BIOINFORMATICS ORIGINAL PAPER

Vol. 31 no. 2 2015, pages 151–157
doi:10.1093/bioinformatics/btu629

Genome analysis

Advance Access publication September 23, 2014

METAINTER: meta-analysis of multiple regression models in genome-wide association studies

Tatsiana Vaitsiakhovich[1,2,*], Dmitriy Drichel[2], Christine Herold[2], André Lacour[2] and Tim Becker[1,2]

[1]Institute for Medical Biometry, Informatics and Epidemiology, University of Bonn and [2]German Center for Neurodegenerative Diseases (DZNE), Sigmund-Freud-Str. 25, D-53105 Bonn, Germany

Associate Editor: John Hancock

ABSTRACT

Motivation: Meta-analysis of summary statistics is an essential approach to guarantee the success of genome-wide association studies (GWAS). Application of the fixed or random effects model to single-marker association tests is a standard practice. More complex methods of meta-analysis involving multiple parameters have not been used frequently, a gap that could be explained by the lack of a respective meta-analysis pipeline. Meta-analysis based on combining *p*-values can be applied to any association test. However, to be powerful, meta-analysis methods for high-dimensional models should incorporate additional information such as study-specific properties of parameter estimates, their effect directions, standard errors and covariance structure.

Results: We modified 'method for the synthesis of linear regression slopes' recently proposed in the educational sciences to the case of multiple logistic regression, and implemented it in a meta-analysis tool called METAINTER. The software handles models with an arbitrary number of parameters, and can directly be applied to analyze the results of single-SNP tests, global haplotype tests, tests for and under gene–gene or gene–environment interaction. Via simulations for two-single nucleotide polymorphisms (SNP) models we have shown that the proposed meta-analysis method has correct type I error rate. Moreover, power estimates come close to that of the joint analysis of the entire sample. We conducted a real data analysis of six GWAS of type 2 diabetes, available from dbGaP (http://www.ncbi.nlm.nih.gov/gap). For each study, a genome-wide interaction analysis of all SNP pairs was performed by logistic regression tests. The results were then meta-analyzed with METAINTER.

Availability: The software is freely available and distributed under the conditions specified on http://metainter.meb.uni-bonn.de

Contact: vait@imbie.meb.uni-bonn.de

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on May 20, 2014; revised on September 12, 2014; accepted on September 16, 2014

1 INTRODUCTION

The need to summarize the results of related genome-wide association studies (GWAS) has encouraged the rapid development of new meta-analytic methods and tools, see e.g. (Liu *et al.*, 2013b; Mägi and Morris, 2010; Tang and Lin, 2013; Willer *et al.*, 2010). Meta-analysis of summary statistics is expected to deliver new susceptibility loci without additional genotype and phenotype data being required (de Bakker *et al.*, 2010). Application of the general fixed effects model (Becker, 1994; Cochran, 1937) to single-marker 1 degree of freedom (df) association tests has become a standard practice. More complex models involving multiple parameters have been used much less frequently, presumably in view of power issues but also because of the absence of a respective meta-analysis pipeline. The need for meta-analysis multiple parameter methods has been recognized in the field (Evangelou and Ioannidis, 2013), and new approaches for performing meta-analysis of rare variant association tests have recently been described in (Liu *et al.*, 2013b).

Fisher's combination test (Fisher, 1932), Stouffer's method (or inverse normal method; Stouffer *et al.*, 1949) and Stouffer's method with weights (Becker, 1994; Lipták, 1959; Zaykin, 2011) are common approaches for combining *p*-values. However, by design, they do not take advantage of additional information available in high-dimensional models. Results of multiple regression analysis, such as parameter estimates, their standard error and covariance, have rarely been used in meta-analysis owing to the complexities underlying the process of their synthesis. In this context, a method for the synthesis of linear regression slopes has been proposed in the educational sciences (Becker and Wu, 2007). The method has been recognized to be useful in Genetic Epidemiology and has been adapted for linear regression 2 df gene–environment tests in (Manning *et al.*, 2011). The method by (Becker and Wu, 2007) involves model parameter estimates and their correlation and provides the overall meta-analytic estimates of regression slopes together with meta-analysis *p*-values.

Here, we elaborate the method by (Becker and Wu, 2007) for multiple logistic regression model. We introduce the analysis tool METAINTER, which, besides Fisher's and Stouffer's methods, implements the method by (Becker and Wu, 2007) for an arbitrary number of model parameters. Thereby, METAINTER enables meta-analysis, for instance, of the single-marker 2 df association test, global haplotype tests and tests for and under gene–gene or gene–environment interaction. The synthesis of regression slopes (Becker and Wu, 2007) relies on the availability of the covariance matrix of the model parameters. Because the covariance matrix is not always provided by genetic analysis tools, we support the analysis framework presented here with an update of our own genetic interaction

*To whom correspondence should be addressed.

analysis tool INTERSNP (Herold *et al.*, 2009) to avoid the potential unavailability of the covariance matrix. We conducted a simulation study for two-single nucleotide polymorphisms (SNP) models and evaluated the data with 8 df logistic regression test under interaction and 4 df logistic regression test for interaction. We show that the method for the synthesis of regression slopes (MSRS) has correct type I error. Moreover, its power is nearly equal to that of the analysis of the joint sample with covariate parameters indicating sub-study membership. Furthermore, we performed a real data analysis of six GWAS of type 2 diabetes (T2D), available from dbGaP (http://www.ncbi.nlm.nih.gov/gap). At first, we conducted a genome-wide interaction analysis (GWIA) with 8 and 4 df logistic regression tests in each study, and then meta-analyzed the results with METAINTER.

2 METHODS

To model genetic effects of genotypes *AA*, *Aa* and *aa* of a single SNP on a quantitative trait, a standard linear regression equation $y=\beta_0+\beta x+\beta_D x_D$ is used, where y is the outcome variable, x, x_D are two genetic predictor variables corresponding to the additive and the dominance effects of a SNP and coded according to the number of copies of the susceptibility allele A. The coding scheme $x=1, 0, -1$ and $x_D=-0.5$, $0.5, -0.5$ is used in our work for the genotypes *AA*, *Aa*, *aa*, respectively (Cordell and Clayton, 2002). The model parameters β_0, β, β_D have to be estimated.

To model genetic effects on a qualitative trait in e.g. case/control studies, a logistic regression model is used. Let p denote the probability of expressing a phenotype, $0<p<1$. In the logistic regression, the logarithm of odds $\log\frac{p}{1-p}=:\text{logit}\,p$ is modeled as $\text{logit}\,p=\beta_0+\beta x+\beta_D x_D$ with the same predictor variables as before.

The logistic regression model for a single locus can easily be extended to a two-SNP model. Moreover, it can be modified for modeling pair-wise statistical interaction. A full logistic regression model, including marginal effects and allowing for pair-wise interaction, is represented by

$$\begin{aligned}\text{logit}\,p=&\beta_0+\beta_1x_1+\beta_2x_2+\beta_{1D}x_{1D}+\beta_{2D}x_{2D}\\&+\gamma_{12}x_1x_2+\gamma_{1,2D}x_1x_{2D}+\gamma_{1D,2}x_{1D}x_2+\gamma_{1D,2D}x_{1D}x_{2D}.\end{aligned}\quad(1)$$

This equation models additive x_i, and dominance effects x_{iD}, $i=1, 2$, of two SNPs, as well as interaction effects between them. The coding scheme is as before. The parameters β_i, β_{iD}, represent the logarithm of the genotype relative risks at the locus i, $i=1, 2$, whereas γ_{12}, $\gamma_{1,2D}$, $\gamma_{1D,2}$, $\gamma_{1D,2D}$ reflect the magnitude of the interaction effects. Testing (1) against the baseline $\text{logit}\,p=\beta_0$ yields an 8 df test *under* interaction, i.e. including marginal effects. Testing (1) against the reduced likelihood $\text{logit}\,p=\beta_0+\beta_1x_1+\beta_2x_2+\beta_{1D}x_{1D}+\beta_{2D}x_{2D}$ yields a 4 df test *for* interaction. For more details on logistic regression models see e.g. (Cordell and Clayton, 2002; Chapman and Clayton, 2007).

2.1 *P-value* combination methods

Methods of combining p-values from independent significance tests have a long history, the most popular of them are listed, for instance, in the review article (Becker, 1994), see also (de Bakker *et al.*, 2010; Zaykin, 2011). When raw data cannot be pooled across studies, meta-analysis based on combination of p-values represents a convenient and feasible approach.

We consider three p-value combination methods. Assume that k studies (sub-studies of the meta-analysis) were conducted to test a particular hypothesis H_1 versus fixed H_0. In two of three methods, the only information required to perform meta-analysis is p-values p_j of each sub-study j and possibly sample size n_j, $j=1,\ldots,k$. It is well-known that meta-analysis generally benefits from weighting. Carefully chosen weights such as, for instance, square root of sample size may improve power of meta-analysis methods, see (Zaykin, 2011) and references therein.

Fisher's method (Fisher, 1932). Test statistic of the Fisher's method has the form $T=-2\sum_{j=1}^{k}\log p_j$ and is χ^2-distributed under H_0 with $2k$ df.

Stouffer's method with weights (Lipták, 1959). A combined p-value is computed as $p=1-\Phi(Z)$, with $Z=\sum_{j=1}^{k}w_j\Phi^{-1}(1-p_j)/\sqrt{\sum_{j=1}^{k}w_j^2}$, where Φ and Φ^{-1} denote the standard normal cumulative distribution function and its inverse, and where w_j is study-specific weight, e.g. $w_j=\sqrt{n_j}$, $j=1,\ldots,k$, see (Zaykin, 2011).

Stouffer's method with weights and effect directions. In case of multiple regression model, the Stouffer's method with weights can be modified to include the information on the consistency of effect directions across sub-studies. Assume that the same regression model is used in all sub-studies, and consider two of them, sub-studies 1 and 2. We assume that a model equation

$$\text{logit}\,Y=\beta_0+\beta_1X_1+\ldots+\beta_PX_P\quad(2)$$

with p predictor variables is tested versus $\text{logit}\,Y=\beta_0$. To compare the effect directions of two studies we suggest the following ad hoc criterion: two studies are said to have the same effect direction if and only if the dot product of two vectors $\left[\hat\beta_{11}/\text{se}(\hat\beta_{11}),\ldots,\hat\beta_{1P}/\text{se}(\hat\beta_{1P})\right]$ and $\left[\hat\beta_{21}/\text{se}(\hat\beta_{21}),\ldots,\hat\beta_{2P}/\text{se}(\hat\beta_{2P})\right]$ in the Euclidian space $\mathbb{R}^P$ is positive. Here, $\hat\beta_{jl}$ are estimates of β_l with the standard error $\text{se}(\hat\beta_{jl})$ in sub-study j, $j=1, 2$, $l=1,\ldots,P$. Our definition is motivated by the geometrical interpretation of the regression models, where the values of the predictor variables and the corresponding outcomes are considered to be a cloud of points in the Euclidian space; for more details see (Pestman, 2009). Note also that the simulation study below supports the validity of the suggested criterion. In case, the model equation (2) is tested versus the restricted model $\text{logit}\,Y=\beta_0+\beta_1X_1+\ldots+\beta_SX_S$, $1\le S<P$, one has to address the predictor estimates $\hat\beta_{jl}$ with $j=1, 2$ and $l=S+1,\ldots,P$.

According to the Stouffer's method with weights and effect directions, a combined p-value can be found as $p=2(1-\Phi(|Z|))$, with

$$Z=\sum_{j=1}^{k}\delta_{j1}w_j\Phi^{-1}(1-p_j/2)/\sqrt{\sum_{j=1}^{k}w_j^2},$$

where $\delta_{j1}=1$, if sub-studies j and 1 have the same direction, and $\delta_{j1}=-1$, otherwise.

2.2 MSRS: method for the synthesis of regression slopes

A new meta-analytic approach based on multivariate generalized least squares estimation has recently been suggested in (Becker and Wu, 2007) to summarize the results of multiple linear regression used in related studies. We modified the method to the case of multiple logistic regression and applied our modification to the meta-analysis of case/control GWAS.

The method in (Becker and Wu, 2007) can briefly be described as follows. Let k be the number of sub-studies, where the multiple linear regression model

$$Y=\beta_0+\beta_1X_1+\ldots+\beta_PX_P\quad(3)$$

was tested versus $Y=\beta_0$ to find the effect of P predictor variables on the outcome variable Y. Note that covariates are allowed to be included. The aim of the meta-analysis is to combine the results of k sub-studies and to obtain the common estimate of the slope vector $\boldsymbol{\beta}=[\beta_1,\ldots,\beta_P]'$. Let $[\hat\beta_{j1},\ldots,\hat\beta_{jP}]'$ be an estimate vector of $\boldsymbol{\beta}$ in sub-study j, $j=1,\ldots,k$. The k sub-studies' estimate vectors are stacked to a vector $\mathbf{b}=[\hat\beta_{11},\ldots,\hat\beta_{1P},\ldots,\hat\beta_{k1},\ldots,\hat\beta_{kP}]'$. The method relies mutually on the availability of the covariance matrix

$\boldsymbol{\Sigma}_j = \mathrm{cov}([\hat{\beta}_{j1}, \ldots, \hat{\beta}_{jP}]')$ or at least its estimate $\hat{\boldsymbol{\Sigma}}_j$ in each sub-study j. Combining k studies, an estimate of the joint covariance matrix $\boldsymbol{\Sigma}$ has the form of a block diagonal matrix $\hat{\boldsymbol{\Sigma}} = \mathrm{diag}\left[\hat{\boldsymbol{\Sigma}}_1, \ldots, \hat{\boldsymbol{\Sigma}}_k\right]$.

The slopes estimates are modeled as $\mathbf{b} = \mathbf{W}\boldsymbol{\beta} + \mathbf{e}$, i.e. as a function of $\boldsymbol{\beta}$ and a design matrix $\mathbf{W}$, where the covariance matrix of $\mathbf{e}$ has to satisfy the condition $\mathrm{cov}(\mathbf{e}) = \boldsymbol{\Sigma}$. The design matrix $\mathbf{W}$ is composed of zeros and ones that identify which slopes are estimated in each sub-study. In case, when all sub-studies examine the same set of P predictors, a stack of k identity matrices of the dimension $P \times P$ serves as $\mathbf{W}$.

The generalized least squares approach results in the common estimates of the slope vector $\boldsymbol{\beta}$ and its covariance matrix:

$$\hat{\boldsymbol{\beta}} = \left[\mathbf{W}'\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{W}\right]^{-1}\mathbf{W}'\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{b}, \mathrm{cov}(\hat{\boldsymbol{\beta}}) = \left[\mathbf{W}'\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{W}\right]^{-1}.$$

With large samples and under typical regularity conditions $\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \mathrm{cov}(\hat{\boldsymbol{\beta}}))$, and the confidence intervals for each element of $\boldsymbol{\beta}$ are $\hat{\beta}_i \pm Z_{1-\alpha/2}\sqrt{C_{ii}}$, where $Z_{1-\alpha/2}$ is the upper tail $1-\alpha/2$ critical value of the standard normal distribution, C_{ii} is the ith diagonal element of $\mathrm{cov}(\hat{\boldsymbol{\beta}})$ matrix, $i = 1, \ldots, P$. Note that in contrast to the presentation in (Becker and Wu, 2007), we avoid here the parameter β_0, as it is not relevant for testing in GWAS.

Within the MSRS framework, there are several tests available:

A **test of the composite hypothesis** $\boldsymbol{\beta} = 0$ is given by

$$T = \hat{\boldsymbol{\beta}}'\left(\mathrm{cov}(\hat{\boldsymbol{\beta}})\right)^{-1}\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}'\left(\mathbf{W}'\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{W}\right)\hat{\boldsymbol{\beta}},$$

which is χ^2-distributed with p df.

A **test of model fit** (or a **test of homogeneity** of model parameters across sub-studies):

$$\beta_{11} = \ldots = \beta_{k1}, \ldots, \beta_{1P} = \ldots = \beta_{kP}.$$

The test statistic is given by

$$T = \left(\mathbf{b} - \mathbf{W}\hat{\boldsymbol{\beta}}\right)'\hat{\boldsymbol{\Sigma}}^{-1}\left(\mathbf{b} - \mathbf{W}\hat{\boldsymbol{\beta}}\right)$$

and is χ^2-distributed with $(k-1)P$ df.

In more general situation, where the model equation (3) is tested versus the restricted model $Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_S X_S$ in each sub-study, $1 \le S < P$, the method can be modified by including the predictor slopes $\beta_{S+1}, \ldots, \beta_P$ in $\boldsymbol{\beta}$, $\mathbf{b}$, $\boldsymbol{\Sigma}$ and thus reducing the dimension of $\mathbf{W}$. Then the test of the composite hypothesis $\beta_{S+1} = \ldots = \beta_P = 0$ has $(P-S)$ df, and the test of model fit $\beta_{1(S+1)} = \ldots = \beta_{k(S+1)}, \ldots, \beta_{1P} = \ldots = \beta_{kP}$ has $(k-1)(P-S)$ df.

Note that it is also possible to include in meta-analysis sub-studies that examine subsets of P predictors. If some sub-studies use fewer than the full set of P predictors, one can still apply the generalized least squares approach to integrate the results into the synthesis; see (Becker and Wu, 2007) for further details.

The method by (Becker and Wu, 2007) can be adjusted to the case of multiple logistic regression. The only change needed is the proper estimation of the covariance matrices themselves. The formula for the estimation of the covariance matrix used for linear regression model in (Becker and Wu, 2007) $\hat{\boldsymbol{\Sigma}}_j = (\mathbf{X}_j'\mathbf{X}_j)\hat{\sigma}_j^2$ is not valid for the case of logistic regression; here $\mathbf{X}_j$ is $n_j \times P$ matrix with the elements X_{jil} being a value of the variable X_l in (3) for the ith individual, n_j is the sample size and $\hat{\sigma}_j^2$ is an estimate of an error variance in sub-study j. We applied an estimation of the covariance matrix for multiple logistic regression models $\hat{\boldsymbol{\Sigma}}_j = (\mathbf{X}_j'\mathbf{V}_j\mathbf{X}_j)^{-1}$ given in (Hosmer and Lemeshow, 2000), where $\mathbf{X}_j$ is defined analogous to the linear regression case but for the model equation (2),

$$\mathbf{V}_j = \mathrm{diag}\left[\hat{\pi}_{j1}(1-\hat{\pi}_{j1}), \ldots, \hat{\pi}_{jn_j}(1-\hat{\pi}_{jn_j})\right],$$

$\hat{\pi}_{ji}$ is the estimated logistic probability computed using the maximum likelihood estimate $\boldsymbol{\beta}$ and the data for the ith individual in sub-study j. Modified MSRS is a powerful method to perform meta-analysis in case/control GWAS and can be used to summarize p-values *and* to obtain the common estimates of model parameters.

3 IMPLEMENTATION

METAINTER is a stand-alone software written in C/C++. It performs meta-analysis of GWAS summary statistics, in particular those obtained by multiple linear or logistic regression models. It is assumed that a unique predefined model is used to analyze the sets of SNP tuples in a series of studies. As input, multiple regression summary statistics of the parti-cipating studies have to be provided in tabulated format. METAINTER supports the output format of the genetic interaction analysis software INTERSNP (Herold *et al.*, 2009) as well as freely defined format. Parameters and options of the meta-analysis have to be specified in a simple configuration file.

METAINTER takes care of some technical issues in the preprocessing of the input data. Suppose that for a C/T SNP the alleles in input file are given in the order C/T for sub-study 1, but given in the order T/C for sub-study 2. In this case, all parameter estimates of log-additive type that depend on the SNP in the underlying regression model in sub-study 2 will be multiplied by –1 to unify the reference. The procedure will be repeated for all SNPs of the tuple under consideration with diverging reference. In addition, all entries of the covariance matrix (if available) that depend on log-additive terms of the SNP will be multiplied by –1. For parameters of entire dominance variation type, no modification is needed. Note that allele reference is automatically handled by INTERSNP.

METAINTER also attempts to resolve strand flips. If a SNP is given as C/T polymorphism in sub-study 1, and as G/A polymorphism in sub-study 2, METAINTER assumes that C ↔G and T ↔A. If the alleles of sub-study 2 occur in the order A/G instead, the SNP will in addition undergo the procedure described in the previous paragraph. C/G polymorphism, of course, will not be flipped by METAINTER. For such polymorphisms, strand consistency across studies has to be established before analysis.

4 SIMULATION STUDY

We simulated genotype data for series of 20 000 cases and 20 000 controls and considered a 'double recessive' disease model. Let i and j be the number of susceptibility alleles at SNP 1 and SNP 2, respectively, $i, j \in \{0, 1, 2\}$. For two-SNP genotypes, we specified a disease penetrance $f_{22} = 0.014$ and a baseline penetrance $f_{ij} = 0.01$ for $i, j \neq 2$. To create a series of disease models, we considered all allele frequency distributions $\{AF_1, AF_2\}$ with $AF_i \in \{0.2, 0.3, \ldots, 0.8\}$, $i = 1, 2$, and with $AF_1 \le AF_2$, the latter for the reason of symmetry. To investigate the global null hypothesis, we also considered the situation $f_{22} = 0.01$. The results were meta-analyzed by all methods described in Section 2. For comparison, we analyzed the entire dataset as well, where covariates indicating sub-study membership were used. Empirical significance level and empirical power level were computed as the portion of the simulated datasets below a pre-specified nominal significance level α. Under the null hypothesis, we considered α-levels of 0.01, 0.001, 0.0001 and

Table 1. Power values of the meta-analysis methods

AF_1/AF_2	Fisher[a]		Stouf[b]		StoufDir[c]		MSRS[d]		JointS[e]	
	8 df	4 df	8 df	4 df	8 df	4 df	8 df	4 df	8 df	4 df
0.2/0.2	0	0	0	0	0	0	0	0	0	0
0.2/0.3	0	0	0	0	0	0	0	0	0	0
0.2/0.4	0	0	0	0	0	0	0	0	0	0
0.2/0.5	0	0	0	0	0	0	0	0	0	0
0.2/0.6	0	0	0	0	0	0	0	0	0	0
0.2/0.7	0	0	0	0	0	0	0	0	0.001	0
0.2/0.8	0	0	0	0	0	0	0.006	0	0.008	0
0.3/0.3	0	0	0	0	0.001	0	0.001	0	0.001	0
0.3/0.4	0	0	0	0	0	0	0	0	0	0
0.3/0.5	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.005	0.001
0.3/0.6	0.003	0	0.007	0	0.007	0	0.026	0.002	0.031	0.002
0.3/0.7	0.035	0	0.045	0	0.064	0.002	0.166	0.002	0.188	0.002
0.3/0.8	0.169	0	0.205	0	0.214	0	0.492	0	0.506	0
0.4/0.4	0	0	0	0	0	0	0.001	0	0.003	0
0.4/0.5	0.013	0.002	0.020	0.002	0.037	0.002	0.090	0.008	0.100	0.009
0.4/0.6	0.157	0.007	0.191	0.010	0.240	0.016	0.476	0.039	0.501	0.042
0.4/0.7	0.568	0.011	0.623	0.015	0.691	0.023	0.890	0.047	0.897	0.051
0.4/0.8	0.908	0.003	0.942	0.008	0.893	0.010	0.995	0.022	0.996	0.024
0.5/0.5	0.245	0.014	0.297	0.020	0.364	0.034	0.623	0.079	0.642	0.084
0.5/0.6	0.786	0.072	0.807	0.086	0.850	0.124	0.965	0.235	0.965	0.240
0.5/0.7	0.984	0.094	0.988	0.130	0.989	0.176	1	0.311	1	0.316
0.5/0.8	1	0.058	1	0.081	0.986	0.099	1	0.188	1	0.191
0.6/0.6	0.786	0.205	0.807	0.247	0.850	0.308	0.965	0.469	0.965	0.473
0.6/0.7	0.984	0.300	0.988	0.357	0.989	0.398	1	0.586	1	0.589
0.6/0.8	1	0.155	1	0.197	0.986	0.231	1	0.421	1	0.420
0.7/0.7	1	0.383	1	0.445	1	0.491	1	0.680	1	0.682
0.7/0.8	1	0.215	1	0.269	0.975	0.249	1	0.494	1	0.497
0.8/0.8	1	0.129	1	0.156	0.932	0.133	1	0.353	1	0.359

[a]Fisher's method.
[b]Stouffer's method with weights.
[c]Stouffer's method with weights and effect directions.
[d]Method for the synthesis of regression slopes.
[e]Power of the regression test in the joint sample analysis.

0.00001. For the power study, we chose $\alpha = 1 \times 10^{-13}$. This is motivated as follows: under the assumption of $n = 1\,000\,000$ independent SNP sites in the human genome (The Wellcome Trust Case Control Consortium, 2007), there are $n(n-1)/2 = 5 \times 10^{11}$ independent SNP pairs in a GWIA. Dividing 0.05 by 5×10^{11} yields $\alpha = 1 \times 10^{-13}$ as criterion for the experiment-wise significance.

To assess the issue of minimal sample size required to provide valid inference of the meta-analysis, we also conducted simulation study for 1000, 3000 and 5000 cases and controls.

5 RESULTS OF THE SIMULATION STUDY

Empirical levels. The results of the simulation study under the null hypothesis are given for the 8 df logistic regression test in Supplementary Table S1. In this table, we present the empirical significance levels for the four meta-analysis methods, the joint sample analysis and the test of homogeneity of model parameters at two nominal levels: $\alpha = 0.01$ for 10 000 simulations and $\alpha = 0.00001$ for 100 000 simulations. In general, all methods behave satisfactory under the null hypothesis. We observe some nominally significant excesses of α, which is to be expected because we investigated 336 situations presented in Supplementary Table S1. Deviation from the nominal level is assessed by the exact binomial test. For $\alpha = 0.01$, the most significant difference ($p = 0.0002$, after Sidak correction with 336 tests $p_{cor} = 0.0650$) is the empirical level 0.0065 obtained by the test of homogeneity of model parameters for a combination of allele frequencies 0.8/0.8. For $\alpha = 0.00001$, the most significant deviation is observed at empirical level of 0.00004. Note that this excess is still only mildly anti-conservative ($p = 0.02$, $p_{cor} = 1.00$). Taking into account multiple comparison in Supplementary Table S1, we conclude that the methods perform well in maintaining the nominal significance level in all settings considered. This is also underlined by the fact that none of the mean empirical levels shows deviation from the respective nominal level. The simulations under the null hypothesis demonstrate the validity of the meta-analysis methods implemented in METAINTER.

Table 2. Correlations r_i between meta-analysis estimates and joint sample analysis estimates of model parameters (8 df test)

Sample size	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8
1000	0.949	0.951	0.937	0.955	0.948	0.949	0.947	0.953
3000	0.985	0.991	0.992	0.991	0.984	0.992	0.993	0.996
5000	0.994	0.997	0.997	0.996	0.995	0.998	0.996	0.995
20 000	0.999	1	0.999	1	0.999	0.999	0.999	1

Power study. Power of all meta-analysis methods considered in our work and implemented in METAINTER is shown in Table 1 for different combinations of allele frequencies AF_1, AF_2, and the significance level $\alpha = 1 \times 10^{-13}$. The primary analysis was performed by the 8 and 4 df logistic regression tests. Power of these tests, used in the joint sample analysis, is given in the last two columns of the table. Naturally, we see the absence of power for low allele frequencies and observe the increase of power with increasing values of allele frequencies, irrespective of the meta-analysis method used. For almost all settings, the ranking of the meta-analysis approaches leads to the following priority list with respect to power: MSRS, the Stouffer's method with weights and effect directions, the Stouffer's method with weights and, finally, the Fisher's method as the least powerful. The most pronounced divergence in power values between MSRS and the joint sample analysis is observed for the allele frequency combinations 0.3/0.7 and 0.4/0.6 for the 8 df test and is rather small in size. From our study, we can conclude that MSRS outperforms substantially all other meta-analysis methods in terms of power. Moreover, power of MSRS is close to that of the analysis of the entire sample.

Furthermore, despite the known problem with low power for small samples (Supplementary Fig. S1), MSRS provides reasonable model parameters estimates. In particular, as shown in Table 2, the correlation between meta-analysis estimates and estimates obtained from the joint sample analysis with the 8 df logistic regression test, as computed for $AF_1 = AF_2 = 0.3$ and based on 100 permutation replicates, is >0.93, indicating the validity of the method even for reasonable sample sizes. These facts demonstrate the usefulness of the information contained in the covariance matrix of model parameters.

In case the covariance matrix is not available and MSRS can not be applied, the Stouffer's method with weights and effect directions represents an effective alternative to perform

Table 3. T2D dbGaP studies

No.	Project	Study	Individuals	SNPs	Pairs[a]
1.	Health Research Vanderbilt U[b]		607	499 350	1.25E11
2.	Health Research Vanderbilt U[b]		1384	919 602	4.22E11
3.	Health Research Northwestern U[c]		1239	495 588	1.23E11
4.	Health Research Northwestern U[c]		267	908 692	4.12E11
5.	GENEVA Diabetes Study	NHS[d]	3435	764 678	2.92E11
6.	GENEVA Diabetes Study	HPFS[e]	2606	787 213	3.10E11

[a]Number of SNP pairs investigated in stage I of the two-stage meta-analysis.
[b]Project Health Research—Vanderbilt University, Northwestern NUgene Project: T2D, National Human Genome Research Institute.
[c]Project Health Research—Northwestern University, Northwestern NUgene Project: T2D, National Human Genome Research Institute.
[d]Nurses Health Study.
[e]Health Professionals Follow-up Study.

meta-analysis. We notice the benefit from differentiating the consistency of effect directions in sub-studies participating in meta-analysis. It results in higher power of the Stouffer's method with effect directions in comparison with the Stouffer's method where the effect directions are neglected.

6 REAL DATA ANALYSIS

We investigated six GWAS datasets of T2D that are available from dbGaP (Mailman *et al.*, 2007), cf. Table 3 and relied on the quality controlled data available there.

Primary analysis of six T2D GWAS was performed by the 8 df logistic regression test under interaction and by the 4 df logistic regression test for interaction. We use both regression models to give examples of the applicability of the meta-analysis approach.

Summary statistics of single-marker GWAS can typically be integrated in one-step meta-analysis. To meta-analyze the results of two-marker GWAS, however, a two-stage procedure is necessary to guarantee computational feasibility. A GWIA involves a large number of SNP pairs (10^{10}–10^{12}). This implies that (i) for running time reasons a fast screening test has to be used in the initial GWIA in all participating studies; (ii) not all results of this GWIA can be stored for lack of hard-disk memory. Therefore, we applied the following two-stage design. In stage I, we conducted a GWIA on all SNP pairs in each study with the 8 df and 4 df logistic regression tests; we did not use any covariates at this stage. In this case, the test statistic can be obtained in closed form (Marchini *et al.*, 2005), and running time depends on how quickly two-SNP genotypes can be counted in cases and controls. The implementation of the Hamming weight method (Wegner, 1960) in INTERSNP (Herold *et al.*, 2012) allowed us to test all SNP pairs from each study, with maximum of 4.22×10^{11} pairs in the second T2D study. It was analyzed by INTERSNP within ~16 h on a dedicated computing node (3 GHz) and required 2 GB working memory. From the stage I analysis, we retained a list of candidate SNP pairs with $p\text{-value} \leq 1 \times 10^{-4}$ in at least one study. After stage I, we merged the candidate lists obtained from each study to form a 'joint' list. In stage II, the joint list was sent back to each study and analyzed with INTERSNP by the 8 and 4 df logistic regression tests. To adjust for potential population stratification, we performed multidimensional scaling as implemented in (Purcell *et al.*, 2007) for each study and retained 3 leading dimensions as covariate parameters for the analysis. After stage II, summary statistics from 2 149 025 (8 df) and 2 950 354 (4 df) candidate SNP pairs were available in all six T2D studies and were meta-analyzed with METAINTER. The meta-analysis took METAINTER 23 min (8 df) and 11 min (4 df).

In Table 4, the top five results of the meta-analysis of six T2D GWAS are shown for both logistic regression tests. In Table 5, three further experiment-wide significant results are given with the indication which test they refer to. In these tables we present the *p*-values of six individual studies, the *p*-values of the meta-analysis performed by four methods implemented in METAINTER and the *p*-values of the test of homogeneity of model parameters across studies; see Section 2.2.

All top SNP pairs for the 8 df logistic regression test contain SNP rs7901695, located within the *TCF7L2* gene. This gene is known to harbor the most important genetic risk factor for T2D, as identified by GWAS (Ali, 2013). Because the 8 df test includes marginal SNP effects, the accumulation of rs7901695 in the top ranking results corroborates the validity and utility of our analysis. The strongest association signals are seen for MSRS method, which coincides with the results of the power study.

Among the top five results for the 8 df logistic regression test, the second SNP of each pair is not located in regions previously implicated in T2D. However, those SNPs can be viewed as candidates for either marginal association or potential interaction with rs7901695.

Table 5 presents two experiment-wide significant SNP pairs for the 8 df logistic regression test. Each pair contains a further SNP previously implicated in T2D. The SNP rs12195232, located 20 kb upstream of *SAYSD1* gene, has been suggested as a T2D risk factor by the standard single SNP analysis in (Cho *et al.*, 2011). Another SNP, rs10012946, is located in the *WFS1* gene. Mutations in this gene are associated with Wolfram syndrome, also called DIDMOAD (diabetes insipidus, diabetes mellitus, optic atrophy and deafness). *WFS1* was also found to be associated with T2D (Voight *et al.*, 2010). The presence of rs12195232 and rs10012946 in experiment-wide significant pairs demonstrates the potential of multiple regression models in general. In single-marker analysis, both SNPs would have gone undetected with the data that were available to us ($p = 0.023$ and $p = 0.012$, respectively). However, in the multiple regression meta-analysis, these SNPs come into focus, while analyzing six comparatively small T2D GWAS. Hence, investigation of interaction between *SAYSD1* and *TCF7L2*, *WFS1* and *TCF7L2* and the marginal effects of both *SAYSD1* and *WFS1* in larger studies would be of interest.

In Table 4, among the top five results of the T2D meta-analysis for the 4 df logistic regression test, none of the SNP pairs reached a genome-wide significant *p*-value. However, SNP pair rs6446498/rs12109266 (rank No. 9), listed in Table 5, represents a plausible

Table 4. Top results of the meta-analysis of six T2D studies

Test	ID_1	Chr_1	Pos_1	ID_2	Chr_2	Pos_2	p_1	p_2	p_3	p_4	p_5	p_6	p-Fisher	p-Stouf	p-StoufDir	p-MSRS	p-JointS	p_H
8 df	rs1864348	1	158741091	rs7901695	10	114754088	6.57E-06	1.57E-03	8.41E-05	1.22E-01	1.08E-02	1.17E-03	1.50E-12	2.12E-13	1.10E-14	3.44E-19	7.19E-19	0.423
	rs1599711	10	71371129	rs7901695	10	114754088	4.88E-06	6.25E-04	3.32E-03	8.66E-02	3.73E-02	1.73E-03	3.97E-11	1.55E-12	5.13E-13	6.43E-19	3.16E-20	0.954
	rs17160788	7	32048171	rs7901695	10	114754088	1.80E-09	9.00E-03	6.48E-03	6.65E-02	8.85E-02	8.97E-05	8.89E-14	6.98E-14	1.17E-12	8.09E-19	2.07E-20	0.283
	rs1834134	2	222857113	rs7901695	10	114754088	6.31E-07	1.72E-04	6.99E-05	4.24E-02	1.05E-02	3.58E-03	2.28E-14	1.22E-15	2.49E-13	9.40E-19	2.85E-18	0.046
	rs1602204	13	82378976	rs7901695	10	114754088	6.92E-08	1.65E-03	1.44E-02	1.07E-01	1.23E-01	8.74E-04	1.53E-11	1.48E-12	6.44E-14	1.17E-18	1.42E-18	0.593
4 df	rs11136094	8	22547052	rs1950646	14	94722497	–	–	–	9.70E-11	–	1.86E-01	4.64E-10	2.86E-10	1.01E-10	6.19E-11	1.25E-10	0.545
	rs10935178	3	135893295	rs7842319	8	95652986	–	–	–	1.98E-11	–	6.18E-01	3.19E-10	1.39E-09	2.09E-10	8.01E-11	2.84E-10	0.607
	rs6800570	3	135909962	rs7842319	8	95652986	–	–	–	3.04E-11	–	6.18E-01	4.83E-10	1.98E-09	3.04E-10	1.17E-10	4.05E-11	0.614
	rs2485944	1	2883858	rs35305	5	58548231	–	–	1.23E-01	1.71E-05	4.55E-06	2.97E-01	1.00E-08	2.40E-09	3.17E-01	1.79E-10	9.09E-11	0.352
	rs4007264	10	55876860	rs11905530	20	17478869	1.66E-05	5.10E-06	–	–	–	–	2.05E-09	7.12E-10	3.69E-10	2.15E-10	6.60E-11	0.388

p-Fisher is p-value of the Fisher's method; p-Stouf is p-value of the Stouffer's method with weights; p-StoufDir is p-value of the Stouffer's method with weights and effect directions; p-MSRS is p-value of the MSRS; p_i is p-value of the sub-study i in the primary analysis, $i = 1, \ldots, 6$; p-JointS is p-value of the regression test in the joint sample analysis; p_H is p-value of a test of homogeneity of model parameters across sub-studies.

Table 5. Three particular results of the meta-analysis of six T2D studies

Test	ID_1	Chr_1	Pos_1	Alleles_1	Gene_1	ID_2	Chr_2	Pos_2	Alleles_2	Gene_2	p_1	p_2	p_3	p_4	p_5	p_6	p-Fisher	p-Stouf	p-StoufDir	p-MSRS	p-JointS	p_H
8 df	rs10012946	4	6293350	T/C	*WFS1*	rs7901695	10	114754088	C/T	*TCF7L2*	7.80E-07	3.15E-04	4.65E-03	6.14E-02	0.137	5.89E-05	6.99E-13	2.96E-14	1.38E-15	1.11E-16	6.29E-17	0.062
	rs12195232	6	39093133	T/C	20 kb to *SAYSD1*	rs7901695	10	114754088	C/T	*TCF7L2*	1.81E-06	3.88E-04	0.014	0.372	0.027	1.39E-03	8.38E-11	7.15E-12	2.77E-13	1.03E-16	6.62E-16	0.368
4 df	rs6446498	4	6400680	C/T	*PPP2R2C*	rs12109266	5	55888260	A/T	*LOC 101928448*	3.98E-07	2.50E-04	-	-	-	-	2.39E-09	2.29E-09	1.16E-09	3.05E-10	1.28E-10	0.331

p-Fisher is p-value of the Fisher's method; p-Stouf is p-value of the Stouffer's method with weights; p-StoufDir is p-value of the Stouffer's method with weights and effect directions; p-MSRS is p-value of the MSRS; p_i is p-value of the sub-study i in the primary analysis, $i = 1, \ldots, 6$; p-JointS is p-value of the regression test in the joint sample analysis; p_H is p-value of a test of homogeneity of model parameters across sub-studies.

candidate. Rs6446498 is located in *PPP2R2C* gene, previously shown to be associated with T2D (Rung *et al.*, 2009). Rs12109266 located near *LOC101928448* gene, which was reported to be associated with body fat distribution using waste circumference in male individuals of African ancestry (Liu *et al.*, 2013a). Moreover, in (Zheng *et al.*, 2013) it was shown that *LOC101928448* interacts with two environmental factors: carbohydrate for fasting insulin and homeostasis model assessment of insulin resistance, making significant contribution to the genotype by environment variance for T2D-related traits. Because the first SNP of the pair rs6446498/rs12109266 is already known as a T2D risk variant, one can argue that the significance threshold of 1×10^{-13} for entirely new pairs is not required here, and that 5×10^{-8} is an appropriate threshold for a test for interaction with a known SNP. In this sense, the pair is significant after adjustment for multiple testing. Replication in independent datasets is nevertheless warranted.

The p-values p_H of the test of homogeneity of model parameters across six T2D GWAS, computed for the predictor slopes only (omitting intercepts), are shown in the last column of Tables 4 and 5. The p-values p_H provide evidence of homogeneity for almost all SNP pairs except for rs1834134/rs7901695 ($p_H = 0.046$).

The meta-analysis by MSRS allows us to estimate not only p-values but effect sizes as well. Table 6 presents the slopes estimates for the 8 df logistic regression model from each of the six T2D GWAS, and shows meta-analysis estimates obtained by MSRS for one SNP pair listed in Table 5. By applying MSRS we are provided with an overall model based on the synthesis of the results in the primary studies, where the interrelationships among predictors are accounted for. For comparison, we included in Table 6 the slopes estimates obtained from the joint sample analysis ('mega-analysis'), see 'JointS' row. The estimates of both approaches are in good concordance.

7 DISCUSSION

New methods to meta-analyze genetic data are obviously of growing demand (Evangelou and Ioannidis, 2013). Meta-analysis of summary statistics of single-marker tests applied in GWAS is an established standard, which however needs to be extended to more complex multiple parameter testing strategies. In context of rare variant analysis, which typically addresses multiple parameters from a genomic region, an important improvement of p-values combination methods has recently been made by (Liu *et al.*, 2013b). The method presented there is based on the correlation structure of single-variant test statistics, and, in terms of power, comes close to power of the rare variant analysis of the combined sample.

The method for the synthesis of linear regression slopes (Becker and Wu, 2007) has been applied to meta-analyze the results of a linear regression gene–environment interaction test (Manning *et al.*, 2011). Here, we presented a general framework that enables application of MSRS to meta-analysis of both linear and logistic multiple regression models. In addition, a test that assesses across-study heterogeneity can easily be formulated within MSRS framework and is a part of our implementation. The method can be used to filter out signals that are driven by

Table 6. Effect estimates in the meta-analysis of six T2D studies

Study	β_1	β_2	β_{1D}	β_{2D}	γ_{12}	$\gamma_{1,2D}$	$\gamma_{1D,2}$	$\gamma_{1D,2D}$
1	−0.13 ± 0.07	0.09 ± 0.10	0.36 ± 0.07	−0.14 ± 0.10	0.04 ± 0.10	−0.32 ± 0.14	0.30 ± 0.15	−0.17 ± 0.19
2	0.09 ± 0.09	−0.12 ± 0.12	0.35 ± 0.09	−0.17 ± 0.12	0.29 ± 0.13	−0.27 ± 0.18	−0.26 ± 0.17	0.16 ± 0.24
3	−0.30 ± 0.23	0.09 ± 0.32	0.49 ± 0.24	−0.39 ± 0.32	0.23 ± 0.34	−0.33 ± 0.46	−0.65 ± 0.48	0.59 ± 0.63
4	0.07 ± 0.07	−0.12 ± 0.09	0.28 ± 0.07	−0.07 ± 0.09	0.18 ± 0.10	−0.02 ± 0.13	−0.28 ± 0.13	0.49 ± 0.18
5	−0.24 ± 0.11	−0.13 ± 0.14	0.37 ± 0.10	0.06 ± 0.14	0.10 ± 0.15	−0.06 ± 0.21	0.05 ± 0.21	0.05 ± 0.28
6	0.10 ± 0.16	−0.46 ± 0.21	0.21 ± 0.16	0.17 ± 0.21	−0.11 ± 0.23	−0.07 ± 0.32	−0.14 ± 0.31	−0.41 ± 0.42
Meta-analysis	−0.03 ± 0.04	−0.08 ± 0.05	0.32 ± 0.04	−0.08 ± 0.05	0.14 ± 0.06	−0.18 ± 0.07	−0.09 ± 0.07	0.17 ± 0.10
JointS	−0.03 ± 0.04	−0.09 ± 0.05	0.31 ± 0.04	−0.08 ± 0.05	0.15 ± 0.05	−0.18 ± 0.07	−0.10 ± 0.07	0.19 ± 0.10

Example for the SNP pair rs10012946/rs7901695.

just a few studies alone, thereby it captures the main advantage of random effects models.

Via application to six T2D GWAS, we could demonstrate that the implementation of MSRS in METAINTER allows efficient computation. Meta-analysis of 2.2 million SNP pairs analyzed in six studies took METAINTER 23 min. We supported the analysis by updating our INTERSNP tool (Herold *et al.*, 2009, 2012). INTERSNP provides now the necessary covariance matrix of parameter estimates in multiple regression models.

The implementation of MSRS in METAINTER allows efficient meta-analysis for multiple linear and logistic regression models. However, the method should be used with caution, if the evidence of strong dependence between predictor variables exists. The phenomenon of multicollinearity can manifest itself, for example, when the inverse covariance matrix has to be computed. If the model is well specified in each sub-study, there should be no strong multicollinearity. Otherwise model reduction can be useful to get around this problem. For GWAS with small samples and for SNPs with low allele frequencies (e.g. both <0.2 or for one SNP <0.05) a test with fewer parameters, for instance, 3 df allelic test under interaction, might be a reasonable choice. MSRS assumes that the multicollinearity issue was checked in sub-studies and that any of them is relatively free from this problem. If the covariance matrix is non-invertible for a particular SNP pair (or SNP, or SNP set) in a sub-study, METAINTER will exclude the sub-study from the meta-analysis for the respective pair.

MSRS is an appropriate generalization of the existing techniques in the following sense: in case of one parameter models, the method is equivalent to the fixed effect model, and in the multiple parameter case its power comes close to that of the analysis of the joint sample, as we have shown in our simulation study. Moreover, power comparison of MSRS with the Fisher's and the Stouffer's methods re-emphasizes the importance of going beyond *p*-value-based meta-analysis for higher-dimensional models.

Funding: T. Vaitsiakhovich was supported by the German Research Foundation (DFG), projects BE 3828/3-2, BE 3828/3-3.

Conflict of interest: none declared.

REFERENCES

Ali,O. (2013) Genetics of type 2 diabetes. *World J. Diabetes*, **4**, 114–123.

de Bakker,P.I. *et al.* (2010) Meta-analysis of genome-wide association studies. *Cold Spring Harb. Protoc.*, **6**, pdb ip71.

Becker,B.J. (1994) Combining significance levels. In: Cooper,H. and Hedges,L.V. (eds) *The Handbook of Research Synthesis*. Russell Sage Foundation, New York, NY, pp. 215–230.

Becker,B.J. and Wu,M.-J. (2007) The synthesis of regression slopes in meta-analysis. *Stat. Sci.*, **22**, 414–429.

Chapman,J. and Clayton,D. (2007) Detecting association using epistatic information. *Genet. Epidemiol.*, **31**, 894–909.

Cho,Y.S. *et al.* (2011) Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. *Nat. Genet.*, **44**, 67–72, doi: 10.1038/ng.1019.

Cordell,H.J. and Clayton,D.G. (2002) A unified stepwise regression procedure for evaluating the relative effects of polymorphisms within a gene using case/control or family data: application to HLA in type 1 diabetes. *Am. J. Hum. Genet.*, **70**, 124–141.

Cochran,W.G. (1937) Problems arising in the analysis of a series of similar experimets. *J. Roy. Statist. Soc.*, **4** (Suppl.), 102–118.

Evangelou,E. and Ioannidis,J.P. (2013) Meta-analysis methods for genome-wide association studies and beyond. *Nat. Rev. Genet.*, **14**, 379–389, doi: 10.1038/nrg3472.

Fisher,R. (1932) *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh.

Herold,C. *et al.* (2009) INTERSNP: genome-wide interaction analysis guided by a priori information. *Bioinformatics*, **25**, 3275–3281, doi: 10.1093/bioinformatics/btp596.

Herold,C. *et al.* (2012) Integrated genome-wide pathway association analysis with INTERSNP. *Hum. Hered.*, **73**, 63–72, doi: 10.1159/000336196.

Hosmer,D.W. and Lemeshow,S. (2000) *Applied Logistic Regression*. 2nd edn. NY: Wiley, New York.

Lipták,T. (1959) On the combination of independent tests. *Publ. Math. Inst. Hungar. Acad. Sci.*, **3**, 171–197.

Liu,C.T. *et al.* (2013a) Genome-wide association of body fat distribution in African ancestry populations suggests new loci. *PLoS Genet.*, **9**, e1003681, doi: 10.1371/journal.pgen.1003681.

Liu,D.J. *et al.* (2013b) Meta-analysis of gene-level tests for rare variant association. *Nat. Genet.*, **46**, 200–204, doi: 10.1038/ng.2852.

Mägi,R. and Morris,A.P. (2010) GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics*, **11**, 288, doi: 10.1186/1471-2105-11-288.

Mailman,M.D. *et al.* (2007) The NCBI dbGaP database of genotypes and phenotypes. *Nat. Genet.*, **39**, 1181–1186.

Manning,A.K. *et al.* (2011) Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP environment regression coefficients. *Genet. Epidemiol.*, **35**, 11–18, doi:10.1002/gepi.20546.

Marchini,J. *et al.* (2005) Genome-wide strategies for detecting multiple loci that influence complex diseases. *Nat. Genet.*, **37**, 413–417.

Pestman,W.R. (2009) *Mathematical Statistics*. 2nd edn. de Gruyter, Berlin.

Purcell,S. *et al.* (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am. J. Hum. Genet.*, **81**, 559–575.

Rung,J. *et al.* (2009) Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. *Nat. Genet.*, **41**, 1110–1115, doi: 10.1038/ng.443.

Stouffer,S. *et al.* (1949) *The American soldier: Adjustment During Army Life*. Vol. 1, Princeton University Press, Princeton, USA.

Tang,Z.Z. and Lin,D.Y. (2013) MASS: meta-analysis of score statistics for sequencing studies. *Bioinformatics*, **29**, 1803–1805.

The Wellcome Trust Case Control Consortium. (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, **447**, 661–678.

Voight,B.F. *et al.* (2010) Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat. Genet.*, **42**, 579–589, doi: 10.1038/ng.609.

Wegner,P. (1960) A technique for counting ones in a binary computer. *Commun. ACM*, **3**, 322, doi:10.1145/367236.367286.

Willer,C.J. *et al.* (2010) METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, **26**, 2190–2191, doi: 10.1093/bioinformatics/btq340.

Zaykin,D.V. (2011) Optimally weighted Z-test is a powerful method for combining probabilities in meta-analysis. *J. Evol. Biol.*, **24**, 1836–1841, doi: 10.1111/j.1420-9101.2011.02297.x.

Zheng,J.S. *et al.* (2013) Genome-wide contribution of genotype by environment interaction to variation of diabetes-related traits. *PLoS One*, **8**, e77442, doi: 10.1371/journal.pone.0077442.