

False discovery vs. True discovery

Speaker: Bernard Devlin

Department of Psychiatry, University of Pittsburgh SOM

The amount of data that is coming from the field of genetics is mounting by the day, and the same is happening in the field of medical imaging. Thus, as we merge more genetics with more brain imaging, the problem of false discoveries - a problem that is already scary enough - is actually getting bigger. This means that one has to be careful when drawing inferences on such a large quantity of data and hypotheses. In other words, if we are going to do hypothesis testing, we have to make a statistical decision, which may imply a necessary tradeoff between true and false discoveries. Considering current methodology trends within the field of Neuroimaging, that is to perform many tests and consequently reducing statistical power, how do we choose a threshold α ? Dr. Devlin thinks that uncorrected testing is definitely a wrong way to go, and suggests controlling for Experiment-wise Type I Error (FWER). This can be accomplished using novel methods such as Random Field Theory or more standard tests such a Bonferroni correction. The fact that in genetics or Neuroimaging some hypotheses are certainly true, gives “principle” to Bonferroni, whose only focus is in the number of false discoveries (FD), but Dr. Devlin suggests that we should aim for both power and principle. According to him, this is possible if we control for the false discovery rate (FDR), a method proposed by Benjamini and Hochberg in 1995 and defined as the expected number of false discoveries divided by the number of discoveries. As well as Bonferroni, FDR automatically controls the overall error rate in the special case that all the null hypotheses are true. However, when some null hypotheses are false, FDR will have better power. Additionally, more power for FDR can be obtain by focusing on dependent tests (i.e. GLM for phenotypes with pair-wise interactions), incorporating prior knowledge such as that derived from gene by gene interaction studies, by weighting the p-value based on prior information of brain regions, among others. FDR-based methods are a good way to control for false positives, generating better power than Bonferroni and making them suitable for genome wide association and linkage studies in the context of genetically complex diseases.