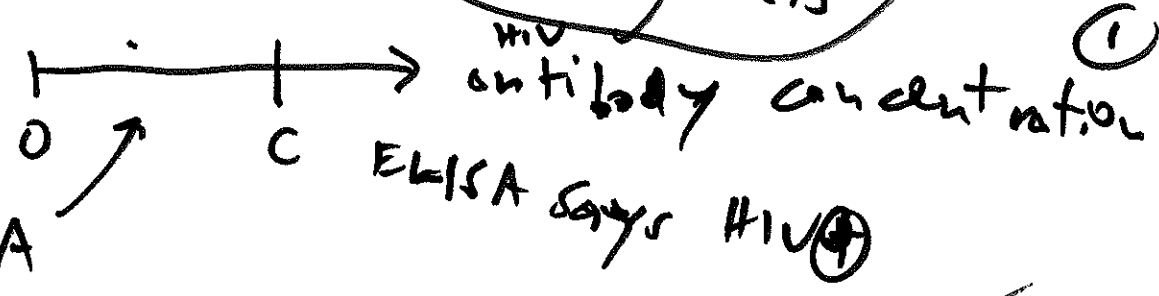


Discussion Section 3

Case study: HIV & AIDS diagnosis

AMS131, 6 Apr 19



Decision: you have to act (based on science)

Science: search for knowledge without action

$A =$  (really is HIV $\oplus$ )      not  $A =$  (really is HIV $\ominus$ )

$\oplus =$  (ELISA says HIV $\oplus$ )      prevalence

$\ominus =$  (ELISA says HIV $\ominus$ )       $P(A) = 0.004$

sensitivity  $P(\oplus | A_+) = 0.96$

specificity  $P(\ominus | \text{not } A) = 0.97$

when blood arrives, <sup>compare</sup> ELISA says  $\oplus$   $\textcircled{2}$

$P(A | \oplus) = ?$   
↑ truth      ↑ data

$\boxed{Q:}$  how does  $P(A | B)$  relate to  $P(B | A)$ ?

almost always not the same:

$P(\text{clouds overhead} | \text{raining}) = \text{high}$   
 $P(\text{raining} | \text{clouds overhead}) = \text{low}$

first  $\boxed{A:}$   
Bayes (1760)

effect  
village:  $\textcircled{\text{some people dying}}$  ;  
 $P(\text{truth} | \text{data})$   
 $P(\text{cause} | \text{effect})$   $\textcircled{\text{data}}$   
harder

cause?  
 $\textcircled{\text{why?}}$   
possible causes

$P(\text{effect} | \text{cause})$  easier  
 $P(\text{data} | \text{truth})$

bad food  
bad water  
bad air  
disease  
 $\textcircled{\text{truth}}$

$$P(A \text{ and } B) = \begin{matrix} P(A) \cdot P(B|A) \\ \parallel \\ P(B) \cdot P(A|B) \end{matrix}$$

So

$$P(A) P(B|A) = P(B) P(A|B)$$

$$P(B|A) = \frac{P(B) P(A|B)}{P(A)}$$

Bayes's Theorem for T/F propositions (sets)

$$P(\text{cause} | \text{effect}) = \frac{P(\text{cause}) P(\text{effect} | \text{cause})}{P(\text{effect})}$$

$$P(\text{truth} | \text{data}) = \frac{P(\text{truth}) P(\text{data} | \text{truth})}{P(\text{data})}$$

a priori      a posteriori

before data       $\xrightarrow{\text{new data arrives}}$       after data      time

posterior info. about truth given data =  $\frac{\text{prior info. about truth} \cdot \text{data info. about truth}}{\text{[normalizing constant]}}$

prevalence      sensitivity

$$P(A|\oplus) = \frac{P(A) \cdot P(\oplus|A)}{P(\oplus)}$$

3 methods to get ①

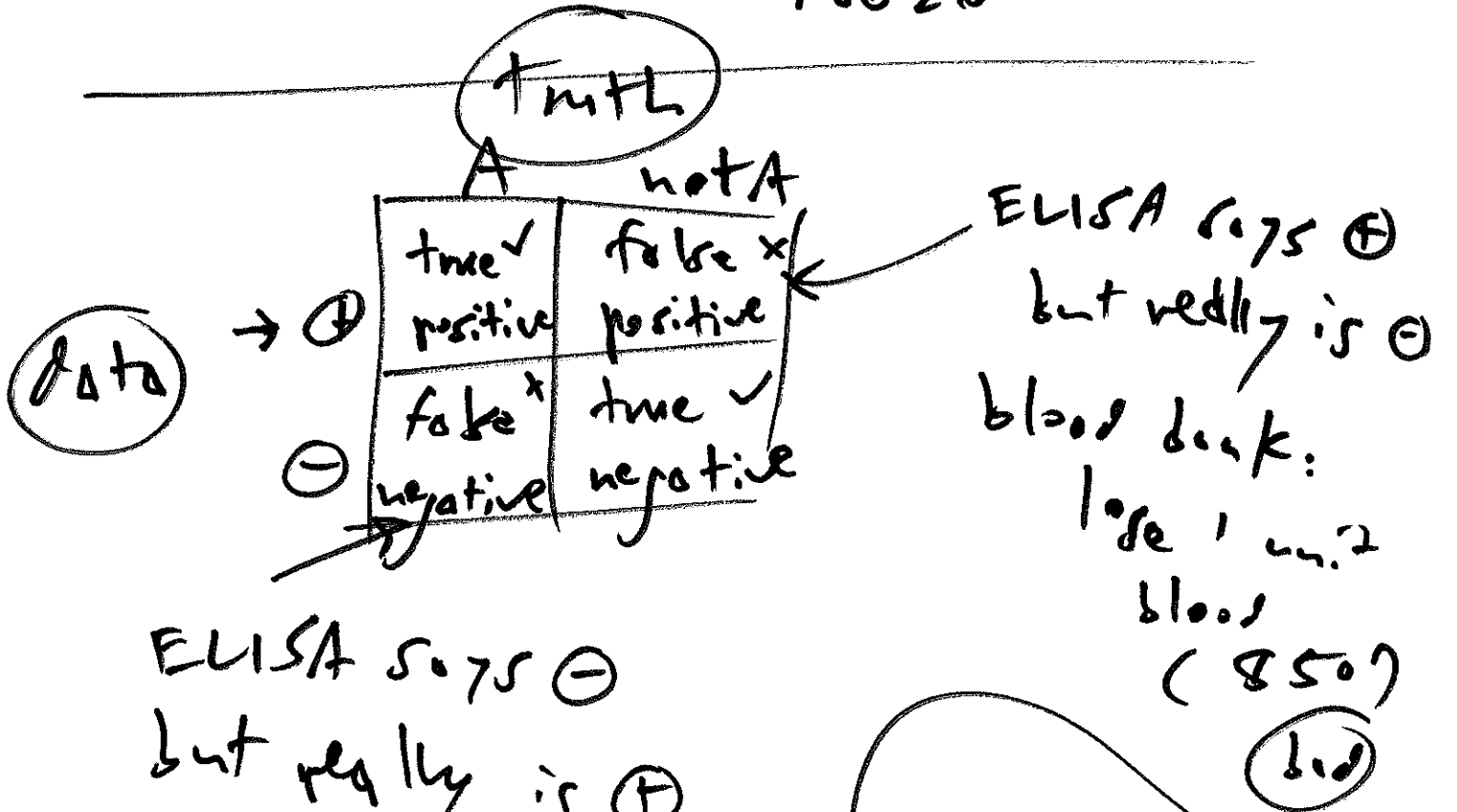
dichotomous → having 2 possible values

		Truth		total
		A	not A	
Data	ELISA (+)	384	2,988	3,372
	soy (-)	16	96,612	96,628
total		400	99,600	100,000

prevalence

$$P(A|\oplus) = \frac{384}{3372} = 11\% (!)$$

$$P(\text{not } A | \ominus) = \frac{96612}{96628} \approx 0.9998 \quad (5)$$



ELISA says ⊖  
but really is ⊕

blood bank:  
somebody gets HIV  
terrible

western blot:  
a lot more WP  
accurate than  
ELISA but more  
expensive

adaptive strategy:

run ELISA on  $\frac{1}{2}$  blood;

if ⊖ announce ⊖; if ⊕

run WP on other  $\frac{1}{2}$  & announce WB result

method 2 (creative laziness) ⑥

$$P(B|A) = \frac{P(B) P(A|B)}{P(A)}$$

$$P(\overset{not}{B}|A) = \frac{P(\overset{not}{B}) P(A|\overset{not}{B})}{P(A)}$$

$$\left( \frac{P(B|A)}{P(\overset{not}{B}|A)} \right) = \left( \frac{P(B)}{P(\overset{not}{B})} \right) \cdot \left( \frac{P(A|B)}{P(A|\overset{not}{B})} \right)$$

posterior odds ratio in favor of B given A = (prior odds ratio in favor of B) (Bayes factor)

↓  
likelihood ratio  
↓  
data odds in favor of B