Gutenberg School of Management and Economics

& Research Unit “Interdisciplinary Public Policy”

Discussion Paper Series

Multiplicity- and dependency-adjusted p-values for control of the family-wise error rate

Jens Stange, Thorsten Dickhaus, Arcadi Navarro, Daniel Schunk

June 2015

Discussion paper number 1505

Johannes Gutenberg University Mainz
Gutenberg School of Management and Economics
Jakob-Welder-Weg 9
55128 Mainz
Germany
wiwi.uni-mainz.de

Contact details

Jens Stange
Weierstrass Institute for Applied Analysis and Stochastics
Mohrenstraße 39
10117 Berlin
Germany

stange@wias-berlin.de

Thorsten Dickhaus
Institute for Statistics
University of Bremen
P. O. Box 330 440
28344 Bremen
Germany

dickhaus@uni-bremen.de

Arcadi Navarro
Institute of Evolutionary Biology, Universitat Pompeu Fabra and
Institucio Catalana de Recerca i Estudis Avancats (ICREA) and
Center for Genomic Regulation (CRG)
Dr. Aiguader 88
08003 Barcelona
Spain

arcadi.navarro@upf.edu

Daniel Schunk
Department of Economics
Johannes-Gutenberg-Universität Mainz
Jakob-Welder-Weg 9
55128 Mainz

daniel.schunk@uni-mainz.de

Multiplicity- and dependency-adjusted p-values for control of the family-wise error rate

Jens Stange[a], Thorsten Dickhaus[b], Arcadi Navarro[c], Daniel Schunk[d]

[a] *Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstraße 39, 10117 Berlin, Germany*
[b] *University of Bremen, Institute for Statistics, P. O. Box 330 440, 28344 Bremen, Germany*
[c] *Institute of Evolutionary Biology, Universitat Pompeu Fabra and Institució Catalana de Recerca i Estudis Avançats (ICREA) and Center for Genomic Regulation (CRG), Dr. Aiguader, 88, 08003 Barcelona, Spain*
[d] *University of Mainz, Department of Economics, 55099 Mainz, Germany*

Abstract

We are concerned with the problem of testing multiple hypotheses simultaneously based on the same data and controlling the family-wise error rate. The multiplicity- and dependency-adjustment method (MADAM) is proposed which transforms test statistics into multiplicity- and dependency-adjusted p-values. The MADAM is closely connected with the concept of the "effective number of tests", but avoids certain inconveniences of the latter. For demonstration, we apply the MADAM to data from a genetic association study by exploiting computational methods for evaluating multivariate chi-square distribution functions.

Keywords: Bonferroni correction, dependency structure, effective number of tests, genetic epidemiology, multiple testing, probability approximations, Šidák correction
2010 MSC: 62J15, 62P10

Email addresses: stange@wias-berlin.de (Jens Stange), dickhaus@uni-bremen.de (Thorsten Dickhaus), arcadi.navarro@upf.edu (Arcadi Navarro), daniel.schunk@uni-mainz.de (Daniel Schunk)

1. Introduction

Dependency plays a crucial role in virtually all modern applications of high-dimensional data analysis, at least for two reasons. On the one hand, data generated with nowadays' high-throughput measurements typically exhibit strong temporal, spatial, or spatio-temporal dependencies due to the underlying (neuro-)biological or technological mechanisms. In biology, linkage disequilibrium for alleles and co-regulation for levels of expression of genes are two prominent examples. Hence, these dependencies have to be taken into account in any realistic statistical model for such data. On the other hand, such dependencies induce an intrinsically low-dimensional structure in the sample and/or the parameter space, thus facilitating or enabling valid statistical inference even for moderate sample sizes.

Here, we focus on the multiple testing context where $M > 1$ null hypotheses $H_1, \ldots, H_M$ are to be tested simultaneously based on one and the same data sample $x \in \mathcal{X}$. We assume that the considered multiple test procedure φ relies on test statistics $T_1, \ldots, T_M$ which are computed from x and compared with multiplicity-adjusted rejection thresholds. In prior work (see [6, 12]) it has been demonstrated that classical multiple testing approaches for control of the family-wise error rate (FWER) like the Bonferroni or the Šidák correction can be improved if the distribution of the vector $\mathbf{T} = (T_1, \ldots, T_M)^\top$ exhibits strong dependencies.[1] The possible relaxation of the necessary correction for multiplicity was described by the concept of the "effective number of tests" of order i, $M_{\text{eff.}}^{(i)}$ for short; see also Section 4.3.3 of [4]. Roughly speaking, $M_{\text{eff.}}^{(i)}$ approximates the number of stochastically independent tests which lead to the same FWER as φ. Hence, $M_{\text{eff.}}^{(i)}$ equals M if all components $T_1, \ldots, T_M$ are stochastically independent, and it equals one if $T_1, \ldots, T_M$ are totally dependent in the sense that all of them essentially assess exactly the same information from the data sample x. Computing $M_{\text{eff.}}^{(i)}$ for $1 \leq i \leq M$ requires knowledge of the i-variate (marginal) distributions of $\mathbf{T}$ which are then utilized in (sum- or product-type) probability approximations of order i. Hence, $M_{\text{eff.}}^{(i)}$ is typically decreasing in i, because more and more information about the dependency structure is exploited.[2]

We may mention here that the term "effective number of tests" has al-

[1] The FWER denotes the probability for at least one type I error among the M individual tests.

[2] Mathematical conditions guaranteeing that $M_{\text{eff.}}^{(i)}$ decreases with i are provided in [6].

ready been used for a longer time and seems to have its origins in the field of genetic epidemiology (see the corresponding references in [6]), but the foundations of this concept have to the best of our knowledge been made mathematically rigorous in [6] for the first time. Methods for computing $M_{\text{eff.}}^{(3)}$ in the genetic epidemiology context have been provided in [12] based on the theory of multivariate chi-square distributions; see also [5].

Although $M_{\text{eff.}}^{(i)}$ describes the quantitative effect of the dependencies in the data x on the FWER behaviour of φ in a transparent and straightforward manner, it has the undesirable property that it depends on the FWER level α. This is both counter-intuitive (the dependency structure is a feature only of the data sample x, not of the parameters of some method to analyze x) and inconvenient in practice, because iterative algorithms are required to match the probability approximation of order i and α for computing $M_{\text{eff.}}^{(i)}$. In the present work, we therefore introduce the multiplicity- and dependency-adjustment method of order i, MADAM$_i$ for short. The MADAM$_i$ transforms the vector $\mathbf{T}$ into a vector of p-values which are adjusted both for multiplicity and for i-th order dependency. Hence, these p-values are typically larger than their unadjusted, marginal counterparts, but smaller than the Bonferroni- or Šidák-corrected marginal p-values. In addition, MADAM$_i$ does not require the specification of α, thus avoiding the undesirable properties of $M_{\text{eff.}}^{(i)}$. However, both methods are closely related by the fact that they exploit the same probability approximations of order i.

The rest of the work is structured as follows. In Section 2, the MADAM is introduced and two different variants of it are illustrated. Section 3 shows how to utilize the MADAM for evaluating genetic association studies, considering a real-data example from this field. We conclude with a discussion in Section 4. Tables displaying the numerical results for the considered real-data example are deferred to Appendix A.

2. Statistical methodology: The MADAM

2.1. Notation and preliminaries

Throughout, we assume a statistical model $(\mathcal{X}, \mathcal{F}, (\mathbb{P}_\vartheta)_{\vartheta\in\Theta})$. The null hypotheses $H_1, \ldots, H_M$ are identified with non-empty subsets of the parameter space Θ. The intersection hypothesis $H_0 = \bigcap_{j=1}^M H_j$ is called the global hypothesis. For a given $\vartheta \in \Theta$, we will denote the index set of true null hypotheses in $\mathcal{H} = \{H_1, \ldots, H_M\}$ by $I_0 \equiv I_0(\vartheta) = \{1 \leq j \leq M : \vartheta \in H_j\}$.

A (non-randomized) multiple test is a measurable mapping $\varphi = (\varphi_j)_{1\le j\le M} : \mathcal{X} \to \{0,1\}^M$ the components of which have the usual interpretation of a statistical test for H_j versus K_j. The family-wise error rate of a multiple test φ is (for a given $\vartheta \in \Theta$) defined as

$$\mathrm{FWER}_\vartheta(\varphi) = \mathbb{P}_\vartheta\left(\bigcup_{j\in I_0(\vartheta)} \{\varphi_j = 1\}\right),$$

and φ is said to (strongly) control the FWER at a pre-specified level $\alpha \in (0,1)$ if $\sup_{\vartheta\in\Theta} \mathrm{FWER}_\vartheta(\varphi) \le \alpha$.

Under this general framework, we make the following assumption.

Assumption 1. *There exists a parameter value $\vartheta^* \in H_0$ such that*

$$\forall \vartheta \in \Theta : FWER_\vartheta(\varphi) \le FWER_{\vartheta^*}(\varphi). \tag{1}$$

Thus we may assume an overall null distribution $\mathbb{P} := \mathbb{P}_{\vartheta^*}$, under which all hypotheses are true, as the worst case with respect to control of the FWER.

2.2. Multiplicity- and dependency-adjusted p-values

We restrict our attention to simultaneous test procedures (STPs) in the sense of [9]. An STP φ is such that $\varphi_j(x) = 1 \iff T_j(x) > c_\alpha$, $1 \le j \le M$, $x \in \mathcal{X}$, for a given real constant c_α which in general depends on the FWER level α. As in Equation (1) of [5], a valid p-value for the marginal test problem H_j versus K_j corresponding to such an STP is given by

$$p_{\mathrm{ideal,j}}(x) = \mathbb{P}\left(\max_{1\le k\le M} T_k > t_j\right) = \mathbb{P}\left(\bigcup_{k=1}^{M}\{T_k > t_j\}\right) = 1-\mathbb{P}\left(\bigcap_{k=1}^{M}\{T_k \le t_j\}\right), \tag{2}$$

where $t_j = T_j(x)$ is the actually observed value of the j-th test statistic for the data sample x. We refer to $p_{\mathrm{ideal,j}}$ as the ideal p-value, because it takes the full joint distribution of $\mathbf{T}$ into account.

Feasible numerical methods for computing the ideal p-values only exist in a limited number of special model classes and for limited ranges of the total number M of tests, except from very time-consuming Monte Carlo approximations. For example, the `R`-package `mvtnorm` computes multivariate t- and normal probabilities up to dimension 1000, but not for higher dimensions. Hence, we propose to approximate $p_{\mathrm{ideal,j}}$ for $1 \le j \le M$ conservatively by

making use of probability bounds. Following Section 4.3 of [4], we refer to an upper bound of the form

$$\forall c \in \mathbb{R} : b^{(i)}(\mathbb{P}, c) \geq \mathbb{P}\left(\bigcup_{k=1}^{M}\{T_k > c\}\right) \tag{3}$$

as a sum-type probability bound of order i (STPB_i), if it takes the marginal distributions of $\mathbf{T}$ up to the i-th order into account. Typically, an STPB_i is obtained from a (higher-order) Bonferroni inequality. Analogously, we call a lower bound of the form

$$\forall c \in \mathbb{R} : \beta^{(i)}(\mathbb{P}, c) \leq \mathbb{P}\left(\bigcap_{k=1}^{M}\{T_k \leq c\}\right) \tag{4}$$

taking the marginal distributions of $\mathbf{T}$ up to the i-th order into account a product-type probability bound of order i (PTPB_i). Based on chain factorization, in [3] the authors considered

$$\beta^{(i)}(\mathbb{P}, c) = \mathbb{P}\left(\bigcap_{k=1}^{i}\{T_k \leq c\}\right)\prod_{k=i+1}^{M}\mathbb{P}\left(T_k \leq c \;\middle|\; \bigcap_{\ell=k-i+1}^{k-1}\{T_\ell \leq c\}\right). \tag{5}$$

It has to be mentioned that the right-hand side of (5) is not always an PTPB_i, because the inequality in (4) may be violated for special dependency structures in $\mathbf{T}$. However, in [12] it was demonstrated that $\beta^{(i)}(\mathbb{P}, c)$ from (5) often yields accurate approximations, even for $i = 3$, and the authors termed it a product-type probability approximation of order i (PTPA_i).

These considerations lead to the following definition of multiplicity- and dependency-adjusted p-values.

Definition 1. *The $MADAM_i$ transforms the values of the test statistics $T_1, \ldots, T_M$ into one of the following multiplicity- and dependency-adjusted p-values.*

$$p_{\Sigma,j}^{MADAM_i}(x) = b^{(i)}(\mathbb{P}, t_j), \tag{6}$$

$$p_{\Pi,j}^{MADAM_i}(x) = 1 - \beta^{(i)}(\mathbb{P}, t_j), \tag{7}$$

for all $1 \leq j \leq M$.

Obviously, neither $p_{\Sigma,j}^{\mathrm{MADAM}_i}$ nor $p_{\Pi,j}^{\mathrm{MADAM}_i}$ depend on α. In practice, however, one can reject H_j in favour of K_j if the j-th multiplicity- and dependency-adjusted p-value does not exceed α. The principle of quantile transformation entails that, under Assumption 1, this decision rule constitutes an FWER-controlling multiple test procedure.

3. An application to genetic data

3.1. Testing genetic association

In genetic association studies, a (potentially very large) number M of genetic markers are simultaneously tested for associations with a given phenotype. In the case that the markers are bi-allelic, they lead to diploid genotypes with three possible realizations per genomic position (locus). Typically, single nucleotide polymorphisms (SNPs) are considered in this context. If, in addition, the phenotype is binary (e. g., a disease indicator), many (2×3) contingency tables have to be evaluated simultaneously. This is a multiple test problem. Here, for illustration, we consider chromosome-wise multiplicity, meaning that the chromosomes are treated as independent units and the methods from Section 2 are applied to each of the 22 autosomes separately (sex chromosomes require a different statistical methodology).

In the sequel, we denote by M_C, $C \in \{1, \ldots, 22\}$, the different numbers of tests (considered loci) for chromosome C. For each $1 \leq j \leq M_C$, an association test based on the contingency table data $x^{(j)}$ (see Table 1) is carried out. In the terminology of Section 2, we thus consider for all $1 \leq j \leq M_C$ the null hypothesis

$$H_j = \{\text{There is no association between the phenotype and locus } j\}.$$

Notice that all quantities in Table 1 depend on the locus j, except for the row sums $n_{1.}$ and $n_{2.}$. This corresponds to the setup of a case-control study design; see [6], [8], and Chapter 9 of [4] for further details.

The null hypothesis H_j can be tested with Pearson's χ^2-test for independence (cf., e. g., Section 3.2.1 of [1]), employing the test statistic T_j, given by

$$T_j(x) = n \sum_{r=1}^{2} \sum_{c=0}^{2} \frac{(x_{rc}^{(j)} - n_{r.} n_{.c}^{(j)}/n)^2}{n_{r.} n_{.c}^{(j)}}, \tag{8}$$

where $x = (x^{(1)}, \ldots, x^{(M_C)})^\top$ denotes the entire data sample.

Genotype	$A_1^{(j)}A_1^{(j)}$	$A_1^{(j)}A_2^{(j)}$	$A_2^{(j)}A_2^{(j)}$	$\sum$
Cases	$x_{10}^{(j)}$	$x_{11}^{(j)}$	$x_{12}^{(j)}$	$n_{1.}$
Controls	$x_{20}^{(j)}$	$x_{21}^{(j)}$	$x_{22}^{(j)}$	$n_{2.}$
$\sum$	$n_{.0}^{(j)}$	$n_{.1}^{(j)}$	$n_{.2}^{(j)}$	n

Table 1: Genotype-phenotype counts at locus j aggregated in a (2×3)-contingency table. In the case of SNPs, the alleles $A_1^{(j)}$, $A_2^{(j)}$ are one of the nucleobases adenine (A), cytosine (C), guanine (G), or thymine (T). Cases correspond to the phenotypic value 1, while controls exhibit the phenotypic value 0.

If H_j is true, T_j is marginally asymptotically (with n tending to infinity) χ^2-distributed with two degrees of freedom. Notice, however, that there exist strong dependencies among the T_j, at least in blocks of markers which are in linkage disequilibrium (LD). Since LD can be regarded as external structural information (cf. [7]), the multivariate methods from Section 2 are a promising approach and typically more powerful than simple Bonferroni- or Šidák-corrections.

3.2. The MADAM for genetic association studies

For an approximation of $p_{\Pi,j}^{\mathrm{MADAM}_i}(x)$ from (7) for $i < M_C$, information about the i-variate (marginal) distributions of $\mathbf{T} = (T_1, \ldots, T_{M_C})^\top$ is required. Due to multivariate central limit theorems (see Section 4 in [6]), it suffices to consider the correlation (i. e., LD) matrix Σ_C of the M_C markers. This LD matrix can either be obtained from publicly available databases or can be estimated from the actual study data. For computational convenience, we propose to replace Σ_C by one of the following schemes.

a) Block thresholding: Submatrices of size $(b \times b)$ along the diagonal are kept, while all other entries are set to 0. This leads to the approximation

$$\tilde{\Sigma}_C = \begin{pmatrix} R_1 & 0 & \cdots & 0 \\ 0 & R_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & R_B \end{pmatrix}, \text{ where } B = M_C/b. \tag{9}$$

Since the inequality

$$\mathbb{P}\left(\bigcap_{i=1}^{M_C}\{T_i \leq x\}\right) \geq \mathbb{P}\left(\bigcap_{i=1}^{b}\{T_i \leq x\}\right)\mathbb{P}\left(\bigcap_{i=b+1}^{2b}\{T_i \leq x\}\right)\cdots\mathbb{P}\left(\bigcap_{(B-1)b+1}^{Bb}\{T_i \leq x\}\right) \tag{10}$$

holds true for all $x \geq 0$ due to the extended Gaussian correlation inequality proven in [11], the approximation

$$1 - \prod_{\ell=1}^{B}\mathbb{P}\left(\bigcap_{i=(\ell-1)b+1}^{\ell b}\{T_i \leq t_j\}\right) \geq p_{\text{ideal,j}} \tag{11}$$

yields a valid p-value. The final approximation $\tilde{p}_{\Pi,j}^{\text{MADAM}_i}(x)$ of $p_{\Pi,j}^{\text{MADAM}_i}(x)$ is obtained by applying (5) to every of the B factors in (11).

b) Neighbourhood thresholding: For every marker j, only one submatrix R_j of dimension $(b \times b)$ belonging to the $b-1$ loci adjacent to j is kept, while all other correlations are set to 0. This leads to the approximated LD matrices

$$\hat{\Sigma}_{C,j} = \begin{pmatrix} I_b & 0 & \cdots & & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & \cdots & R_j & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & & \cdots & & I_b \end{pmatrix}, \; j = 1, \ldots, M_C, \tag{12}$$

where I_b denotes the identity matrix in dimension $(b \times b)$. Again this approximation induces a valid p-value, because

$$\begin{aligned} p_{\text{ideal,j}} &\leq 1 - \prod_{i=1}^{j-b/2}\mathbb{P}(T_i \leq t_j)\;\mathbb{P}\left(\bigcap_{i=j-b/2+1}^{j+b/2}\{T_i \leq t_j\}\right)\prod_{i=j+b/2+1}^{M_C}\mathbb{P}(T_i \leq t_j) \\ &= 1 - \mathbb{P}\left(\bigcap_{i=j-b/2+1}^{j+b/2}\{T_i \leq t_j\}\right) F_{\chi^2_2}(t_j)^{M_C-b}. \end{aligned} \tag{13}$$

The final approximation $\hat{p}_{\Pi,j}^{\text{MADAM}_i}(x)$ of $p_{\Pi,j}^{\text{MADAM}_i}(x)$ is obtained by applying (5) to the probability expression in (13).

Obviously, the p-value $\tilde{p}_{\Pi,j}^{\mathrm{MADAM}_i}(x)$ yields a closer approximation of $p_{\mathrm{ideal,j}}$ than $\hat{p}_{\Pi,j}^{\mathrm{MADAM}_i}(x)$, because more information is kept. On the other hand, for every j one has to apply (5) B times in order to compute $\tilde{p}_{\Pi,j}^{\mathrm{MADAM}_i}(x)$, while one single application of (5) suffices to compute $\hat{p}_{\Pi,j}^{\mathrm{MADAM}_i}(x)$.

Remark 1. *The $MADAM_1$ is equal to the Šidák-correction which is typically used in case of stochastically independent test statistics. Thus, for both approximation schemes a) and b) from above it holds*

$$\tilde{p}_{\Pi,j}^{MADAM_1}(x) = \hat{p}_{\Pi,j}^{MADAM_1}(x) = p_{\check{S},j}(x) := 1 - F_{\chi_2^2}(t_j)^{M_C}.$$

3.3. Data analysis

For a numerical demonstration, we consider here a study comprising genotype data of $n = 2{,}729$ individuals. The number of markers under consideration varies between the chromosomes, ranging from $M_1 = 58{,}528$ SNPs on chromosome 1 to $M_{22} = 9{,}563$ SNPs on chromosome 22. In this study, most markers are found on the second chromosome with $M_2 = 61{,}103$. Further, for each individual six different behavioural phenotypes were assessed in the study. The data are stored in `PLINK`-formatted files (see [10]). Therefore, the first steps of data analysis were performed with the open-source software `PLINK`. For instance, with `PLINK` the pairwise correlations between markers were estimated. To this end, the definition of genotypic correlations as in [13] or Chapter 10 in [14] was used. Further computations were then performed with `MATLAB`, e. g., computation of the test statistics T_j. For the computation of the p-value approximations $\tilde{p}_{\Pi,j}^{\mathrm{MADAM}_3}(x)$ and $\hat{p}_{\Pi,j}^{\mathrm{MADAM}_3}(x)$, we employed `MATLAB` routines for the evaluation of two- and three-dimensional χ^2-distribution functions, which were developed in [12]. For the results reported in Appendix A we set the block size to $b = 100$ for scheme a), and we used block sizes $b = 100$ and $b = 200$ in scheme b). In certain cases, it may be possible that the dependency in the data extends beyond this block size. This would then lead to a slightly conservative multiple test procedure, meaning that the FWER level α is not exhausted. Every table in Appendix A contains the results for one of the six phenotypes. These tables illustrate the gain in power which is possible by applying the MADAM_3, compared with a univariate Šidák-correction.

4. Discussion

We have demonstrated how to apply sum- and product-type approximations of joint probabilities for the computation of multiplicity- and dependency-

adjusted p-values for control of the FWER. As these p-values incorporate parts of the correlation structure in the data, this leads to a better exhaustion of the nominal significance level, and thus to a more powerful multiple test procedure than common generic methods, which are typically conservative (not exhausting the FWER level α). Compared to previous work on effective numbers of tests, the main advantage of the MADAM is that it can be applied without relying on a pre-specified value of α, which also facilitates the computations (no iterative algorithms are necessary).

Since the methodology of effective numbers of tests has its origins in the field of genetic epidemiology and is to our knowledge mainly applied there, we illustrated the MADAM on such type of data. The p-values displayed in Tables A.1 to A.6 are adjusted for chromosome-wise multiplicity and block dependency. In some genetic association studies, however, one is interested in the genome-wide association test problem. In this context, one has to deal with very large values of $M \sim 10^5$ or $M \sim 10^6$, and FWER control is considered a too conservative criterion, even if multivariate methods are applied. Instead, for problems with such massive multiplicity, control of the false discovery rate (FDR, cf. [2]) has become a standard criterion. The development of multivariate methods controlling the FDR constitutes a vivid field of modern mathematical statistics. How to apply the MADAM in the context of FDR control is an interesting and challenging direction for future research. In this, bounds or approximations for expectations of ratios of dependent random variables are needed. A hybrid two-stage approach for the analysis of whole-genome or genome-wide association studies was recommended in [8] (see also the references in this article for earlier developments). In the first (screening) step of such a two-stage analysis, all M markers are tested for association employing a non-stringent type I error measure like the FDR in order to identify candidate SNPs. In the second (validation) step, these candidate SNPs are then tested on an independent data sample. In this confirmatory step the FWER is the appropriate type I error measure, and the MADAM can be applied on the reduced set of candidate markers which typically has an order of magnitude of 10^3, as considered in Section 3.

Acknowledgments

Support by the ERC "Foundations of Economic Preferences (FEP)" and by the Deutsche Forschungsgemeinschaft via grants DI 1723/3-1 und SCHU 2828/2-1 is gratefully acknowledged.

References

[1] Agresti, A., 2013. Categorical data analysis. John Wiley & Sons.

[2] Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Stat. Methodol. 57 (1), 289–300.

[3] Block, H. W., Costigan, T., Sampson, A. R., 1992. Product-type probability bounds of higher order. Probab. Eng. Inf. Sci. 6 (3), 349–370.

[4] Dickhaus, T., 2014. Simultaneous Statistical Inference with Applications in the Life Sciences. Springer-Verlag Berlin Heidelberg.

[5] Dickhaus, T., Royen, T., 2015. A survey on multivariate chi-square distributions and their applications in testing multiple hypotheses. Statistics 49 (2), 427–454.

[6] Dickhaus, T., Stange, J., 2013. Multiple point hypothesis test problems and effective numbers of tests for control of the family-wise error rate. Calcutta Statistical Association Bulletin 65 (257-260), 123–144.

[7] Dickhaus, T., Stange, J., Demirhan, H., 2014. On an extended interpretation of linkage disequilibrium in genetic case-control association studies. WIAS Preprint No. 2029, Weierstrass Institute for Applied Analysis and Stochastics Berlin, available at http://www.wias-berlin.de/preprint/2029/wias_preprints_2029.pdf.

[8] Dickhaus, T., Strassburger, K., Schunk, D., Morcillo-Suarez, C., Illig, T., Navarro, A., 2012. How to analyze many contingency tables simultaneously in genetic association studies. Stat. Appl. Genet. Mol. Biol. 11 (4), Article 12.

[9] Gabriel, K., 1969. Simultaneous test procedures - some theory of multiple comparisons. Ann. Math. Stat. 40, 224–250.

[10] Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., De Bakker, P. I., Daly, M. J., et al., 2007. Plink: a tool set for whole-genome association and population-based linkage analyses. The American Journal of Human Genetics 81 (3), 559–575.

[11] Royen, T., 2014. A simple proof of the Gaussian correlation conjecture extended to some multivariate gamma distributions. Far East J. Theor. Stat. 48 (2), 139–145.

[12] Stange, J., Loginova, N., Dickhaus, T., 2014. Computing and approximating multivariate chi-square probabilities. WIAS Preprint No. 2005, Weierstrass Institute for Applied Analysis and Stochastics Berlin, available at http://www.wias-berlin.de/preprint/2005/wias_preprints_2005.pdf.

[13] Wellek, S., Ziegler, A., 2008. A genotype-based approach to assessing the association between single nucleotide polymorphisms. Human heredity 67 (2), 128–139.

[14] Ziegler, A., König, I. R., Pahlke, F., 2010. A Statistical Approach to Genetic Epidemiology: Concepts and Applications, with an e-learning platform. John Wiley & Sons.

Appendix A. Tables

In Tables A.1 to A.6 the results for the most significant SNPs for each of the six phenotypes are displayed. Hereby, "Id" denotes the rs-identifier of the SNP, C is the corresponding chromosome with number of SNPs equal to M_C, T refers to the value of the chi-square test statistic, and p_{loc} denotes the marginal unadjusted p-value.

Id	C	M_C	T	p_{loc}
rs17009384	3	50864	25.068	3.6023632e-06
rs41368544	6	46044	24.242	5.4446072e-06
rs17076797	6	46044	23.920	6.3963029e-06
rs2683561	10	40184	23.906	6.4411680e-06
rs730242	16	22704	23.082	9.7226976e-06
rs1322990	9	35148	22.782	1.1298956e-05
rs6940980	6	46044	22.571	1.2554344e-05
rs9320543	6	46044	22.525	1.2844489e-05
rs4129267	1	58528	22.282	1.4507236e-05
rs9488718	6	46044	22.237	1.4832580e-05

Id	$p_{\check{\text{S}},j}$	$\hat{p}_{\Pi,j}^{\text{MADAM}_3}$ $(b = 100)$	$\hat{p}_{\Pi,j}^{\text{MADAM}_3}$ $(b = 200)$	$\tilde{p}_{\Pi,j}^{\text{MADAM}_3}$
rs17009384	0.1674241	0.1673261	0.1672423	0.1169289
rs41368544	0.2217381	0.2216367	0.2214924	0.1569391
rs17076797	0.2551052	0.2549915	0.2548290	0.1816032
rs2683561	0.2280479	0.2279864	0.2278656	0.1611671
rs730242	0.1980790	0.1978343	0.1976757	0.1431898
rs1322990	0.3277587	0.3276230	0.3274678	0.2378089
rs6940980	0.4390120	0.4388389	0.4386177	0.3243108
rs9320543	0.4464568	0.4462961	0.4460599	0.3303719
rs4129267	0.5721941	0.5719520	0.5717569	0.4296207
rs9488718	0.4948785	0.4947035	0.4944592	0.3704433

Table A.1: Results for the first phenotype.

Id	C	M_C	T	p_{loc}
rs3769489	2	61103	27.292	1.1847522e-06
rs6993816	8	40827	26.289	1.9557600e-06
rs2376823	8	40827	26.173	2.0734337e-06
rs17019407	3	50864	25.240	3.3052476e-06
rs9310805	3	50864	25.209	3.3561351e-06
rs7530457	1	58528	25.007	3.7128416e-06
rs3754801	2	61103	24.150	5.6995712e-06
rs1528989	7	39982	23.699	7.1416379e-06
rs3754802	2	61103	23.614	7.4540591e-06
rs1528992	7	39982	23.535	7.7533459e-06

Id	$p_{\check{\text{S}},j}$	$\hat{p}_{\Pi,j}^{\text{MADAM}_3}$ $(b = 100)$	$\hat{p}_{\Pi,j}^{\text{MADAM}_3}$ $(b = 200)$	$\tilde{p}_{\Pi,j}^{\text{MADAM}_3}$
rs3769489	0.0698338	0.0698100	0.0697732	0.0477690
rs6993816	0.0767432	0.0766966	0.0766562	0.0527880
rs2376823	0.0811682	0.0811173	0.0810779	0.0558606
rs17019407	0.1547462	0.1546838	0.1545886	0.1078663
rs9310805	0.1569312	0.1568707	0.1567687	0.1094256
rs7530457	0.1953160	0.1951931	0.1950919	0.1346670
rs3754801	0.2940860	0.2939937	0.2938540	0.2085338
rs1528989	0.2483902	0.2482458	0.2481095	0.1775373
rs3754802	0.3658483	0.3657439	0.3655750	0.2632018
rs1528992	0.2665497	0.2663758	0.2662585	0.1911207

Table A.2: Results for the second phenotype.

Id	C	M_C	T	p_{loc}
rs16872525	7	39982	25.753	2.5571790e-06
rs7628096	3	50864	22.707	1.1728766e-05
rs13000805	2	61103	22.659	1.2015884e-05
rs11216411	11	37115	22.330	1.4158464e-05
rs12794686	11	37115	22.263	1.4645703e-05
rs3757142	6	46044	22.220	1.4964335e-05
rs3757146	6	46044	22.220	1.4964335e-05
rs3757148	6	46044	22.220	1.4964335e-05
rs11683516	2	61103	22.060	1.6204048e-05
rs17731	10	40184	21.804	1.8422075e-05

Id	$p_{\check{\text{S}},j}$	$\hat{p}_{\Pi,j}^{\text{MADAM}_3}$ $(b = 100)$	$\hat{p}_{\Pi,j}^{\text{MADAM}_3}$ $(b = 200)$	$\tilde{p}_{\Pi,j}^{\text{MADAM}_3}$
rs16872525	0.0971883	0.0971335	0.0970476	0.0678973
rs7628096	0.4493057	0.4491249	0.4489139	0.3313491
rs13000805	0.5201161	0.5199587	0.5197999	0.3880328
rs11216411	0.4087375	0.4084143	0.4080645	0.2957558
rs12794686	0.4193338	0.4190808	0.4188688	0.3041531
rs3757142	0.4979336	0.4977042	0.4974477	0.3730113
rs3757146	0.4979336	0.4977117	0.4974542	0.3730113
rs3757148	0.4979336	0.4977055	0.4974467	0.3730113
rs11683516	0.6284694	0.6282879	0.6281430	0.4836979
rs17731	0.5230194	0.5228212	0.5225375	0.3933791

Table A.3: Results for the third phenotype.

Id	C	M_C	T	p_{loc}
rs4683625	3	50864	27.487	1.0745823e-06
rs13317804	3	50864	26.880	1.4560467e-06
rs9831276	3	50864	23.632	7.3854128e-06
rs4447734	3	50864	23.438	8.1376285e-06
rs11660040	18	21992	22.719	1.1657151e-05
rs7537401	1	58528	22.651	1.2062033e-05
rs11071658	15	21535	22.441	1.3395667e-05
rs7565497	2	61103	21.256	2.4232394e-05
rs6127200	20	19075	20.975	2.7878889e-05
rs6752766	2	61103	20.537	3.4712161e-05

Id	$p_{\check{\text{S}},j}$	$\hat{p}_{\Pi,j}^{\text{MADAM}_3}$ $(b = 100)$	$\hat{p}_{\Pi,j}^{\text{MADAM}_3}$ $(b = 200)$	$\tilde{p}_{\Pi,j}^{\text{MADAM}_3}$
rs4683625	0.0531907	0.0531663	0.0531312	0.0365894
rs13317804	0.0713844	0.0713546	0.0713059	0.0491987
rs9831276	0.3131594	0.3130476	0.3128598	0.2243635
rs4447734	0.3389422	0.3388155	0.3386058	0.2440722
rs11660040	0.2261410	0.2258627	0.2255321	0.1599944
rs7537401	0.5063709	0.5061836	0.5059677	0.3733004
rs11071658	0.2505964	0.2503414	0.2501048	0.1805887
rs7565497	0.7725193	0.7723564	0.7721810	0.6269534
rs6127200	0.4124519	0.4117301	0.4110675	0.3040814
rs6752766	0.8800948	0.8799602	0.8798123	0.7556091

Table A.4: Results for the fourth phenotype.

Id	C	M_C	T	p_{loc}
rs10829295	10	40184	30.995	1.8604455e-07
rs7083092	10	40184	30.995	1.8604455e-07
rs893218	17	16934	29.958	3.1235982e-07
rs7213761	17	16934	29.958	3.1235982e-07
rs7069754	10	40184	29.131	4.7240208e-07
rs4575326	12	35701	26.763	1.5436871e-06
rs4678160	3	50864	26.633	1.6471182e-06
rs5766192	22	9563	25.687	2.6430735e-06
rs1252069	1	58528	24.759	4.2044287e-06
rs7799805	7	39982	24.347	5.1647975e-06

Id	$p_{\check{\text{S}},j}$	$\hat{p}_{\Pi,j}^{\text{MADAM}_3}$ $(b = 100)$	$\hat{p}_{\Pi,j}^{\text{MADAM}_3}$ $(b = 200)$	$\tilde{p}_{\Pi,j}^{\text{MADAM}_3}$
rs10829295	0.0074481	0.0074409	0.0074329	0.0051403
rs7083092	0.0074481	0.0074413	0.0074334	0.0051403
rs893218	0.0052755	0.0052678	0.0052592	0.0037344
rs7213761	0.0052755	0.0052679	0.0052589	0.0037344
rs7069754	0.0188040	0.0187855	0.0187669	0.0129383
rs4575326	0.0536201	0.0535933	0.0535609	0.0375648
rs4678160	0.0803656	0.0803373	0.0802975	0.0554468
rs5766192	0.0249590	0.0249108	0.0248544	0.0181650
rs1252069	0.2181382	0.2180146	0.2179217	0.1510069
rs7799805	0.1865733	0.1864735	0.1863761	0.1320074

Table A.5: Results for the fifth phenotype.

Id	C	M_C	T	p_{loc}
rs17053752	6	46044	32.771	7.6527821e-08
rs9397537	6	46044	30.069	2.9559073e-07
rs6685470	1	58528	29.366	4.1996667e-07
rs2240291	7	39982	29.284	4.3767486e-07
rs12349952	9	35148	29.197	4.5704580e-07
rs10978953	9	35148	29.188	4.5917396e-07
rs17168107	7	39982	28.979	5.0968202e-07
rs9384020	6	46044	28.703	5.8516935e-07
rs10973251	9	35148	28.336	7.0293605e-07
rs16945357	16	22704	27.453	1.0929746e-06

Id	$p_{\check{\text{S}},j}$	$\hat{p}_{\Pi,j}^{\text{MADAM}_3}$ $(b = 100)$	$\hat{p}_{\Pi,j}^{\text{MADAM}_3}$ $(b = 200)$	$\tilde{p}_{\Pi,j}^{\text{MADAM}_3}$
rs17053752	0.0035174	0.0035152	0.0035136	0.0024504
rs9397537	0.0135180	0.0135073	0.0134978	0.0093328
rs6685470	0.0242802	0.0242713	0.0242584	0.0163655
rs2240291	0.0173469	0.0173338	0.0173231	0.0120415
rs12349952	0.0159359	0.0159100	0.0158855	0.0110753
rs10978953	0.0160095	0.0159834	0.0159589	0.0111263
rs17168107	0.0201719	0.0201573	0.0201438	0.0140009
rs9384020	0.0265838	0.0265626	0.0265445	0.0183368
rs10973251	0.0244041	0.0243886	0.0243655	0.0169556
rs16945357	0.0245096	0.0244535	0.0244100	0.0173930

Table A.6: Results for the sixth phenotype.