SISG 2019: Module 11
Genetic Epidemiology

Housekeeping

- Introductions
- Access to website/internet access
- Evaluations
- Schedule

Day	Time	Lead	Topics	Details
Wed	1:30-2:00	Alie/Sara	Class Intro	Intro to class/Review Agenda and topics Student introductions
	2:00-3:00	Sara	Overview of Genetic Epi	Intro to Epidemiology and Genetic Epidemiology
	BREAK			
	3:30-5:00	Alie	Human genetic variation	Types of genetic variation, linkage disequilibrium.

Thursday	8:30-10:00	Alie	Population Genetics	Hardy-Weinberg Equilibrium, population structure, admixture mapping.
	BREAK			
	10:30-12:00	Sara	Family-based Studies	Linkage Analysis, family-based association studies.
	BREAK			
	1:30-3:00	Alie	Association Studies	Sequencing, genotyping, imputation, association analyses.
	BREAK			
	3:30-5:00	Sara	Association Studies	GWAS, rare variants.

Friday	8:30-9:15	Sara	Mendelian Randomization	Concept, methods
	9:15-10:00	Alie	Pharmacogenetics	Pathways of pgx
	BREAK			
	10:30-12:00	Sara	Gene-Environment Interactions	Definitions, methods, practical issues
	BREAK			
	1:30-3.00	Alie	Bioethics and Implementation	PPV, NPV, sensitivity, specificity, principles of bioethics, genetic epidemiology in law
	BREAK			
	3:30-4:30	Sara	Risk prediction	Methods, applications
	4:30-5:00	Alie/Sara	Wrap-up	

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.

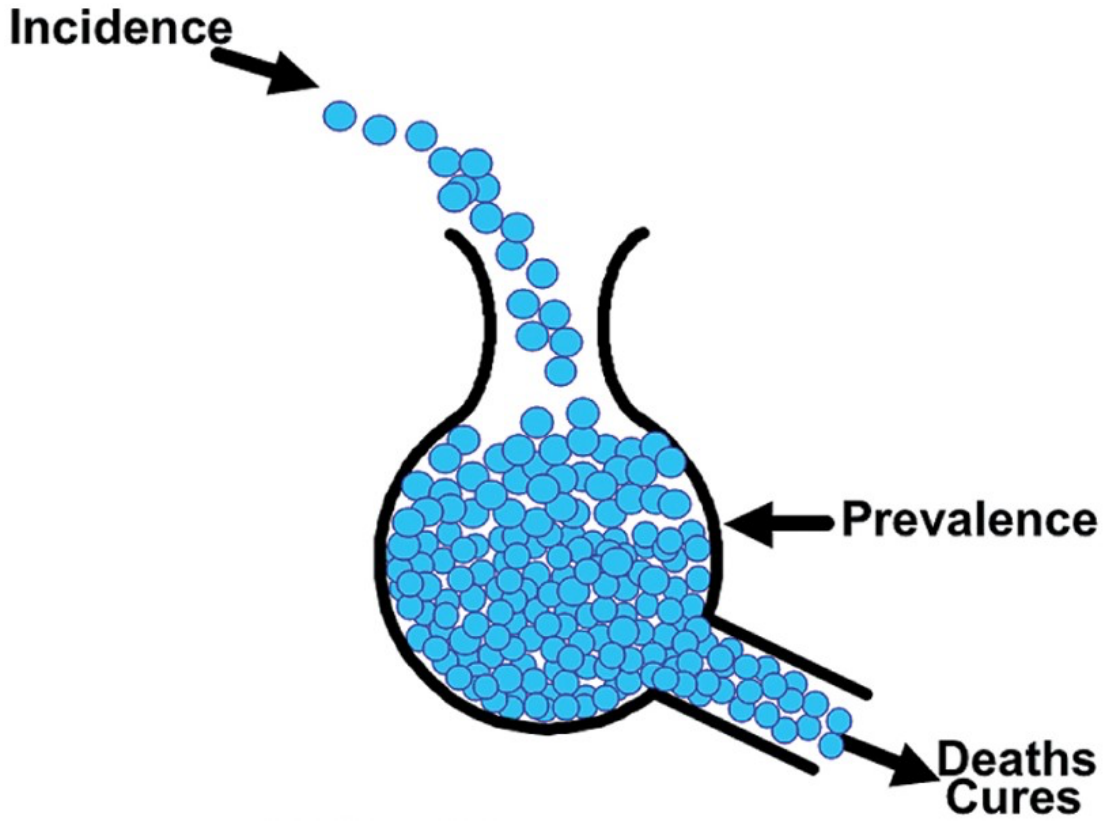
Major goals in Epidemiology

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

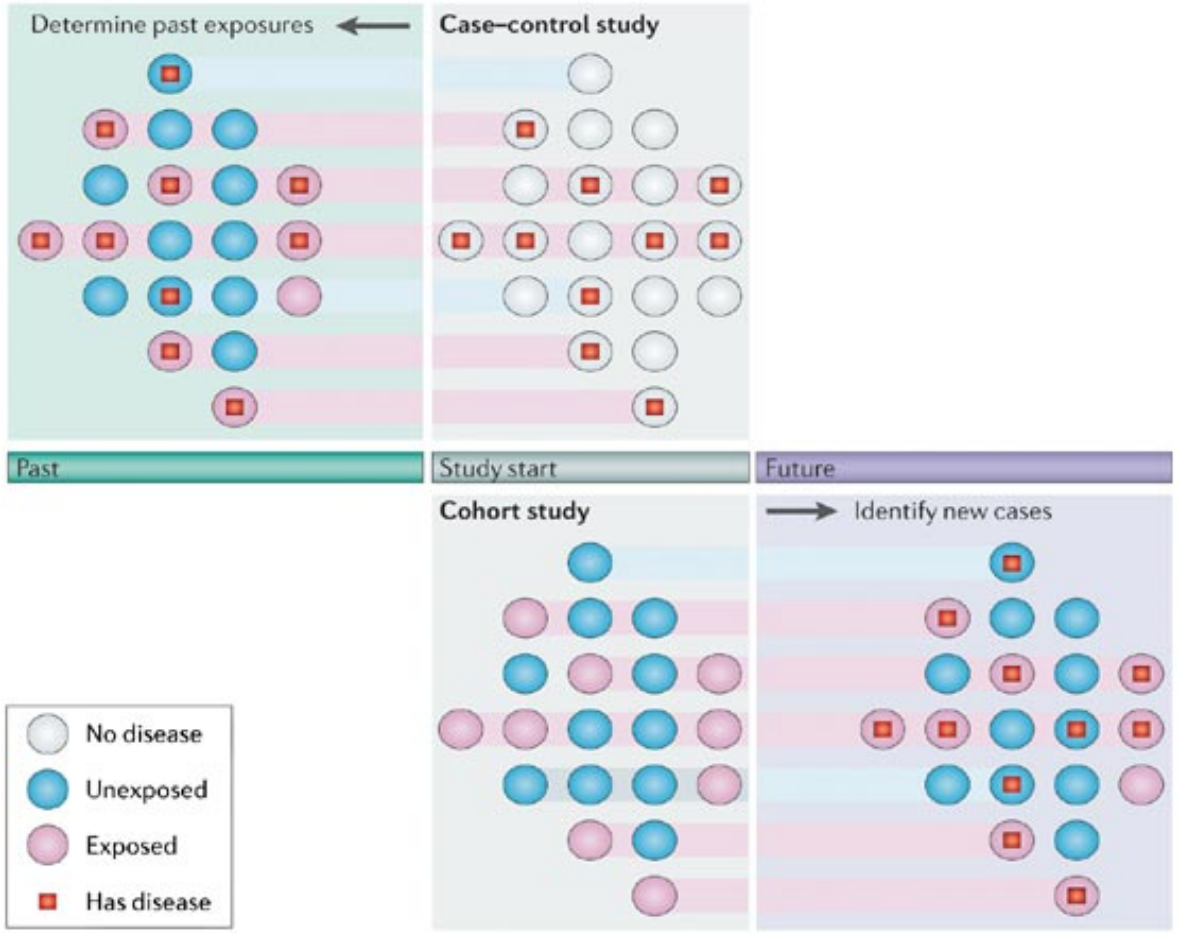
- To use this knowledge to prevent and treat disease

Key concepts in Epidemiology

- Incidence
  - Number of new cases in a population during a fixed time period
    - The reported number of new prostate cancer cases in United States during 2015 was 183,529.
- Prevalence
  - Number of existing cases in a population at a given time
    - In 2015, there were an estimated 3,120,176 men living with prostate cancer in the United States.



Cohort vs. case-control studies



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

it is difficult to assess rare exposures due to small sample sizes

The 2x2 Table For Count Data

		Disease status		
		Cases	Controls	Total
Exposure	Exposed	a	b	a+b
	Not Exposed	c	d	c+d
Total		a+c	b+d	a+b+c+d

Main Measures of Association in Epidemiology

- Relative Risk

  measure of the *relative probability of developing disease* given exposure status

- Odds Ratio

  measure of the *relative odds of exposure* given disease status

Relative Risk (RR) For Count Data

- Relative probability of developing disease given exposure status
- Used in cohorts
- Also known as risk ratio
- If no association RR=1

	Cases	Controls	Total
Exposed	a	b	a+b
Not Exposed	c	d	c+d
	a+c	b+d	a+b+c+d

$$\text{RR} = \frac{a/(a+b)}{c/(c+d)} = \frac{\text{(Incidence of Disease in Exposed)}}{\text{(Incidence of Disease in Unexposed)}}$$

Odds Ratio (OR) For Count Data

- Relative odds of exposure given disease status
- Used primarily in case-control studies
- Good *estimate* of RR
- If no association OR=1

	Cases	Controls	Total
Exposed	a	b	a+b
Not Exposed	c	d	c+d
	a+c	b+d	a+b+c+d

$$\textbf{OR} = \frac{a/c}{b/d} = \frac{a*d}{b*c} = \frac{\text{(Odds of Exposure among Cases)}}{\text{(Odds of Exposure among Controls)}}$$

Confidence Intervals and p-values

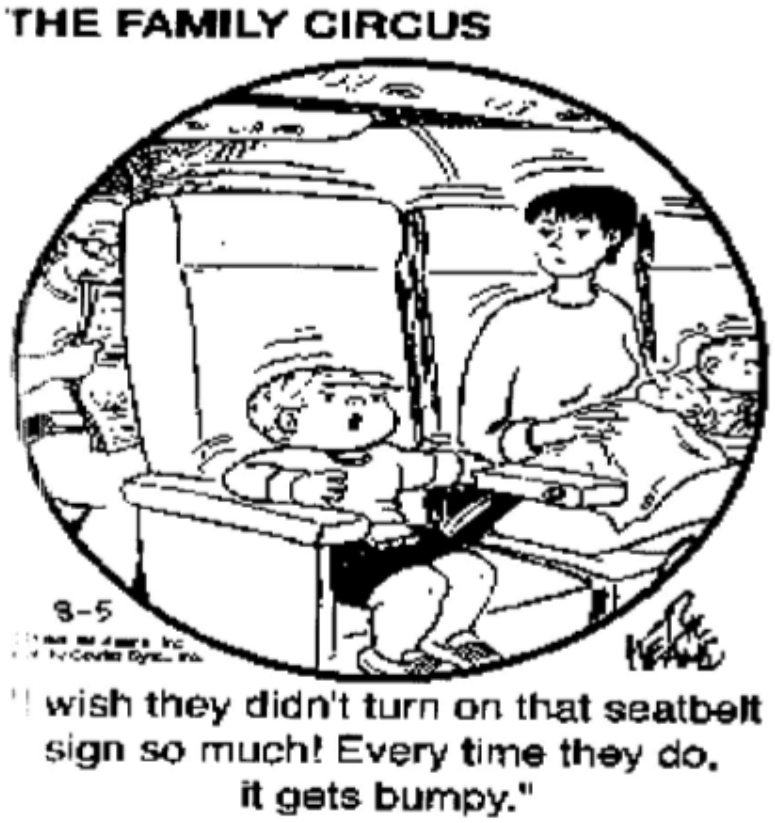
- Relative risks and odds ratios give information on the *magnitude* of association

- Important to consider *precision and statistical significance*, along with estimate of magnitude of association.

- Statistical software will in addition to relative risks and odds ratios provide estimate of confidence intervals and p-values

Association and Causality

- An exposure and outcome are *associated* if there is a differential distribution:
  - Incidence of outcome differs for exposed and unexposed group (cohorts); or
  - Prevalence of exposure differs between cases and controls (case-control study)
- An exposure is *causal* for the outcome if the presence (or absence) of the exposure directly or indirectly influences whether the outcome occurs.



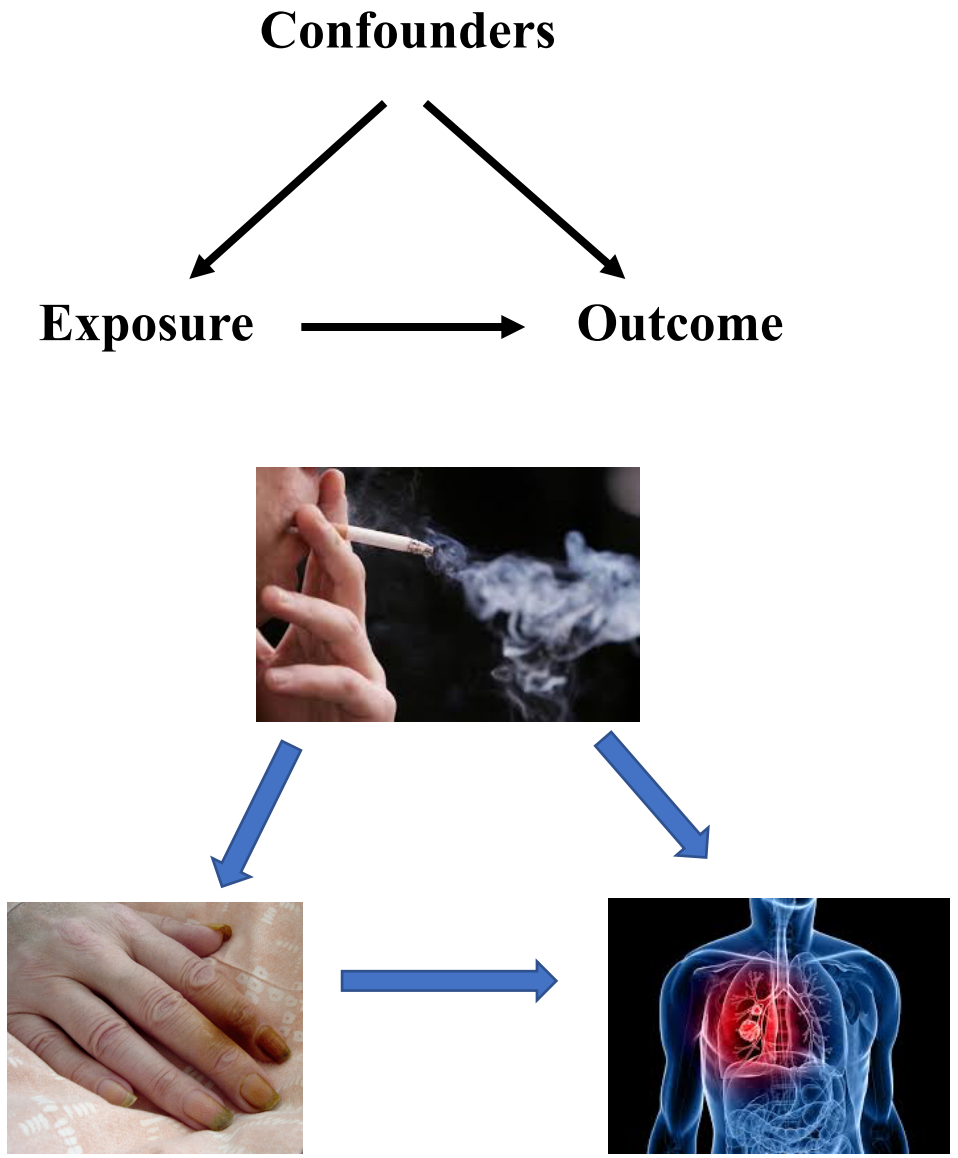
Sources of bias in epidemiology

- **Selection Bias**
  - Arises when cases and controls are coming from different source populations (e.g. female cases, male 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, she may look more carefully at the x-ray).
- **Recall bias**
  - Accuracy and completeness of exposures, life style behaviors etc (e.g. cases might be more motivated to complete a questionnaire accurately).

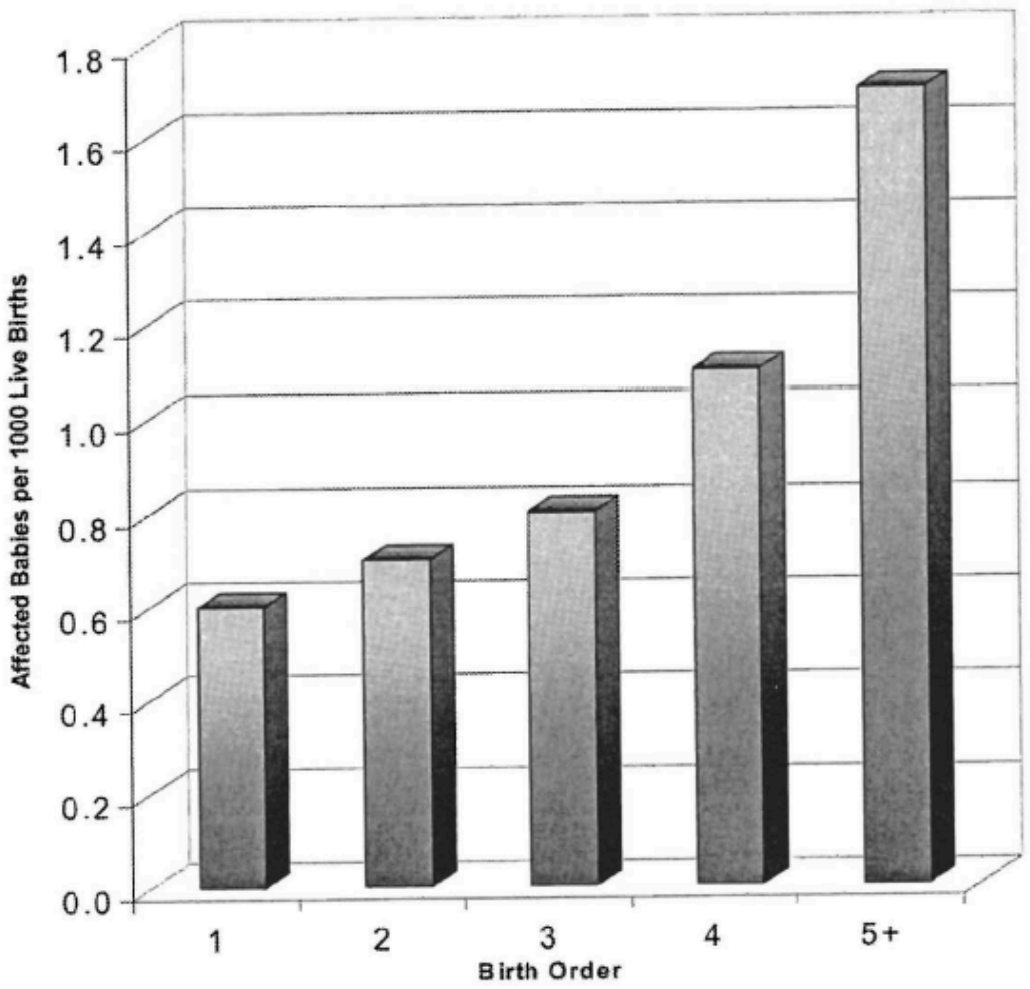
Confounding

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

- Confounding can lead to false positive findings.



Confounding example: Birth order and Down syndrome



Data from Stark and Mantel (1966)

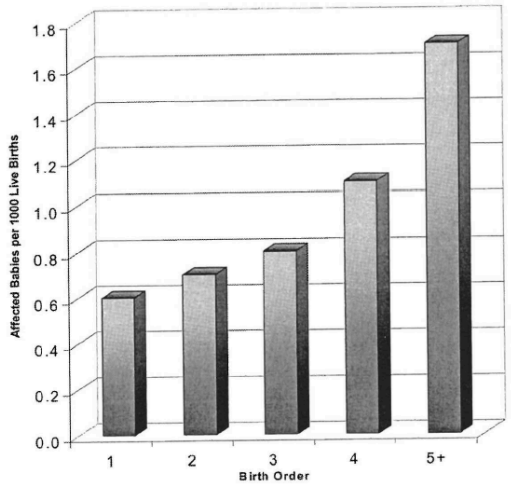
Source: Rothman 2002

Confounding example: Birth order and Down syndrome

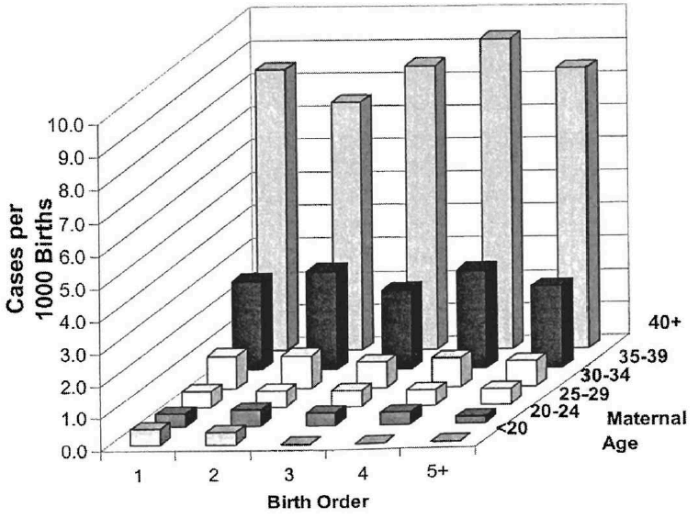
- Later order children have higher risk
  - Maternal age is associated with birth order
  - Maternal age is associated with Down Syndrome

- Stratifying on maternal age, there is no longer evidence of an association between birth order and Down syndrome

Association between birth order and Down syndrome



Data from Stark and Mantel (1966) Source: Rothman 2002



Data from Stark and Mantel (1966) Source: Rothman 2002

Summary

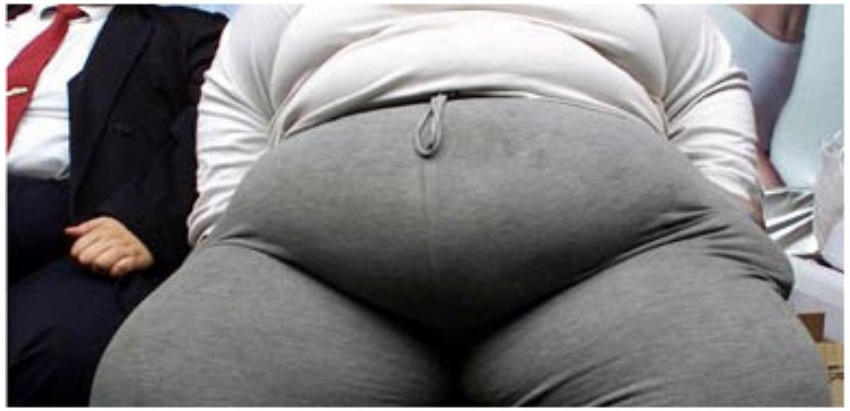
- Epidemiology is the study of the distribution and determinants of health-related outcomes in populations
- Study design is a key component of epidemiology
- Relative risks and odds ratios are commonly used to measure association
- It is important to consider and address bias in epi studies
- Understanding confounding is important when conducting association studies

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.*

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**

Newsroom

SCHOOLS | FOR FACULTY AND STAFF | TOPICS

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

HUTCH NEWS

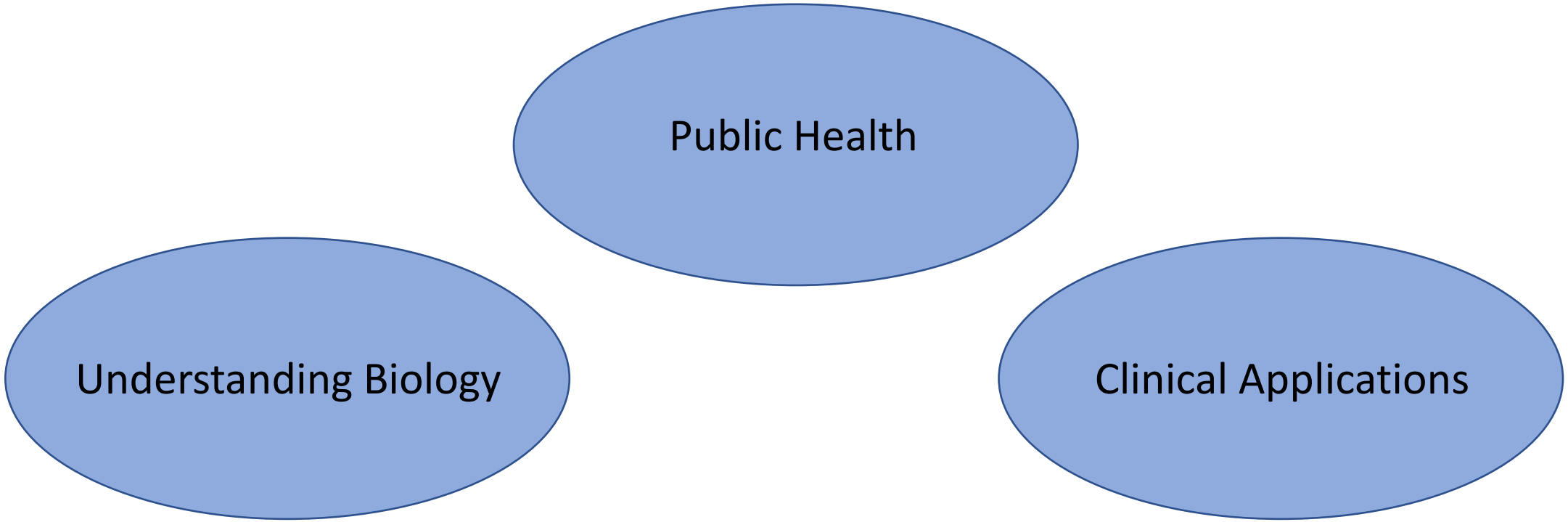
Does aspirin prevent colorectal cancer? Depends on your DNA

Fred Hutch researchers move closer to cracking the code on how genes and environmental factors influence colorectal cancer risk

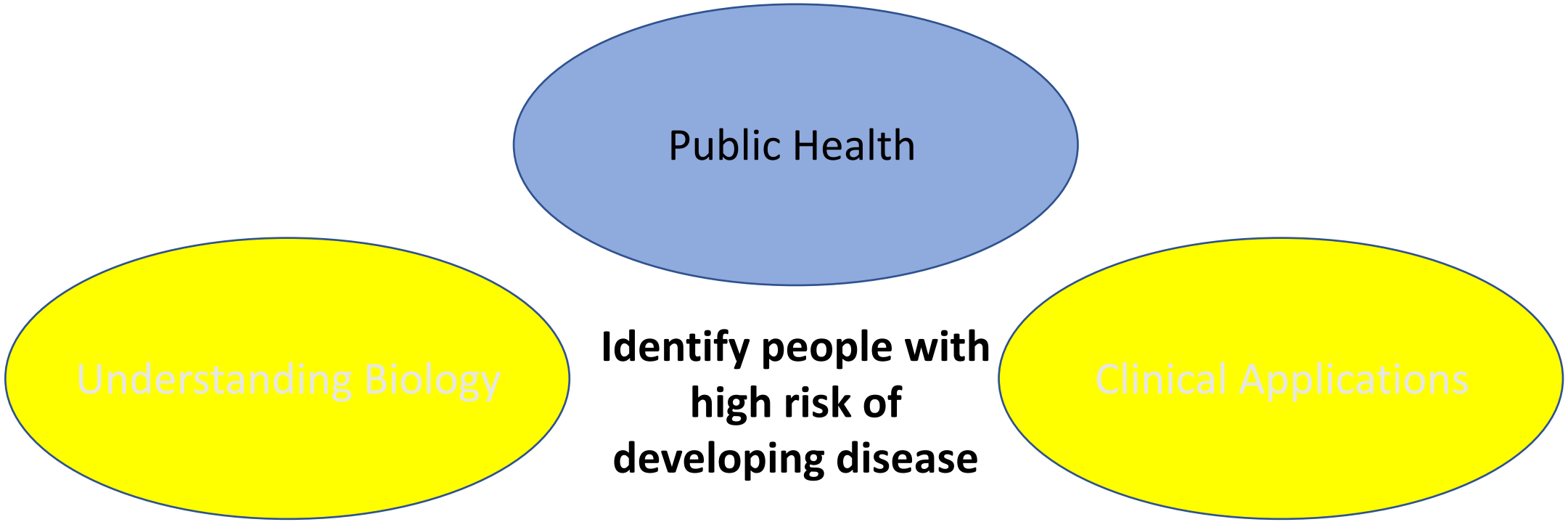
March 17, 2015 | By Diane Mapes / Fred Hutch News Service

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

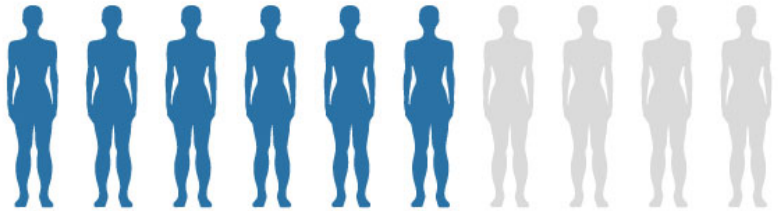
Loren Capelli

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.

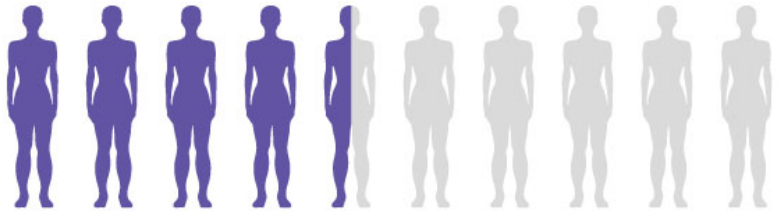
NATIONAL CANCER INSTITUTE
CHANCES OF DEVELOPING BREAST CANCER BY AGE 70

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.



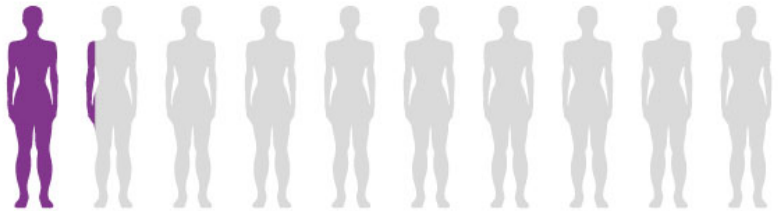
MUTATED BRCA1

55-65%



MUTATED BRCA2

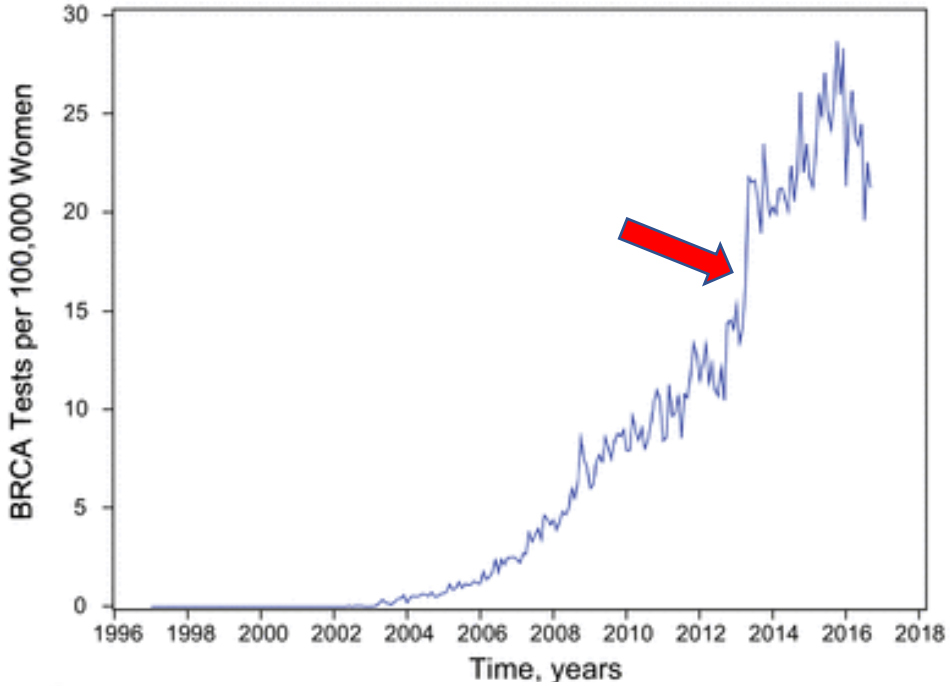
45%

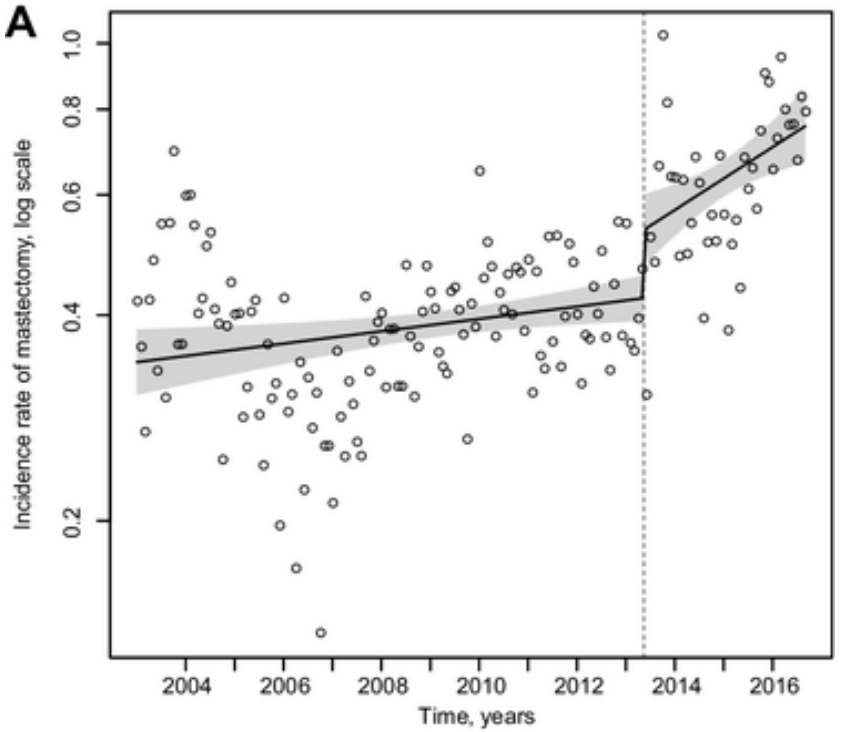


NORMAL BRCA

12%

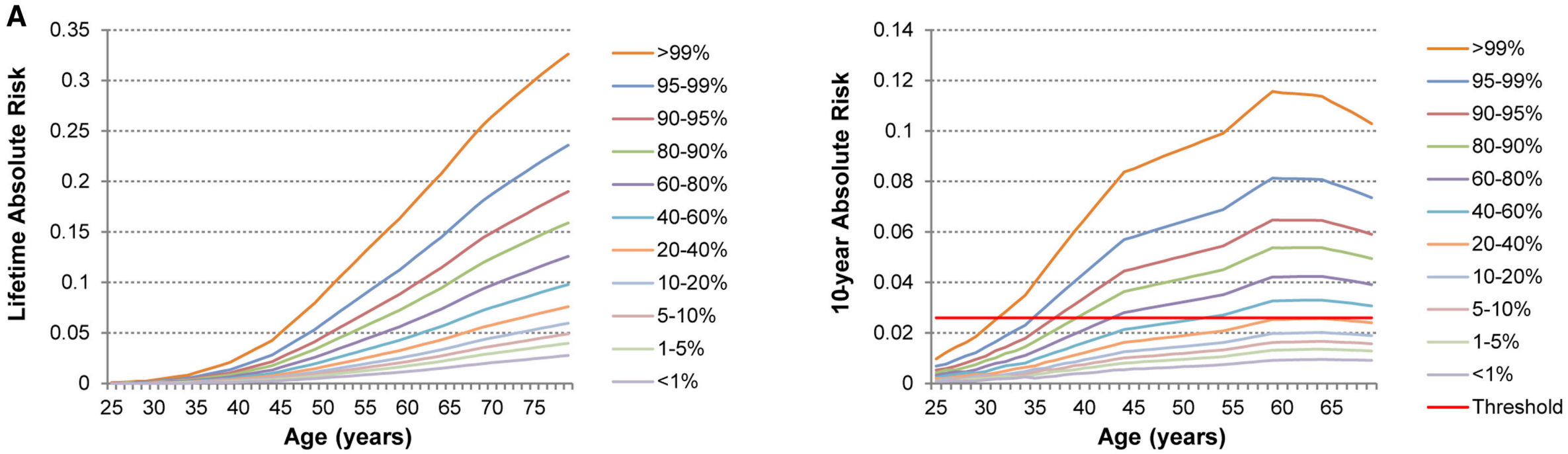
www.cancer.gov/brca-fact-sheet





Liede et al. BCRT 2018

Association between a genetic risk score and breast cancer risk in women of European origin



"The average 10-year absolute risk of breast cancer for a 47-year-old woman (the age at which women become eligible to enter the UK breast cancer screening program) in the general population is 2.6%. However, the 19% of women with highest genetic risk will attain this level of risk by age 40 years"

Mavaddat et al. AJHG 2019

23andMe Granted First FDA Authorization for Direct-to-Consumer Genetic Test on Cancer Risk

March 6, 2018

Authorization allows 23andMe to report on BRCA1- and BRCA2-related genetic risk for breast, ovarian and prostate cancer

Mountain View, California – March 6, 2018 – 23andMe, Inc., the leading personal genetics company, today received the first-ever FDA authorization for a direct-to-consumer genetic test for cancer risk. The authorization allows 23andMe to provide customers, without a prescription, information on three genetic variants found on the BRCA1 and BRCA2 genes known to be associated with higher risk for breast, ovarian and prostate cancer.

"Being the first and only direct-to-consumer genetics company to receive FDA authorization to test for cancer risk without a prescription is a major milestone for 23andMe and for the consumer," said Anne Wojcicki, 23andMe CEO and co-founder. "We believe it's important for consumers to have direct and affordable access to this potentially life-saving information. We will continue pioneering a path for greater access to health information, and promoting a more consumer-driven, preventative approach to health care."

23andMe will report on three variants in the BRCA1 and BRCA2 genes associated with a significantly higher risk of breast and ovarian cancer in women, and breast cancer in men. The variants may also be associated with an increased risk for certain other cancers. These variants are

23andMe Receives FDA Clearance for Genetic Health Risk report that looks at a Hereditary Colorectal Cancer Syndrome

January 22, 2019 By 23andMe under Health and Traits

23andMe received FDA clearance to report on the two most common genetic variants influencing what is called MUTYH-associated polyposis (MAP), a hereditary colorectal cancer syndrome.

This new clearance is part of 23andMe's ongoing efforts to work with the FDA to offer additional Genetic Health Risk reports as part of its Health + Ancestry Service.

"We are committed to giving people affordable and direct access to important health information that can impact their lives," said Anne Wojcicki, 23andMe CEO and co-founder.

The MAP report is not yet available, but as soon as it is ready for release, it will be offered to new Health + Ancestry Service customers as well as existing 23andMe Health + Ancestry Service customers genotyped on the company's most recent platforms. Like the 23andMe BRCA1/BRCA2 (Selected Variants) Genetic Health Risk report, customers first must choose whether or not they want to receive this information.

23andMe Offers New Genetic Report on Type 2 Diabetes

March 10, 2019 By 23andMe under Health and Traits

23andMe's new Type 2 Diabetes report* offers customers insight into their likelihood of developing one of the most prevalent, pernicious, and yet preventable health conditions in the United States.

This new report will impact more 23andMe customers than any other Health Predisposition report the company has ever released. In part, this is because the likelihood of developing type 2 diabetes is so high in the general population.

"Diabetes is a significant health issue in the United States that is expected to impact nearly half of the population," said Anne Wojcicki, CEO and Co-Founder of 23andMe. "When customers learn about their genetic likelihood of developing type 2 diabetes, we believe there is an opportunity to motivate them to change their lifestyle and ultimately to help them prevent the disease."

Rising Rates of Diabetes

In the United States alone, where rates of obesity have ratcheted up in recent decades, a staggering one in three Americans have prediabetes, according to estimates from the U.S. Centers for Disease Control and Prevention. Worse, 90 percent don't know they have it. The impact is even greater among certain

What's a Polygenic Score?

The genes we inherit from our mother and father influence our risk of disease. While that's pretty straightforward, the role genetics plays is a bit more complicated.

All health predisposition reports*

REPORT	GENE(S)	VARIANTS	RELEVANT ETHNICITIES
BRCA1/BRCA2 (Selected Variants)	BRCA1 and BRCA2	3	Ashkenazi Jewish
Age-Related Macular Degeneration	ARMS2 and CFH	2	European
Alpha-1 Antitrypsin Deficiency	SERPINA1	2	European
Celiac Disease	HLA-DQB1 and HLA-DQA1	2	European
Familial Hypercholesterolemia	LDLR and APOB	24	All populations
G6PD Deficiency	G6PD	1	African
Hereditary Hemochromatosis (HFE-Related)	HFE	2	European
Hereditary Thrombophilia	F2 and F5	2	European
Late-Onset Alzheimer's Disease	APOE	1	All populations
Parkinson's Disease	LRRK2 and GBA	2	European, Ashkenazi Jewish, North African Berber
Type 2 Diabetes (Powered by 23andMe Research)	n/a	Model with 1000+	European, Hispanic/Latino, African, East Asian, and South Asian

The New York Times

Don't Count on 23andMe to Detect Most Breast Cancer Risks, Study Warns

More than 10 million people have signed up for 23andMe. Many are initially drawn in by ancestry data, but later opt in for health risk tests. George Frey/Reuters

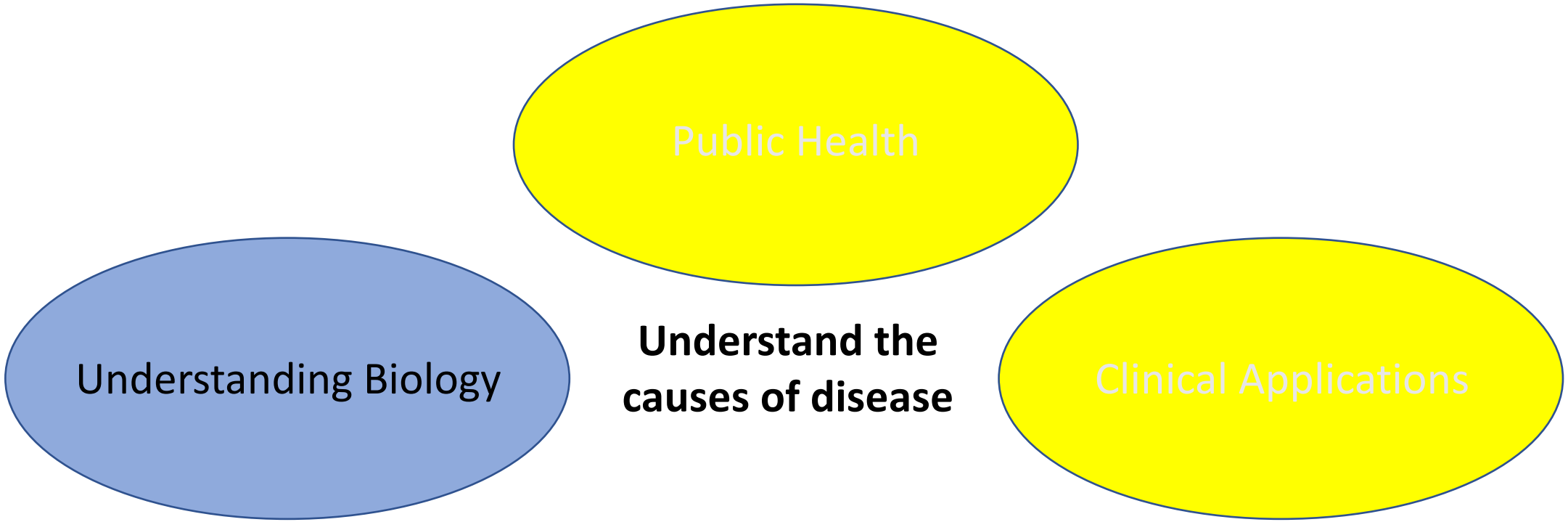
By Heather Murphy

April 16, 2019

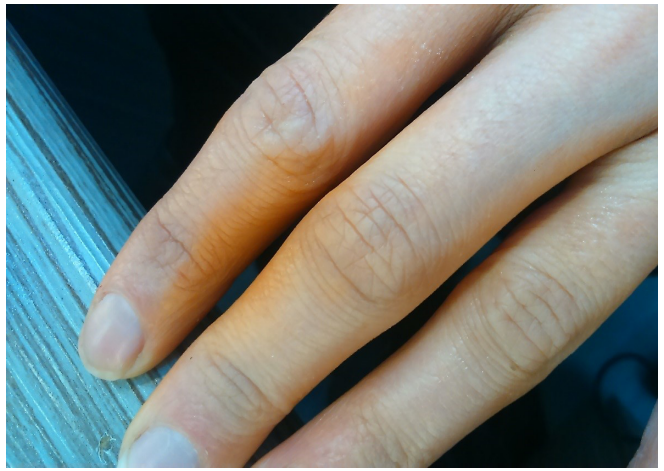
49

"In about 5,000 subjects, analysts identified at least one variant known to significantly increase an individual's risk of breast and ovarian cancer. Among the Ashkenazi Jews in the positive group, 81 percent had one of the three founder mutations, suggesting that 23andMe's test could be helpful for them. Among the rest, 94 percent carried variants that would have failed to be detected by 23andMe."

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



“Association does not imply causation”



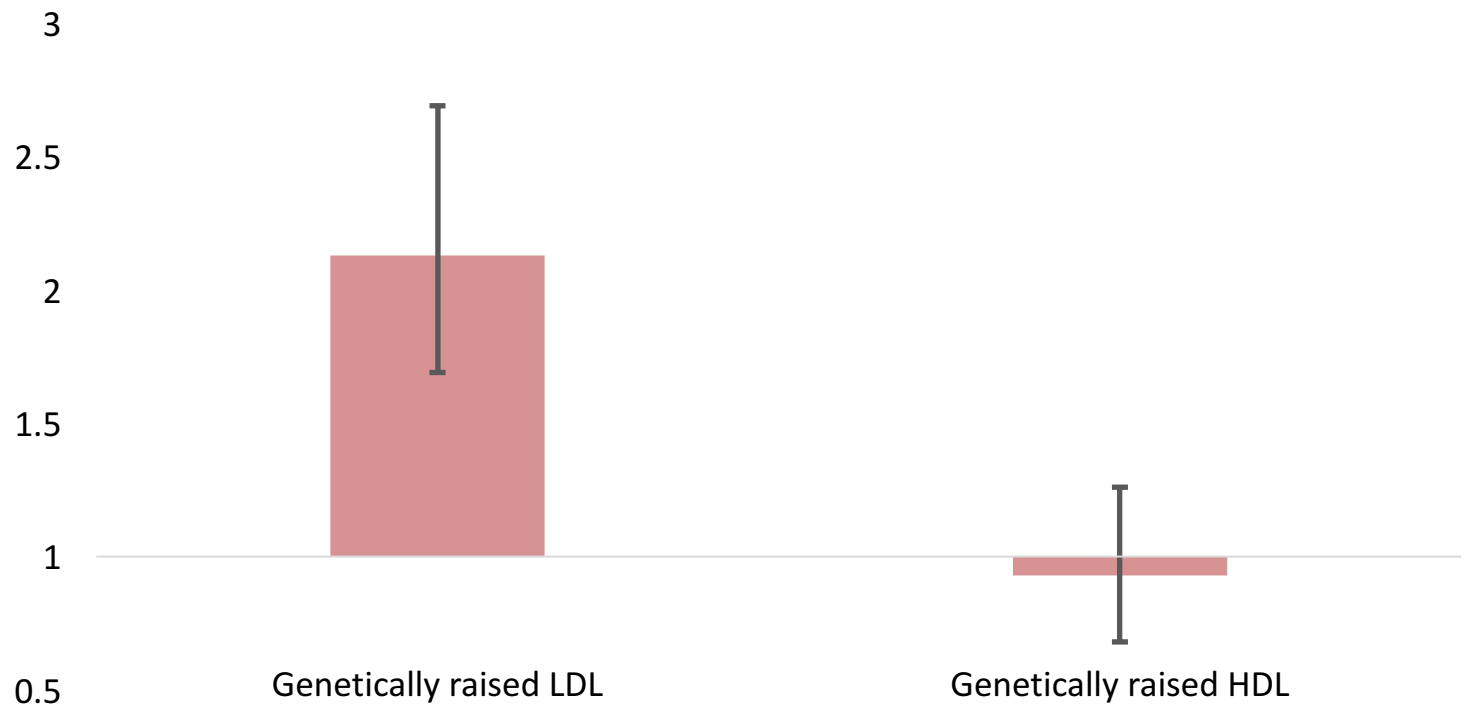
HDL (“Good”) Cholesterol and Myocardial Infarction (MI)

- ⬆ 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

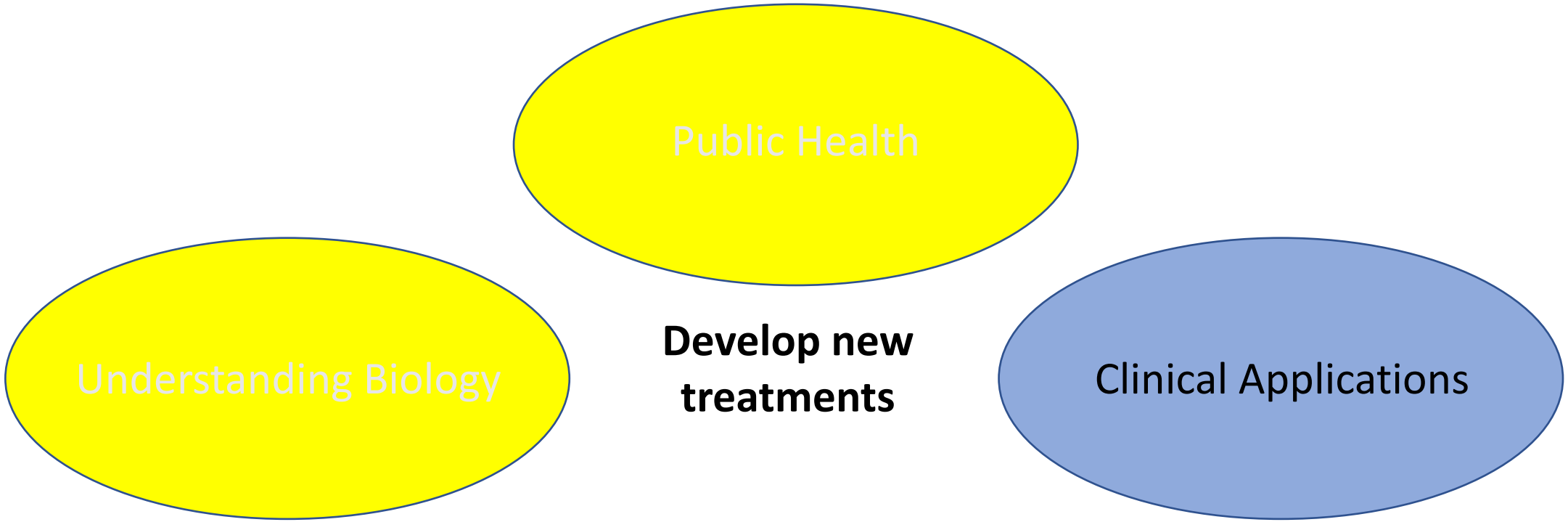
RESEARCH ARTICLES

HEART DISEASE

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

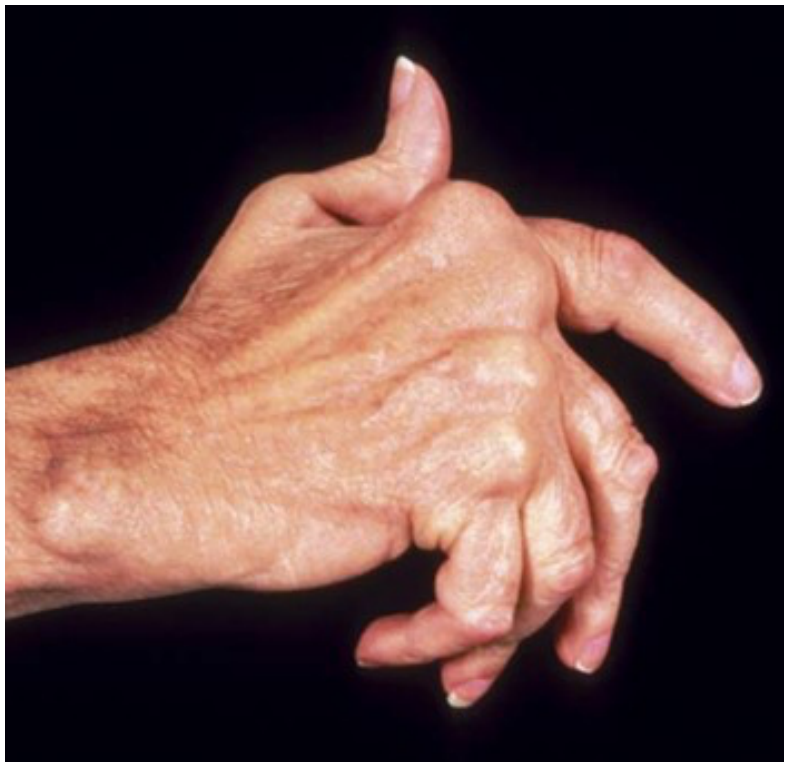
Zanoni et al, Science 2016

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.

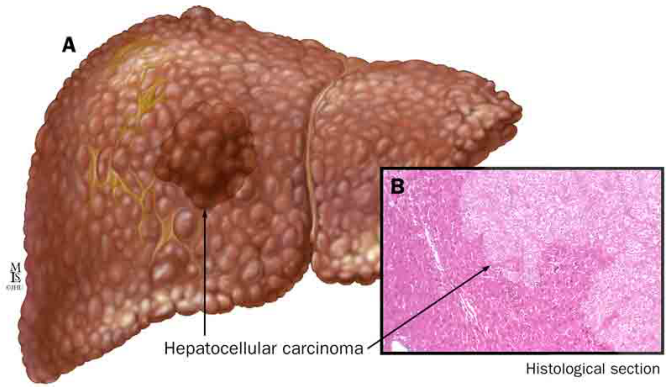


A study of 10 million genetic variants in 29,880 RA cases and 73,758 controls

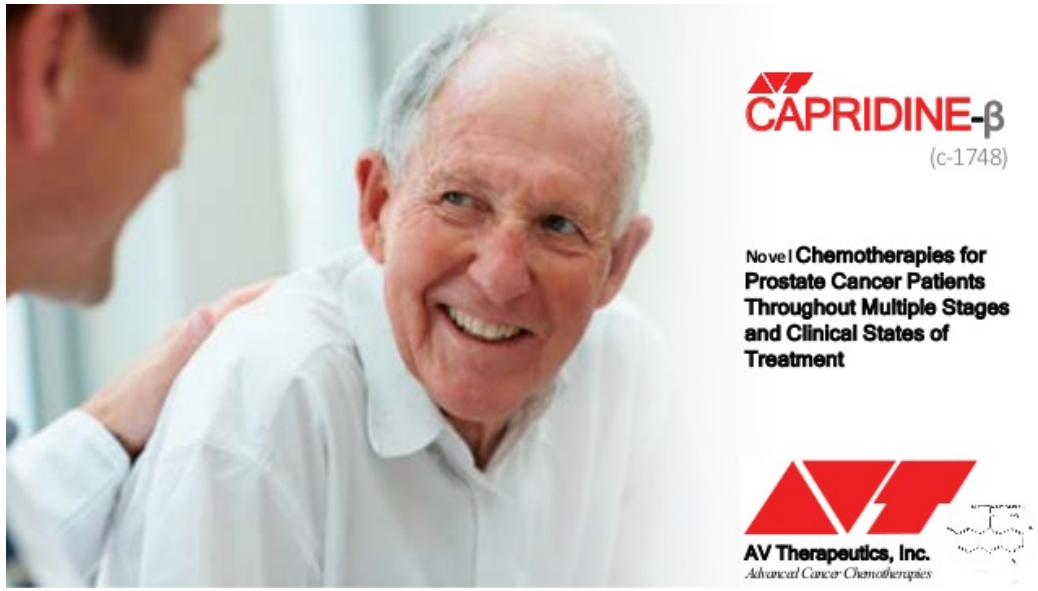
Okada, Nature, 2014

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

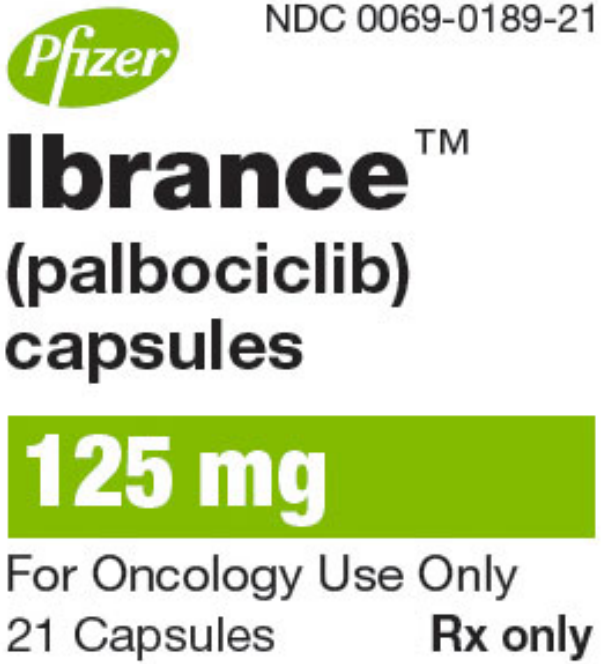


Lymphoma/Leukemia/Liver cancer



Cancers/psoriasis

Breast cancer

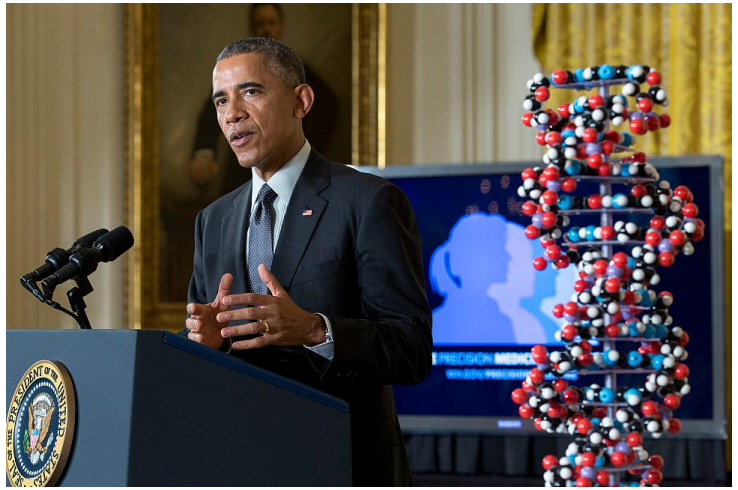


Okada, Nature, 2014

We are entering the era of “precision medicine”

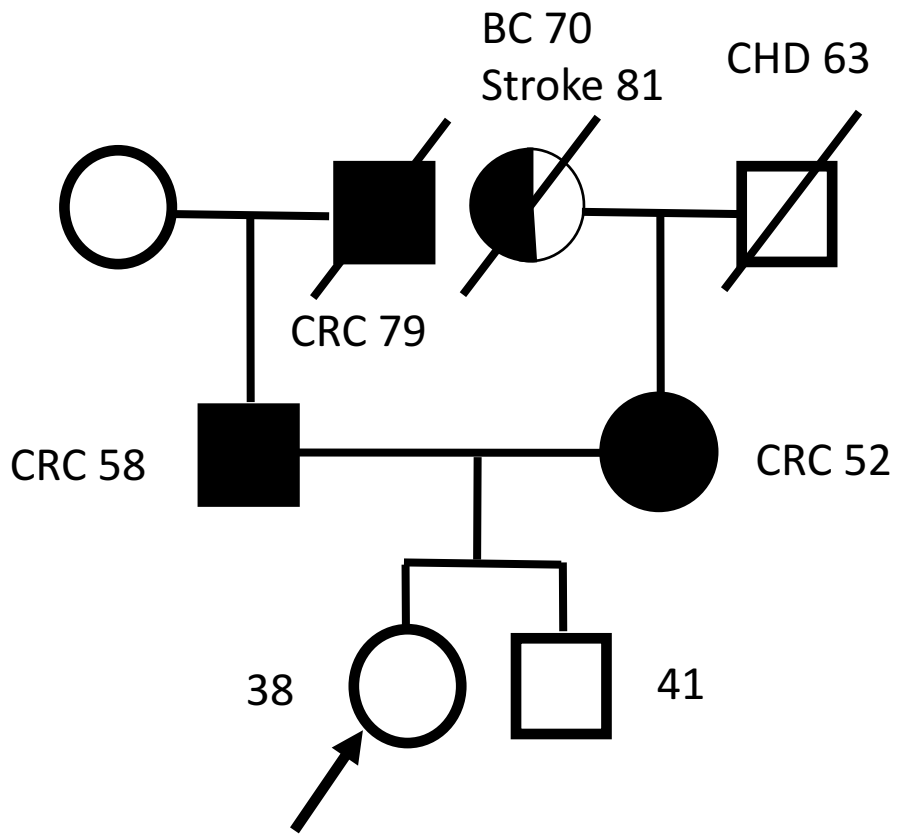
“I want the country that eliminated polio and mapped the human genome to lead a new era of medicine — one that delivers the right treatment at the right time…[]…So tonight, I’m launching a new Precision Medicine Initiative to bring us closer to curing diseases like cancer and diabetes, and to give all of us access to the personalized information we need to keep ourselves and our families healthier.

President Obama, State of the Union. Jan 20, 2015.



“Precision medicine is an emerging approach to treating illnesses that takes into account a patient’s individual genetic make-up as well as molecular subtypes of diseases to improve the chances of successful treatment.”

http://www.whitehouse.gov



Key

Male

Affected male

Deceased male

Female

Affected Female

Deceased Female

Affected female with different cancer