SISG 2020: Module 11

Session 1: Introduction

Each person, please introduce yourself to the other members in your group:

1. Your name and pronouns. Your position (student, researcher) and affiliation (what University or institute).
2. What are your strengths in your training so far? (i.e., is your background in genetics, biostatistics, law?)
3. What prompted you to take this course? What are you hoping to learn?

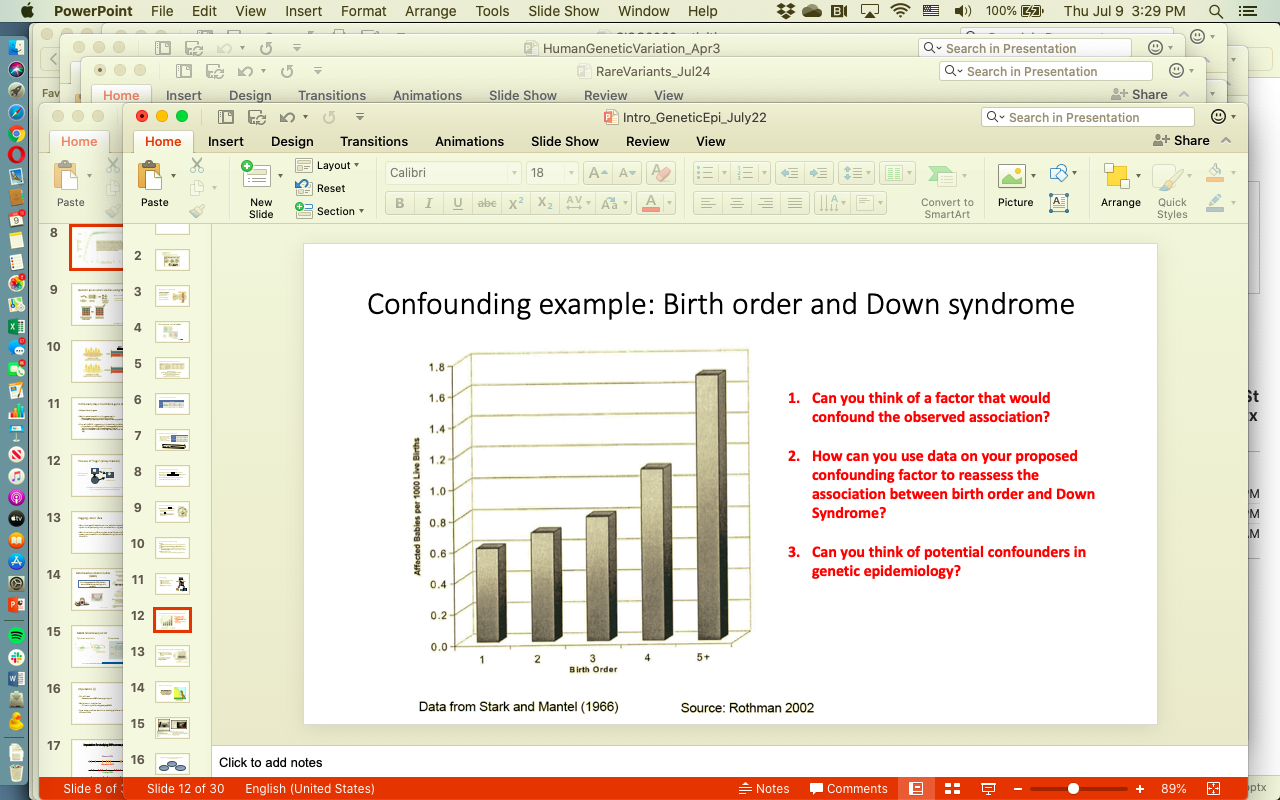
Once everyone is introduced, discuss in your group:

1. Why do we study the role of genetics in human disease?

SISG 2020: Module 11

Session 2: Introduction to Genetic Epidemiology

The following questions are based on this figure:

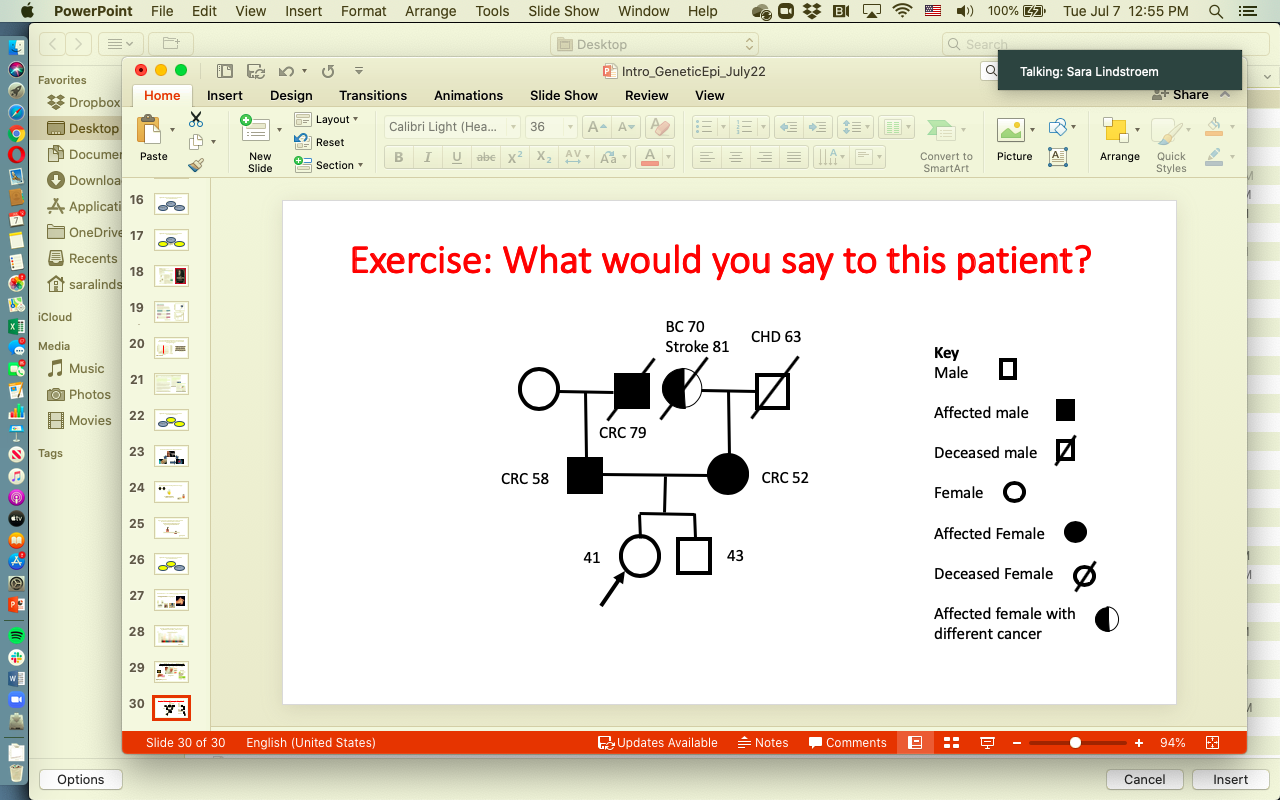


1. One of the most significant factors confounding the observed association is the maternal age. It is associated with both the exposure (Birth order) and the outcome (Incidence of Down Syndrome)

2. One approach is to stratify the study population based on the maternal age and investigate the association of interest within each stratum.

3. When studying the genotype-phenotype association in genetic epidemiology research, the most significant confounder is population structure, which represents the difference in allele frequency between the subpopulations.

Assume you are a medical professional and the 41 yr old female patient identified with the arrow visits you for a general health consultation, She is concerned about her risk for colorectal cancer (CRC) based on her family history. What would you tell her based on her family pedigree?

****

The parents of the patient have an early-onset of colorectal cancer, which suggests that she may have a higher risk of the disease due to inherited reason. She should be suggested with healthier life-style, early screening, and additional preventive approach.

SISG 2020: Module 11

Session3: Human Genetic Variation

1. A recent study sequenced the genome of 2,504 individuals and identified 84.7 million SNPs (single nucleotide polymorphisms) between the participants. On average, each individual carried 3.5-4.3 million SNPs each. About 0.5% of those SNPs were in coding regions of genes. Remember, 1.5% of the genome is in a coding region. Why might only 0.5% of variants be in coding regions compared to what would be expected if SNPs were randomly allocated throughout the genome?

Heterogeneity of SNPs distribution across the genome suggests the coding regions are more conservative. The polymorphisms in coding region may lead to severe functional consequences (missense mutations, nonsense mutations, frameshift mutations), and may be eliminated before they get the chance to be passed to the next generation.

1. Match the genetic term with the definition:
2. Nonsense
3. Heterozygous
4. Exon
5. Allele
6. Synonymous
7. Missense
8. Non-coding region
9. Haplotype
10. Autosomal
11. Phenotype
12. Genotype
13. Frameshift
14. Intron
15. Homozygous

d\_\_\_ Alternative forms of a gene or DNA base.

k\_\_\_ Genetic makeup of an individual at a particular DNA location based on both alleles.

b\_\_\_ Genotype consisting of two different alleles at a particular location.

e\_\_\_ DNA base change that does not change the translated amino acid.

n\_\_\_ Genotype consisting of two of the same alleles at a particular location.

j\_\_\_ Observable characteristics resulting from a genotype.

i\_\_\_ Concerning the 22 pairs of chromosomes that are not sex chromosomes.

m\_\_\_ Portion of gene that does not code for amino acids and appears in between exons.

l\_\_\_ Insertion or deletion mutation that changes the whole subsequent sequence of amino acids by changing the 3-codon groups for generating amino acids.

c\_\_\_ Portion of gene that encodes amino acids.

g\_\_\_ Section of DNA that does not become protein.

a\_\_\_ Substitution of a single DNA base that causes a stop in protein production.

f\_\_\_ DNA base change that changes the translated amino acid.

h\_\_\_ Set of DNA variations at several positions that are inherited together.

1. Look up “rs7412” in dbSNP (https://www.ncbi.nlm.nih.gov/snp/). <https://www.ncbi.nlm.nih.gov/snp/rs7412>
   1. What are the DNA bases identified at this location? C/T
   2. What gene is this SNP located in? *APOE*
   3. What is the effect of this SNP on the amino acid sequence? Missense Mutation
   4. Click on the “frequency” tab to the left. What is the frequency of the minor allele (less common allele) in the 1000 Genomes study overall? How do these frequencies differ by ancestral subgroup within this study? General population: 92.5% C/7.5% T