2018 SISG Practical

Session #2: Association Analysis

1. Download file 'JW_Practical' from SISG folder. This contains results for diseases from the GWAS website.

2. Read this file into R (or just use Excel for now).

3. Select from this file two diseases you find interesting, and for each disease, pick out a couple of SNPs.

4. Look at the papers for these SNPs.
    a. What statistical analysis approach was used? If regression, what type of regression model (e.g., logistic)?
    b. How did the authors address population stratification? Was this sufficient?

5. Interpret the results for these SNPs.
    a. What genetic model was assumed in coding of the SNPs?
    b. Write out an interpretation of the odds ratio (OR). How much does the risk increase for carrying one SNP? How about for carrying two SNPs?
    c.   Interpret the p-value?
    d. Can you figure out the standard error and 95% CI for the ORs?

6. A recent consortia made all of their summary statistics available (i.e., logistic regression coefficients), reporting the absolute value of each coefficient. What is the limitation of such data? How might you rectify this limitation?

7. A case/control study has been conducted and a SNP genotyped. Compute the odds ratios for the table below. Compute chi square tests for codominant, dominant and recessive modes of inheritance. Discuss the results in terms of plausibility of a model.

	X=0	X=1	X=2
Cases	500	350	120
Controls	521	270	130