Genomic Screening and the Reverend Bayes

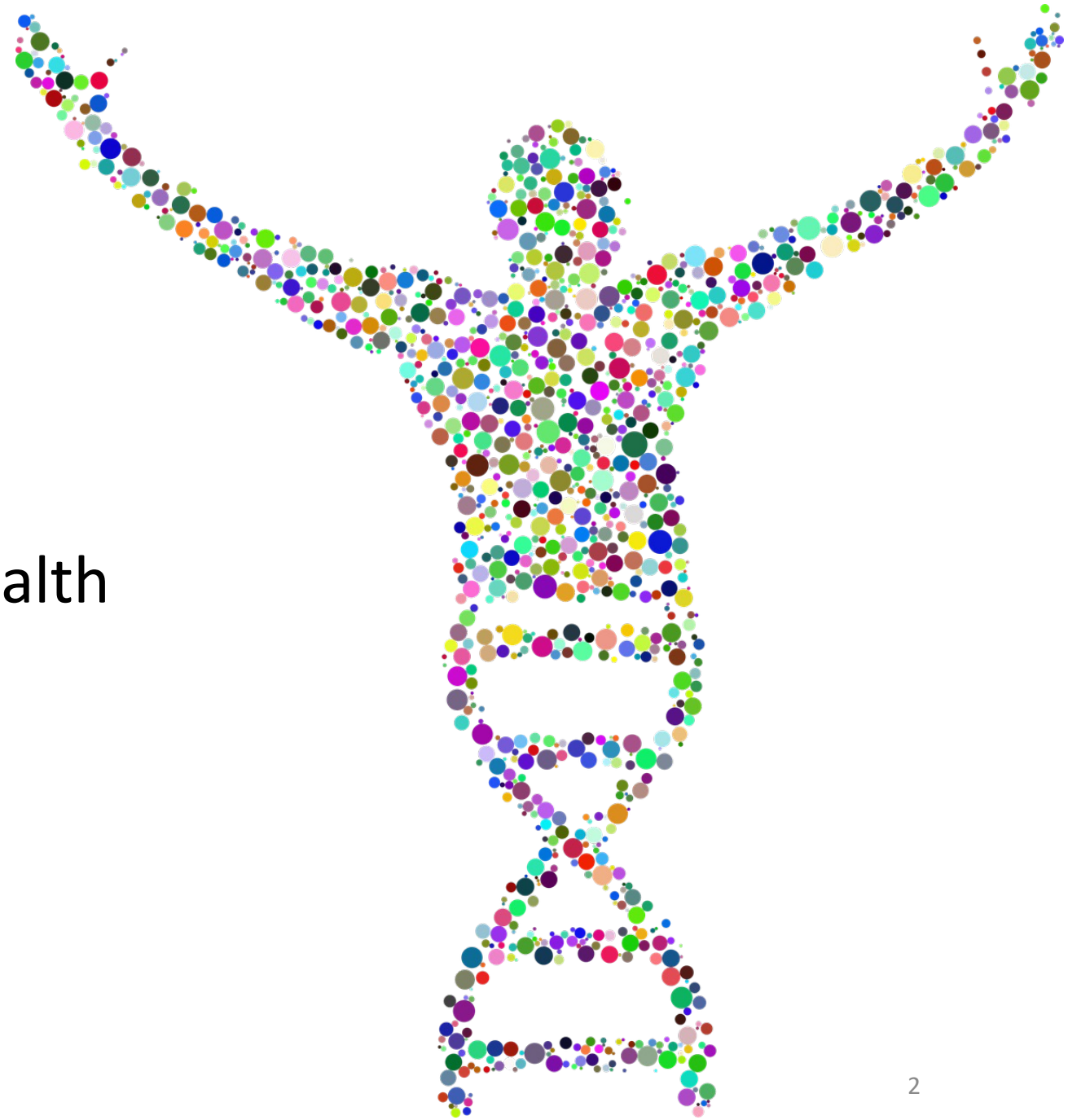
Les Biesecker

Center for Precision Health Research

NHGRI

November 8, 2023

Genomes → Health

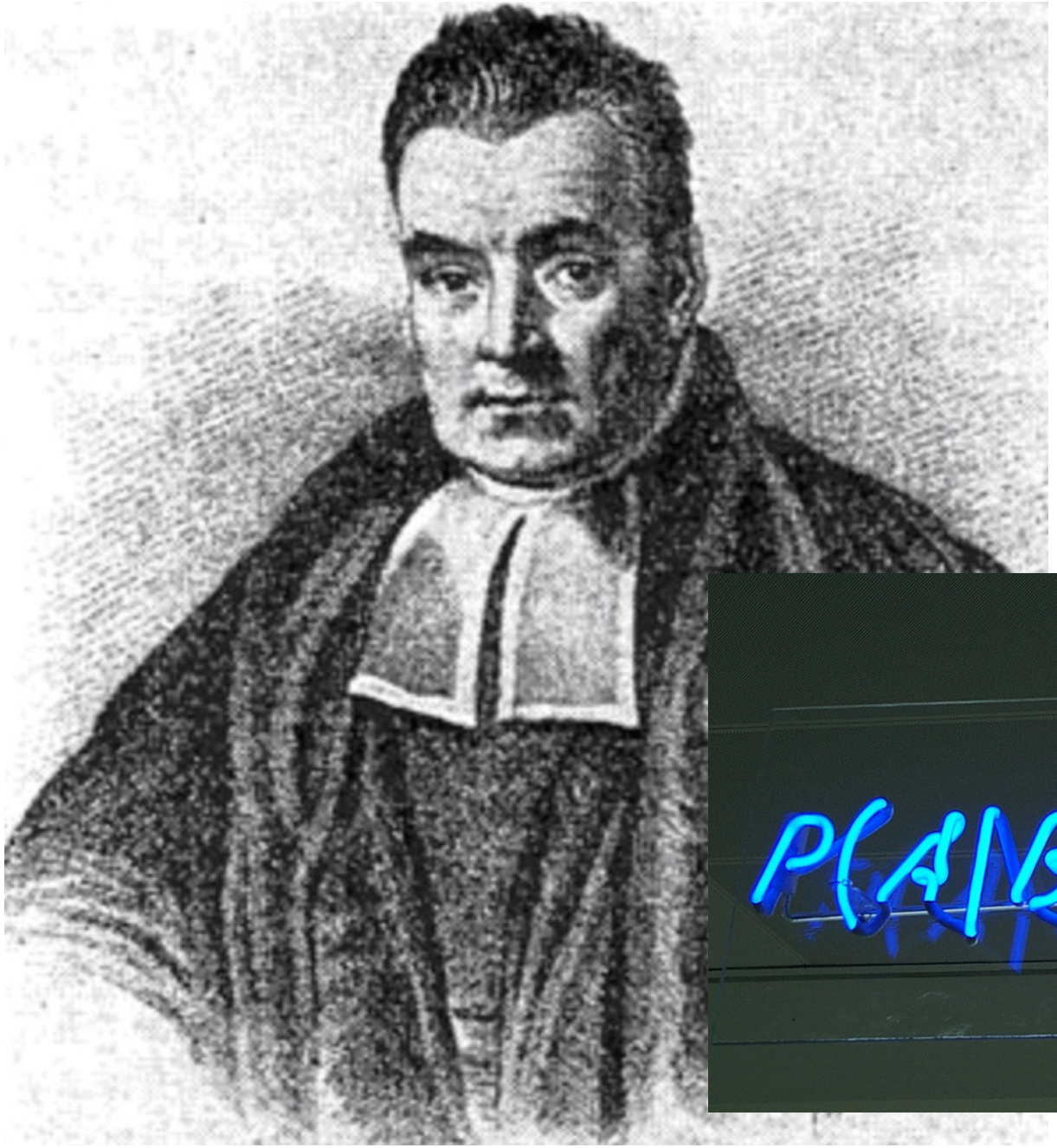


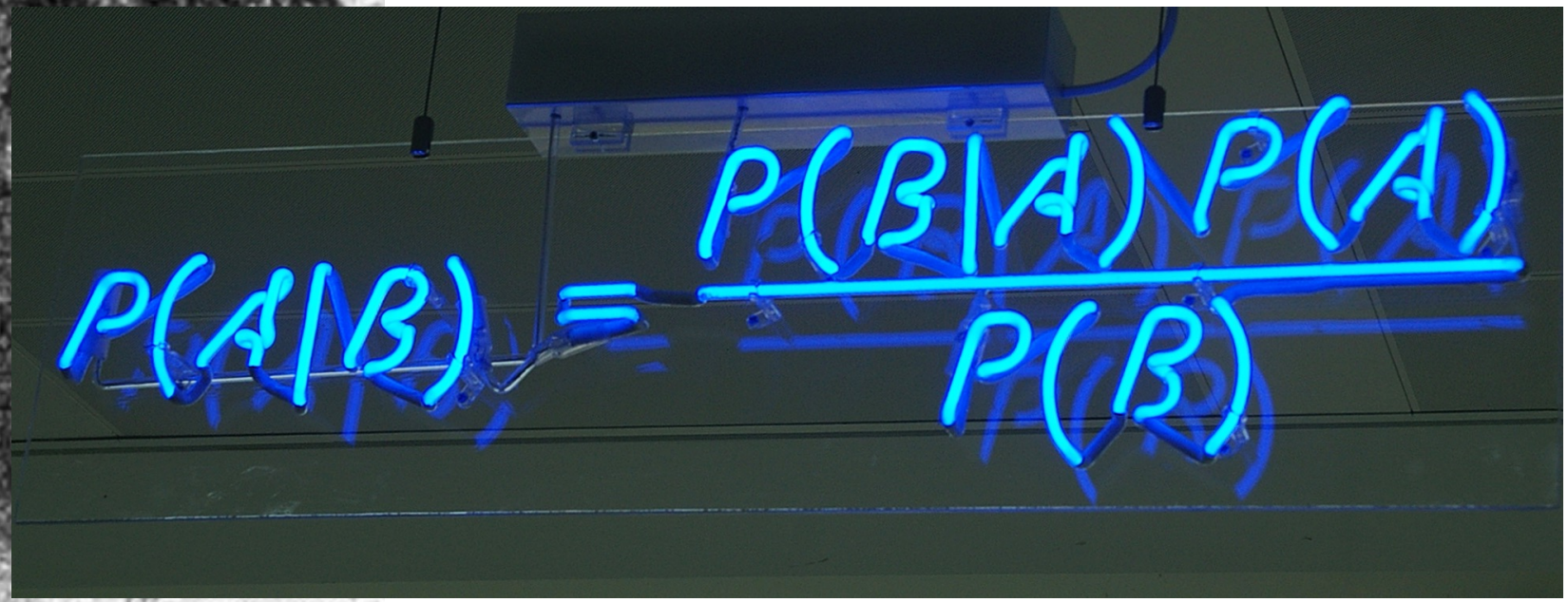
Genomics is *Not* Exceptional

- Genomic testing, with few exceptions, performs like every other medical test
- It has sensitivity, specificity, PPV, & NPV
- Like any other test, the PPV depends on the testing scenario

Genetics Practice is Exceptional

- We generally use genomic testing in scenarios where prior probability of disease is high
  - Ironically, in this setting, the testing little changes diagnosis & management
- When testing in low prior probability settings, the game is different




P(A|B) = P(B|A) P(A) / P(B)

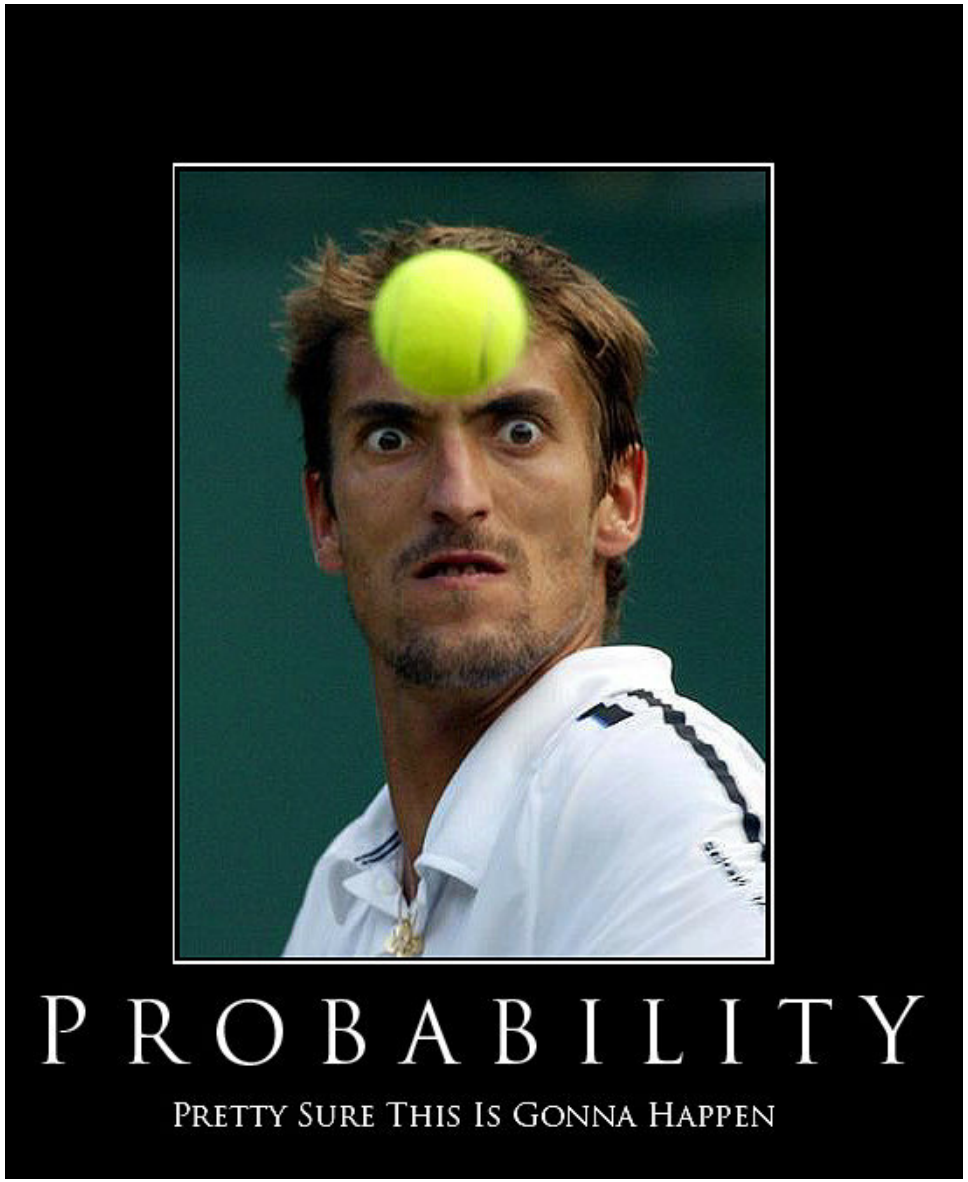
Math…

- P(A)
- The probability of A
- P(A|B)
- The probability that A is true *if* B is true

How To Think Clearly About Screening

- Analytic validity is a probability
- Clinical validity is a probability
  - Pathogenicity
- Clinico-molecular diagnosis is a probability
- Penetrance is a probability
- Expressivity is a probability

- We need a formal, probabilistic model of genetic diagnosis



A Few Concepts & Definitions

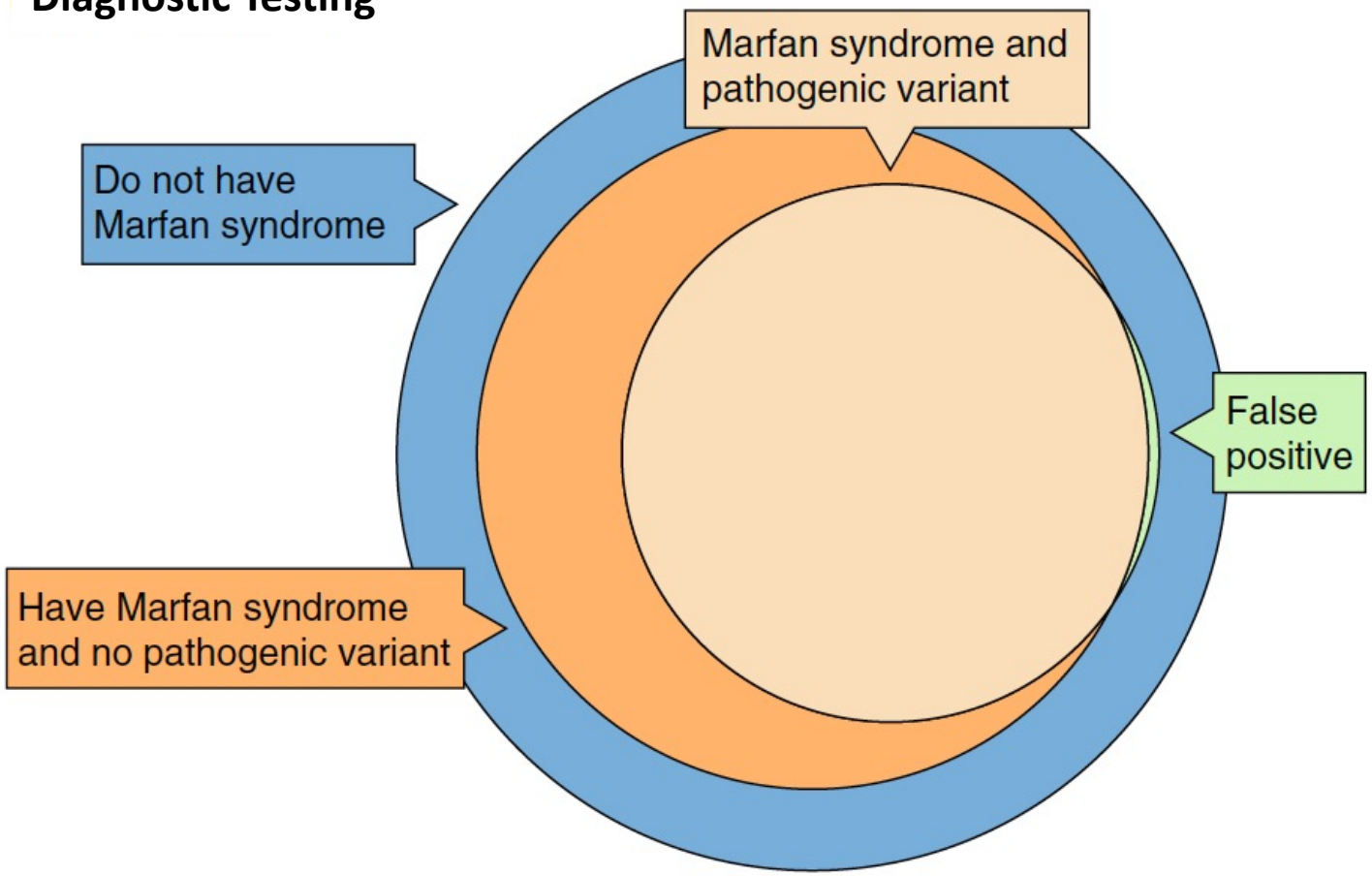
- All genetic disease is an increased susceptibility to abnormal phenotype
  - Having the disease but no manifestations is non-penetrance
  - *Not having the disease ≠ non-penetrance*
- Nearly all variants have a *probability* of pathogenicity
  - They are not certain to be causative
  - A few have probability of pathogenicity 100%
    - For these, harboring the variant ≡ has the disease ≠ has the phenotype

Steps

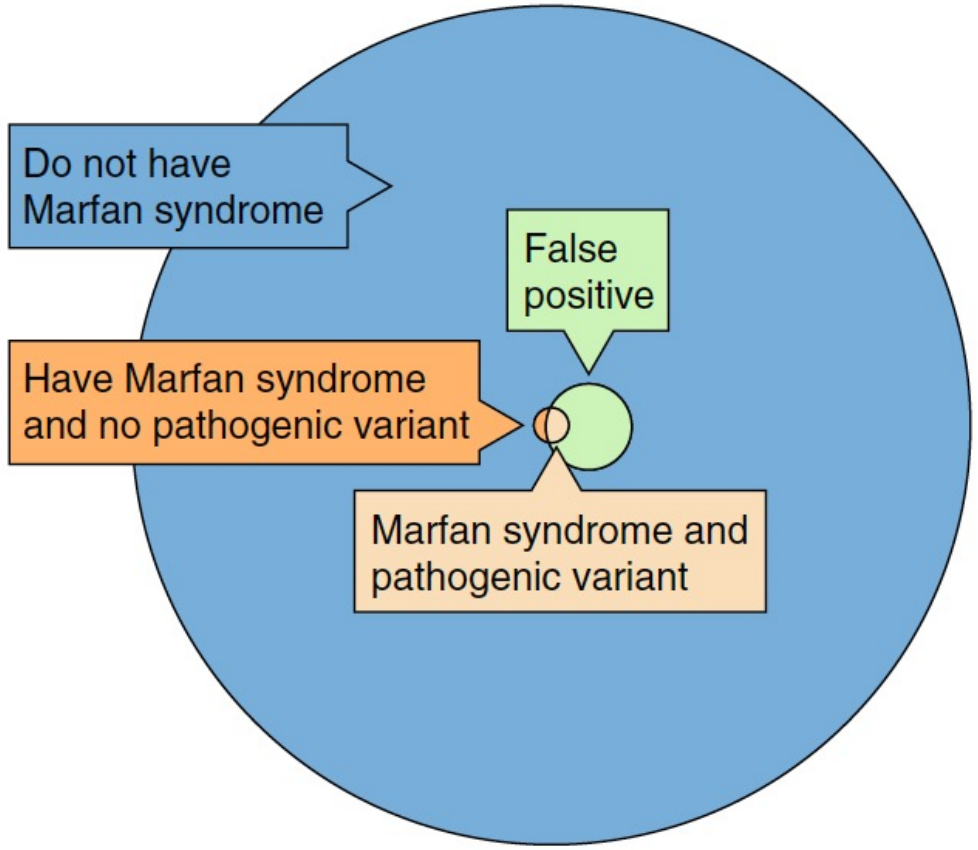
- Assess probability of pathogenicity of the variant
  - P(Path|Evid)
  - Predictors plus historical data on variant
- Clinical interpretation of person with variant
  - P(Disorder|Phen)
  - Clinico-molecular diagnosis
    - Harboring a variant ≠ having disorder
- For those with disorder but without phenotype
  - P(Phen|Disorder)

Graphical Math – Marfan Syndrome

Diagnostic Testing



Screening Testing

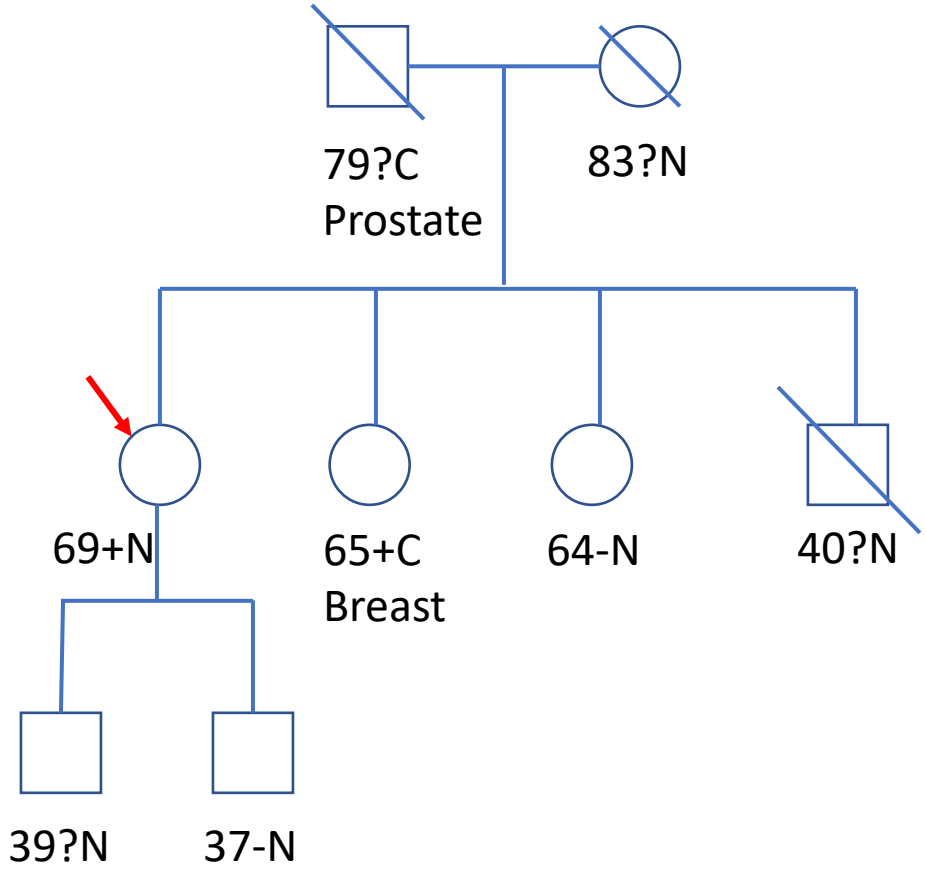


Genome Medicine 2019 11(1):75

Example Scenario

37 yo man trio exome sequenced (for for neuro disorder)

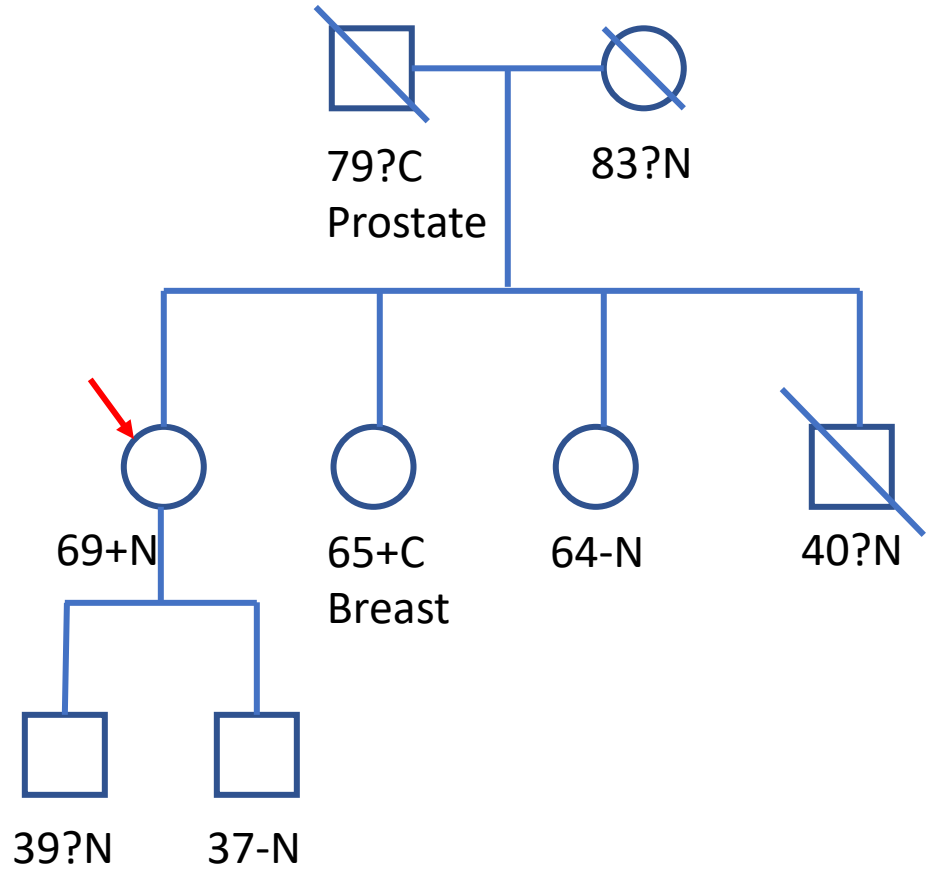
<u>Mother</u> has Likely Pathogenic *BRCA2* variant



‘69+N’ annotation:
69 = age in yrs
+ = harbors variant
- = known to not have variant
? = variant status unknown
N = No cancer
C = had cancer

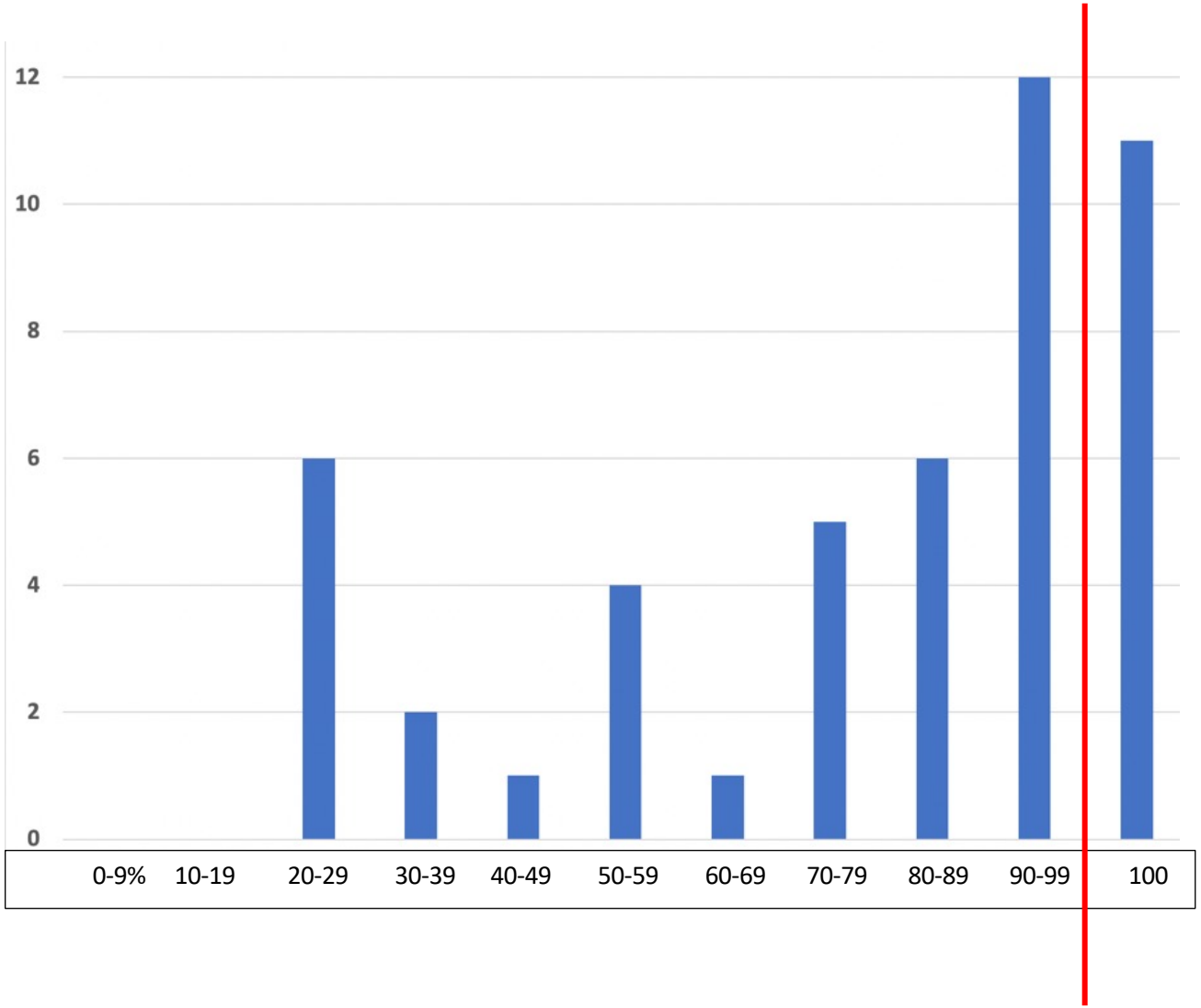
Clinico-Molecular Diagnostic Probabilities (CMD)

- Just the test result
  - 65% probability of CMD
- Testee is phenotype negative @ 69
  - 47% probability of CMD
- Include test positive relatives
  - 85% probability of CMD
- Include all relatives with genotype probability
  - 83% probability of CMD



Pedigree-Based Posterior Probability of Clinico-Molecular Diagnosis

- *BRCA1* & *BRCA2* n=48
- Common AJ variants
  - Not Bayesian as pathogenicity ~100%
- Others range from 23.5-99.98%



A Probabilistic Model of Population Screening

- Robust variant classification to determine P(Path|Evid)
- Practical methods to determine P(Diagnosis|Pheno)
- Patient decision making support
- Defined care pathways & CDS

Kill Determinism

Two Closing Thoughts

- Must assess risk precisely
- Cannot assess risk accurately
- The numerical risk may not be the primary determinant of care management decisions

Second Closing Thought

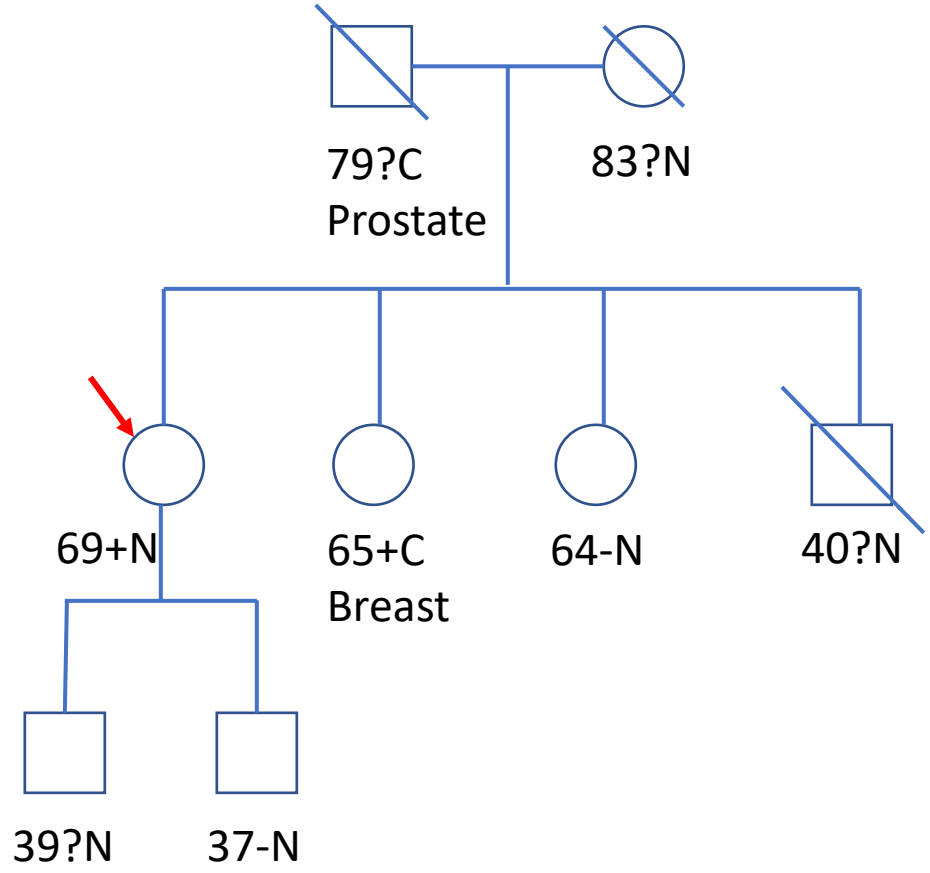
- The larger challenge is to change our mindset from one of non-directiveness to management
- The challenge will no longer be consoling & adaptation to diagnosis
- It will be to motivate people who don’t have manifestations of disease to engage in desired health behaviors
  - Without requiring hours of genetics professional care


The Forefront
of Genomics®

Bayes Calculation 1

- Prior probability she has HBOC 1/400
- If one has HBOC, 75% chance you will identify P/LP variant in *BRCA1* or *2*
- If one does not have HBOC, 0.1% chance you would harbor a P/LP *BRCA* variant (false positive)
- Posterior probability of disease: 65%

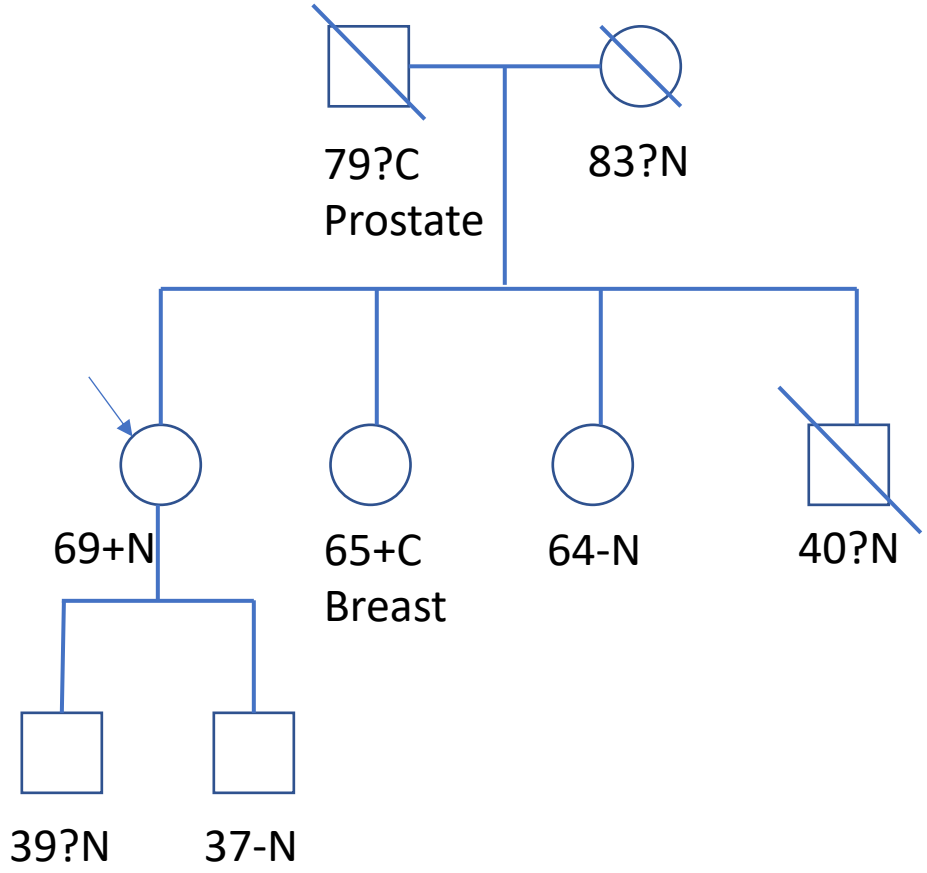
	+HBOC	-HBOC
Prior	0.0025	0.9975
Conditional	0.75	0.001
Joint	0.001875	0.0009975
Posterior	0.653	0.347



Bayes Calculation 2

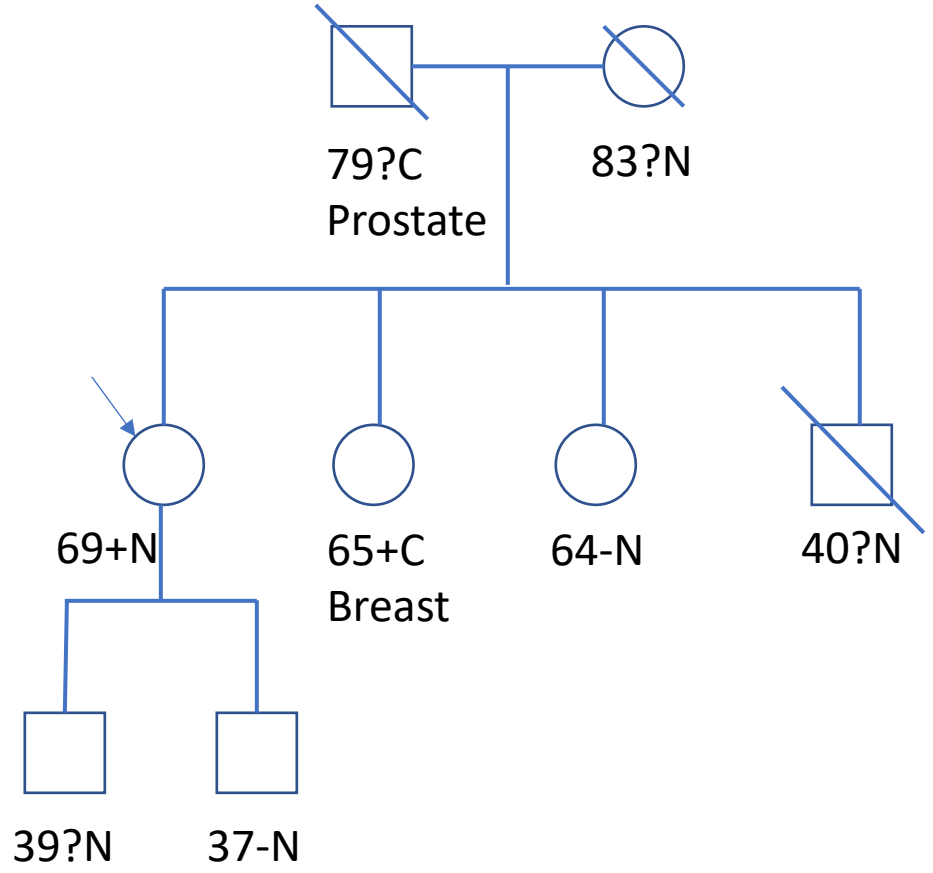
- Prior probability II-1 has HBOC 65%
- If the proband has HBOC, the likelihood that she would be cancer-free is 43.2% (ASK2ME)
- If the proband does not have HBOC, the likelihood she would be cancer free is 90.6%
- Posterior probability of HBOC is 47.3%

	+HBOC	-HBOC
Prior	0.652	0.347
Conditional	0.432	0.906
Joint	0.282	0.314
Posterior	0.473	0.527



Bayes Calculation 3

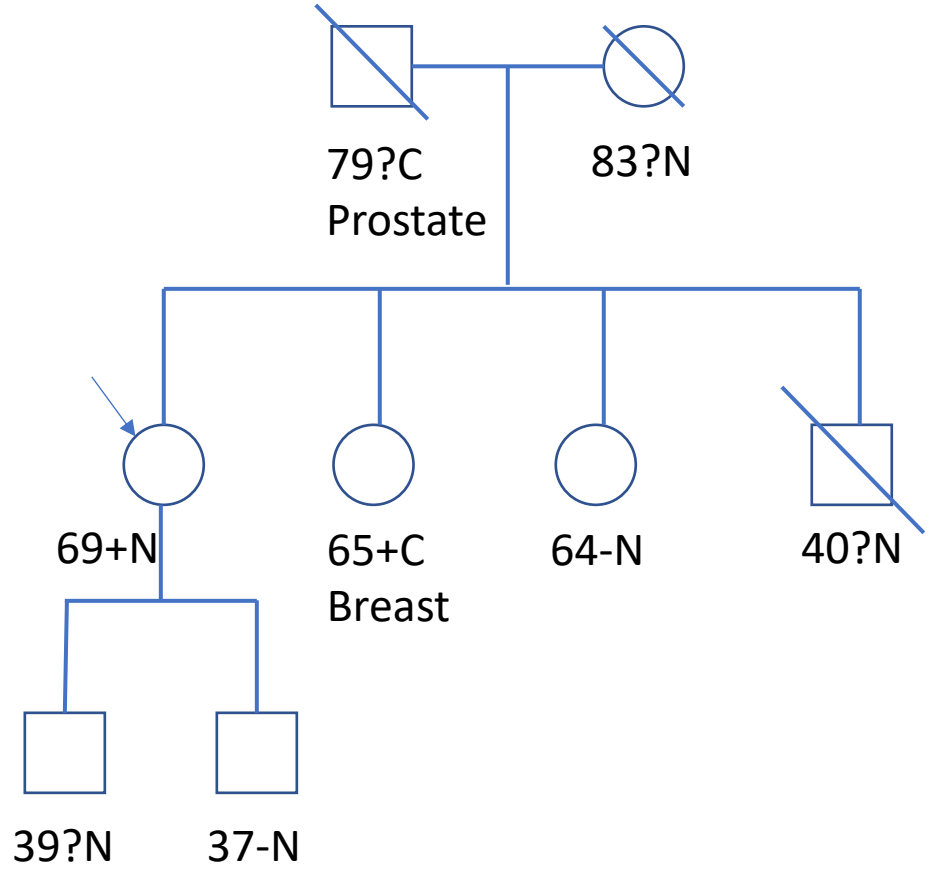
- Instead of calculation 2, consider as conditional all the individuals who are genotyped & +
- 69+N .432 / .906
- 65+C .433 / .068
- 64-N & 37-N (not relevant - maybe)



	+HBOC	-HBOC
Prior	0.652	0.348
Conditional	0.432*.433	0.906*.068
Joint	0.122	0.021
Posterior	0.853	0.147

Bayes Calculation 4

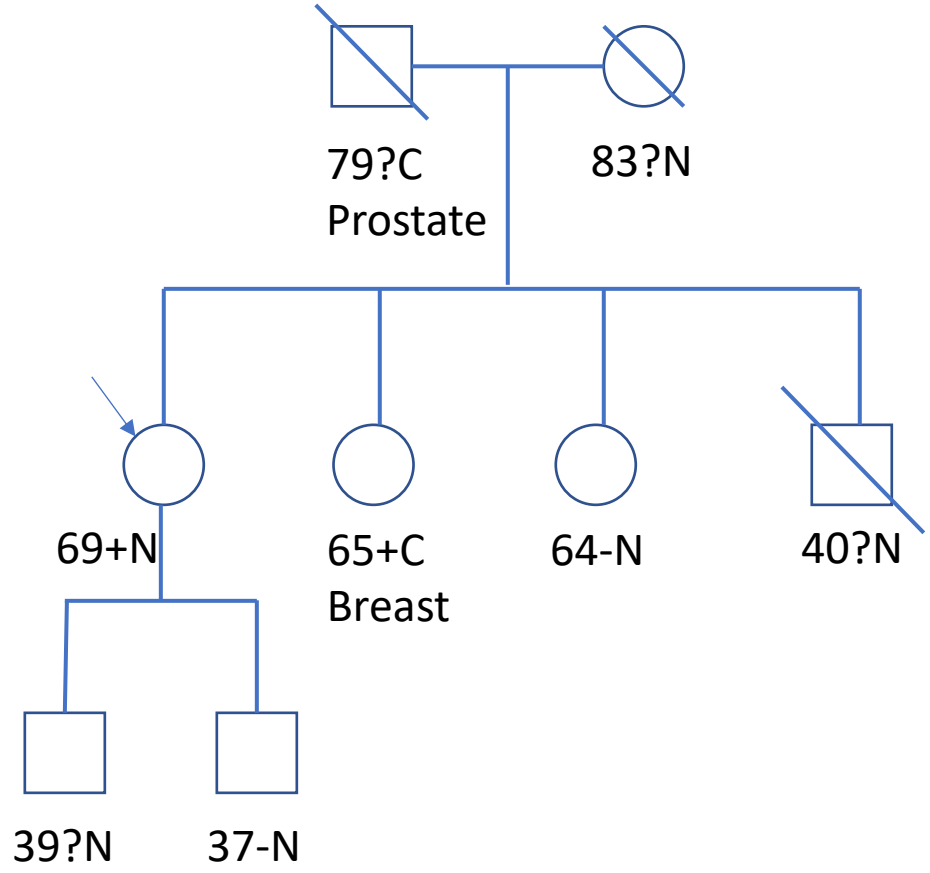
- Instead of calculation 3, consider as conditional all the individuals who are genotyped & + & use Mendelian rules for ungenotyped people
- $((R^{+}-R^{-})*0.5^{\wedge}N)+R^{-}$
- R^{+} = cancer risk of person who has HBOC
- R^{-} = cancer risk of person who doesn't have HBOC
- N = # of meioses from person with variant
- For conditional probability of cancer, taking into account relationship to genotype + individual :
- Father ((.1892-.1197)*1/2)+.1197 = .1544
- Mother 1-((.7117-.1506)*1/2)+.1506 = .5685
- Brother 1-((.0083-.0002)*1/2)+.0002 = .9957
- Son 1-((.0077-.0002)*1/2)+.0002 = .9960



Bayes Calculation 4

- Instead of calculation 2, consider all
- 69+N .432 / .906
- 65+C .433 / .068
- Father ((.1892-.1197)*1/2)+.1197 = .1544 / .1167
- Mother 1-((.7117-.1506)*1/2)+.1506 = .5685 / .8494
- Brother 1-((.0083-.0002)*1/2)+.0002 = .9957 / .9998
- Son 1-((.0077-.0002)*1/2)+.0002 = .9960 / .9998

	+HBOC	-HBOC
Prior	0.652	0.348
Conditional	0.432*.433*.1544*.5685*.9957*.9960	0.906*.068*.1167*.8494*.9998*.9998
Joint	0.0106	0.00212
Posterior	0.833	0.167



Secondary Findings Family Evaluation

Statement of the Problem

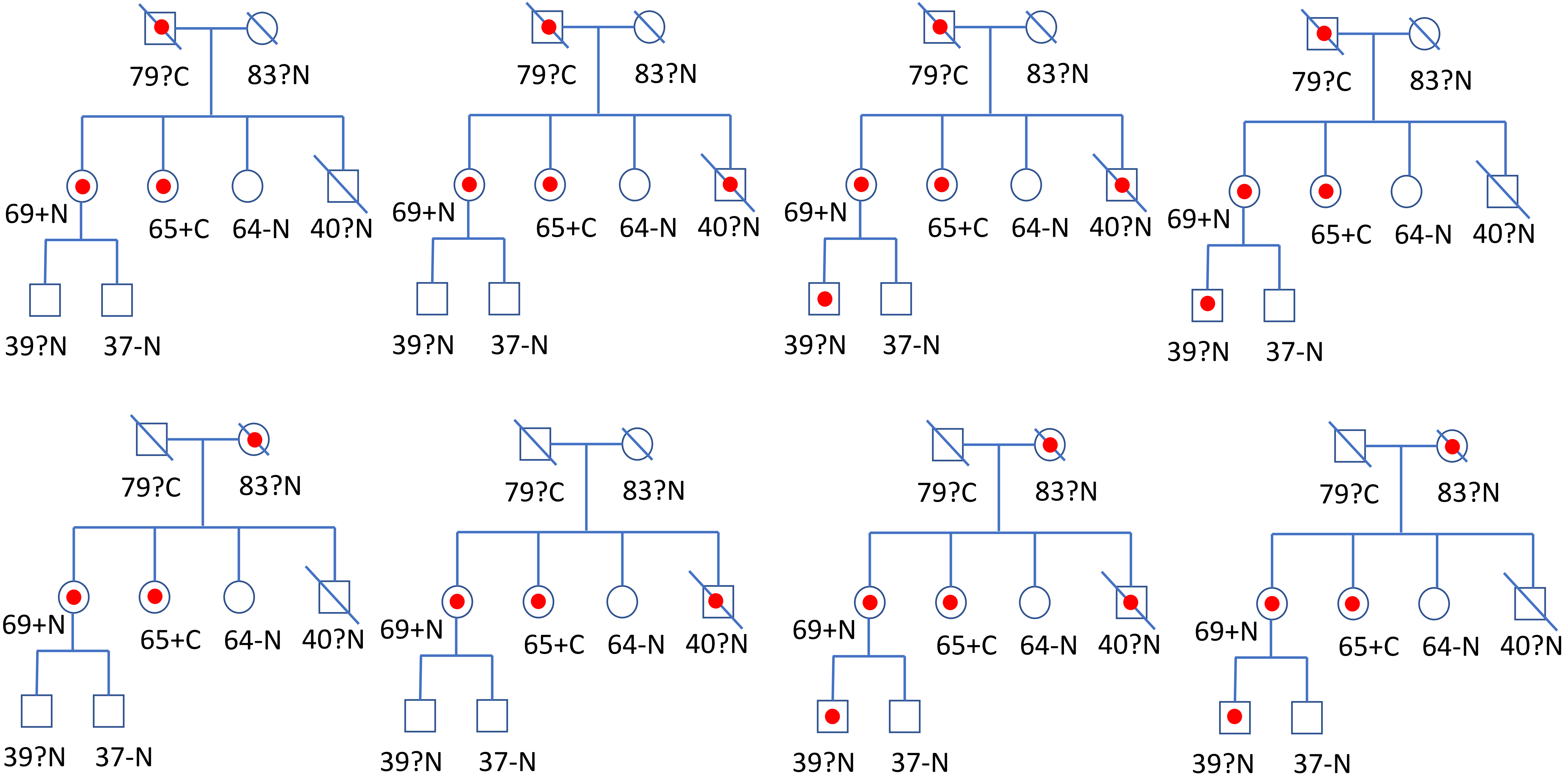
- Secondary Findings (SF) are genomic testing results that are returned in the absence of an indication for testing
  - Similar to population screening, very different from diagnostic testing
    - In general, the testee is not known have the associated phenotype *a priori*
    - Prior probability of disease ≈population risk (1/400-1/50,000)
    - In contrast, diagnostic testing prior probability of disease can be high 10->90%
  - Nearly all variants have a posterior probability of pathogenicity of <100%
    - Based on variant predictions and prior case evidence unrelated to family at hand
  - Given pathogenicity <100% and low prior probability of disease, Bayes says this results in a lower posterior probability of disease (vs diagnostic testing)
    - i.e., it enriches for truly benign variants that we currently (erroneously) believe to have high likelihood of pathogenicity

Nitty gritty

- All mendelian genetic diseases have to be defined as a state of increased propensity for disease manifestations
  - Lynch syndrome is the state of having increased liability to colon and endometrial cancer – whether or not you have cancer
  - Increased liability is very similar to penetrance
    - For some mendelian genetic diseases, penetrance is essentially 100%
    - For some it is quite low (5-10%)
- Different way to say this is that you can have a disease even if you don’t have a clinical manifestation of disease (you are nonpenetrant)
  - We must distinguish someone who has the disorder but is nonpenetrant from someone who doesn’t have the disorder

Nitty gritty

- If you harbor a variant with pathogenicity of 100%, then by definition anyone who harbors that variant has that disease
- If you harbor a variant with pathogenicity <100%, then there is a *likelihood* that you have the disease
  - The likelihood that you have the disease depends on your phenotype
    - If you have a manifestation of the disease, it is higher
    - If your family members have a manifestation, it is higher



I think that these are the only (reasonably) possible family genotypes and that all eight are equally likely

Limitations/Deficiencies

- Calculation of “cancer free” probably not correct
- Individuals over 85 treated as 85
- How to handle individuals with two cancers (357901)
- Inheritance pattern calculations (0.5^N) does not take into account skewing to one parent or the other based on affection status
- Does not take into account dependencies (if dad has variant, mom does not)
- Lumped P & LP variants together
- If a person has a proph surgery for an organ, how to take that into account when they are phenotype negative

Common Misconception

- “We performed population genomic ascertainment and observed that the penetrance was much lower than in phenotypic ascertainment”

- This is wrong
  - They are measuring both false positive

Closing Thought

- Stupid question
  - “Would you like to know if you have a high risk of developing cancer?”

Closing Thought

- Stupid question
  - “Would you like to know if you have a high risk of developing cancer?”
- Thoughtful question
  - “If you had a high risk of cancer would you rather know it and reduce it or would you rather ignore it?”

Isaac Asimov

- “Uncertainty that comes from knowledge is not the same as uncertainty that comes from ignorance.”

Interpretation in Pathology

- 115 practicing pathologists reviewed 240 breast biopsy slides
- Truth = “Consensus-derived reference”

Consensus	Pathologist Interpretation		
	Concordance Rate (95% CI)	Discordance Rate for Overinterpretation (95% CI)	Discordance Rate for Underinterpretation (95% CI)
Benign, no Atypia	87 (85-89)	13 (11-15)	
Atypia	48 (44-53)	17 (15-21)	35 (31-39)
DCIS	84 (82-86)	3 (2-2)	13 (12-15)4 (3-6)
Invasive Carcinoma	96 (94-97)		4 (3-6)

Elmore JG et al JAMA 2015;313:1122

Hat tip: Bob Nussbaum

Interpretation in Pathology

- 115 practicing pathologists reviewed 240 breast biopsy slides
- Truth = “Consensus-derived reference”

Compared with the consensus-derived reference, overall concordance per slide was 75% (95% CI, 73%-77%)

Consensus	Pathologist Interpretation		
	Concordance Rate (95% CI)	Discordance Rate for Overinterpretation (95% CI)	Discordance Rate for Underinterpretation (95% CI)
Benign, no Atypia	87 (85-89)	13 (11-15)	
Atypia	48 (44-53)	17 (15-21)	35 (31-39)
DCIS	84 (82-86)	3 (2-2)	13 (12-15)4 (3-6)
Invasive Carcinoma	96 (94-97)		4 (3-6)

Elmore JG et al JAMA 2015;313:1122

Hat tip: Bob Nussbaum

Interpretation in Pathology

- 115 practicing pathologists reviewed 240 breast biopsy slides
- Truth = "Consensus-derived reference"

Compared with the consensus-derived reference, overall concordance per slide was 75% (95% CI, 73%-77%)

Consensus	Pathologist Interpretation		
	Concordance Rate (95% CI)	Discordance Rate for Overinterpretation (95% CI)	Discordance Rate for Underinterpretation (95% CI)
Benign, no Atypia	87 (85-89)	13 (11-15)	
Atypia	48 (44-53)	17 (15-21)	35 (31-39)
DCIS	84 (82-86)	3 (2-2)	13 (12-15)4 (3-6)
Invasive Carcinoma	96 (94-97)		4 (3-6)

Elmore JG et al JAMA 2015;313:1122

Hat tip: Bob Nussbaum

Uncertainty at another level: 0.5%-2.5% of histology reports are the result from some other patient…

How to calculate

	Has Marfan	Does not have Marfan
Prior	.75	.25
Conditional	.7	.001
Joint	.525	.00025
Posterior	.525/(.525+.00025) ≈ .9995	.00025/(.525+.00025) ≈ .0005

Context matters – enormously – V2

- Pediatrician orders exome on a toddler re autism
- No variant for the autism is identified
- There is a secondary finding of a pathogenic variant in *FBN1*
- This toddler has no apparent features of Marfan syndrome
- She is adopted, so she has no known family history
- What is the likelihood the toddler has Marfan syndrome?

Context matters – enormously – V2

- Pediatrician orders exome on a toddler re autism
- No variant for the autism is identified
- There is a secondary finding of a pathogenic variant in *FBN1*
- This toddler has no apparent features of Marfan syndrome
- She is adopted, so she has no known family history
- What is the likelihood the toddler has Marfan syndrome?
- ~8%

How to calculate V2

	Has Marfan	Does not have Marfan
Prior	.00013	.99987
Conditional	.7	.001
Joint	.000091	≈.001
Posterior	.000091/(.000091+.001) ≈ .085	.001/(.000091+.001) ≈ .915

Back to variant classification…

- Let’s start with an easy one
- *GLI3* c.444C>A; p.Y148*
- GLI3 zinc finger transcription factor
  - Assoc w Greig cephalopolysyndactyly, Pallister-Hall syndrome, various polydactyly, etc.
- Putative loss of function variant