0.807	0.228	0.957	0.974	0.995	1.000

<sup>a</sup> Score test

<sup>b</sup> Score test including phenotypic information of un-genotyped relatives

**Table 4** Results of association testing for the X-linked SNPs

Marker	Position	al1	al2	$N_0^a$	$N_1^a$	EA $F_0^b$	EA $F_1^b$	ST <sup>c</sup>	ST <sub>fam</sub> <sup>d</sup>
<b>rs12840872</b>	X:4942237	T	C	741	930	0.458	0.510	<b>8.08E–06</b>	<b>1.51E–07</b>
rs11094824	X 22212620	A	G	741	925	0.452	0.491	1.64E–02	2.86E–04
rs10855652	X:86416874	G	A	741	933	0.443	0.488	1.18E–02	5.51E–04

Bold numbers indicate approximate genome-wide significance

<sup>a</sup> Number of controls and cases

<sup>b</sup> Allele frequency of al2 in controls and cases

<sup>c</sup> Score test

<sup>d</sup> Score test including positive family history

genotyped using Illumina660W (Rotterdam, Netherlands) and their partners were genotyped using Illumina660W or OmniExpress (Estonina Biocentre, Genotyping Core Facility, Estonina). GenomeStudio was used for genotyping calling algorithm. Sample call rate was >95 %, and SNP exclusions criteria were Hardy–Weinberg equilibrium  $p$  value <  $10^{-4}$ , SNP call rate <95 %, and minor allele frequency <1 %. The number of the overlapping SNPs that passed quality controls in both samples was 296 K. To increase the overall coverage of the genome to 2.5 million SNPs, we imputed autosomal SNPs with HapMap (Haplotype Mapping Project, <http://www.hapmap.org>) release 22, build 36 of the CEU sample. The imputation program IMPUTE2 ([http://mathgen.stats.ox.ac.uk/impute/impute\\_v2.html](http://mathgen.stats.ox.ac.uk/impute/impute_v2.html)) was used.

For association of X-linked SNPs, we consider an affected sibling pair (ASP) and control design: 933 long-lived subjects from 420 nonagenarian sibling pairs are served as cases and 741 partners of offspring as controls. Additional phenotypic information (current age or age at death) for un-genotyped (or deceased) family members (parents or additional siblings) was included: 404 long-lived relatives. Here, to take into account different expectancies in the different birth cohorts, we used life tables for the Dutch population to define the longevity phenotype, the top 10 % of the specific birth cohorts [17]. In Table 4 the results are shown using an arbitrary threshold of  $p$  value <0.0001. From ca. 5900 SNPs on X chromosome tested, no SNP was found to be genome-wide significant. To determine approximate genome-wide significance thresholds as described in [16], the ratio of effective and actual number of tests (0.2) was used to estimate the number of the tests and to correct for multiple testing, namely  $p$  value =  $1\text{E}–05$ . Only rs12840872 exceeded the corrected significance level ( $p$  value =  $8.47\text{E}–06$ ). Further ST<sub>fam</sub> test that incorporates positive family history gave smaller  $p$  values, than the ST test.

Imputing SNPs that are not directly genotyped but are present on a reference panel such as the HapMap usually results in not one imputed value but three probabilities of the possible genotype value, 0, 1, or 2 for each individual. To deal with genotype uncertainty, one can choose the “best” genotype—genotype with the largest posterior probability, or one can use expected genotype counts (genotype dosages) as in (6). By calculating the variance of the score, there are again two options: incorporating uncertainty as in our method, or ignoring uncertainty, which is equivalent to setting  $\Sigma_{\text{loss}} = 0$  in (8). We consider three scenarios: (i) “best” guess of genotype and  $\Sigma_{\text{loss}} = 0$ , (ii) genotype dosage and  $\Sigma_{\text{loss}} = 0$ , and (iii) genotype dosage and incorporating

**Table 5** Comparison of methods to deal with uncertainty caused by genotype imputation

	$U_{\bar{X}}$ <sup>a</sup>	$\text{Var}_{X_{obs}} U_{\bar{X}}$ <sup>b</sup>	ST	$R_T^2$ <sup>c</sup>
(i) “best” guess & $\Sigma_{\text{loss}} = 0^d$	3.24	219.71	4.66E–01	1.00
(ii) genotype dosage & $\Sigma_{\text{loss}} = 0^d$	20.64	18.6	1.71E–06	1.00
(iii) genotype dosage & $\Sigma_{\text{loss}} \geq 0$	20.64	16.53	3.87E–07	0.78

<sup>a</sup> Equation (7)<sup>b</sup> Equation (8)<sup>c</sup> Post-analysis information measure as in Eq. (9)<sup>d</sup> Ignoring uncertainty

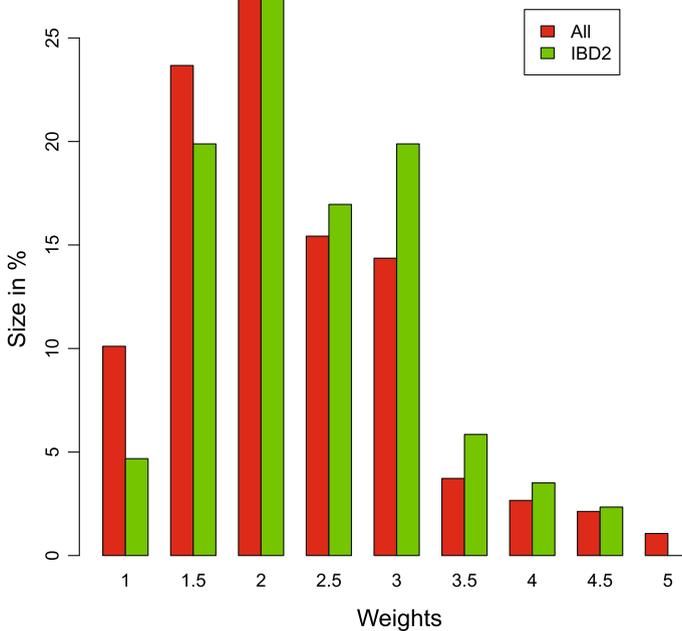
imputation uncertainty,  $\Sigma_{\text{loss}} \geq 0$  (8). For extensive simulation studies comparing these three approaches that deal with uncertainty in analysis of imputed genotypes, we refer to [20,47].

For illustration purpose, we chose an extreme case: a SNP, rs17183864 (chromosome 6), with MAF 1 % and the post-imputation measure  $r^2 = 0.3$  of MACH [24] or *info* of IMPUTE, which is the default threshold for ensuring imputation quality. As shown in Table 5, selecting the most likely genotype performs poorly, even though this specific SNP passed the pre-set imputation accuracy threshold,  $r^2 = 0.3$ . The results also shows that incorporating imputation uncertainty can be a more powerful method for testing imputed SNPs with corresponding post-analysis quality measure  $R_T^2 = 0.78$  in (9).

## 4.2 The Genetics, Arthrosis, and Progression (GARP) Study

The Genetics osteoARthritis and progression (GARP) study consists of 187 Caucasian sibling pairs and four trios of Dutch origin affected by symptomatic and radiographic OA at multiple sites [31]. Osteoarthritis (OA) is a common degenerative disease of the articulating joints with a considerable, but complex, genetic component. By performing a genome-wide linkage scan and combined linkage and association, the iodothyronine-deiodinase enzyme type 2 (D2) gene (*DIO2*) was identified as an osteoarthritis susceptibility gene. The common coding variant (rs225014; Thr92Ala) in the *DIO2* gene showed significant association. Information about the number of siblings and parents with similar symptoms was available: 30 % of the genotyped affected siblings have no missing (un-genotyped) affected siblings, 30 % have one missing affected sibling. The maximum number of missing affected siblings is 8 (one family). Regarding affected parents, 16 and 60 % of the ASPs had two and one affected un-genotyped parents, respectively. Using this extra information, Callegaro et al. proposed the allele-sharing statistics for genetic linkage analysis to account for the family history, and this considerably increased the evidence of linkage in the surrounding of the *DIO2* susceptibility locus [7].

The question arises whether this strategy could successfully be adapted to a genetic association study. We consider a case–control association study: the 380 cases from the GARP study and the control population from the Leiden Longevity Study described in the previous section [17]. The controls consist of 1671 offspring of nonagenarian



**Fig. 1** Weights in  $ST_{fam}$  for all GARP ASPs compared to ASPs with 2 IBD (Color figure online)

siblings (in 420 families) and 744 partners of the offspring. After exclusion of patients with OA and unknown status, 1947 subjects were served as controls.

Suggestive evidence for linkage was observed on chromosome 14q32.11 (75–95 cM; mean informativity = 0.50, [31]), wherein the genes *DIO2* (78 cM), *FLRT2* (82 cM), and *CALM1* (89 cM) were located. First, the linkage region of 75–95 cM was converted to the physical map, 79147631–95189589 bp, using a build GRCh37 genetic map ([ftp://ftp.ncbi.nlm.nih.gov/hapmap/recombination/2011-01\\_phaseII\\_B37/](ftp://ftp.ncbi.nlm.nih.gov/hapmap/recombination/2011-01_phaseII_B37/)). The GWAS data for both the GARP and LLS study were available, and the genotype imputation was performed using IMPUTE2 ([http://mathgen.stats.ox.ac.uk/impute/impute\\_v2.html](http://mathgen.stats.ox.ac.uk/impute/impute_v2.html)) based on the reference panel from the February 2012 release of the 1000 Genomes project (<ftp://ftp.1000genomes.ebi.ac.uk/>) and genome build 37. For the respective linkage region ca. 200K SNPs were available. We performed the association analysis, and the following filtering was employed: MAF in controls > 1%, IMPUTE info > 0.4, and  $R_2^T > 40\%$ . After filtering out insertions and deletions (indels), the resulting number of the SNPs tested was 57557. For the two study designs—(1) GARP cases ( $n = 380$ ) versus controls and (2) GARP cases with 2 IBD ( $n = 152$ ) versus controls—we compared the score test (ST) to the test that includes positive family history ( $ST_{fam}$ ). The weights in  $ST_{fam}$  in the expression (11) of the two samples are compared in Fig. 1. For the population with 2 IBD, larger weights are given corresponding to the increasing family size of affected members (Fig. 1). Using approximate genome-wide significance threshold of  $p$  value =  $1E-05$  as in the previous analysis [16], the  $p$  values in boldface are the significant SNPs with corresponding test that passed this threshold in Table 6. The SNPs with significant testing were reported, and

**Table 6** Results of association testing for the linkage region (75–95 cM)

Marker	chr	Position	al1	al2	EAF <sup>a</sup>		Info <sup>b</sup>	GARP vs. controls			
					Cases	Controls		ST <sup>c</sup>	$R_T^2$ <sup>d</sup>	ST <sub>fam</sub> <sup>e</sup>	$R_{T,fam}^2$
rs72688979	14	79826540	C	T	0.023	0.012	0.989	7.97E–03	59.9	<b>6.72E–08</b>	42.8
rs148126909	14	79826567	C	T	0.023	0.012	0.989	7.97E–03	59.9	<b>6.73E–08</b>	42.8
rs73337429	14	92866956	T	G	0.067	0.032	0.979	3.31E–05	75.3	<b>4.28E–06</b>	79.2
rs113235844	14	92876680	T	C	0.060	0.026	0.987	<b>8.53E–06</b>	79.3	1.03E–05	79.0
rs113272510	14	92894256	A	G	0.057	0.024	0.985	3.97E–06	73.6	<b>3.52E–06</b>	69.9

Bold numbers indicate approximate genome-wide significance

<sup>a</sup> Frequency of allele 2 (al2)

<sup>b</sup> Imputation quality measure by IMPUTE

<sup>c</sup> Score test of affected siblings and control design

<sup>d</sup> Relative efficiency measure (9) that reflects imputation uncertainty with respect to the association parameter

<sup>e</sup> Score test that incorporates positive family history

in this dataset except one SNP, testing that incorporates positive family history gave smaller  $p$  values.

Next we revisited the common coding variant, rs225014, in the *DIO2* gene. In [31], joint linkage and association analysis using GARP ASPs was performed to identify SNPs that explain the observed linkage signal by linkage and association modeling in pedigrees (LAMP, <http://csg.sph.umich.edu/LAMP>) [23]. A significant predisposing association with the C allele of *DIO2* SNP rs225014 was obtained ( $p$  value = 0.006). Moreover, allele frequencies in sibling pairs sharing two alleles identical by descent (IBD) at the *DIO2* locus (indicating those subjects that contribute to the linkage) were compared with allele frequencies of random controls—a random sample of unrelated subjects aged 55–65 years of the Rotterdam study [15]; the C allele of rs225014 was found significantly associated ( $p$  value = 0.025). For confirmation and replication in independent UK, Dutch, and Japanese OA studies, significant recessive association of the C-C haplotype of the *DIO2* SNPs rs12885300 and rs225014 with women with advanced symptomatic hip OA was found.

In the current association study as shown in Table 7, the minor allele frequency (MAF) of the cases (0.368) was almost the same as the MAF in the controls (0.364); observe that we have here a different control population, namely from LLS. Additionally, allele frequencies in sibling pairs sharing two alleles identical by descent (IBD) at the *DIO2* locus ( $n = 152$  in 85 families)—indicating those subjects that contribute to the linkage—were compared to those in controls, and MAF in cases increased to 0.428, which led to a smaller  $p$  value. For this specific SNP the magnitude of the  $p$  values of the two tests was comparable, which might indicate that a larger number of cases are needed to detect association.

## 5 Discussion

When testing for weak genetic effects, besides the efforts to increase sample size, it is desirable to develop statistical methods that more effectively detect an association.

**Table 7** Results of association study for rs225014

Design	<i>N</i> cases	<i>N</i> controls	MAF cases	MAF controls	ST <sup>a</sup>	ST <sub>fam</sub> <sup>b</sup>
(1) GARP vs. controls	380	1947	0.368	0.364	9.96E−01	9.08E−01
(2) GARP (IBD = 2) vs. controls	152	1947	0.428	0.364	1.08E−01	9.90E−02

<sup>a</sup> Score test of affected siblings and control design

<sup>b</sup> Score test that incorporates positive family history

Through simulation we showed that sampling genotyped cases from the high-risk families increases the power, even if only one case was sampled. To further increase the efficiency of the study, the multiplex-case and control design was considered, in which genotype frequencies between cases, each of whom was sampled from multiple affected families, and unrelated controls were compared. The primary reason for choosing this design would be that familial cases are enriched for genetic factors and therefore may be more informative for genetic research, particularly in the presence of genetic heterogeneity and phenocopies. For genetic linkage analysis Callegaro et al. proposed allele-sharing statistics using information on family history [7]. In the same spirit, we investigated the use of readily available family information for genetic association within a framework of a score test.

The first issue related to the proposed test is the ascertainment of the cases. The retrospective likelihood is considered to adjust for ascertainment by conditioning on the phenotypes of family members. For a general phenotype, a score statistic to test for an additive effect of a diallelic locus on phenotype is the genotype–phenotype covariance; the score statistic can be conveniently applied in both prospective and retrospective manners [10,28]. The score test has a flexible structure that allows testing of multiplicative, dominant, and recessive effects of specific genotype features on (disease) phenotype. Testing recessive effect is straightforward in a score test by using different dosage score  $(0, 0, 1)^T$  opposed to  $(0, 1, 2)^T$  for testing an additive model.

Secondly, the methods should allow for familial correlations due to sampling-related individuals. The variance of the score statistic can be readily modified to account for familial relationships based on kinship coefficients. To detect recessive effects of a SNP, however, the calculation of correlation coefficient depends on kinship coefficients as well as allele frequency. Another important extension is to construct a test for X-linked SNPs in mixed-sex-related samples. Combining the two statistics obtained by applying an allele-based test to males and females separately is not a valid test for family data. In this work we adapted the score test using X-linked correlation matrices.

Thirdly, to incorporate the extra phenotypic information, the weighting scheme similar to allele-sharing statistics for genetic linkage analysis of [7] is applied to the association testing. When the type of families is uniform in all families—for example each family consists of 2 sibling pairs and 2 additional phenotype information of ungenotyped parents, the weight is uniform and is equivalent to that of the (ordinary) score test. Using various weights other than the value 0 or 1, the proposed score test tends to the continuous weights in that for the quantitative trait.

As shown in Sect. 3 incorporating positive family history in the test statistic appeared to be a powerful strategy. Just by selecting one case from each family with positive

family history the power can be increased. This strategy of selecting cases with positive family history might be advantageous, when disease variant allele is rare and residual variance is small. In particular, for detection of recessive effects of rare variants, the increase in the power was remarkable under the ascertainment schemes that required a minimum of two affected individuals in each family. This finding supports previous results that greater power was achieved by sampling cases from multiplex pedigrees [34]. We also have shown improvement of the association results by including the number of affected relatives who were not genotyped in the score statistic through a real data example. Especially, the power of association using rare SNPs can be much improved by adding phenotypic information of un-genotyped family members. Currently the quality of genotype imputation using family data and its impact on the actual analysis is not yet clear [8]. When possible, utilizing additional information of extensive family members from the previous linkage studies is a viable option opposed to imputing un-genotyped family members.

Another benefit of employing our methods would be the use of the weight to pinpoint the extreme families for further investigation. The weights  $y$  implemented in the (ordinary) score statistic depend only on the individual's case–control status, 0 or 1, whereas the weights now can vary depending on the relationship configurations as well as on the phenotype of individuals. Applying these weights to GARP data showed that larger weights were assigned to ASPs with 2 IBD (Fig. 1). Extrapolating this idea, more efficient weights can be constructed as in [17]. For families selected for excessive survival, they computed the family-specific standard mortality ratio's (SMRs) to describe lifespan distributions of each generation within a family. Instead of discriminating between cases and controls based on these values they can be directly used as weights, which will induce more variabilities in the test statistic. Another extension would be the joint testing of multiple SNPs as described in [18]. Although our test is intended mainly for the multiplex cases, we have shown that this can equally be applied for related controls. A concern regarding the use of family-based samples for conducting case–control association analysis is the potential for population stratification effects. For well-designed studies this should be of minor concern; this effect can be controlled using genomic controls [14]. In addition, it is possible to use principal component analysis to correct for population stratification [32]. Top PCs can be modeled as covariates with the original outcome, and the residual, the result of removing any stratification effect on the outcome under the null, can be used for further analysis. When family structure is more complicated than siblings or parents such as in an isolated founder population, instead of using *naive* estimator of genotype frequency the best linear unbiased estimator [30] can be used in our score test, and the weight incorporating positive family history can be modified in the similar manner [40].

In the Genome-Wide Association Studies (GWAS) era, genotype imputation has become an essential tool. The imputation of genotypes allows investigators to test association at un-genotyped genetic markers and to combine results across studies that rely on different genotyping platforms. Combining the GWAS results for meta-analysis, using family data under strong ascertainment and using case–control data, can be inefficient and difficult. Here, our test can be of great use, since the score statistics from the individual studies can be easily combined. Another aspect of popularity of imputation is that more and more SNPs are imputed using external information

of reference haplotypes from the HapMap and 1000 Genomes Project. The ever-increasing number of tests to be performed calls for a tractable flexible approach such as score test.

In conclusion, incorporating positive family history in the test statistic appeared to be a powerful strategy. Especially, the power of association using rare SNPs can be much improved by adding phenotypic information of un-genotyped family members.

**Acknowledgments** The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7-Health-F5-2012) under Grant Agreement no. 305280 (MIMOMics). This study was supported by the Program Grant NOW 917.66.344 from the Netherlands Organization for Scientific Research, and by grant from the Innovation-Oriented Research Program on Genomics (Senter-Novem IGE05007), the Centre for Medical Systems Biology, the Netherlands Consortium for Healthy Ageing (Grant 050-060-810), all in the framework of the Netherlands Genomics Initiative, Netherlands Organization for Scientific Research (NWO), and BBMRI-NL (Biobanking and Biomolecular Resources Research Infrastructure).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

1. Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30(1):97–101. doi:[10.1038/ng786](https://doi.org/10.1038/ng786)
2. Abney M (2009) A graphical algorithm for fast computation of identity coefficients and generalized kinship coefficients. *Bioinformatics* 25(12):1561–1563. doi:[10.1093/bioinformatics/btp185](https://doi.org/10.1093/bioinformatics/btp185)
3. Antoniou AC, Easton DF (2003) Polygenic inheritance of breast cancer: implications for design of association studies. *Genet Epidemiol* 25(3):190–202. doi:[10.1002/gepi.10261](https://doi.org/10.1002/gepi.10261)
4. Armitage P (1955) Tests for linear trends in proportions and frequencies. *Biometrics* 11(3):375–386
5. Bourgain C, Hoffjan S, Nicolae R, Newman D, Steiner L, Walker K, Reynolds R, Ober C, McPeck MS (2003) Novel case-control test in a founder population identifies p-selectin as an atopy-susceptibility locus. *Am J Hum Genet* 73(3):612–626. doi:[10.1086/378208](https://doi.org/10.1086/378208)
6. Browning SR, Briley JD, Briley LP, Chandra G, Charnecki JH, Ehm MG, Johansson KA, Jones BJ, Karter AJ, Yarnall DP, Wagner MJ (2005) Case-control single-marker and haplotypic association analysis of pedigree data. *Genet Epidemiol* 28(2):110–122. doi:[10.1002/gepi.20051](https://doi.org/10.1002/gepi.20051)
7. Callegaro A, Meulenbelt I, Kloppenburg M, Slagboom PE, Houwing-Duistermaat JJ (2010) Allele-sharing statistics using information on family history. *Ann Hum Genet* 74(6):547–554. doi:[10.1111/j.1469-1809.2010.00602.x](https://doi.org/10.1111/j.1469-1809.2010.00602.x)
8. Chen MH, Huang J, Chen WM, Larson MG, Fox CS, Vasani RS, Seshadri S, O'Donnell CJ, Yang Q (2012) Using family-based imputation in genome-wide association studies with large complex pedigrees: the framingham heart study. *PLoS One* 7(12):e51589. doi:[10.1371/journal.pone.0051589](https://doi.org/10.1371/journal.pone.0051589)
9. Chow JC, Yen Z, Ziesche SM, Brown CJ (2005) Silencing of the mammalian x chromosome. *Ann Rev Genom Hum Genet* 6:69–92. doi:[10.1146/annurev.genom.6.080604.162350](https://doi.org/10.1146/annurev.genom.6.080604.162350)
10. Clayton D (2008) Testing for association on the x chromosome. *Biostatistics* 9(4):593–600. doi:[10.1093/biostatistics/kxn007](https://doi.org/10.1093/biostatistics/kxn007)
11. Cochran W (1954) Some methods of strengthening the common chi-square test. *Biometrics* 10:417–451
12. Cox DR, Hinkley DV (1974) *Theoretical statistics*. Chapman and Hall, London
13. Dempster A, Laird NM, Rubin DB (1977) Maximum-likelihood estimation from incomplete data via the em algorithm. *J R Stat Soc* 39:1–38
14. Devlin B, Roeder K (1999) Genomic control for association studies. *Biometrics* 55(4):997–1004
15. Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA (1991) Determinants of disease and disability in the elderly: the rotterdam elderly study. *Eur J Epidemiol* 7(4):403–422
16. Hoggart CJ, Clark TG, De Iorio M, Whittaker JC, Balding DJ (2008) Genome-wide significance for dense snp and resequencing data. *Genet Epidemiol* 32(2):179–185. doi:[10.1002/gepi.20292](https://doi.org/10.1002/gepi.20292)

17. Houwing-Duistermaat JJ, Callegaro A, Beekman M, Westendorp RG, Slagboom PE, van Houwelingen JC (2009) Weighted statistics for aggregation and linkage analysis of human longevity in selected families: the leiden longevity study. *Stat Med* 28(1):140–151. doi:[10.1002/sim.3421](https://doi.org/10.1002/sim.3421)
18. Kim S, Morris NJ, Won S, Elston RC (2010) Single-marker and two-marker association tests for unphased case-control genotype data, with a power comparison. *Genet Epidemiol* 34(1):67–77. doi:[10.1002/gepi.20436](https://doi.org/10.1002/gepi.20436)
19. Knight S, Uh HW, Martinez M (2009) Summary of contributions to gaw group 15: family-based samples are useful in identifying common polymorphisms associated with complex traits. *Genet Epidemiol* 33(Suppl 1):S99–104. doi:[10.1002/gepi.20480](https://doi.org/10.1002/gepi.20480)
20. Kutalik Z, Johnson T, Bochud M, Mooser V, Vollenweider P, Waeber G, Waterworth D, Beckmann JS, Bergmann S (2011) Methods for testing association between uncertain genotypes and quantitative traits. *Biostatistics* 12(1):1–17. doi:[10.1093/biostatistics/kxq039](https://doi.org/10.1093/biostatistics/kxq039)
21. Lasky-Su J, Won S, Mick E, Anney RJL, Franke B, Neale B, Biederman J, Smalley SL, Loo SK, Todorov A, Faraone SV, Weiss ST, Lange C (2010) On genome-wide association studies for family-based designs: an integrative analysis approach combining ascertained family samples with unselected controls. *Am J Hum Genet* 86(4):573–580. doi:[10.1016/j.ajhg.2010.02.019](https://doi.org/10.1016/j.ajhg.2010.02.019)
22. Li CC, Sacks L (1954) The derivation of joint distribution and correlation between relatives by the use of stochastic matrices. *Biometrics* 10:347–360
23. Li M, Boehnke M, Abecasis GR (2005) Joint modeling of linkage and association: identifying snps responsible for a linkage signal. *Am J Hum Genet* 76(6):934–949. doi:[10.1086/430277](https://doi.org/10.1086/430277)
24. Li Y, Abecasis G (2006) Mach 1.0: rapid haplotype reconstruction and missing genotype inference. *Am J Hum Genet* 79:S2290
25. Liang KY, Zeger S (1988) Longitudinal data analysis using generalized linear models. *Biometrika* 73:13–33
26. Loley C, Ziegler A, Koenig IR (2011) Association tests for x-chromosomal markers a comparison of different test statistics. *Hum Hered* 71:23–36. doi:[10.1159/000323518](https://doi.org/10.1159/000323518)
27. Louis TA (1982) Finding the observed information matrix when using the em algorithm. *J R Stat Soc* 2:226–233
28. Mantel N (1963) Chi-square tests with one degree of freedom; extensions of the mantel-haenszel procedure. *J Am Stat Assoc* 58:690–700
29. Marchini J, Howie B, Myers S, McVean G, Donnelly P (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 39(7):906–913. doi:[10.1038/ng2088](https://doi.org/10.1038/ng2088)
30. McPeck MS, Wu X, Ober C (2004) Best linear unbiased allele-frequency estimation in complex pedigrees. *Biometrics* 60(2):359–367. doi:[10.1111/j.0006-341X.2004.00180.x](https://doi.org/10.1111/j.0006-341X.2004.00180.x)
31. Meulenbelt I, Min EA (2008) Identification of *dio2* as a new susceptibility locus for symptomatic osteoarthritis. *Hum Mol Genet* 17:1867–1875. doi:[10.1093/hmg/ddn082](https://doi.org/10.1093/hmg/ddn082)
32. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38:904–909
33. Rabinowitz D, Laird N (2000) A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. *Hum Hered* 50:211–213
34. Risch N, Teng J (1998) The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases i. dna pooling. *Genome Res* 8(12):1273–1288
35. Sasiemi PD (1997) From genotypes to genes: doubling the sample size. *Biometrics* 53(4):1253–1261
36. Slager SL, Schaid DJ (2001) Evaluation of candidate genes in case-control studies: a statistical method to account for related subjects. *Am J Hum Genet* 68(6):1457–1462. doi:[10.1086/320608](https://doi.org/10.1086/320608)
37. Slager SL, Schaid DJ, Wang L, Thibodeau SN (2003) Candidate-gene association studies with pedigree data: controlling for environmental covariates. *Genet Epidemiol* 24(4):273–283. doi:[10.1002/gepi.10228](https://doi.org/10.1002/gepi.10228)
38. Sung YJ, Wang L, Rankinen T, Bouchard C, Rao DC (2011) Performance of genotype imputations using data from the 1000 genomes project. *Hum Hered* 73(1):18–25. doi:[10.1159/000334084](https://doi.org/10.1159/000334084)
39. Thomas DC (2004) *Statistical methods in genetic epidemiology*. Oxford University Press, New York
40. Thornton T, McPeck MS (2007) Case-control association testing with related individuals: a more powerful quasi-likelihood score test. *Am J Hum Genet* 81(2):321–337. doi:[10.1086/519497](https://doi.org/10.1086/519497)
41. Uh HW, Deelen J, Beekman M, Helmer Q, Rivadeneira F, Hottenga JJ, Boomsma DI, Hofman A, Uitterlinden AG, Slagboom PE, Boehringer S, Houwing-Duistermaat JJ (2011) How to deal with

- the early gwas data when imputing and combining different arrays is necessary. *Eur J Hum Genet* 20(5):572–576. doi:[10.1038/ejhg.2011.231](https://doi.org/10.1038/ejhg.2011.231)
42. Uh HW, Houwing-Duistermaat JJ, Putter H, van Houwelingen HC (2009) Assessment of global phase uncertainty in case-control studies. *BMC Genet* 10:54. doi:[10.1186/1471-2156-10-54](https://doi.org/10.1186/1471-2156-10-54)
  43. Uh HW, van der Wijk HJ, Houwing-Duistermaat JJ (2009) Testing for genetic association taking into account phenotypic information of relatives. *BMC Proc* 3 Suppl 7:S123
  44. Wallace C, Chapman JM, Clayton DG (2006) Improved power offered by a score test for linkage disequilibrium mapping of quantitative-trait loci by selective genotyping. *Am J Hum Genet* 78(3):498–504. doi:[10.1086/500562](https://doi.org/10.1086/500562)
  45. Whittemore AS (1995) Logistic regression of family data from case-control studies. *Biometrika* 82:57–67
  46. Zheng G, Joo J, Zhang C, Geller NL (2007) Testing association for markers on the x chromosome. *Genet Epidemiol* 31(8):834–843. doi:[10.1002/gepi.20244](https://doi.org/10.1002/gepi.20244)
  47. Zheng J, Li Y, Scheet GRAP (2011) A comparison of approaches to account for uncertainty in analysis of imputed genotypes. *Genet Epidemiol* 35(2):102–110
  48. Zhu Y, Xiong M (2012) Family-based association studies for next-generation sequencing. *Am J Hum Genet* 90(6):1028–1045. doi:[10.1016/j.ajhg.2012.04.022](https://doi.org/10.1016/j.ajhg.2012.04.022)