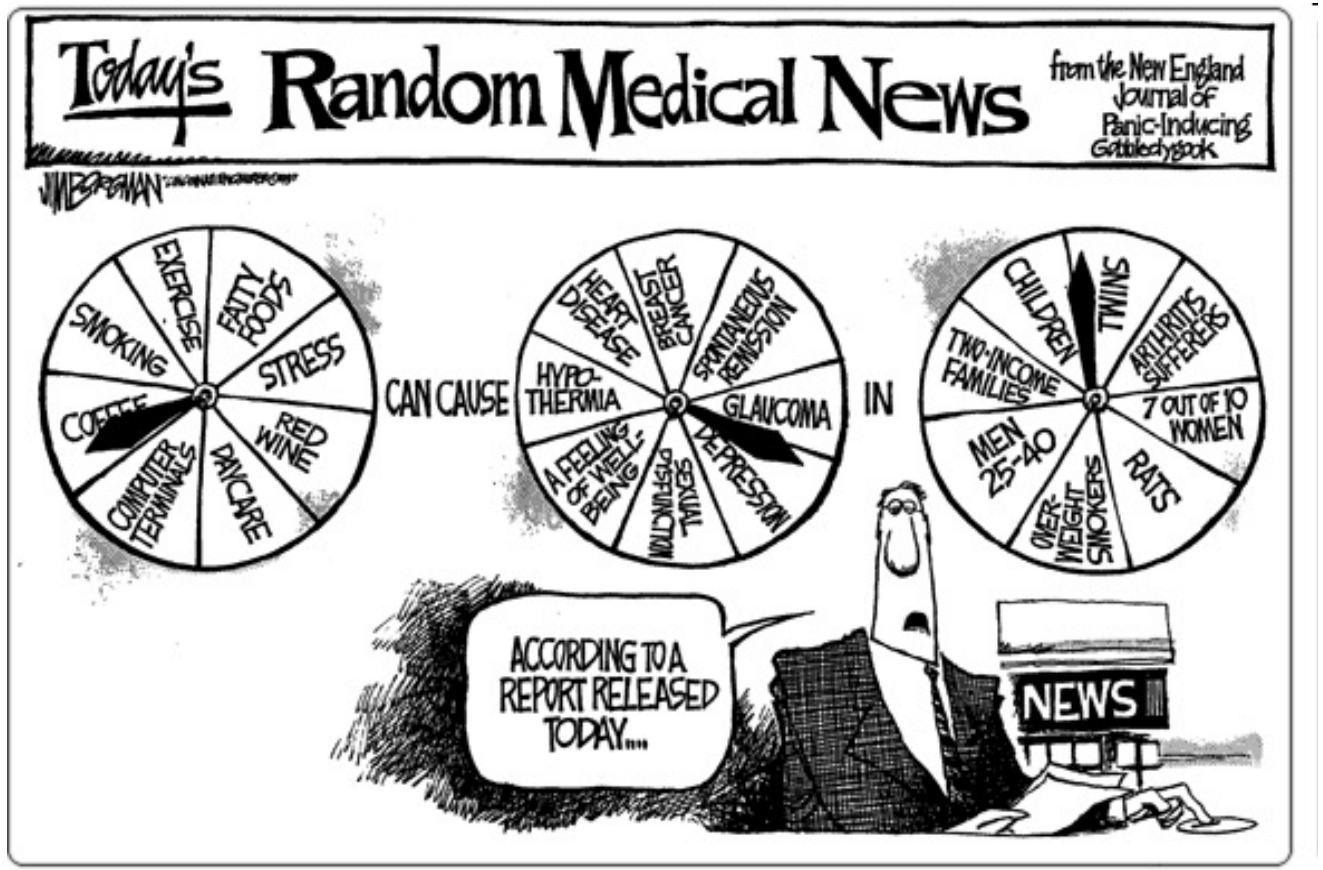
Session 2:
Introduction to Epidemiology and Genetic Epidemiology

ep·i·de·mi·ol·o·gy

/ˌepəˌdēmēˈäləjē/

noun

the branch of medicine that deals with the incidence, distribution, and possible control of diseases and other factors relating to health.

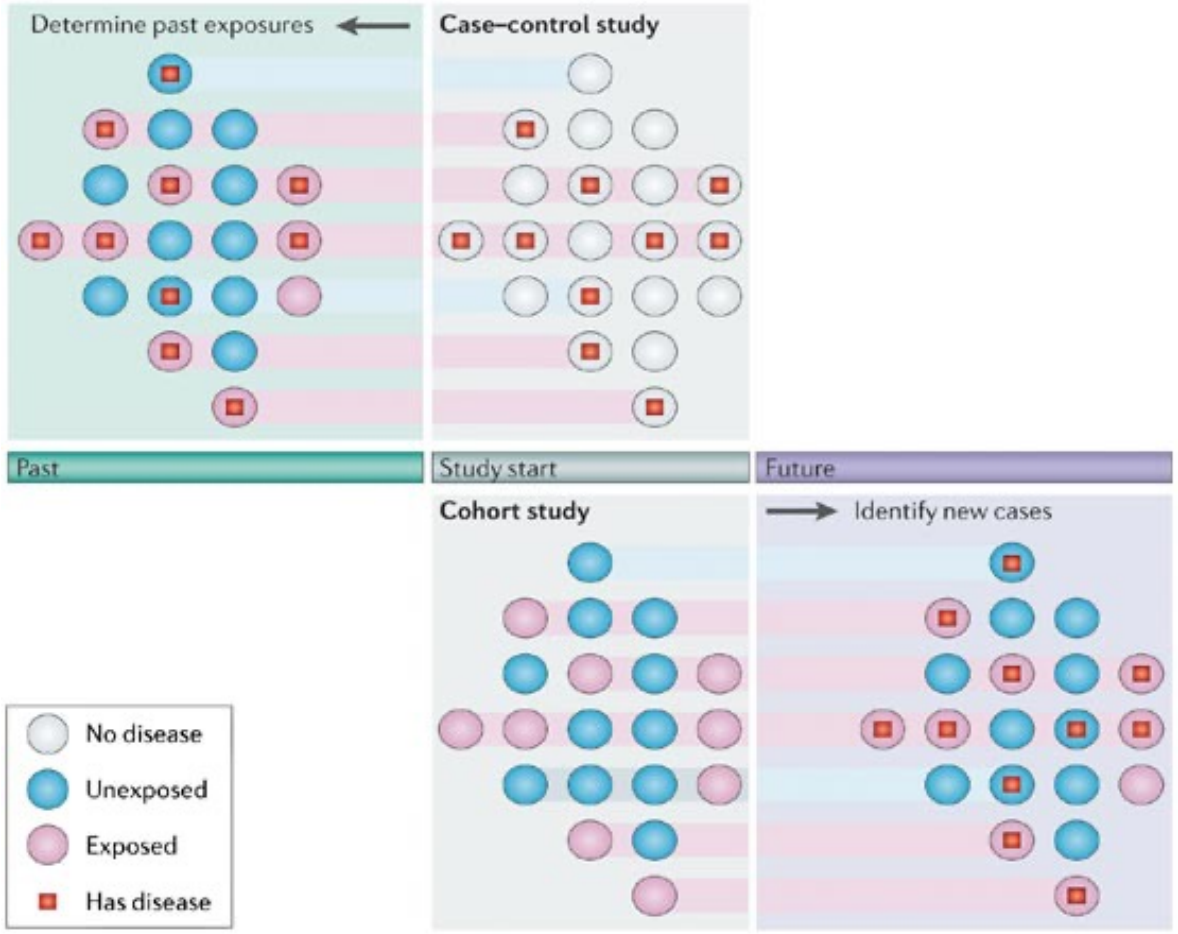


Major goals in Epidemiology

- To obtain an *unbiased & precise* estimate of the true effect of an exposure or intervention on outcome in the population at risk

- To use this knowledge to prevent and treat disease

Cohort vs. case-control studies



Copyright © 2006 Nature Publishing Group
Nature Reviews | Genetics

Manolio *et al. Nature Reviews Genetics* **7**, 812–820 (October 2006) | doi:10.1038/nrg1919

nature REVIEWS GENETICS

Estimated incidence rates in cohorts

Disease incidence per 100,000 per year (%)	Disease examples	Number of incident cases in 5 years for different cohort sizes		
		200,000	500,000	1,000,000
10 (0.01)	Parkinson disease, schizophrenia	91	228	457
50 (0.05)	Colorectal cancer, renal failure	456	1,141	2,282
100 (0.10)	Breast cancer, hip fracture	912	2,279	4,559
200 (0.20)	Diabetes, stroke, heart failure	1,820	4,550	9,100
500 (0.50)	Myocardial infarction, all cancers	4,524	11,309	22,618
3,000 (3.00)	Cataracts, hypertension	25,858	64,644	129,289

Estimated numbers of incident cases available after 5 years of follow-up across the entire age range in the US population are shown, assuming an attrition rate of 3% per year. Data are taken from the Incidence and Prevalence Database.

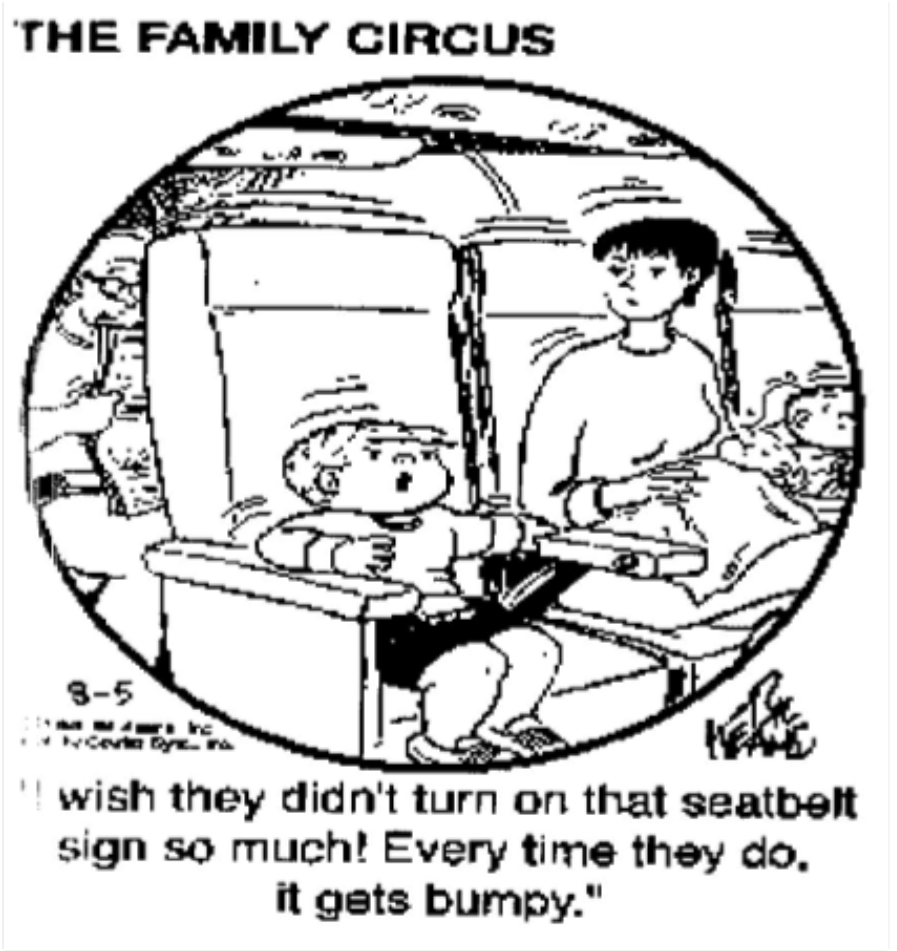
Manolio. *Nature Reviews Genetics* 2006

Compared to cohorts, case-control studies are cheap, fast and powerful

However, case-control studies suffer from several drawbacks:
the need to identify appropriate controls
they are more sensitive to recall bias

Association and Causality

- An exposure and outcome are *associated* if there is a differential distribution:
  - The prevalence of exposure differs between cases and controls.
  - An exposure is *causal* for the outcome if the presence (or absence) of the exposure directly or indirectly influences whether the outcome occurs.



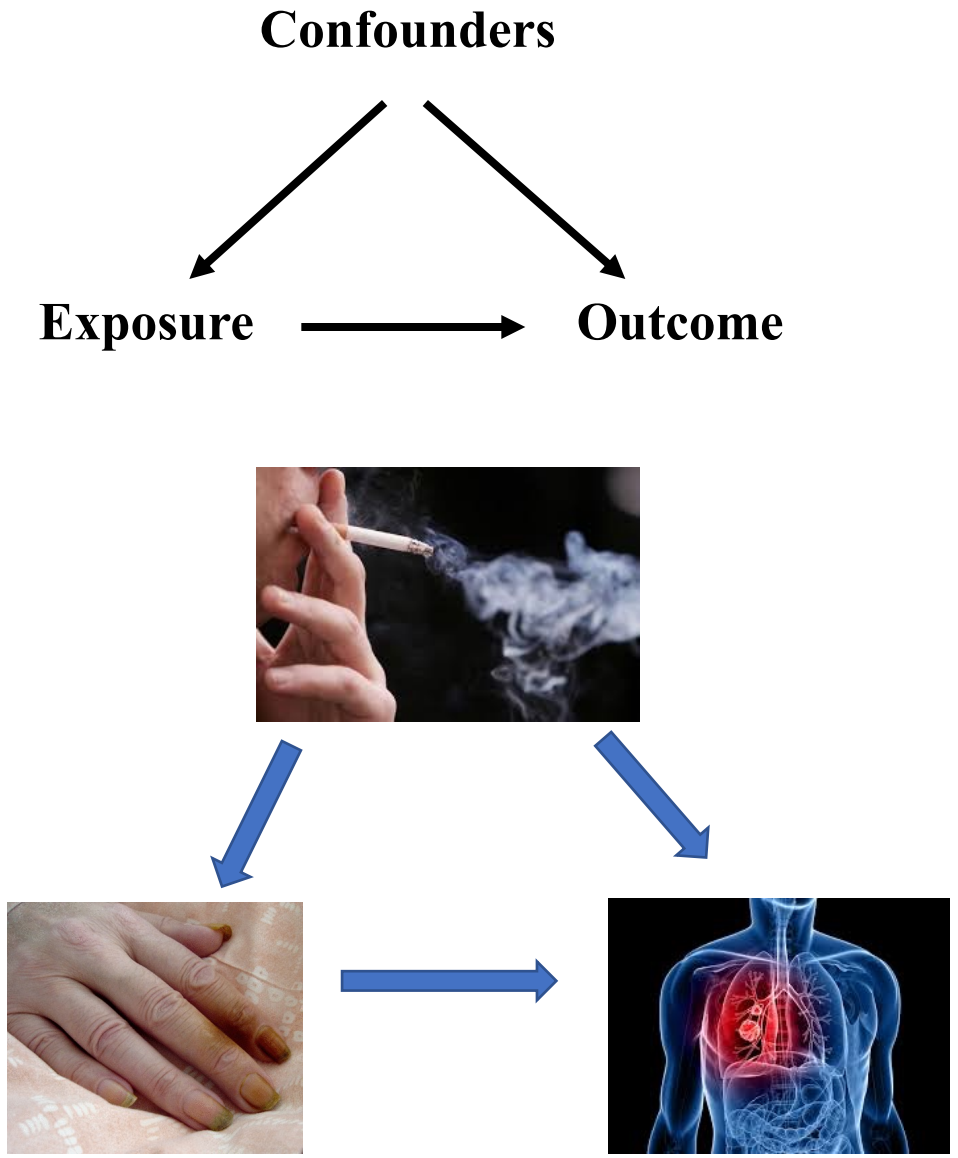
Sources of bias in epidemiology

- **Selection Bias**
  - Arises when cases and controls are coming from different source populations (e.g., pediatric cases, adult controls)
- **Survival bias**
  - When cases are recruited some time after they were diagnosed. Might lead to a milder form of disease. This is especially true for aggressive/fatal disease (e.g., pancreatic cancer, heart attack)
- **Diagnostic bias**
  - If the investigator determining the outcome knows whether the person was exposed or not to the risk factor under study (e.g., if the radiologist knows that a potential pulmonary disease patient smokes, they may look more carefully at the x-ray).
- **Recall bias**
  - Accuracy and completeness of exposures, life-style behaviors,… (e.g., cases might be more motivated to complete a questionnaire accurately).

Confounding

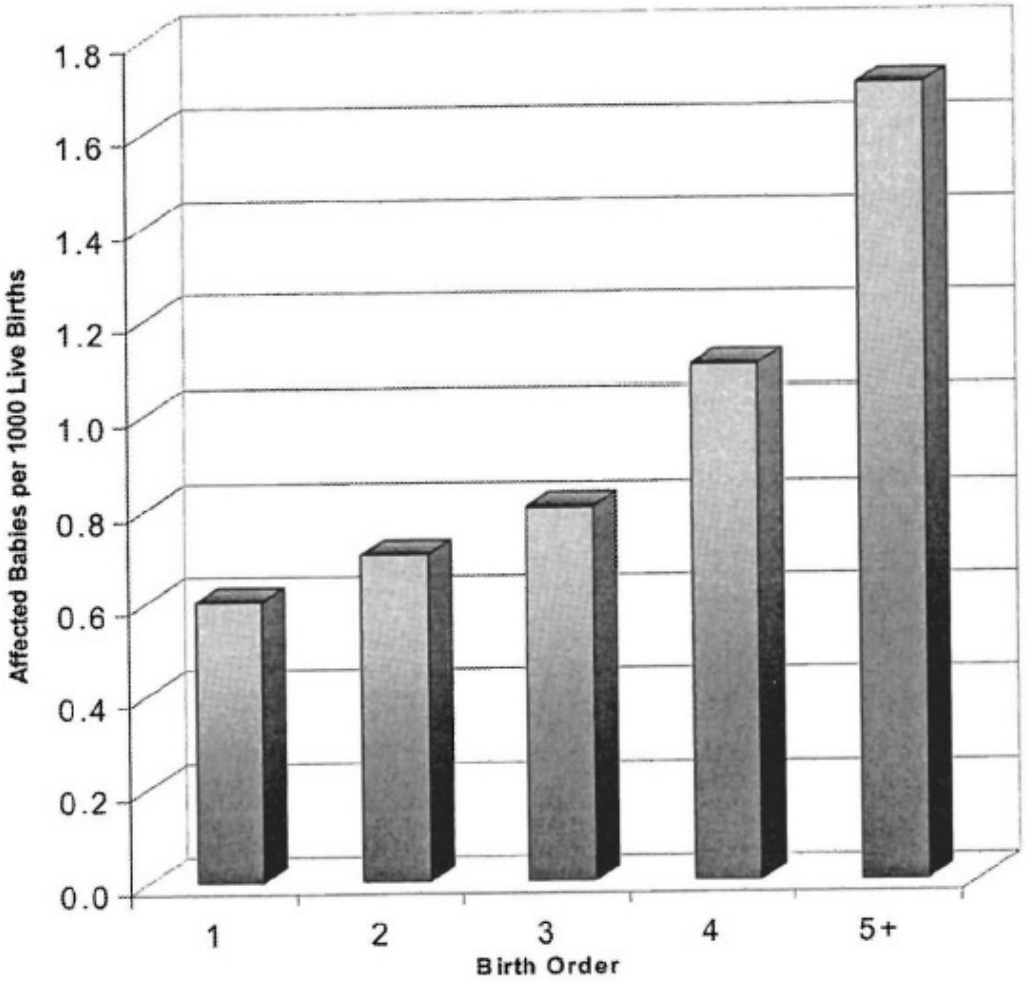
- A confounder is often defined as a factor that is:
  ① A risk factor for the disease
  ② Associated with the exposure
  ③ Not a direct result of the exposure

- Confounding can lead to false positive findings.



BREAKOUT ACTIVITY

Confounding example: Birth order and Down syndrome



Data from Stark and Mantel (1966)

Source: Rothman 2002

1. **Can you think of a factor that would confound the observed association?**
2. **How can you use data on your proposed confounding factor to reassess the association between birth order and Down Syndrome?**
3. **Can you think of potential confounders in genetic epidemiology?**

Genetic Epidemiology

Genetic epidemiology *is the study of the role of* ***genetic*** *factors in determining health and disease in families and in populations, and the interplay of such* ***genetic*** *factors with environmental factors.*

Ed was unlucky enough to find the needle in the haystack!

April 13, 2007

'Fat' gene found by scientists

Mark Henderson, Science Editor

A gene that contributes to obesity has been identified for the first time, promising to explain why some people easily put on weight while others with similar lifestyles stay slim.

Smoking addiction gene found

Scientists say a gene makes people more likely to get hooked on tobacco, causing them to smoke more, making it harder to quit, and leading more often to deadly lung cancer. **Full story**

Newsweek: Differing conclusions

Researchers make human-cow embryos

Science wishy-washy on water benefits | **Vote**

Recommend 414 Share Tweet

Brain-Aging Gene Discovered

Genetic variant accelerates normal brain aging in older people by up to 12 years

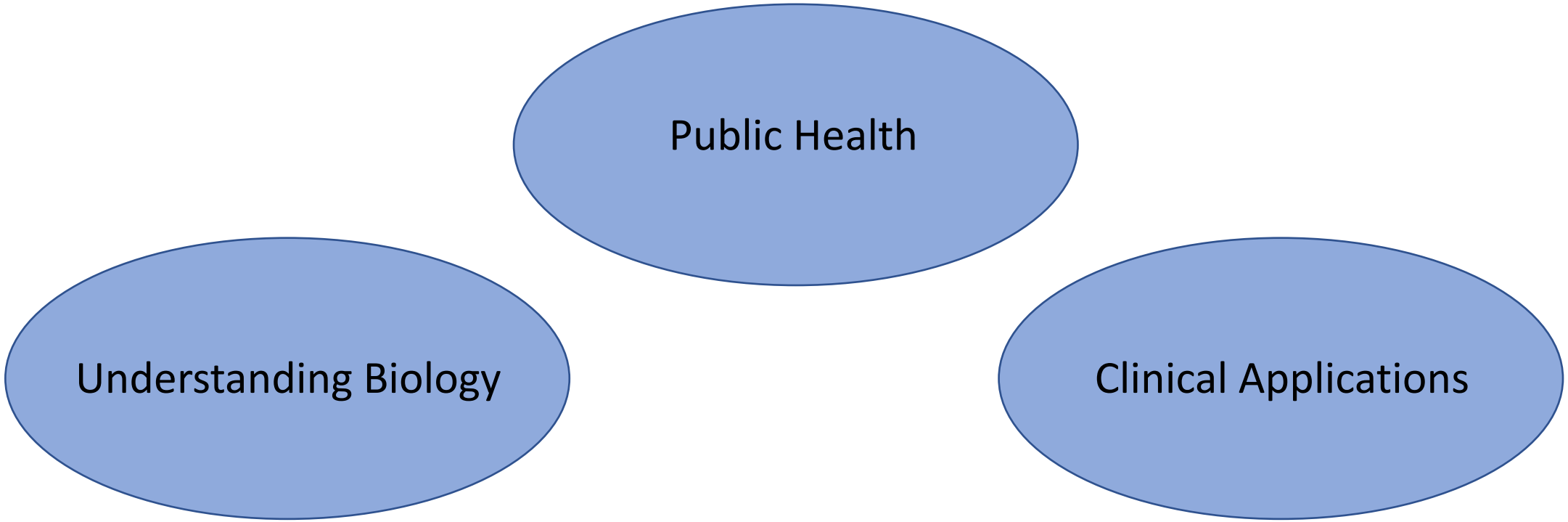
March 15, 2017

Posted in: Neurology / Medicine

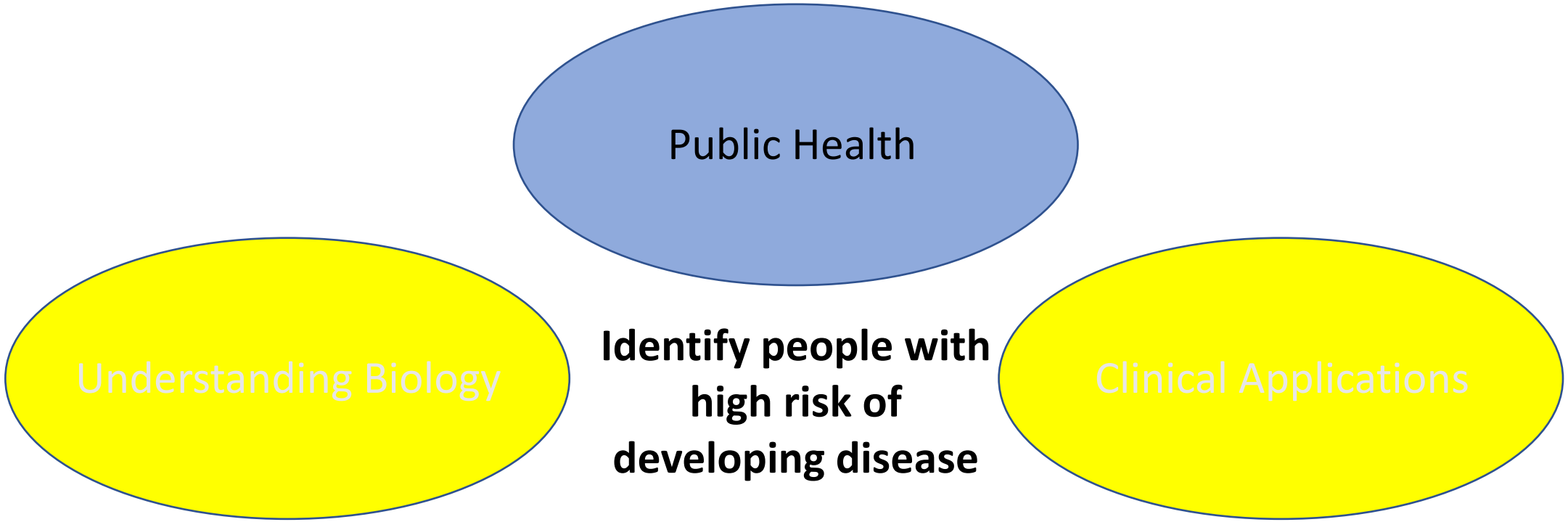
Does aspirin prevent colorectal cancer? Depends on your DNA

Why do we want to study how our genome is involved in disease?

Why do we want to study how our genome is involved in disease?



Why do we want to study how our genome is involved in disease?



The New York Times

The Opinion Pages

OP-ED CONTRIBUTOR

My Medical Choice

By ANGELINA JOLIE

Published: May 14, 2013 | 1712 Comments

LOS ANGELES

Enlarge This Image

Loren Capelli

MY MOTHER fought cancer for almost a decade and died at 56. She held out long enough to meet the first of her grandchildren and to hold them in her arms. But my other children will never have the chance to know her and experience how loving and gracious she was.

We often speak of "Mommy's mommy," and I find myself trying to explain the illness that took her away from us. They have asked if the same could happen to me. I have always told them not to worry, but the truth is I carry a "faulty" gene, BRCA1, which sharply increases my risk of developing breast cancer and ovarian cancer.

FACEBOOK
TWITTER
GOOGLE+
SAVE
E-MAIL
SHARE
PRINT
REPRINTS

Enough Said
Now Playing

NATIONAL CANCER INSTITUTE
CHANCES OF DEVELOPING BREAST CANCER BY AGE 70

Specific inherited mutations in the BRCA1 and BRCA2 genes increase the risk of breast and ovarian cancers. Testing for these mutations is usually recommended in women without breast cancer only when the person's individual or family history suggests the possible presence of a harmful mutation in BRCA1 or BRCA2. Testing is often recommended in younger women newly diagnosed with breast cancer because it can influence treatment decisions and have implications for their family members.

MUTATED BRCA1

55-65%

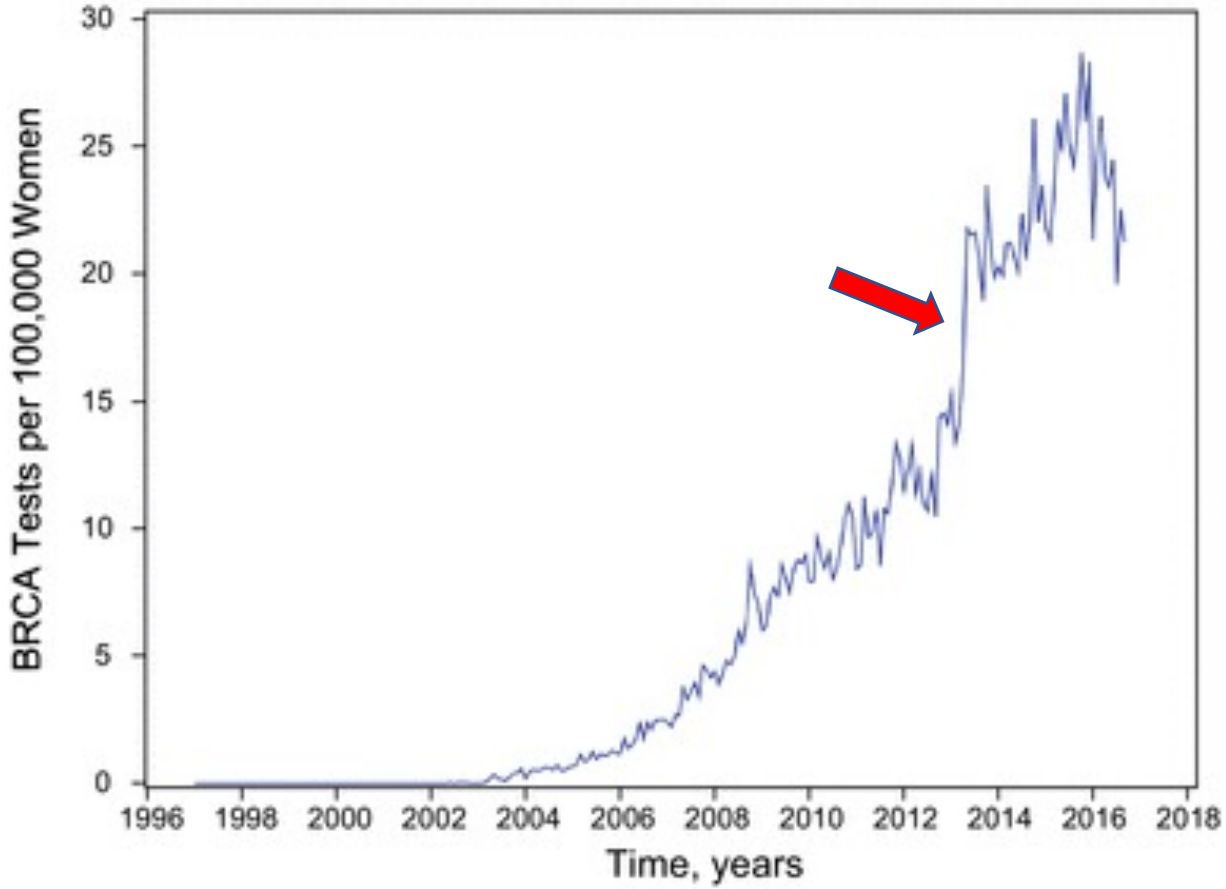
MUTATED BRCA2

45%

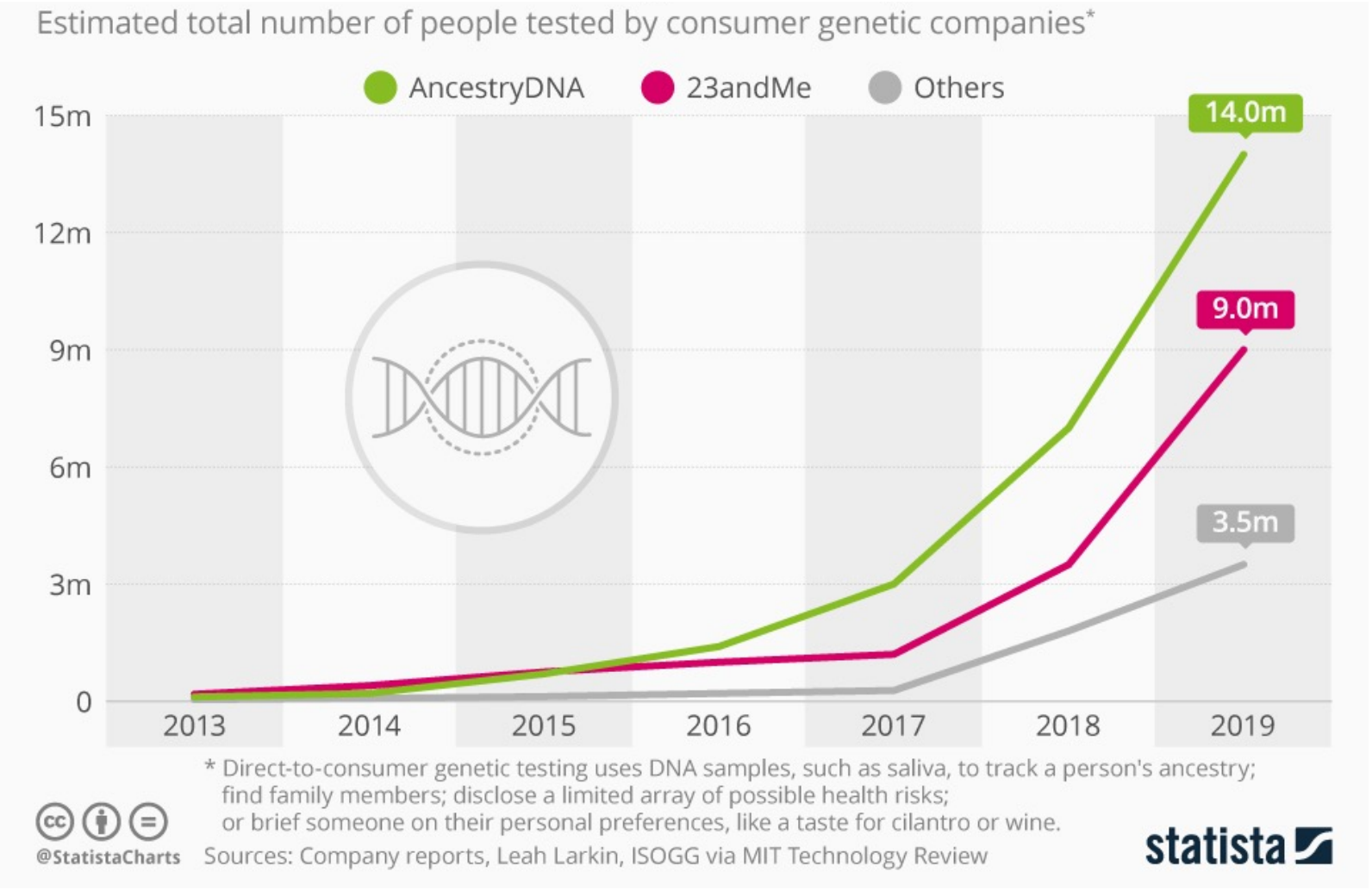
NORMAL BRCA

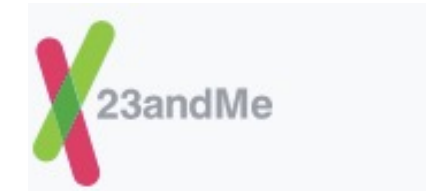
12%

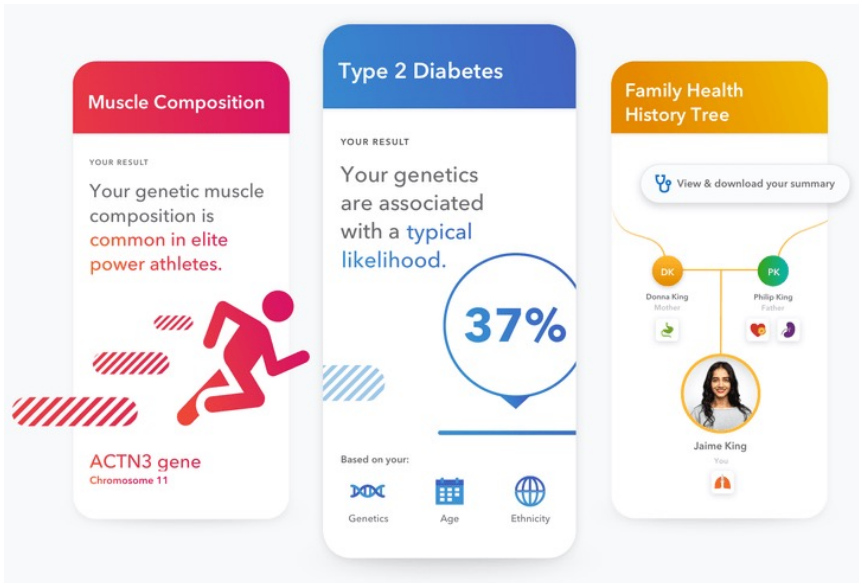
www.cancer.gov/brca-fact-sheet



Liede et al. BCRT 2018

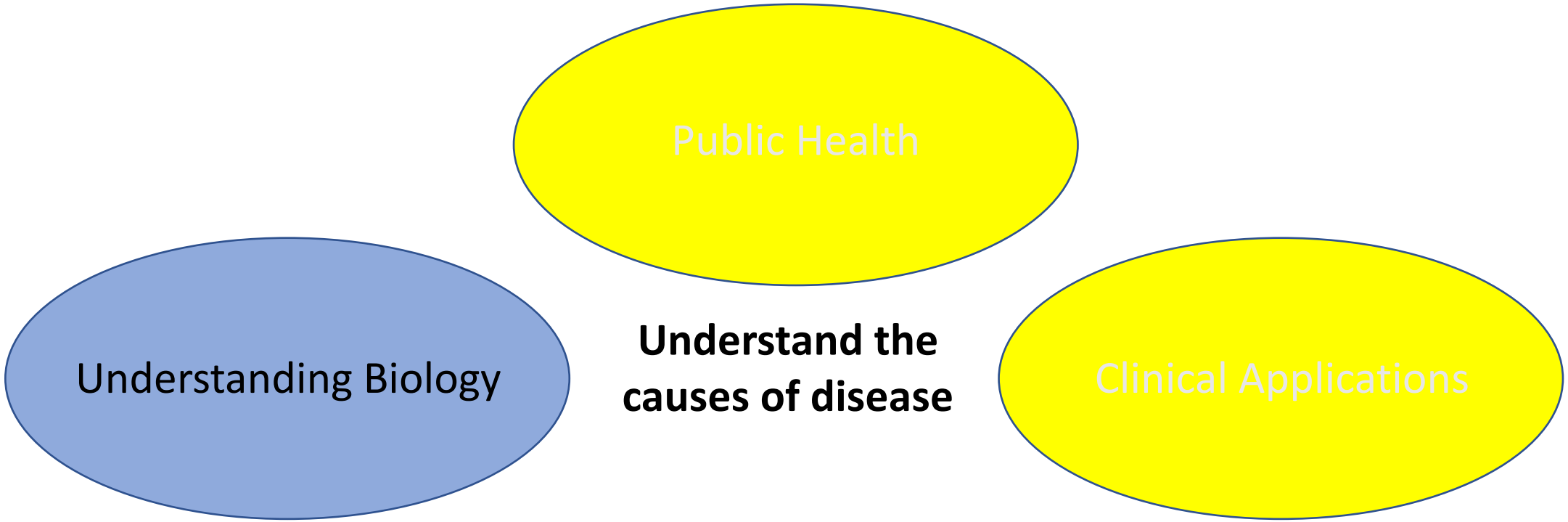






https://www.statista.com/chart/17023/commercial-genetic-testing/

Why do we want to study how our genome is involved in disease?



“Association does not imply causation”

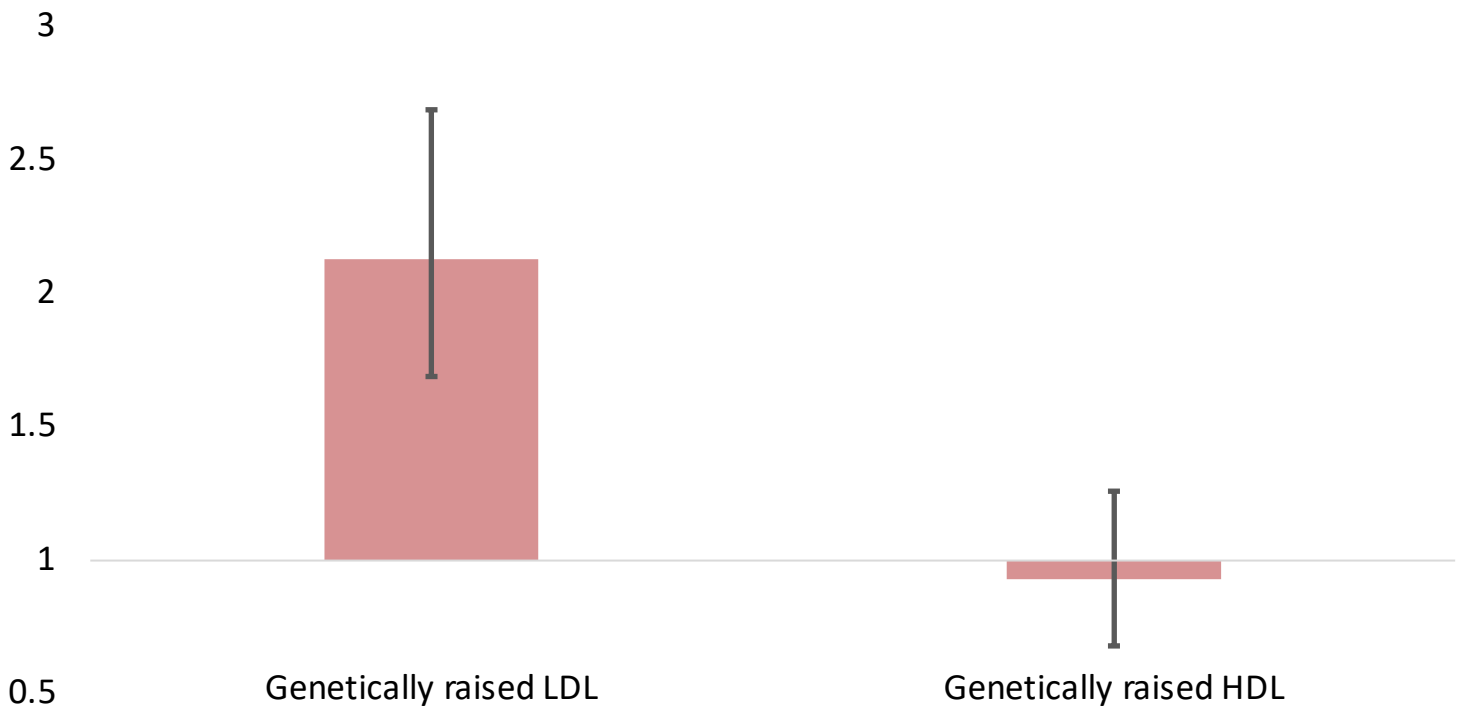
HDL (“Good”) Cholesterol and Myocardial Infarction (MI)

- ⬆ HDL -> ⬇ MI risk

Increasing HDL concentrations might help decrease cardiovascular disease risk.

People who carry gene variants that increase HDL do not have a lower risk of MI

Since HDL is correlated with exercise, weight loss, diet (nuts, fish) it is likely that these lower your risk for MI rather than HDL itself



Voight et al, Lancet 2012

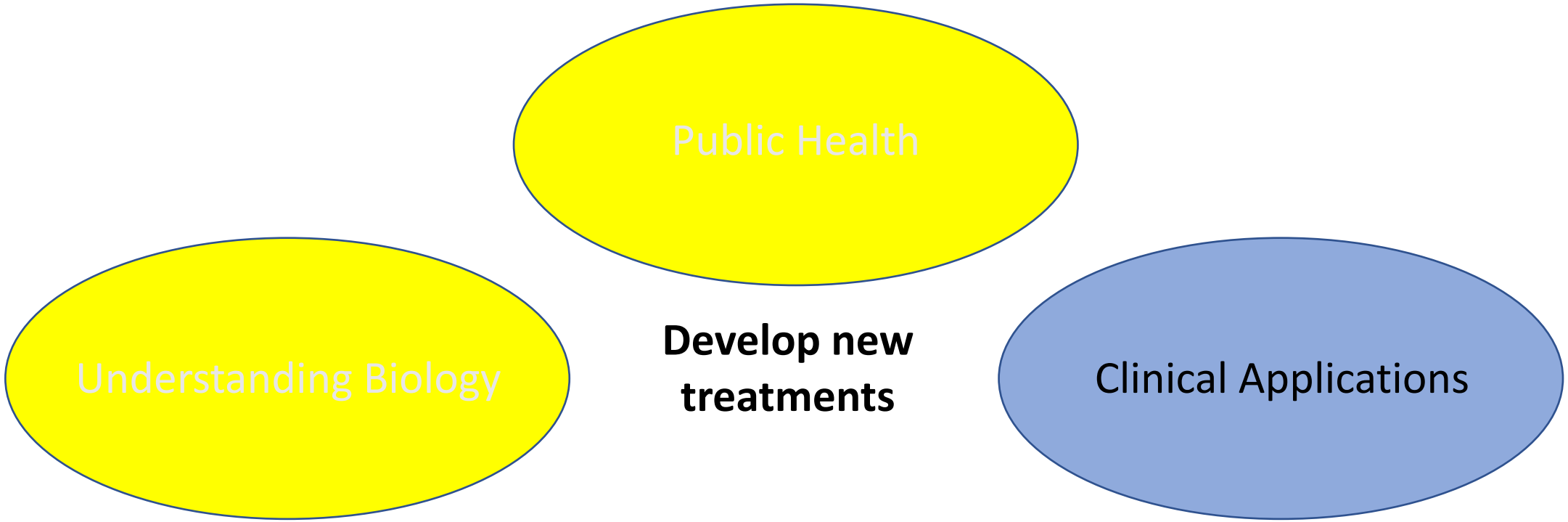
RESEARCH ARTICLES

HEART DISEASE

Rare variant in scavenger receptor BI raises HDL cholesterol and increases risk of coronary heart disease

Zanoni et al, Science 2016

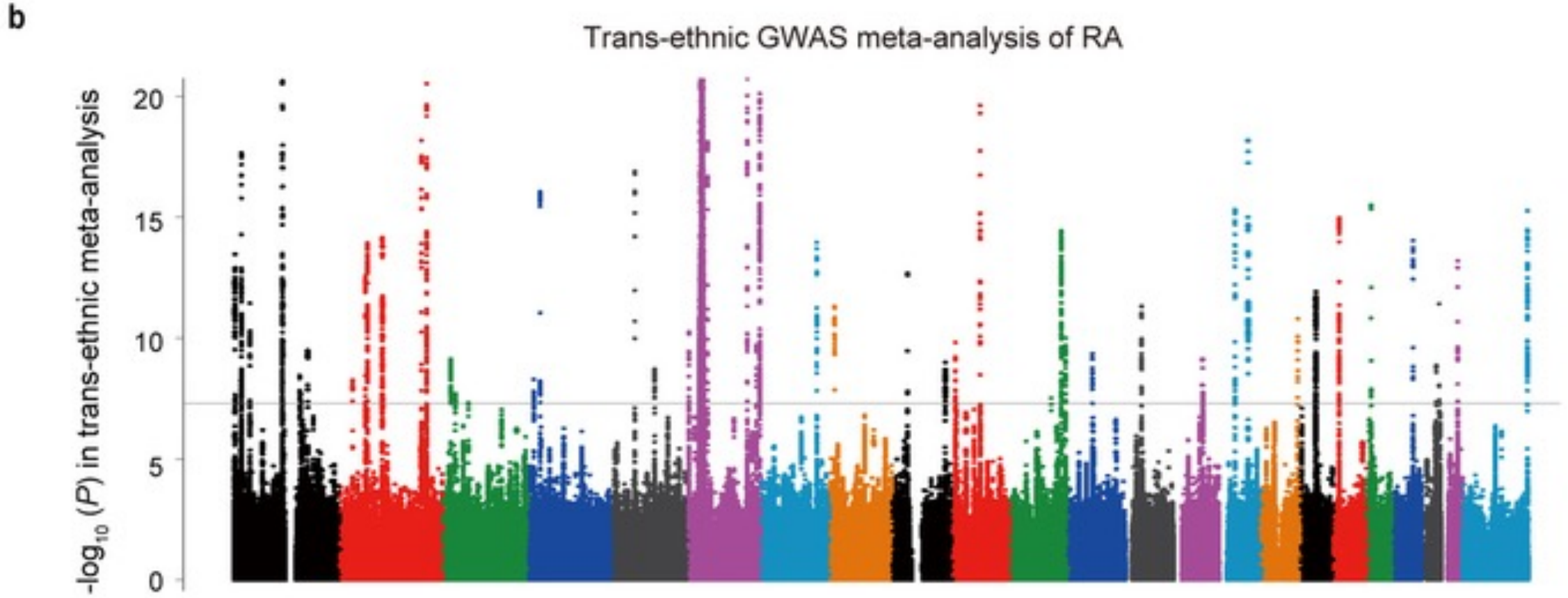
Why do we want to study how our genome is involved in disease?



Rheumatoid Arthritis – an inflammatory, crippling, incurable disease

- In 2005, an estimated 1.5 million (0.6%) of US adults age ≥ 18 had RA.

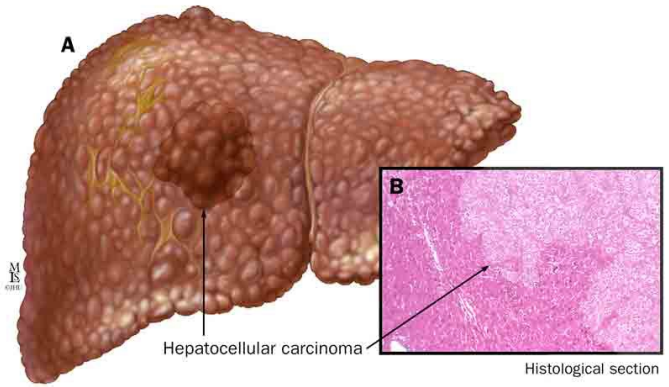
A study of 10 million genetic variants in 29,880 RA cases and 73,758 controls



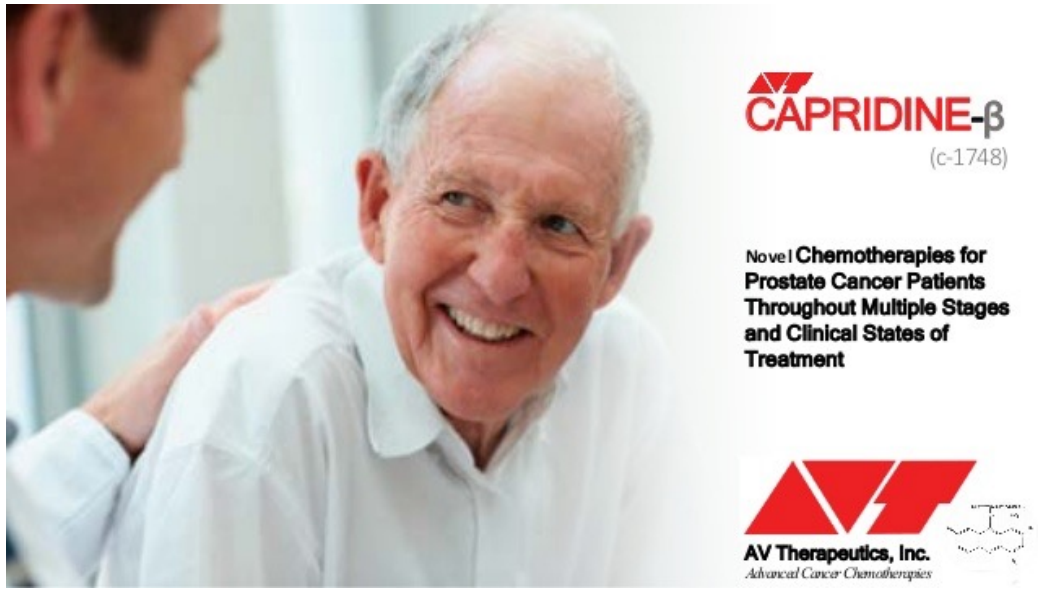
Okada, Nature, 2014

Identified genes are targets of approved therapies for RA, and further suggest that drugs approved for other diseases may be repurposed for the treatment of RA

flavopiridol

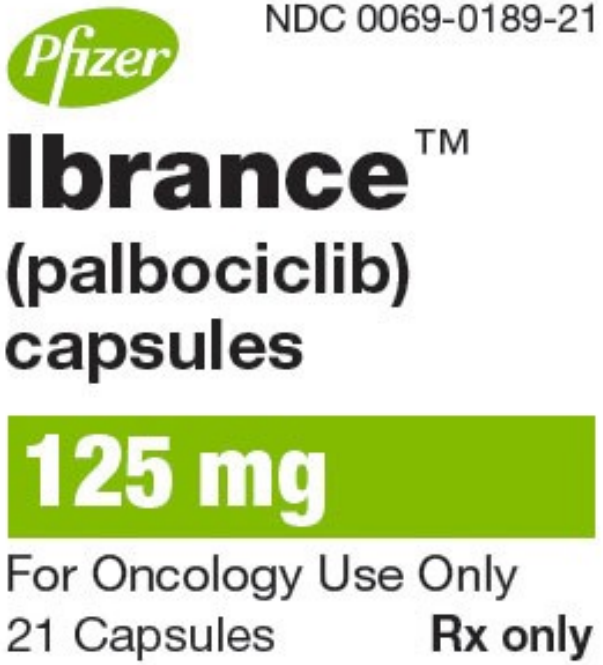


Lymphoma/Leukemia/Liver cancer



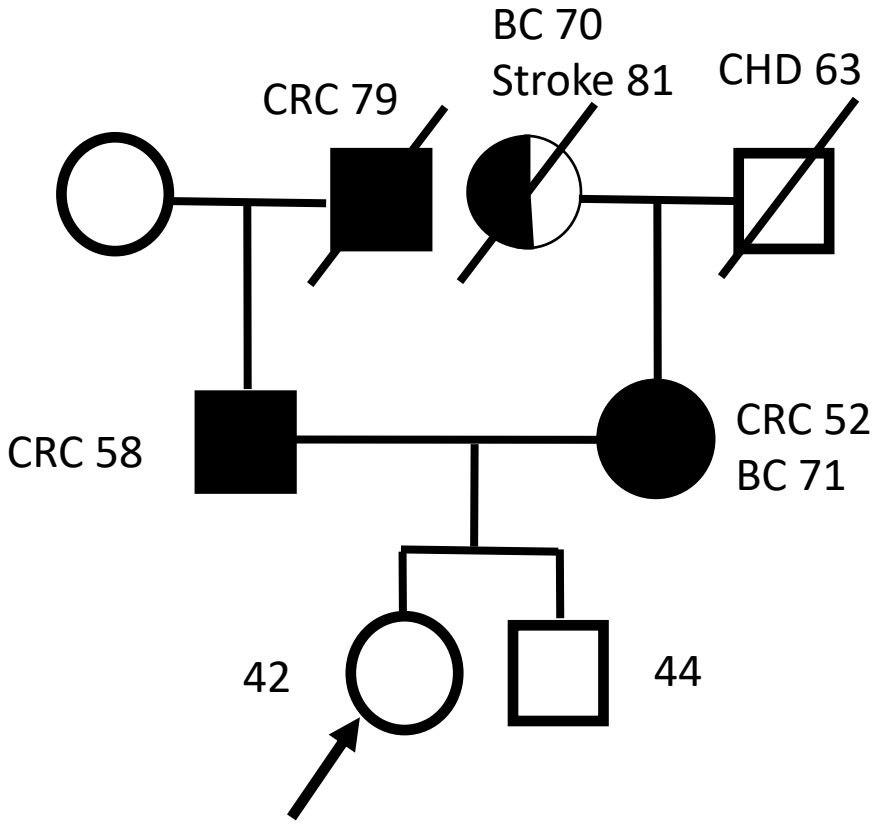
Cancers/psoriasis

Breast cancer



Okada, Nature, 2014

What would you say to this patient? (Answers in the chat)



Key

Male

Affected male

Deceased male

Female

Affected Female

Deceased Female

Affected female with different cancer

CRC – colorectal cancer

BC – breast cancer