Genetic risk calculations: Dominant & X-linked recessive

26.10.2007

GE02 day 5 part 1

Yurii Aulchenko

Erasmus MC Rotterdam

Overview

- Reminder on Bayes
- Risk under rare X linked recessive mutation model
- Risk under dominant model

Total probability and Bayes' formulae

Two sets of events are considered:

- "Hypothesis" H_i for which *a prioi* probabilities, $P(H_i)$ are known. These hypotheses must be mutually exclusive and cover all possible outcomes. E.g. genotypes in some person may be "hypotheses".
- Event(s) of interest, A, e.g. disease. For this event, conditional probabilities, $P(A|H_i)$, are known

Total probability & Bayes' formulae

Total probability (of event A)

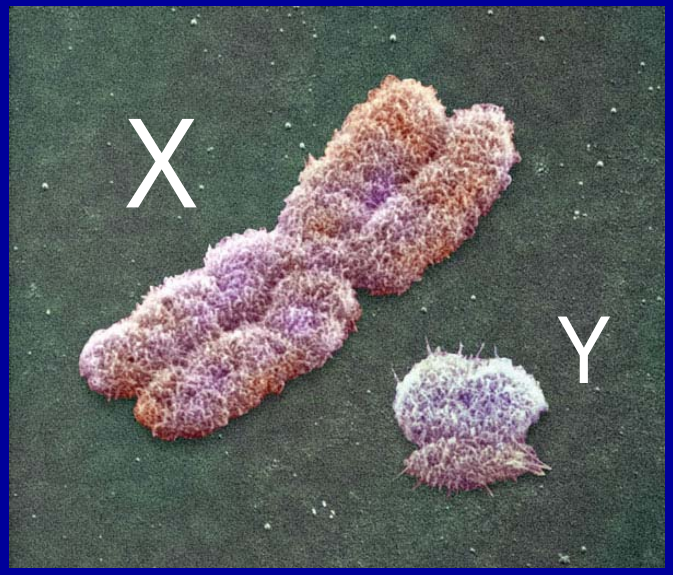
$$P(A) = \sum_i P(A|H_i)P(H_i)$$

Probability of hypothesis H_i, given A

$$P(H_i \mid A) = \frac{P(A, H_i)}{P(A)} = \frac{P(A|H_i)P(H_i)}{P(A)} = \frac{P(A|H_i)P(H_i)}{\sum_i P(A|H_i)P(H_i)}$$

On X and Y

- Female: XX, Male: XY
- Y:
  - “small and useless”
  - “genetic wasteland”
  - “junk”
  - …
  - “Poor men! They miss a chromosome!”
- Not quite – Y contains genes which make man a man!



Rare X-linked recessive

Penetrances:

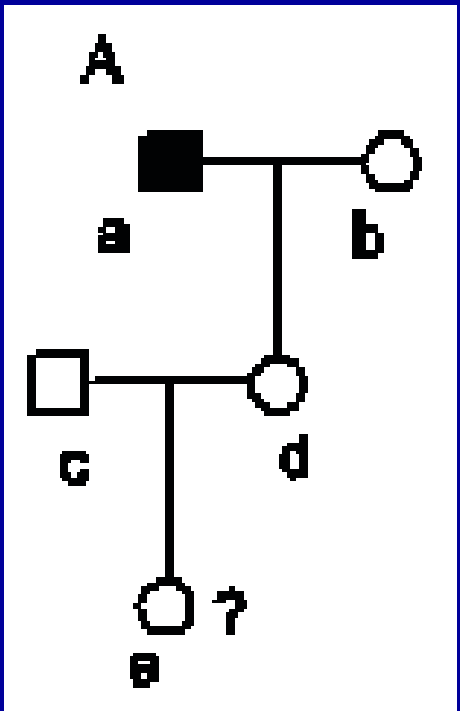
$$P(A|DD)=P(A|DY)=1$$
$$P(DX)=P(XX)=P(XY)=0$$

If frequency is q

- Disease prevalence in women $P(D|w) = q^2$
- Disease prevalence in men $P(D|m) = q$
- RR for men compared to women is 1/q
- Say, rare is q=1/1000, then RR=1000 (!)

Probability to be a carrier (1)

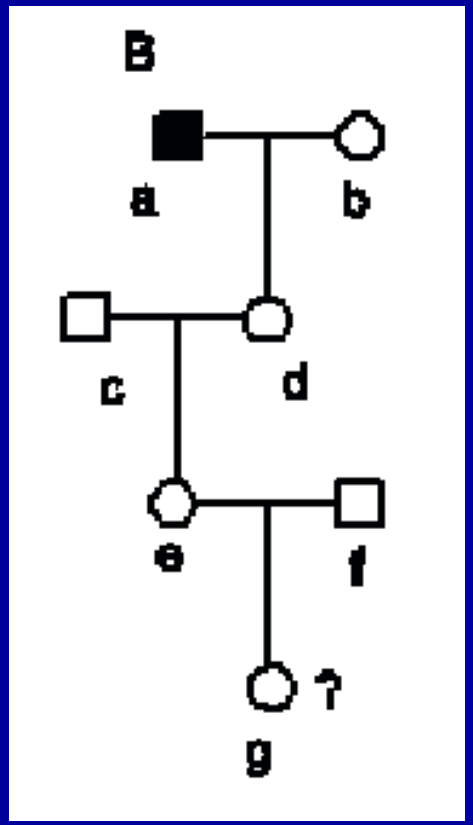
- Grandfather “a” of a girl “e” is affected
- What is probability that she is a carrier (DX)?



- “d” is obligate carrier (DX)
- Chance of transmission of D, and normal X, from “d” to “e” is ½
- The chance for “e” to be a carrier is ½

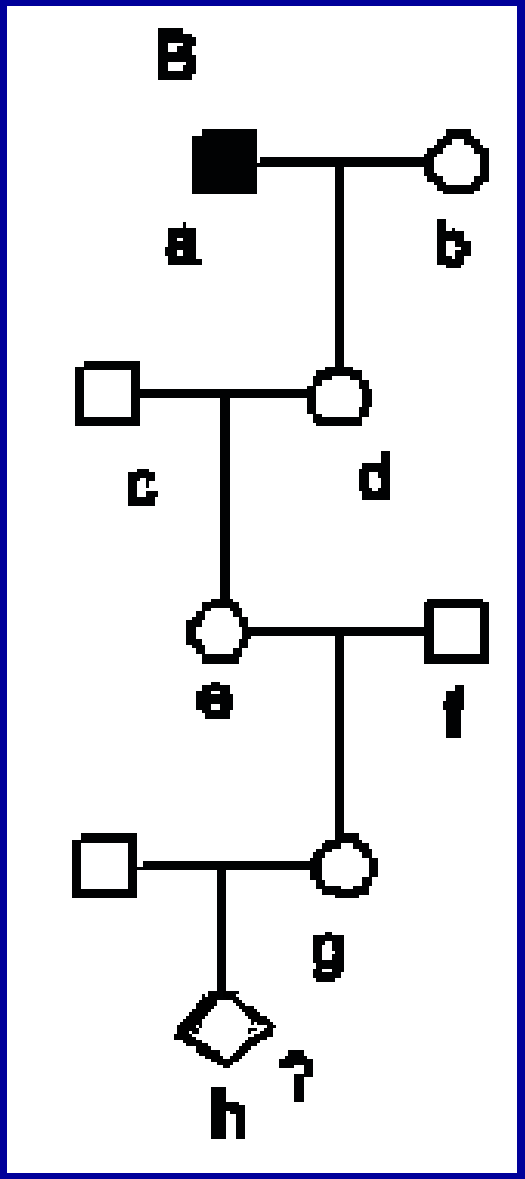
Probability to be a carrier (2)

- Grandfather “a” of a girl “g” is affected
- What is probability that she is a carrier (DX)?



- “d” is obligate carrier (DX)
- P(“e” is DX) = ½
- If “e” is DX, probability that D will be transmitted to “g” is ½
- The chance for “g” to be a carrier is ½ ½ = ¼

Risk for a (boy-)child



- The chance for “g” to be a carrier is ¼
- If “g” is carrier, she will transmit “D” with probability of ½
- Risk that a **boy** would get the disease is ¼ ½ = 1/8
- Risk that **a** child would get the disease is ¼ ½ ½ = 1/16

We are applying total probability!

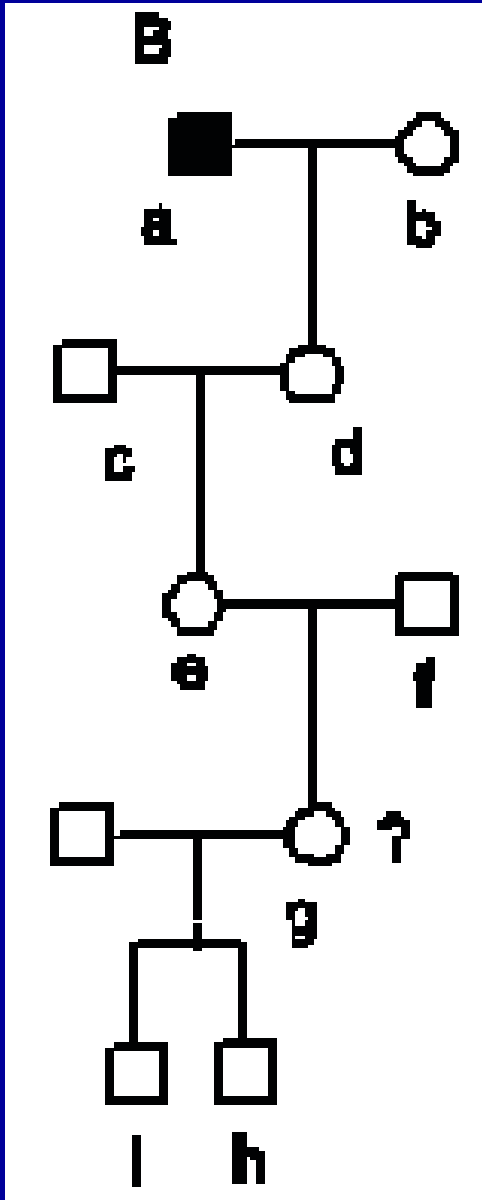
- Mutually exclusive hypotheses, covering all possibilities:
  - H1: “g” is carrier (DX)
  - H2: “g” is not carrier (XX)
- Prior probabilities of hypotheses
  - P(H1) = ¼
  - P(H2) = ¾
- Conditional probabilities of event given hypotheses
  - P(boy=A|H1) = ½
  - P(boy=A|H2) = 0

- P(boy=A) =

  P(boy=A|H1) P(H1) + P(boy=A|H2) P(H2)

  ½ x ¼ + 0 x ¾ = 1/8

Probability to be a carrier (3)



$$P(g = DX \mid data = \text{"two sons are OK"}) =$$

$$\frac{P(g = DX, data)}{P(data)} = \frac{P(data \mid g = DX)P(g = DX)}{P(data)} =$$

$$\frac{P(data \mid g = DX)P(g = DX)}{\sum_{g_g = DX, XX} P(data \mid g_g)P(g_g)} =$$

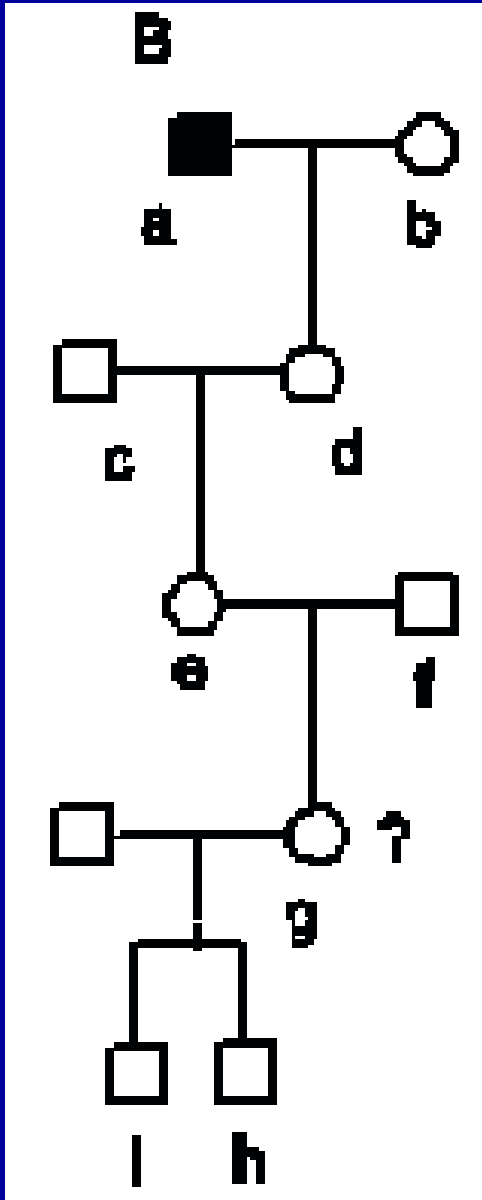
$$\frac{P(data \mid g = DX)P(g = DX)}{P(data \mid g = DX)P(g = DX) + P(data \mid g = XX)P(g = XX)} =$$

$$\frac{(1/2 \cdot 1/2) \cdot 1/4}{(1/2 \cdot 1/2) \cdot 1/4 + (1 \cdot 1) \cdot 3/4} = \frac{1/16}{1/16 + 12/16} = 1/13 = 0.077$$

- Information on 2 healthy sons decreases initial “carriership” risk estimate by 0.25/0.077 = 3.25 times

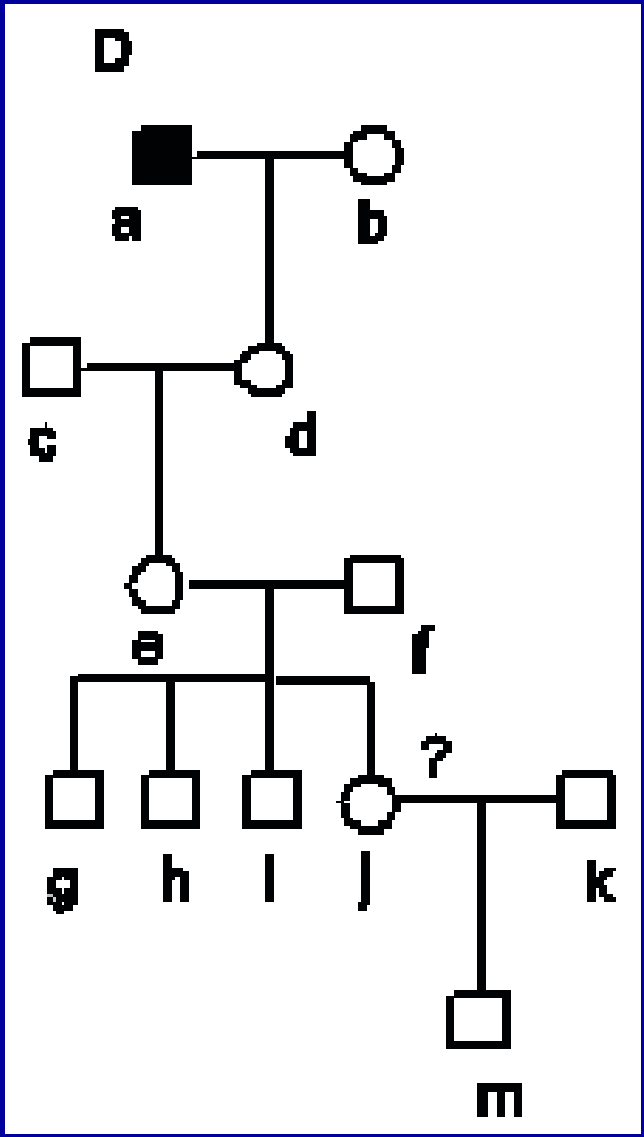
Tabular way

$$P(H_i \mid data) = \frac{P(H_i, data)}{P(data)}$$



Probability	Hypothesis	
	H_1: "g" is DX	H_2: "g" is XX
Prior, $P(H_i)$	¼	¾
Conditional, $P(data\|H_i)$	½ ½ = ¼	1
Joint, $P(data\|H_i)P(H_i)$	¼ ¼ = 1/16	1 ¾ = ¾
Total, P(data)	1/16 + ¾ = 13/16	
Posterior, $P(H_i\|data)$	(1/16)/(13/16) =1/13	(3/4)/(13/16) =12/13
Risk for the next boy	1/13 ½ = 1/26	

Risk for a boy-child of “j”



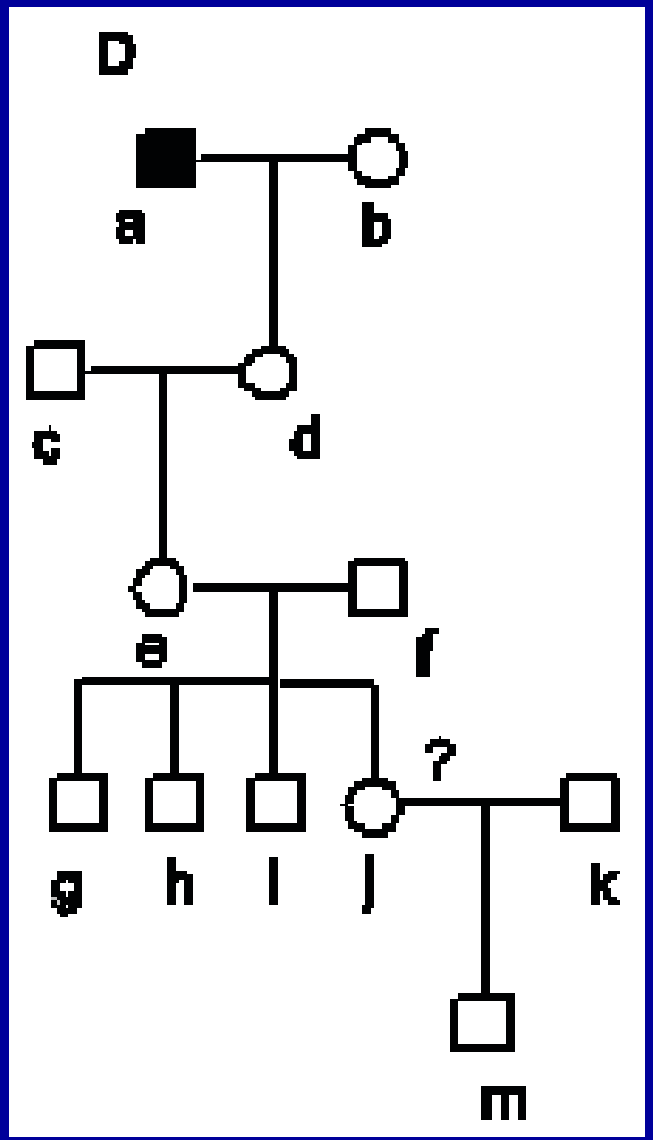
Hint: act in steps (“peeling”)

(1) Compute posterior probability that “e” is carrier, given she has 3 healthy sons

(2) Use (1) to compute the new “prior” – that “j” is carrier

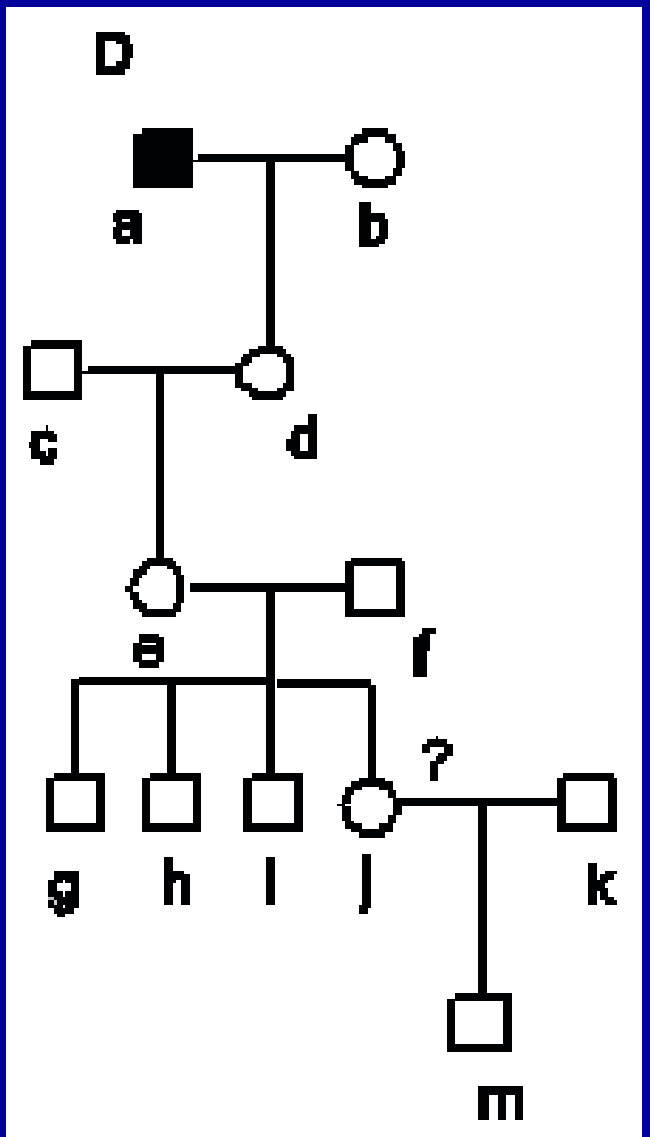
(3) Compute posterior probability that “j” is carrier, given she has 1 healthy son

Probability that “e” is carrier



Probability	Hypothesis	
	H_1: “e” is DX	H_2: “e” is XX
$P(H_i)$	½	½
$P(data\|H_i)$	$½^3 = 1/8$	1
$P(data\|H_i)P(H_i)$	½ 1/8 = 1/16	1 ½ = ½
$P(data)$	1/16 + ½ = 9/16	
$P(H_i\|data)$	(1/16)/(9/16)=**1/9**	1 – 1/9 = 8/9

Probability that “j” is carrier



Probability	Hypothesis	
	H_1: “j” is DX	H_2: “j” is XX
$P(H_i)$	½ **1/9** = 1/18	1 – 1/18 = 17/18
$P(data\|H_i)$	½	1
$P(data\|H_i)P(H_i)$	½ 1/18 = 1/36	1 17/18 = 17/18
$P(data)$	1/36 + 17/18 = 35/36	
$P(H_i\|data)$	(1/36)/(35/36)=**1/35**	34/35
Risk for boy	½ 1/35 = **1/70**	

P(“j”=DX|data)=1/35

Risk for boy-child = 1/70

Overview

- Reminder on Bayes
- Risk under rare X linked recessive mutation model
- **Risk under dominant model**

Dominant disease

- Risk allele D with frequency q

- P(Affected|DD) = P(Affected|DN) = 1
- P(Affected|NN) = 0

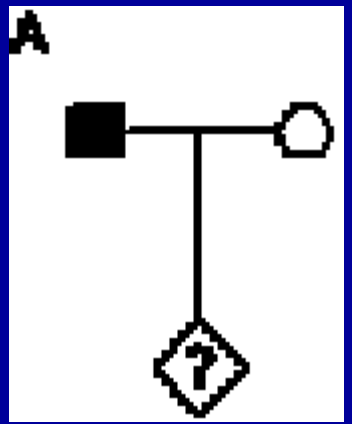
Rare mutation (q=0.001)

What is the risk of the child to have the disease?

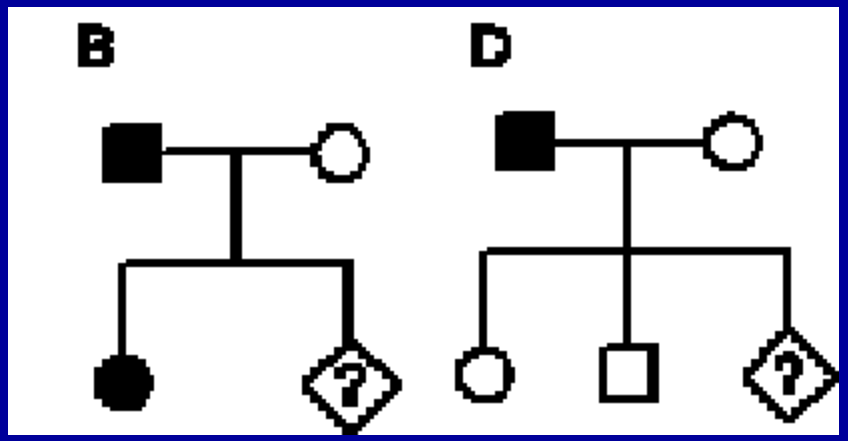
- Because the mutation is so rare, you can assume that any affected person has genotype DN

- Say, in 2000 affected, 1 is DD and 1999 are ND.

DN



NN



P(?=A|f=A) = 50%

P(?=A|data)=50%!

Mutation is not so rare (q=0.1)

- You can not assume that “all” affected are heterozygous
- Frequency of DD among affected = [Bayes] =

$$P(DD \mid aff) = \frac{P(aff, DD)}{P(D)} = \frac{P(aff \mid DD) \cdot P(DD)}{P(D)} =$$

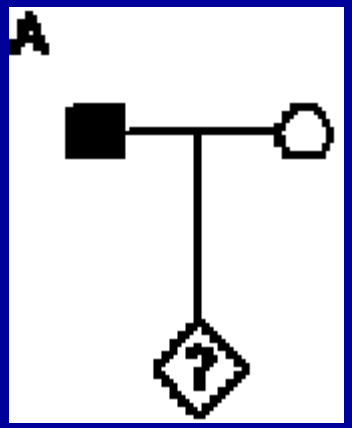
$$\frac{P(aff \mid DD) \cdot P(DD)}{P(aff \mid DD) \cdot P(DD) + P(aff \mid ND) \cdot P(ND) + P(aff \mid NN) \cdot P(NN)} =$$

$$\frac{1 \cdot q^2}{1 \cdot q^2 + 1 \cdot 2 \cdot (1-q) \cdot q + 0 \cdot (1-q)^2} = \frac{0.1^2}{0.1^2 + 2 \cdot (1-0.1) \cdot 0.1} = 0.053$$

- ~ Among 100 affected 5 are DD and 95 are ND

What is the risk for the child?

P(DN|aff)=0.947
P(DD|aff)=0.053



P(NN|unaff)=1

P(?=affected|data) = 1*0.053 + ½ 0.947 = 0.526

Formal derivation

$$P(?=A \mid f=A) = P(?=ND \mid f=A) = \frac{P(?=ND, f=A)}{P(f=A)} =$$

$$\frac{P(?=ND, f=A)}{\sum_{g_f=NN,ND,DD} P(f=A \mid g_f)P(g_f)} = \frac{\sum_{g_f=NN,ND,DD} P(?=ND, f=A \mid g_f)P(g_f)}{\sum_{g_f=NN,ND,DD} P(f=A \mid g_f)P(g_f)} =$$

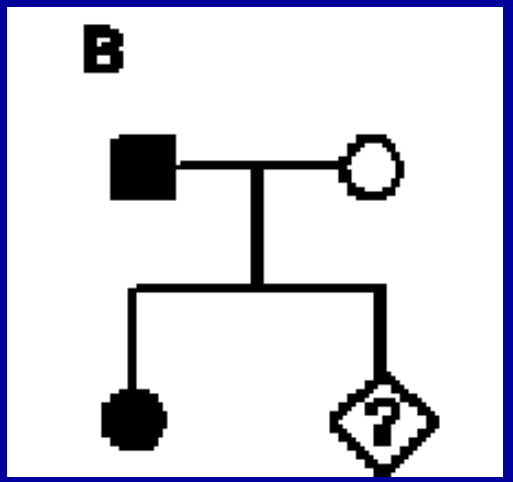
$$\frac{\sum_{g_f=NN,ND,DD} P(?=ND \mid g_f)P(f=A \mid g_f)P(g_f)}{\sum_{g_f=NN,ND,DD} P(f=A \mid g_f)P(g_f)} =$$

$$\frac{P(?=ND|f=NN)P(f=A|f=NN)P(f=NN)+P(?=ND|f=ND)P(f=A|f=ND)P(f=ND)+P(?=ND|f=DD)P(f=A|f=DD)P(f=DD)}{P(f=A|f=NN)P(f=NN)+P(f=A|f=ND)P(f=ND)+P(f=A|f=DD)P(f=DD)} =$$

$$\frac{0\cdot 0\cdot 0.9^2 + 0.5\cdot 1\cdot(2\cdot 0.1\cdot(1-0.1)) + 1\cdot 1\cdot 0.1^2}{0\cdot 0.9^2 + 1\cdot(2\cdot 0.1\cdot(1-0.1)) + 1\cdot 0.1^2} = \frac{0.09+0.01}{0.18+0.01} = \frac{0.1}{0.19} = 0.526$$

Pedigree B

- Father and one child are affected
- What is the risk for the next child?
- Still 53%?



- If we knew P(f=DD|data) & P(f=ND|data) then the risk for “?” is computed as

P(?=A|data) = ½ P(f=ND|data) + 1 P(f=DD|data)

P(f=ND|data)

- Use Bayes formula:

$$P(f = ND \mid data) = P(f = ND \mid f = A, s = A) =$$

$$\frac{P(f = ND, f = A, s = A)}{P(f = A, s = A)} = \frac{P(f = ND, s = A)}{P(f = A, s = A)} =$$

$$\frac{P(f = ND) \cdot 0.5}{P(f = A, s = A)} = \frac{0.18 \cdot 0.5}{P(f = A, s = A)} =$$

$$\frac{0.18 \cdot 0.5}{0.18 \cdot 0.5 + 0.01} = \frac{0.09}{0.1} = 0.9$$

P(f=DD|data)

$$P(f = ND \mid data) = P(f = ND \mid f = A, s = A) =$$

$$\frac{P(f = DD, f = A, s = A)}{P(f = A, s = A)} = \frac{P(f = DD, s = A)}{P(f = A, s = A)} =$$

$$\frac{P(f = DD)}{P(f = A, s = A)} = \frac{0.1^2}{P(f = A, s = A)} =$$

$$\frac{0.1^2}{0.18 \cdot 0.5 + 0.01} = \frac{0.01}{0.1} = 0.1$$

- Shortcut: given phenotype, father can not be NN
- Thus P(f=DD|data) = 1 - P(f=DD|data) = 1 - 0.9 = 0.1

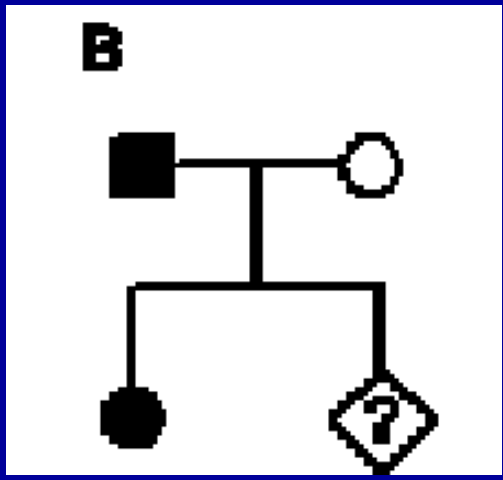
Pedigree B: risk for next boy

- Risk for the next child is

  P(?=A|data) =

  ½ P(f=ND|data) + P(f=DD|data) =

  ½ 0.9 + 0.1 = 0.55



Hypothesis:	Father is DD	Father is ND
Prior, P(g)	0.053	0.947
Conditional, P(X|g)	1	½
Joint, P(X|g) P(g)	0.053	0.4735
Posterior	0.053/(0.053+0.4735)=0.1	0.9
Risk for next boy	0.1 + 0.9 ½ = 0.55	