**Supplementary Note: Deriving optimal P-value weights**

**A1: Notation**

Our weights rely on the standard normal probability density function and the normal cumulative distribution function . We will denote a random variable that is normal with mean and variance as .

We will derive the weights for one-sided tests. By flipping signs if necessary we can assume that we are testing for negative prior effect sizes.

**A2: Statistical Model**

We assume that our association test statistics are standard normal with standardized effect sizes and variance 1: , where the index ranges from 1 to J . This can always be achieved by standardizing the test statistics. Further, we model the effect sizes as normal with prior effect size and variance : . Then it follows that the marginal distribution of is also normal, with parameters that are only functions of and : . We also assume that the prior standard errors are strictly positive, i.e. , which corresponds to the realistic setting where there is some uncertainty in the prior information about . The case is treated in Roeder and Wasserman (2009)1, and the extension to positive variances is our main methodological contribution. We denote the marginal variance . The P-value *Pi* for the one-sided test versus equals .

**A3: Optimization Problem**

Many different approaches can be used to assign weights from a pleiotropic data set to be used in a second data set. For example, a candidate gene approach simply assigns equal weight to the candidate genes and zero to all other genes, and the conditional false discovery method assigns a weight according to the p-value in the pleiotropic data set. We want to find the weights that maximize the expected number of significant hits, given the prior information. We optimize the weights for the weighted Bonferroni procedure. This procedure can be defined for an arbitrary weight vector , and for a fixed level it declares significant those SNPs whose p-values are small enough that . If the weights are non-negative and average to 1, then the weighted Bonferroni procedure controls the family-wise error rate at the level .

Hence the number of significant hits as a function of the weights can be written formally as

Here denotes the indicator function, which equals 1 if the statement between the parentheses is true and equals 0 otherwise. The expected number of significant hits, as a function of , then equals by linearity of expectation:

The expectation is taken over two sources of uncertainty: The distribution of the test statistics from the current data, and the distribution of the effect size estimated from the prior data. The above function can be written explicitly in terms of the normal cumulative distribution function. Thus, we have formulated the optimization objective explicitly. With this notation the optimization objective is to maximize the above function, subject to the constraints

Next we discuss an efficient solution for the optimization problem.

**A4: Solution of the Optimization Problem**

Our solution of the optimization problem relies on Lagrange duality, which is a standard tool in numerical optimization. We refer the reader to any of a number of standard reference books on numerical optimization, eg.2. In short, in Lagrangian duality one reduces a constrained optimization problem to an unconstrained penalized optimization problem. Under certain conditions, if one has zero duality gap, then the penalized problems can be used to obtain the solution of the original problem. In our case the last step consists of a simple one-dimensional line search.

For convex problems, it is often easy to show that Lagrange duality works2. However, our problem is not convex, and for such problems zero duality gap holds only in a handful of cases, and relies on a detailed case-by-case analysis for each individual problem. We undertake such an analysis below, and for the full details we refer to the methodology paper Dobriban et al 20154..

To summarize: Even though the optimization problem is not convex, we can solve it efficiently via a one-dimensional line search. The key reason is that under certain conditions the problem has zero duality gap.

To explain this formally, we need to introduce some additional notation. We define a function , which will turn out to give the critical values for our test. is a function of the prior mean and standard error as well as of a so-called dual variable which is of key importance in the optimization procedure. The function is given explicitly by the formula:

This formula arises from the analysis of the Lagrangian, the penalized version of the objective function mentioned above. The condition under which we can maximize the objective explicitly is given by an inequality involving the prior parameters

This is a mild condition, which holds in all cases involving GWAS data that we have seen. It can also be checked easily. After some analysis, one can see that it is enough that some prior effect sizes are large in magnitude; the details are presented in Dobriban et al 20154.

Then we can state the following:

**Theorem.** If the above condition holds, then the optimal weights are, for a certain constant , . The constant is the unique value for which , and it can be found via a one-dimensional line search.

We present a proof of this statement in Dobriban et al 20154. It leads to an efficient algorithm for finding the optimal weights: 1) check the condition, 2) perform a one-dimensional line search starting at for the unique value such that the values sum to .

The one-dimensional line search is easy, because the weights depend monotonically on . Specifically, the function is monotonically decreasing in . Hence there are many standard numerical algorithms for the line search. The bisection method is the simplest one, and it has well-known convergence guarantees, se eg. Burden and Faires3. By numerical experiments we have also found that the classical Newton’s method is effective, and can be faster than the bisection method. However, it is not known theoretically that it converges, because it lacks the strong convexity required by most analyses. In our implementation, the user can choose either method.

We provide an open-source and free implementation of our method written in Matlab and R. Both implementations are available at <https://github.com/dobriban>. We also developed an R package, pweight, implementing p-value weighting methods and iGWAS. Pweight is available from CRAN:  <https://cran.r-project.org/web/packages/pweight/>.

References

1. Roeder, K. & Wasserman, L. Genome-Wide Significance Levels and Weighted Hypothesis Testing. *Stat. Sci.*  **24,** 398–413 (2009).

2. Boyd, S. & Vandenberghe, L. *Convex Optimization*. (Cambridge University Press, 2004).

3. Burden, R. L. & Faires, J. D. *Numerical analysis*. (Thomson : Brooks/Cole, 2005).

4. Dobriban, E., Fortney, K., Kim, S. & Owen, A. *Biometrika*. in press (2015).