GWAS Interpretation
By: Alicia Martin

From the class GWAS with data combined between last year and this year, we have the following data for the bitter phenotype with the most significantly correlated SNP (p=6e-5, N=58)

```
bitter    X713598
             CC CG GG NULL
  bitter_no   9  4  1    1
  bitter_yes  1 29 14    1
```

From the genotypes, we can construct the following contingency table:

		Can taste bitter?	
		Yes	No
Allele Counts	C	31	22
	G	57	6

From this contingency table, we can calculate an odds ratio and likelihood ratio.

Odds Ratio

Definition: odds ratio – a measure of effect size, describing the strength of association or non-independence between two binary data values.

In a case control study, this is the ratio between the fraction with the risk variant versus non-risk variant in the groups of affected versus the controls, i.e. expressed in terms of probabilities conditional on the affection status:

OR = (a × d)/(b × c) when you have the following contingency table:

		Treatment	
		Placebo	Drug
Event	Occurs	a	b
	Does not occur	c	d

Therefore, the OR is:

OR = (31 * 6) / (22 * 57)
OR = 0.15

Likelihood Ratio

Given a risk allele, what is the likelihood of having the disease?

LR = P(A | Trait) / P(A | No trait)
P(A | Trait) = P(Trait & A) / P(Trait)

In our example:

P(G | bitter taste) = 57/(57 + 31)
P(G | No trait) = 6/(6+22)

LR = [57/(57 + 31)] / [6/(6+22)]
LR = 3.02

This is the likelihood of tasting bitter given a G allele. What does this mean in terms of GG and GC? You can use P(GG | Trait) / P(GG | No trait) and look at GG versus GC and CC using the genotypes table from last week.

Relative Risk

The relative risk (RR) is the probability that a member of an exposed group will develop a disease relative to the probability that a member of an unexposed group will develop that same disease. This calculation requires that we know something about the overall incidence of the trait.

R_{CC} = P(trait | CC)/population prevalence of trait
R_{CC} = P(affected | CC genotype)/population affected prevalence