Reply

Human Heredity

Hum Hered 2005;60:61–62
DOI: 10.1159/000087919

Received: July 3, 2005
Accepted: July 5, 2005
Published online: August 31, 2005

Evaluation of Nyholt's Procedure for Multiple Testing Correction – Author's Reply

Dale R. Nyholt

Queensland Institute of Medical Research, Brisbane, Australia

I welcome Salyakina et al.'s (2005) [1] thorough evaluation of my approach to estimate the effective number of independent single-nucleotide polymorphisms (SNPs) from a set of correlated SNPs [2], and I hope their paper and this response will stimulate more researchers to think seriously about their analysis methods. I emphatically agree with the author's main conclusion that my procedure (SNPSpD) is a useful exploratory tool and is not a *replacement* for using a permutation test.

The development of high throughput and inexpensive SNP genotyping platforms coupled with the availability of numerous SNP databases and their high frequency throughout the genome has lead to the enthusiastic utilization of SNP markers in genetic trait association studies. Indeed, an NCBI PubMed search using (SNP, association, gene, 2005) indicates approximately 1000 studies testing SNPs for association to a trait will be published in 2005 alone. Furthermore, this field is shifting toward a more (thorough) gene-based approach in which many variants within a candidate gene (region) are considered (see for example [3]).

However, although both the number of studies being performed and the development of powerful statistical analysis techniques are proceeding at an astounding rate, the reality exists that the knowledge and computational skills of many researchers who are performing these studies has not kept pace. Therefore, access to simple approximate approaches such as SNPSpD, which address the issue of multiple testing within a set of correlated SNPs, still have an important role to play. Indeed, this is supported by the recommendation of Salyakina et al. (2005) that 'any significant result obtained using Nyholt's procedure which is not already significant using Šidák's original correction should be confirmed by a permutation test'.

I almost agree with this recommendation, on the basis that a significant (or nearly significant) association using the SNPSpD procedure which is not already significant using Šidák's (or Bonferroni's) original correction, would indicate substantial non-independence (i.e., strong linkage disequilibrium, LD) among the tested SNPs and one may therefore benefit from performing permutations. That is, I acknowledge my approach is slightly conservative in the presence of strong LD.

Specifically, as in my original publication [2] and noted on the SNPSpD webpage (http://genepi.qimr.edu.au/general/daleN/SNPSpD/) my approach will be conservative in the presence higher order inter-marker LD and in particular, SNPs in perfect LD ($r = 1$). As Salyakina et al. (2005) note, the latter is only true when you have a combination of perfectly ($r = 1$) and imperfectly ($r < 1$) correlated SNPs. Although the note on my webpage did not explicitly make this distinction, it does not change the interpretation of my recommendation to exclude all SNPs but one from any set of SNPs that are in perfect LD (i.e., remove 'redundant' SNPs). It is therefore unfortunate

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

0001–5652/05/0601–0061$22.00/0

Accessible online at:
www.karger.com/hhe

Dr. Dale Nyholt
Queensland Institute of Medical Research
Post Office Royal Brisbane Hospital
Brisbane QLD 4029 (Australia)
Tel. +61 7 33620258, Fax +61 7 33620101, E-Mail daleN@qimr.edu.au

Salyakina et al. (2005) did not follow this recommendation in their evaluation of the SNPSpD procedure, as their failure to do so would clearly have biased their results. Despite this short coming, the results of Salyakina et al. (2005) indicate my procedure produces type I errors very close to the desired 5%.

To address this conservativeness and assist users of the SNPSpD web-based procedure I have updated the code to now restrict analysis to individuals who are genotyped at all loci and automatically remove redundant SNPs.

Finally, I do not agree that a significant finding using the SNPSpD procedure must be *confirmed* via permutation and I am somewhat puzzled by the finding of Salyakina et al. (2005) that my procedure 'may also be anti-conservative when there is strong LD between SNPs'. I can think of no theoretical justification for why my procedure would *underestimate* the approximate number of independent SNPs (M_{eff}), nor is one provided by Salyakina et al. (2005). That is, although it seems reasonable to describe dependencies between SNPs by pairwise correlations, because LD associated with alleles from three or more markers decays more rapidly than that from two markers [4, 5], higher order SNP correlations which are not captured via pairwise measures could only contribute to an *overestimate* of M_{eff}. Instead, I wonder whether their few 'anti-conservative' findings result from a peculiarity in their simulations and/or the assumption that their simulated minimum p values correspond to a beta distribution for the SNP sets in question.

In closing, I wish to make clear that I never intended the SNPSpD approach to *replace* permutation procedures. Instead, I envisaged the procedure to be utilized by a wide variety of researchers who will appreciate its convenience and useful measure of SNP independence. Results from SNPSpD enable researchers to make an informed decision on whether future analyses, namely permutations, would be beneficial. Furthermore, besides from using M_{eff} estimates to adjust significance thresholds after SNP genotyping, one can incorporate M_{eff} estimates calculated from HapMap data [6] at the study design stage, to obtain a less conservative estimation of the number of planned independent tests for use in power calculations.

Acknowledgement

Dr. Nyholt is supported by an NHMRC RD Wright Research Fellowship.

References

1 Salyakina D, Seaman SR, Browning BL, Dudbridge F, Muller-Myhsok B: Evaluation of Nyholt's procedure for multiple testing correction. Hum Hered 2005;60:19–25.

2 Nyholt DR: A simple correction for multiple testing for SNPs in linkage disequilibrium with each other. Am J Hum Genet 2004;74:765–769.

3 Neale BM, Sham PC: The future of association studies: Gene-based analysis and replication. Am J Hum Genet 2004;75:353–362.

4 Bennett JH: On the theory of random mating. Ann Eugen 1954;18:311–317.

5 Meng Z, Zaykin DV, Xu CF, Wagner M, Ehm MG: Selection of genetic markers for association analyses, using linkage disequilibrium and haplotypes. Am J Hum Genet 2003;73:115–130.

6 The International HapMap Consortium: The international HapMap project. Nature 2003; 426:789–796.