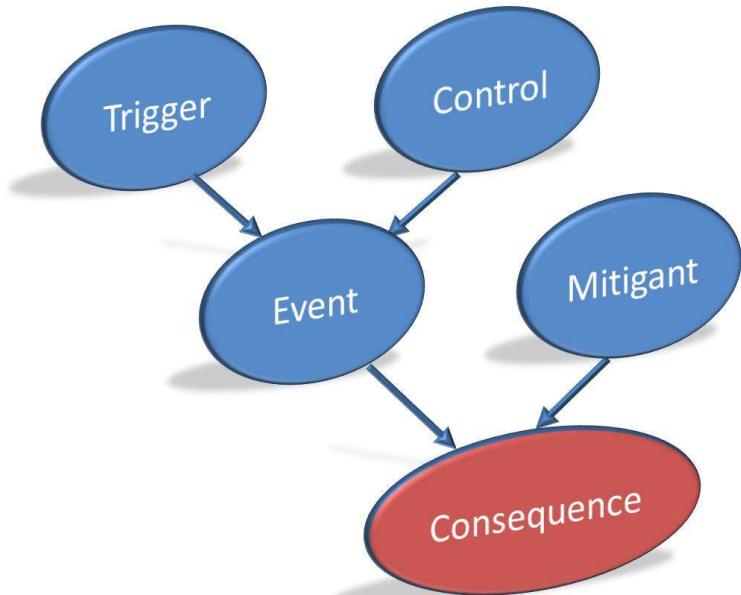


## Risk and Decision Making for Data Science and AI

# Week 11: Bayes and Risk Assessment in the Law and Forensics

Norman Fenton  
@ProfNFenton



### **Scenario 1**

A murder is committed in a crowded pub. The victim is known to have put up a fight and DNA from one other person is found on the victim. The DNA profile matches Jimmy and, according to the DNA expert analyst:

*“This DNA evidence very strongly supports the hypothesis that it was Jimmy’s DNA on the victim. In fact, it is over a billion times more likely under this hypothesis than under the hypothesis that it came from a person other than Jimmy”*

On a scale of 0 to 100 how sure are you that:

- the DNA came from Jimmy?
- Jimmy was the murderer?

### **Scenario 2**

A murder is committed in a crowded pub. The victim is known to have put up a fight and DNA from at least two other people is found on the victim. A DNA profile that matches Jimmy’s is in the DNA mixture and, according to the DNA expert analyst :

*“This DNA evidence very strongly supports the hypothesis that Jimmy was one of the people whose DNA was on the victim. In fact it is over a billion times more likely under this hypothesis than under the hypothesis that Jimmy was not a contributor”*

On a scale of 0 to 100 how sure are you that:

- Jimmy’s DNA was on the victim?
- Jimmy was the murderer?

# Overview

- 1. Cases**
- 2. Bayes and the Law: barrier and fallacies, DNA evidence**
- 3. Probative value of evidence**
- 4. Limitations of the Likelihood ratio**
- 5. The need for causal models (BNs) in evidence evaluation**
- 6. Likelihood ratio: the worst part (DNA mixture evidence)**
- 7. Conclusions and way forward**

# **R vs Levi Bellfield, Sept 07 – Feb 08 convicted of murder, 2011 convicted of murder of Millie Dowler**



Manhunt watched by 9 million

Published: Thu 17 Jan 2019



**Levi Bellfield** (now named **Yusuf Rahim** after converting to Islam in prison)

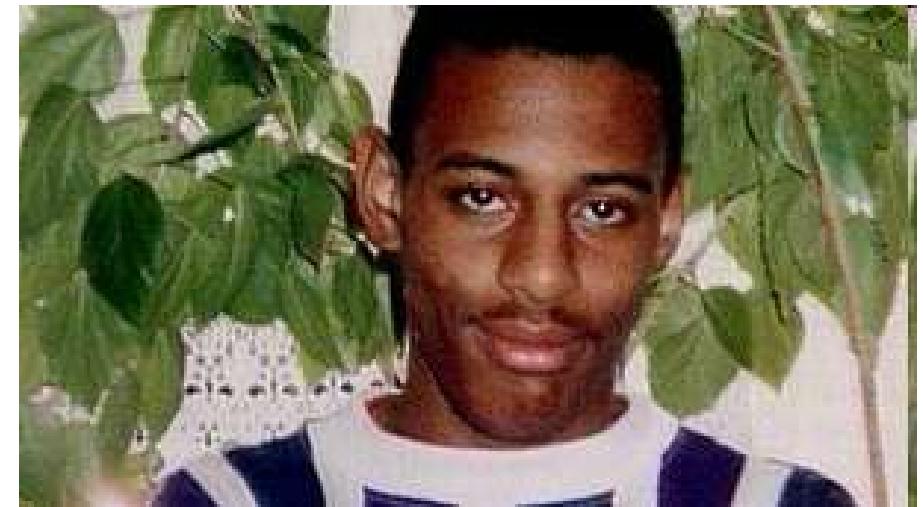


**Amelie Delagrange**



**Marsha McDonnell**

# **R v Gary Dobson 2011 (2<sup>nd</sup> trial) convicted of murder**



**Stephen Lawrence**

<https://probabilityandlaw.blogspot.com/2012/01/prosecutor-fallacy-in-stephen-lawrence.html>

## **R v Ben Geen, 2014**



**Nurse - convicted of murder of 2 patients**

<https://probabilityandlaw.blogspot.com/2016/03/statistics-of-coincidences-ben-geen.html>

# R v K, Croydon Riots 2011

K charged with robbery  
during riot



# Lewenstein family v Dutch Government Restitutions Committee, 2018

Family claimed Kandinsky painting belonged to them

<https://probabilityandlaw.blogspot.com/2018/12/the-case-of-kandinsky-painting-and.html>

Support The Guardian

Contribute → Subscribe →

The Guardian

Sign in

News Opinion Sport Culture Lifestyle

World ▶ Europe US Americas Asia Australia Middle East Africa Inequality Cities More

Netherlands

## Dutch art panel's ruling against Jewish family criticised as 'step back'

Committee backed museum in battle over Kandinsky painting obtained before Nazis invaded



A detail from Painting with Houses (1909) by Wassily Kandinsky. Photograph: World History Archive/Alamy

**Daniel Boffey in Brussels**

Wed 5 Dec 2018 05.00 GMT

This article is over 1 month old

# State v Dominguez 2019

Previous murder conviction – released after DNA analysis deemed problematic

The San Diego Union-Tribune

## Murder case that highlighted DNA-analysis controversy ends with plea to reduced charge, release

The case highlighted the clash between powerful new forensic tools developed by private companies for police and prosecutors and rights of defendants to examine how the tools work

By GREG MORAN  
DEC. 6, 2019  
3:53 PM

A man who was once sentenced to 50 years to life in prison for a 2008 murder in a San Diego park walked out of custody Friday after pleading guilty to a reduced charge, ending a long-running case that challenged the use of a powerful new DNA testing technology.

In an agreement with the San Diego County District Attorney's Office, Florencio Jose Dominguez, 41, pleaded guilty to a single charge of manslaughter in the 2008 death of 15-year-old Moises Lopez. The agreement calls for Dominguez to get an 11-year sentence, but he has been in prison for a longer time than that, when credits for the time he has already served are factored in.

As a result, Superior Court Judge Charles G. Rogers ordered Dominguez to be released immediately Friday.

[Facebook](#) [Twitter](#) [Email](#)

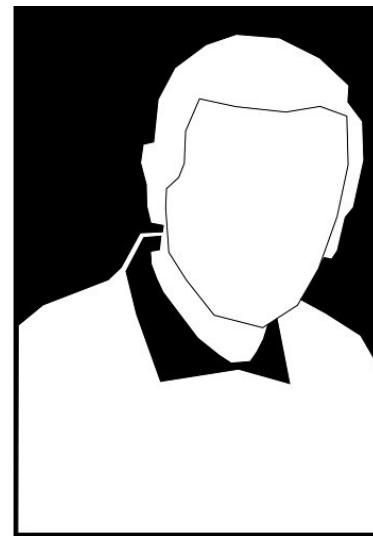


<https://probabilityandlaw.blogspot.com/2019/12/problems-with-dna-mixed-profile.html>

## **R v LW since 2010**

**Convicted of rape of half-sister**

**Low template “DNA match” at issue**



# **R v APR et al, R v MR**

## **Drugs on Banknotes (statistical analysis)**

**R v APR et al, Liverpool Crown Court,  
Nov 2016**

**R V MR, Snaresbrook Crown Court, May-  
August 2016**



**R v C, 2020**

## **Gunshot residue evidence**

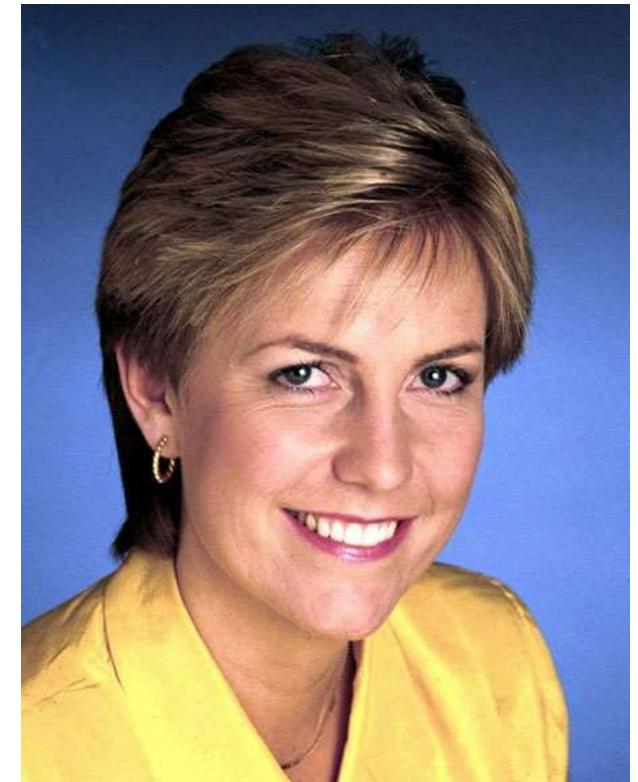
**How probative is the evidence?**



# **R v Barry George, 2001-2007**

originally convicted of  
murder

probabilistic issue  
with probative value  
of gunpowder residue  
evidence



**Jill Dando**

<https://probabilityandlaw.blogspot.com/2013/09/barry-george-case-new-insights-on.html>

<https://probabilityandlaw.blogspot.com/2019/01/new-paper-shows-how-important-evidence.html>

## R v Sally Clark

Originally convicted of murder of both her children

Prosecution expert argued  $p(\text{SIDS}) = 1/8543$  and hence that

$$P(\text{'both SIDs'}) = P(\text{SIDS}) \times P(\text{SIDS}) = 1/72,982,849$$

Concluded “ $P(\text{innocence})$  about 1 in 73 million”

But this makes two fundamental statistical errors

1. Wrongly assumes two brother's dying of SIDS are independent events
2. Fails to take account of prior probability of murder given an infant death; only 8% of infant deaths are murder, hence if SIDS and murder are the only possibilities,  
 $P(\text{SIDS for a dead infant}) = 92\%$



<https://probabilityandlaw.blogspot.com/2014/01/sally-clark-revisited-another-key.html>

# **R v KG, 2020**

**Parasomnia:**  
**Convicted of rape – but was  
the defendant asleep during  
non-consensual sex?**



# BAYES and the law: barriers and fallacies

How would you answer these questions  
(method not actual answer)?

**What is  $723539016321014567$  divided by  
 $9084523963087620508237120424982$ ?**

**What is the area of a field whose length is approximately 100 metres and whose width is approximately 50 metres?**

**Assume 1% of the population has coronavirus. A screening test has a true positive rate of 99% and a true negative rate of 95%. What's the probability a person testing positive has the virus?**

# Court of Appeal Rulings (about the use of Bayes)

**“The task of the jury is to evaluate evidence and reach a conclusion not by means of a formula, mathematical or otherwise, but by the joint application of their individual common sense and knowledge of the world to the evidence before them”**

**(R v Adams, 1995)**

**“..no attempt can realistically be made in the generality of cases to use a formula to calculate the probabilities. .. it is quite clear that outside the field of DNA (and possibly other areas where there is a firm statistical base) this court has made it clear that Bayes theorem and likelihood ratios should not be used”**

**(R v T, 2010)**

# Evidence from a crime scene

Some of those who were at  
the scene of the crime



Fred



Police discover  
shoeprint of person who  
committed the crime –  
it's size 13

Nationally only about 1 in  
a 100 men are size 13



Fred is size 13

*What is the probability Fred is innocent?*

Which of these statements are correct?  
(you can choose more than 1)

1. The probability of this evidence given Fred is innocent is 1 in 100  $p(E | \text{Innocent})$
2. The probability Fred is innocent given this evidence is 1 in 100  $p(\text{Innocent} | E)$
3. The statements in 1 and 2 are equivalent
4. Neither 1 or 2 is correct

'This evidence' is that the guilty person is size 13 - same as Fred

*The 'prosecution fallacy' is to (wrongly) assume 3*

i.e.  $p(\text{Innocent} | E) = p(E | \text{Innocence}) = 1/100$

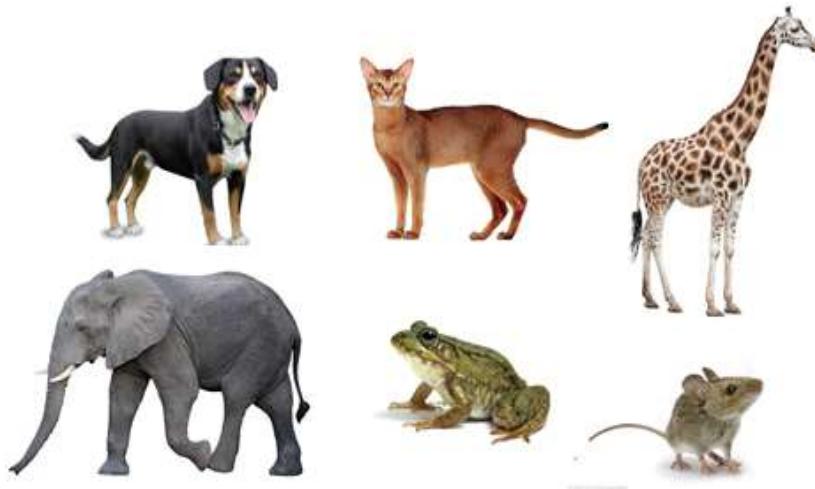
# Prosecutor's fallacy (transposed conditional)

All cows are 4-legged animals

Can we conclude all 4-legged animals  
are cows?

Let H be the assertion: “Animal is a cow”  
(our Hypothesis)

Let E be the assertion: “Animal has 4  
legs” (our Evidence)



The probability of E given H, written  $P(E|H)$ , is equal (or close to) to 1  
The probability of H given E, written  $P(H|E) = ??$

*The prosecutor's fallacy is to assume  $P(H|E) = P(E|H)$*

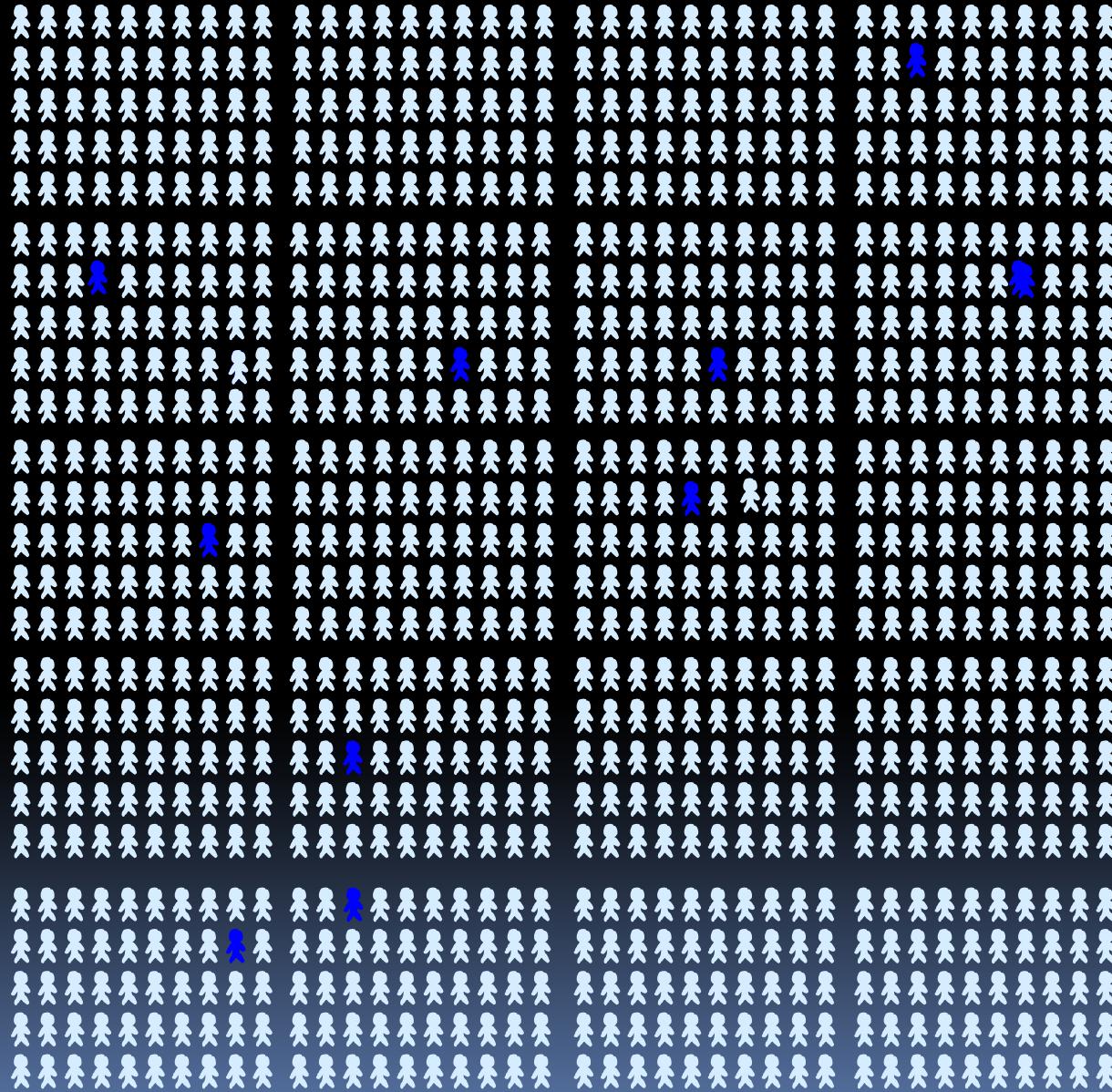


**Fred has size 13**



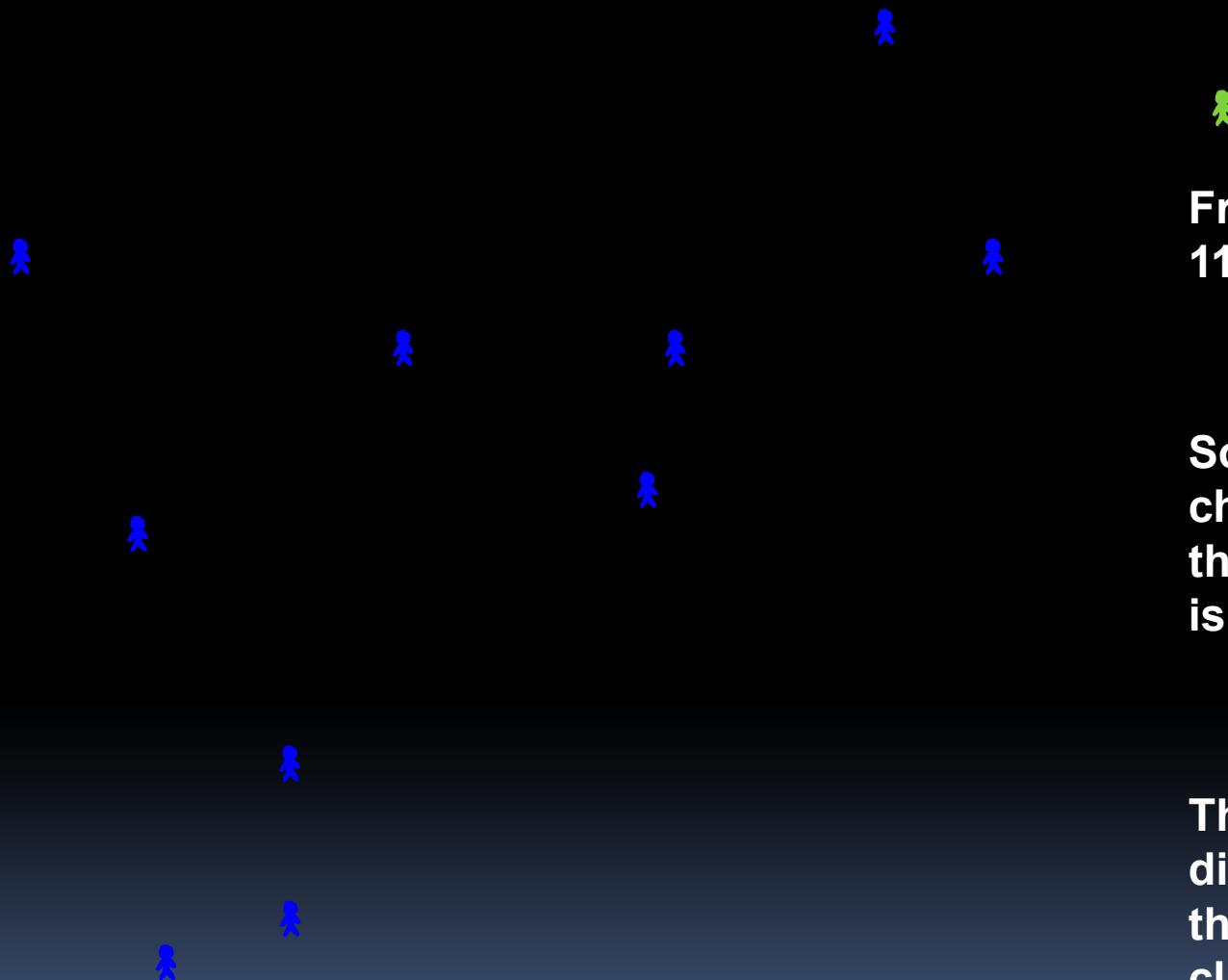
**Fred has size 13**

**Imagine 1,000  
other people  
also at scene**



**Fred has size 13**

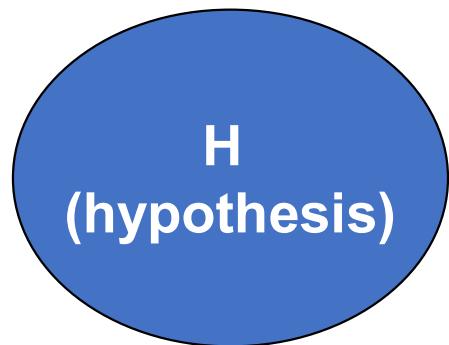
**About 10  
out of the  
1,000 people  
have size 13**



# The Legal Approach = Bayes Theorem (?)

We have a prior probability  $P(H)$  (probability Fred innocent is 1/1001)

We now get some evidence  $E$  (Fred has same shoe size 13 as killer)



Cannot expect  
domain experts to  
engage with this

We want to know the 'posterior' probability

But we know  $P(E|H)$  (= 1/100)

$$P(H|E) = \frac{P(E|H)*P(H)}{P(E)} = \frac{P(E|H)*P(H)}{P(E|H)*P(H) + P(E|\text{not } H)*P(\text{not } H)}$$

$$P(H|E) = \frac{\frac{1}{100} * \frac{1000}{1001}}{\frac{1}{100} * \frac{1000}{1001} + \frac{1}{99} * \frac{1}{10001}} = \frac{0.00999}{0.00999 + 0.000999} \approx 0.91 = 91\%$$

## How the fallacy is also stated

---

“The chances of finding this evidence in an innocent man are so small that you can safely disregard the possibility that this man is innocent”



Ahh.. but DNA evidence is different?

Very low ‘random match’ probabilities

$P(DNA\ match \mid Innocent) = 1/\text{billion}$  ... but same error

‘Low template’ DNA matches have high random match probabilities ...  $P(DNA\ match \mid Innocent) = 1/100$

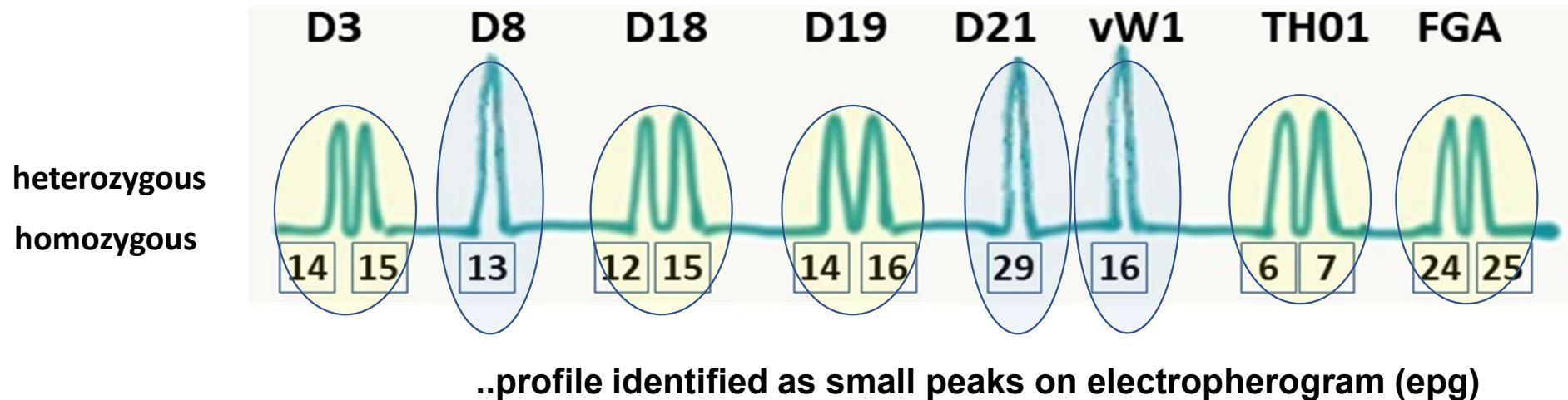
Probability of contamination/testing errors not considered:  
DNA ‘statistics’ are meaningless without this

Principle applies to ALL types of forensic match evidence

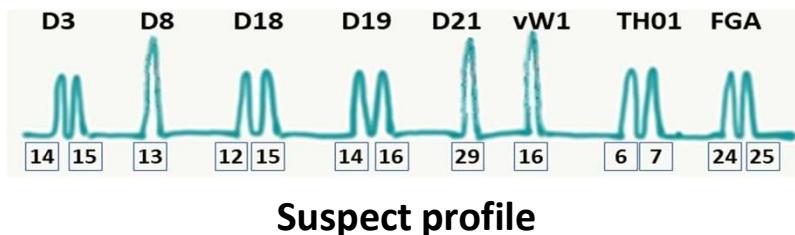
# DNA profiles

Locus	D3	D8	D18	D19	D21	vW1	TH01	FGA	...
Allele	14,15	13,13	12,15	14,16	29,29	16,16	6,7	24,25	...

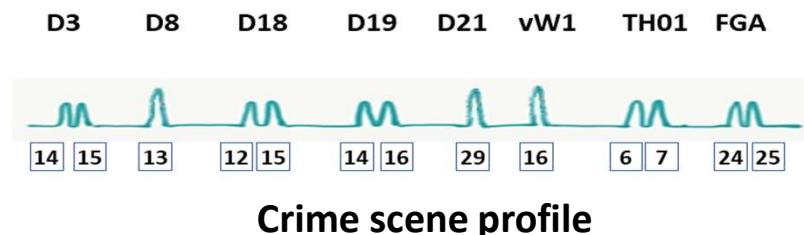
An individual DNA profile



# DNA match evidence



Suspect profile



Crime scene profile

Each genotype at each locus has a population frequency that is estimated from a DNA database (typically genotypes are prevalent in between 5% and 20% of the population).

Forensic experts assume the loci are ‘independent’ and so compute a ‘random match probability’ of an unrelated person having the same matching profile by multiplying the loci probabilities.

So, e.g. if each genotype was prevalent in 10% of the population and we had a match on 9 loci then the rmp is  $1/10^9$  which is one in a billion, i.e.

“The probability of finding this evidence in a person unrelated to the suspect is 1 in a billion”

The error can be very  
tricky to spot....

---

“The chances that this DNA comes from another person (unrelated to the defendant) is less than one in a billion”



# Tip of the Fallacies Iceberg

- Defendant fallacy
- Confirmation bias fallacy
- Base rate neglect
- Treating dependent evidence as independent
- Coincidences fallacy
- Various evidence utility fallacies
- Cross admissibility fallacy
- ‘Crimewatch UK’ fallacy

Fenton, N.E. and Neil, M., 'Avoiding Legal Fallacies in Practice Using Bayesian Networks', Australian Journal of Legal Philosophy 36, 114-151, 2011

# The probative value of evidence

'Probative value':

*When does evidence  $E$  “support” hypothesis  $H$ ?*

**when the probability of  $H$  being true increases after we find  $E$**

i.e.  $P(H | E) > P(H)$

**(or equivalently ‘posterior odds’ of  $H$  increase over the ‘prior odds’ of  $H$ )**

***the bigger the increase the more  $E$  supports  $H$***

**Example:**  $H$  ‘coin is double-headed’.  $E$ : each of 5 tosses is a Head.

If  $P(H)=0.5$  then by Bayes  $P(H|E)=32/33$

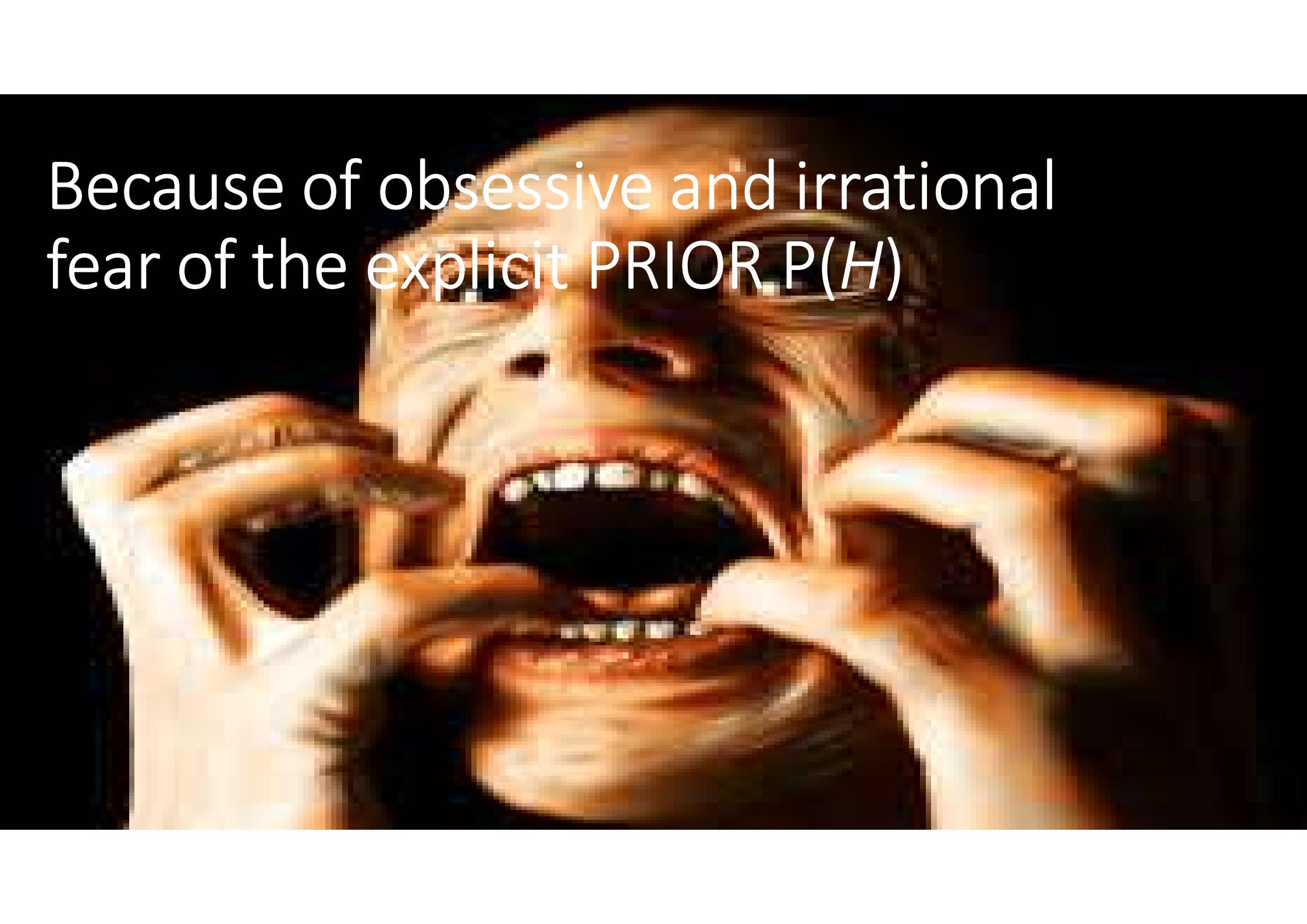
As an exercise in Bayes Theorem  
you should check this!

Hence: A simple formal definition of *probative value* of evidence

The ratio  $R$ : 
$$\frac{P(H|E)}{P(H)}$$

- $R > 1$  means  $E$  supports  $H$  (*and the bigger  $R$  is the more  $E$  supports  $H$* )
- $R < 1$  means  $E$  supports *not*  $H$
- $R = 1$  means  $E$  is **neutral** for  $H$

Why do we never see this definition used?



Because of obsessive and irrational  
fear of the explicit PRIOR  $P(H)$

# DNA evidence likelihood ratios

DNA sample from  
crime scene (this  
is the evidence E)

Locus	D3S1358	vWA	D16S539
Alleles found	15, 16	17, 17	9, 11
Suspect alleles	15, 16	17, 17	9, 11

DNA from suspect

We have a ‘match’

Suppose  $H$  is the prosecution hypothesis: “the crime scene DNA comes from suspect”

$$LR = \frac{\text{Probability of } E \text{ given } H}{\text{Probability of } E \text{ given not } H} = \frac{1}{1/1000} = 1000$$

*This is an estimate based on many assumptions*

If we find a ‘match’ on 14 loci then the LR is approximately  $10^{14}$

Even if we assume a prior as low as  $1/(10 \text{ billion})$  it follows from Bayes Theorem that the posterior probability of  $H$  is very close to 1 (odds are  $10^5$  to 1 in favour of  $H$ )

# The Likelihood ratio (LR)

**Probability of evidence if prosecution hypothesis is true**

**Probability of evidence if defence hypothesis is true**

$$\frac{P(E | H_p)}{P(E | H_d)}$$

For our DNA example the numerator is assumed to be 1 and the denominator is  $1/(1 \text{ billion})$

*So LR = 1 billion*



This evidence very strongly supports the prosecution hypothesis. In fact it is a billion times more likely under the prosecution hypothesis than the under the defence hypothesis

# Likelihood Ratio (LR) as a measure of probative value

Bayes Theorem:  
Posterior odds of  $H$  = LR x Prior odds of  $H$

$$\frac{\text{Posterior odds}}{\text{Prior odds}} = \text{LR} = \frac{P(H|E)}{P(not H|E)} = \frac{P(E|H)}{P(E|not H)} \times \frac{P(H)}{P(not H)}$$

# Likelihood Ratio (LR) as a measure of probative value

Bayes Theorem:  
Posterior odds of  $H = LR \times$  Prior odds of  $H$

**LR > 1:** means  $E$  supports prosecution hypothesis

(as  $P(H|E) > P(H)$  in this case)

**LR < 1:** means  $E$  supports defence hypothesis

(as  $P(\text{not } H|E) > P(\text{not } H)$  in this case)

**LR = 1:** means  $E$  has no probative value

(as  $P(H|E) = P(H)$  in this case)

*So it seems to work in  
exactly the same way as  
'my' ratio  $P(H|E)/P(H)$*

# Likelihood Ratio Example

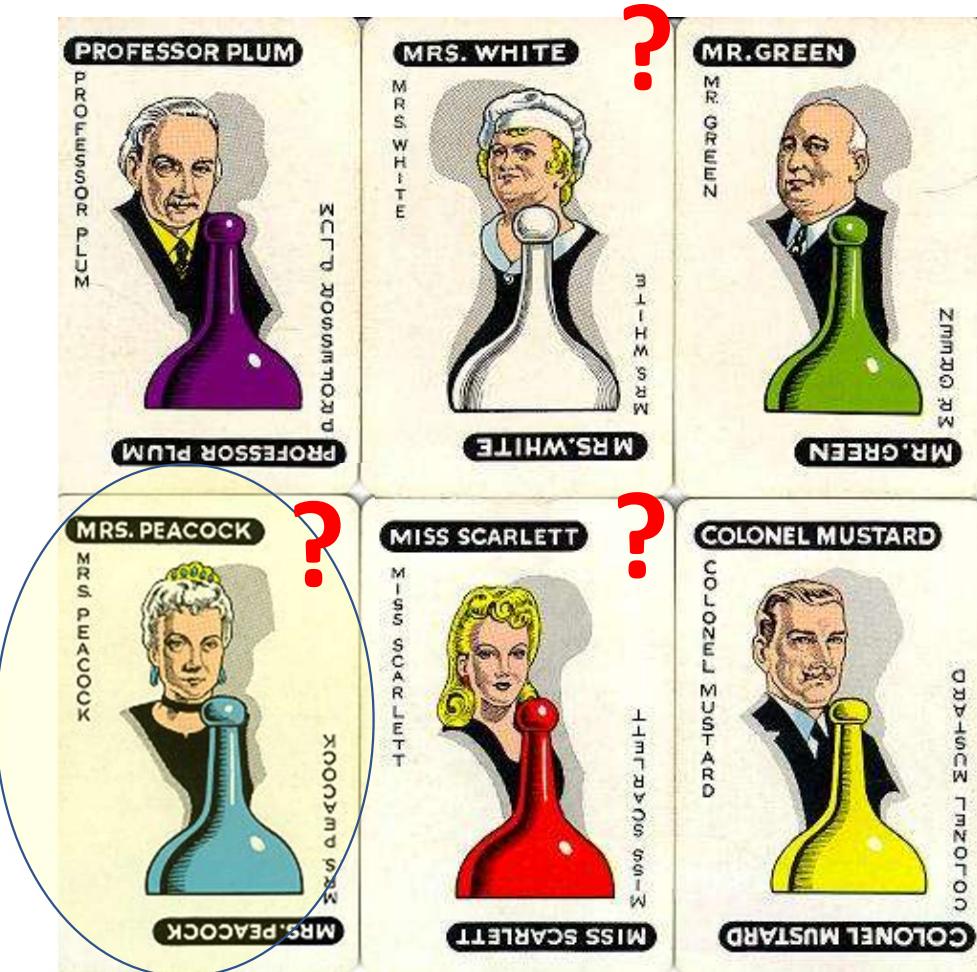
	Prior odds	Likelihood ratio	Posterior Odds
Prosecutor	1	100	1
Defence	1000	1	10

---

Prosecutor	1	100	25
Defence	4	1	1

So the LR is the same but different priors for  $H$  produce 'opposite' outcomes

# Example: Cluedo Who Dunnit?



$H$ : "Mrs Peacock guilty"

$E$ : "The murderer was female"

*What is the LR of  $H$  against not  $H$ ?  
(menti.com)*

$$P(E | H) = 1$$

$$P(E | \text{not } H) = 2/5$$

$$\text{LR} = 5/2 = 2.5$$

(so definitely probative in support of  $H$ )

In this case we know the prior  $P(H)=1/6$ , i.e. odds are 1 to 5  
So the posterior odds are:

$$\frac{1}{5} \times \frac{5}{2} = \frac{1}{2}$$
 i.e. odds of 1 to 2, which is a probability of 1/3

# Likelihood Ratio: The Good

Simple formula for probative value of evidence

No need to explicitly consider prior for  $H$

Forces forensic experts to consider the likelihood of both the prosecution hypothesis and the defence hypothesis

*...but none of these are exactly what they seem in practice*

# The limitations of the Likelihood ratio

# Likelihood Ratio: The Bad (part 1)

The LR only works as a measure of probative value when the defence hypothesis is the negation of the prosecution hypothesis (i.e. the hypotheses are “mutually exclusive and exhaustive”)

## Likelihood Ratio: The Bad (part 2)

If the hypotheses are not mutually exclusive and exhaustive then it is possible that

- $LR > 1$  but the evidence supports the defence hypothesis
- $LR = 1$  but the evidence is still probative

# LR>1 ....but the evidence supports the defence hypothesis

A lottery has 10 tickets numbered 1 to 10

Fred buys 3 tickets and gets numbers 3, 4 and 5.

Jane buys 2 tickets and gets numbers 1 and 6

The winning ticket is drawn but is blown away in the wind. However, a totally reliable eye-witness asserts that the winning ticket was a number between 4 and 10.

Fred claims he must have won and sues the organisers.

The prosecution hypothesis  $H_f$  is “Fred won the raffle” (i.e  $H_f$  : “winning ticket was 3, 4, or 5”).

Fred’s lawyer provides the following argument to support the claim:

We have two alternative hypotheses. Either Fred won the lottery ( $H_f$ ) or Jane won the lottery ( $H_j$ ).

We have the evidence E that the winning ticket was a number between 4 and 10.

$P(E | H_f) = 2/3$  because if Fred won then there is a 2/3 chance the winning number was 4 or 5

$P(E | H_j) = 1/2$  because if Jane won then there is a 1/2 chance the winning number was 6

Hence, the LR is 2/3 divided by 1/2 which is equal to 4/3. As the  $LR > 1$ , the evidence supports  $H_f$

BUT: While the evidence supports  $H_f$  over  $H_j$  it does NOT support  $H_f$

The defence hypothesis is    *not*  $H_f$    “winning ticket = 1,2,6,7,8,9, or 10”

$P(E | \text{not } H_f) = 5/7$

Hence, LR of  $H_f$  against *not*  $H_f$  is 2/3 divided by 5/7 which is equal to 14/15. As the  $LR < 1$  the evidence supports *not*  $H_f$

**The probability of  $H_f$  drops from a prior of 0.3 to a posterior of 0.286 after getting the evidence E**

# LR=1 ...but the evidence is probative

A lottery has 10 tickets numbered 1 to 10

Fred buys 3 tickets and gets numbers 3, 4 and 5.

Jane buys 3 tickets and gets numbers 1, 2 and 6

The winning ticket is drawn but is blown away in the wind. However, a totally reliable eye-witness asserts that the winning ticket was less than 7.

Fred claims he must have won and sues the organisers, arguing that the evidence supports  $H_f$  is “Fred won the raffle”

This time the Defence lawyer argues the evidence provides no probative value to support  $H_f$  as follows

We have two alternative hypotheses. Either Fred won the lottery ( $H_f$ ) or Jane won the lottery ( $H_j$ ).

We have the evidence E that the winning ticket was a number less than 7.

$P(E | H_f) = 1$  because if Fred won then it is certain the winning number was less than 7

$P(E | H_j) = 1$  because if Jane won then it is certain the winning number was less than 7

Hence the  $LR = 1$  proving the evidence has no probative value

BUT: While the evidence provides no support for  $H_f$  over  $H_j$  it **does** support  $H_f$

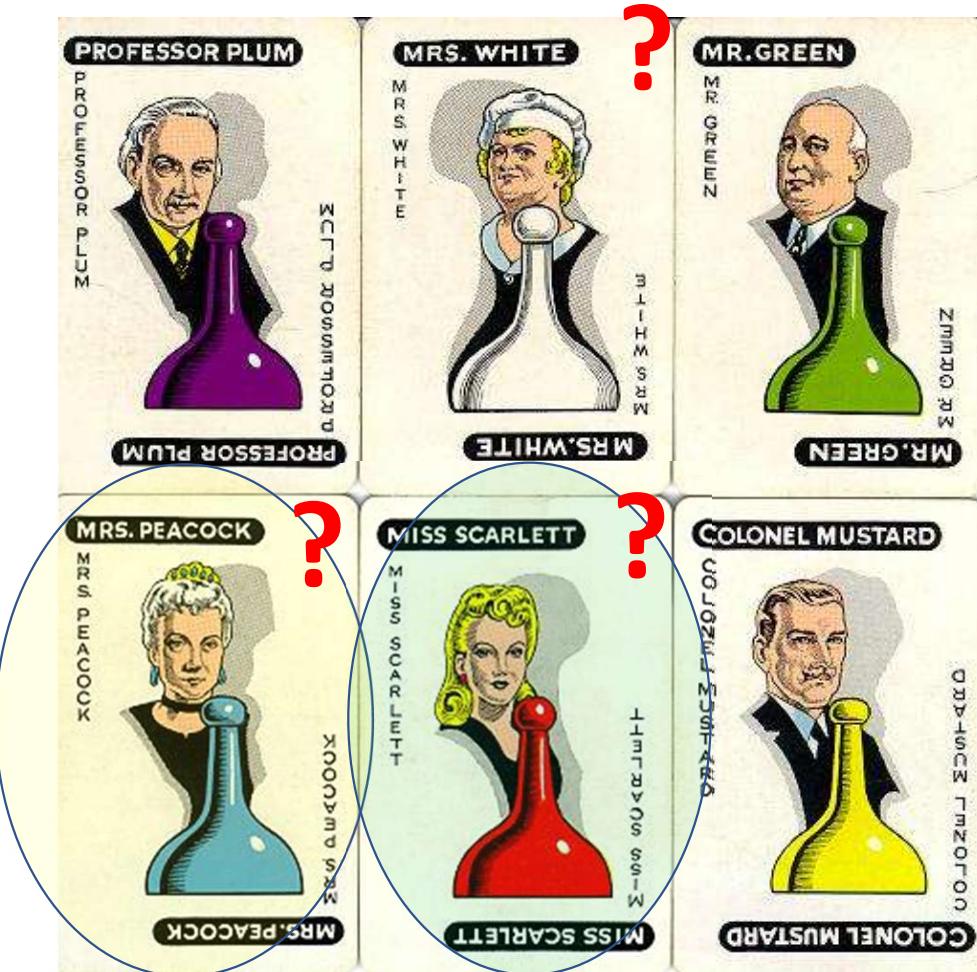
$not H_f$  is the hypothesis “winning ticket = 1,2,6,7,8,9, or 10”

so  $P(E | not H_f) = 3/7$

Hence, LR of  $H_f$   $not H_f$  is 1 divided by 3/7 which is equal to 7/3. As the  $LR > 1$  the evidence supports  $H_f$

**The probability of  $H_f$  increases from a prior of 0.3 to a posterior of 0.5 after getting the evidence E**

# Example: Cluedo Who Dunnit? (cherry-picking non- exhaustive hypotheses)



$H$ : "Mrs Peacock guilty"

$E$ : "The murderer was female"

But now....

...defence hypothesis is:

$H_d$ : "Miss Scarlet guilty"

Menti: what is the LR?

$P(E | H) = 1$  (as before)

$P(E | H_d) = 1$

$LR=1$  ("no probative value in support of  $H$ ")

but although the posterior odds of  $H$  against  $H_d$  are unchanged the posterior probability of  $H$  has increased from  $1/6$  to  $1/3$  as before

# Likelihood Ratio: The Bad (in DNA evidence)

The LR is used routinely in practice when presenting DNA evidence as follows:

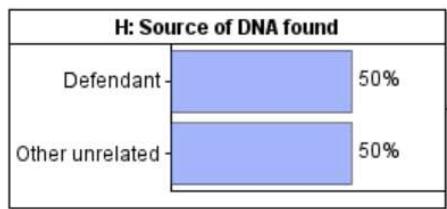
Prosecution hypothesis: “*DNA from crime scene belongs to the defendant*”

Defence hypothesis: “*DNA from crime scene belongs to a person unrelated to the defendant*”

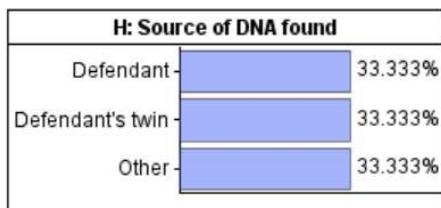
***These hypotheses are NOT exhaustive***

# Example: Twins (assumes 1/1000 DNA rmp)

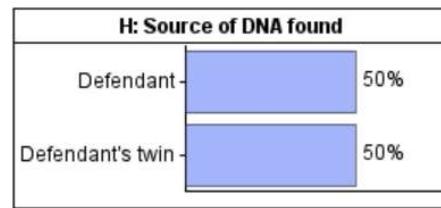
## Prior



**E: DNA found matches defendant**



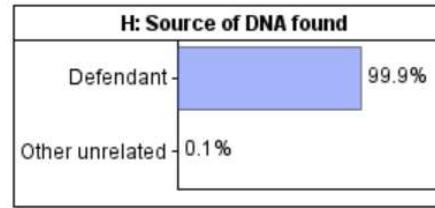
**E: DNA found matches defendant**



**E: DNA found matches defendant**

## Posterior

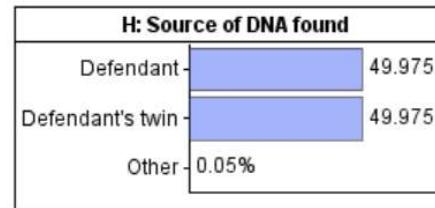
### Highly probative



**E: DNA found matches defendant**

Scenario 1 : True

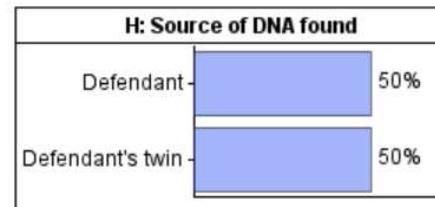
No probative value over twin but reasonably probative



**E: DNA found matches defendant**

Scenario 1 : True

No probative value



**E: DNA found matches defendant**

Scenario 1 : True

## How convincing is a high LR for non-exhaustive hypotheses?



**Evidence E:**  
**I come home to find  
a broken vase on the  
floor, but nothing  
else fallen/broken**



H<sub>p</sub>: “Burglars broke in”  
H<sub>d</sub>: “Earthquake”

$$P(E | H_p) = 1/100$$
$$P(E | H_d) = 1/100,000$$

Critical ambiguity

$$\text{So } LR = 1000$$

“The evidence strongly supports the burglary hypothesis because **it is** 1000 times more likely than the alternative”

***Demonstrates danger of LR  
for non-exhaustive hypotheses***

# Suspect accused of involvement in a bombing

The evidence:

*Suspect tests positive for explosive substances on hands (this test is 95% accurate when explosive substances present)*



*But the test also has 95% chance of positive results if suspect handled playing cards*

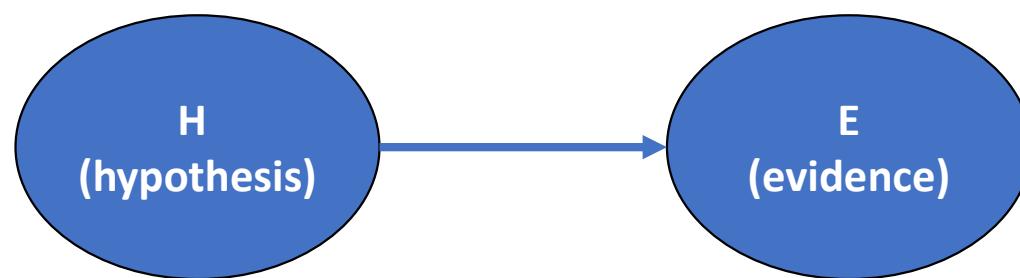
Is the evidence probative?

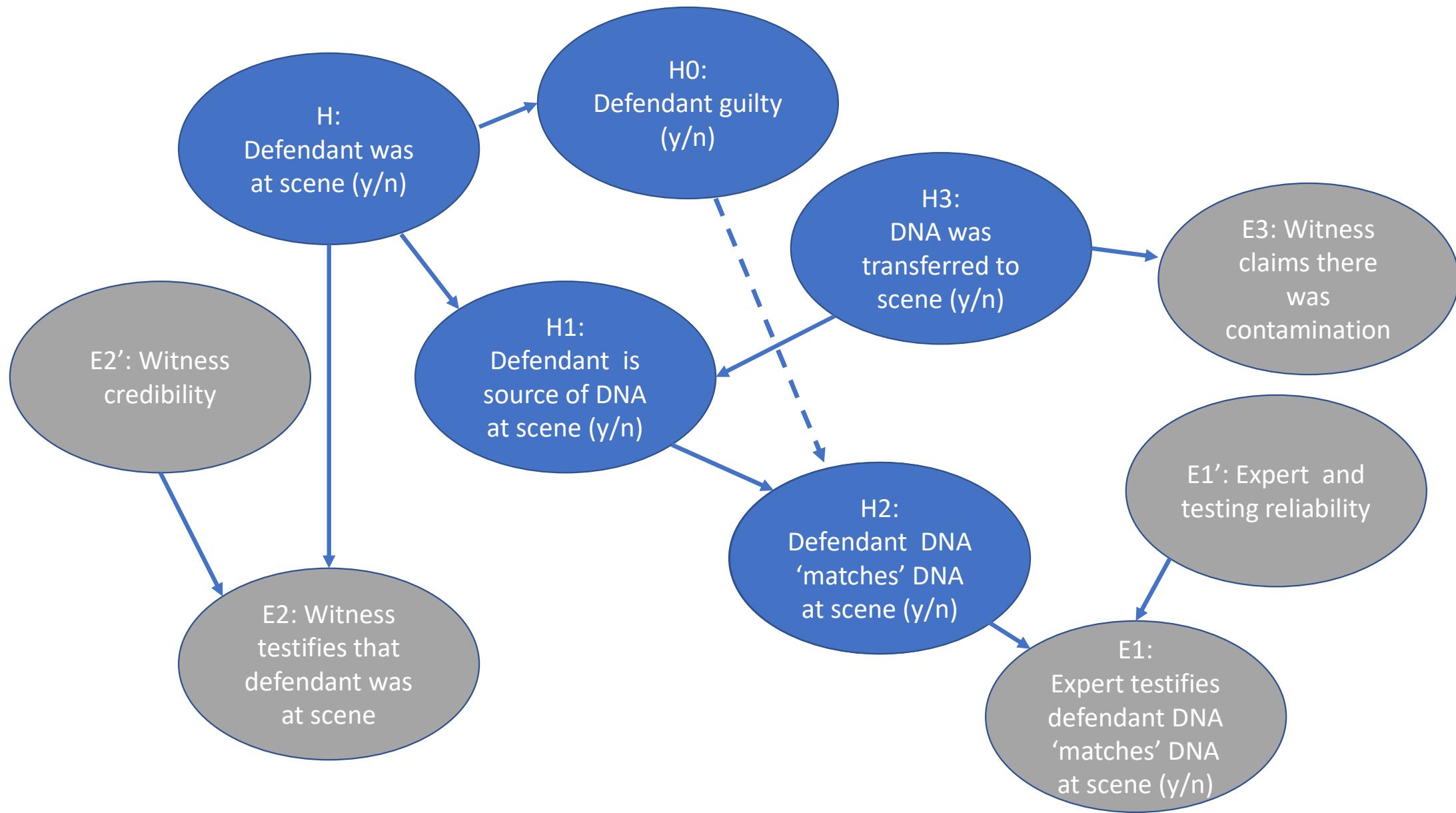


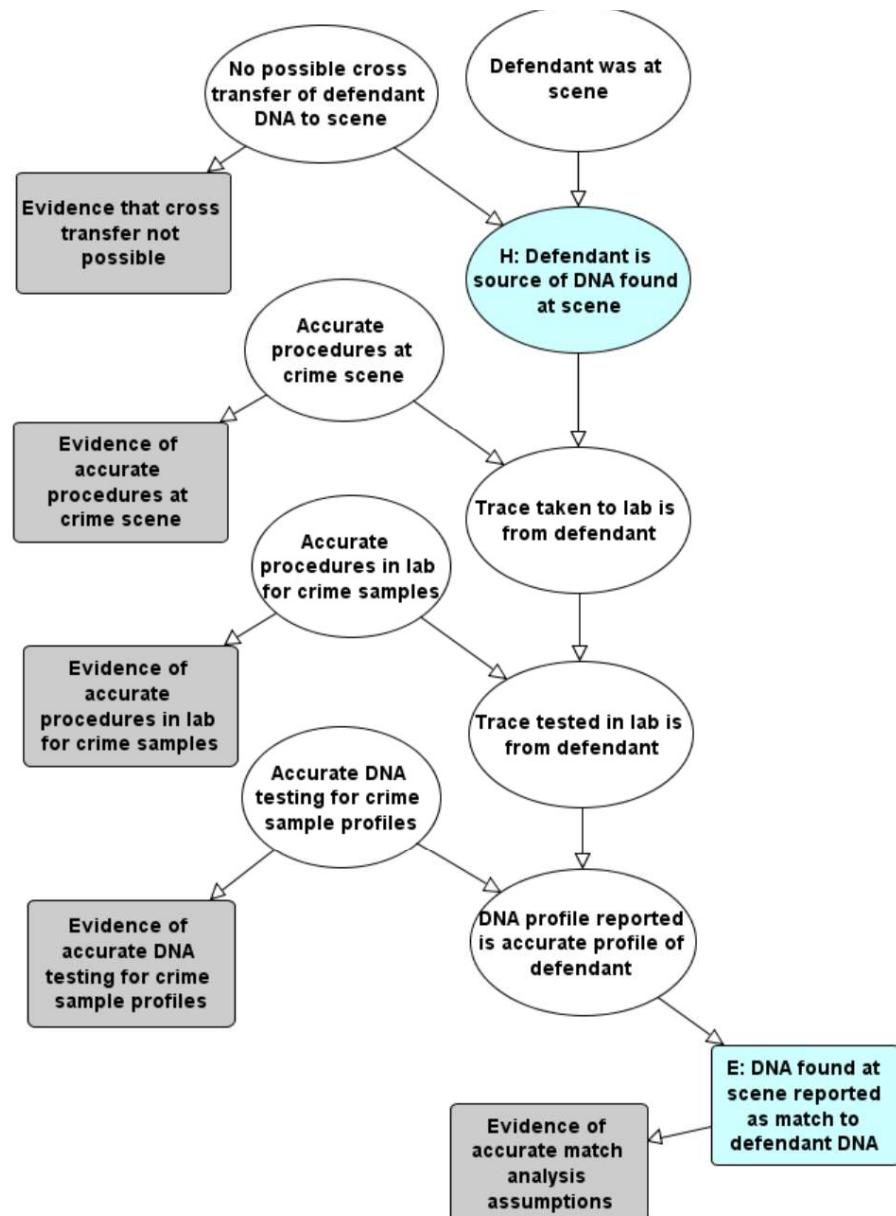
The need for Causal  
Models (Bayesian  
networks) in evidence  
evaluation

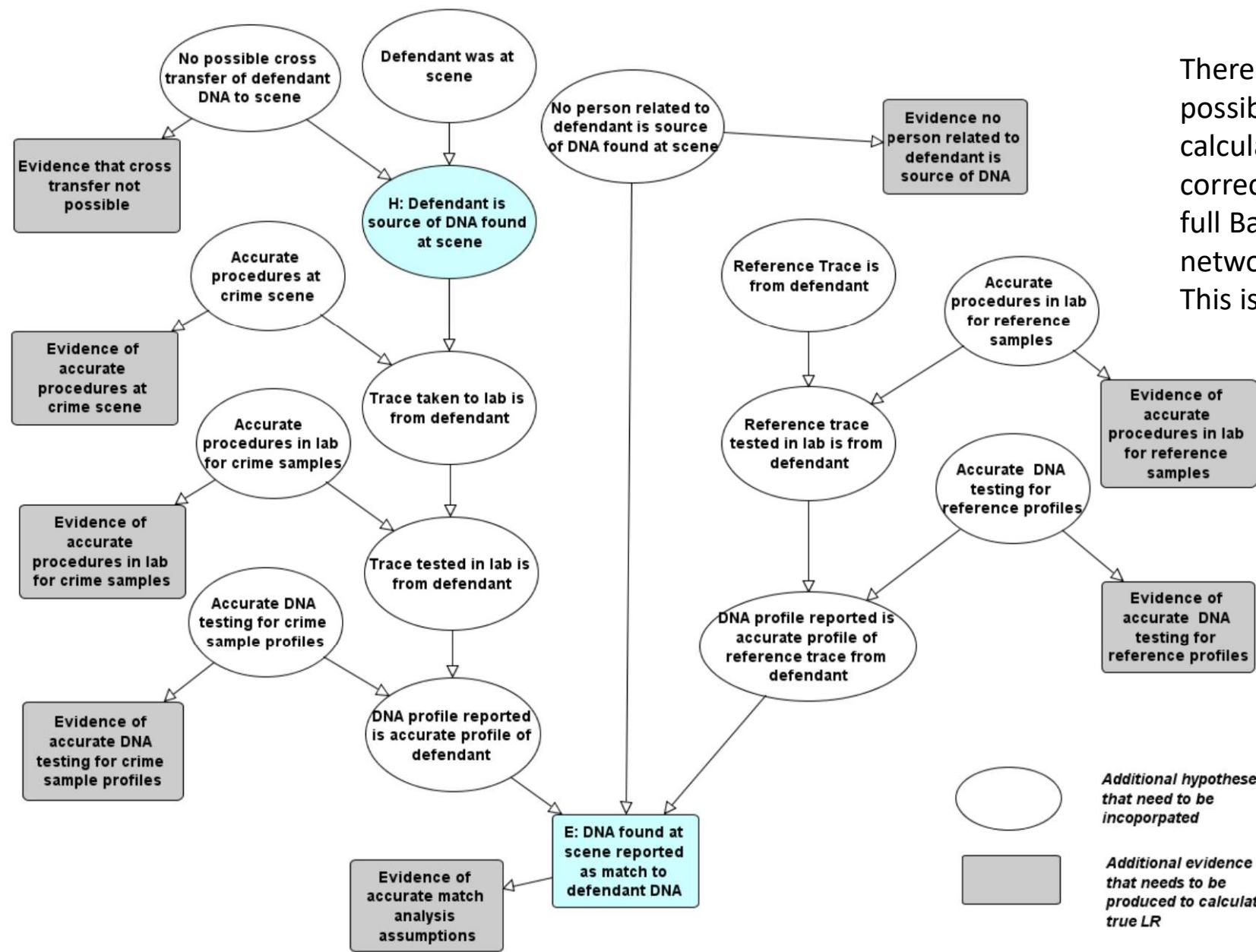
## **ISSUE 3: The Likelihood Ratio is normally not calculated correctly**

Wrongly encourages experts to over-simplify the evidence by combining multiple hypotheses into a single hypothesis and a single piece of evidence









There is no possibility of calculating the correct LR without full Bayesian network inference. This is NEVER done

*Additional hypotheses that need to be incorporated*

*Additional evidence that needs to be produced to calculate true LR*

<https://arxiv.org/abs/2106.05328>



Cornell University

the

arXiv.org > stat > arXiv:2106.05328

Search...  
Help | Advanced

Statistics > Applications

[Submitted on 9 Jun 2021]

## Calculating the Likelihood Ratio for Multiple Pieces of Evidence

Norman Fenton, Martin Neil

When presenting forensic evidence, such as a DNA match, experts often use the Likelihood ratio (LR) to explain the impact of evidence . The LR measures the probative value of the evidence with respect to a single hypothesis such as 'DNA comes from the suspect', and is defined as the probability of the evidence if the hypothesis is true divided by the probability of the evidence if the hypothesis is false. The LR is a valid measure of probative value because, by Bayes Theorem, the higher the LR is, the more our belief in the probability the hypothesis is true increases after observing the evidence. The LR is popular because it measures the probative value of evidence without having to make any explicit assumptions about the prior probability of the hypothesis. However, whereas the LR can in principle be easily calculated for a distinct single piece of evidence that relates directly to a specific hypothesis, in most realistic situations 'the evidence' is made up of multiple dependent components that impact multiple different hypotheses. In such situations the LR cannot be calculated . However, once the multiple pieces of evidence and hypotheses are modelled as a causal Bayesian network (BN), any relevant LR can be automatically derived using any BN software application.

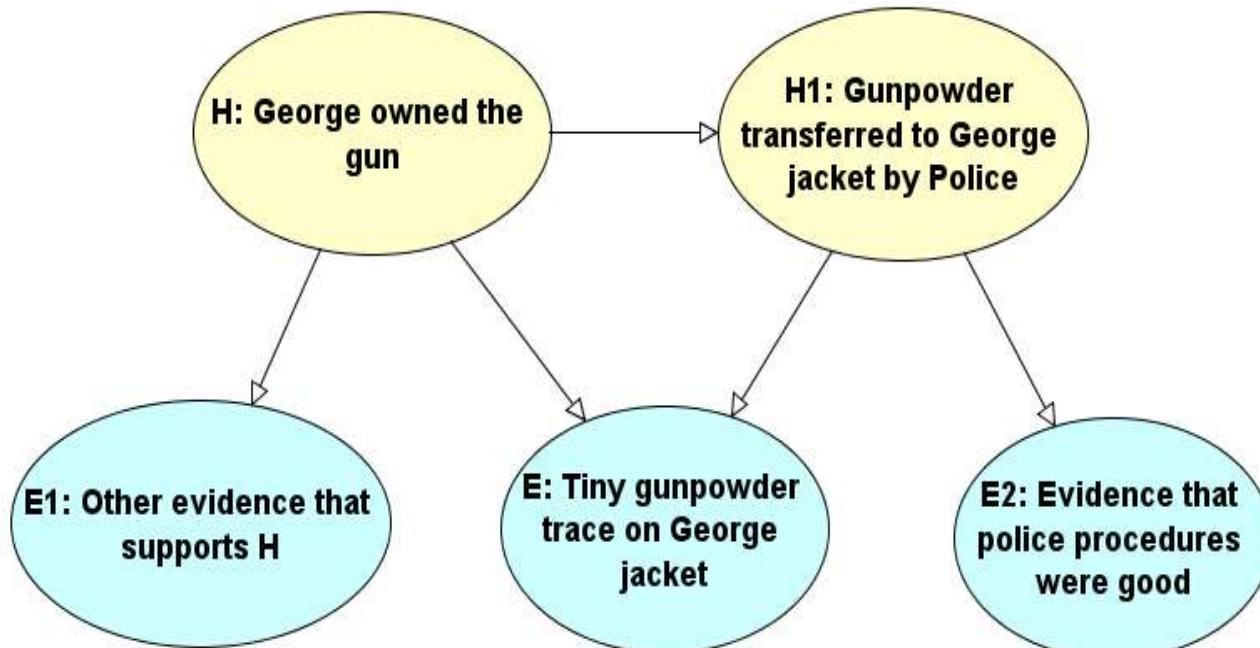
Comments: 27 pages, 12 figures

Subjects: [Applications \(stat.AP\)](#)

Cite as: [arXiv:2106.05328 \[stat.AP\]](#)

(or [arXiv:2106.05328v1 \[stat.AP\]](#) for this version)

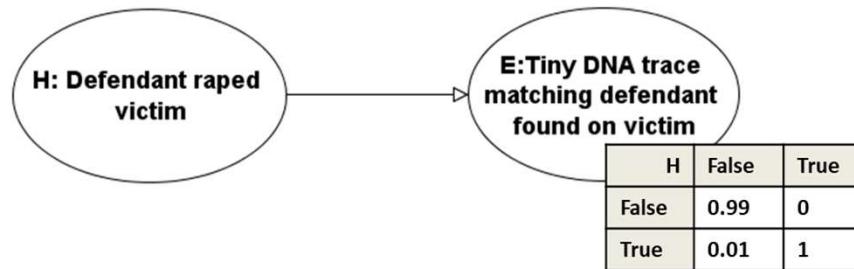
# The Barry George case



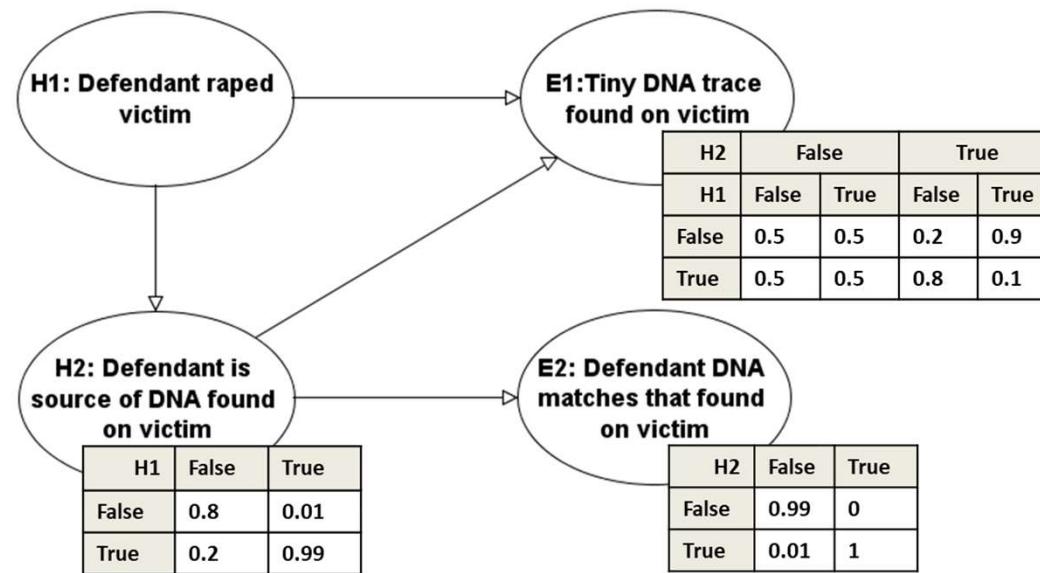
Fenton, N. E., D. Berger, D. Lagnado, M. Neil and A. Hsu, (2014). "When 'neutral' evidence still has probative value (with implications from the Barry George Case)", *Science and Justice*, 54(4), 274-287

# Example: Revealing full BN model reverses impact of evidence

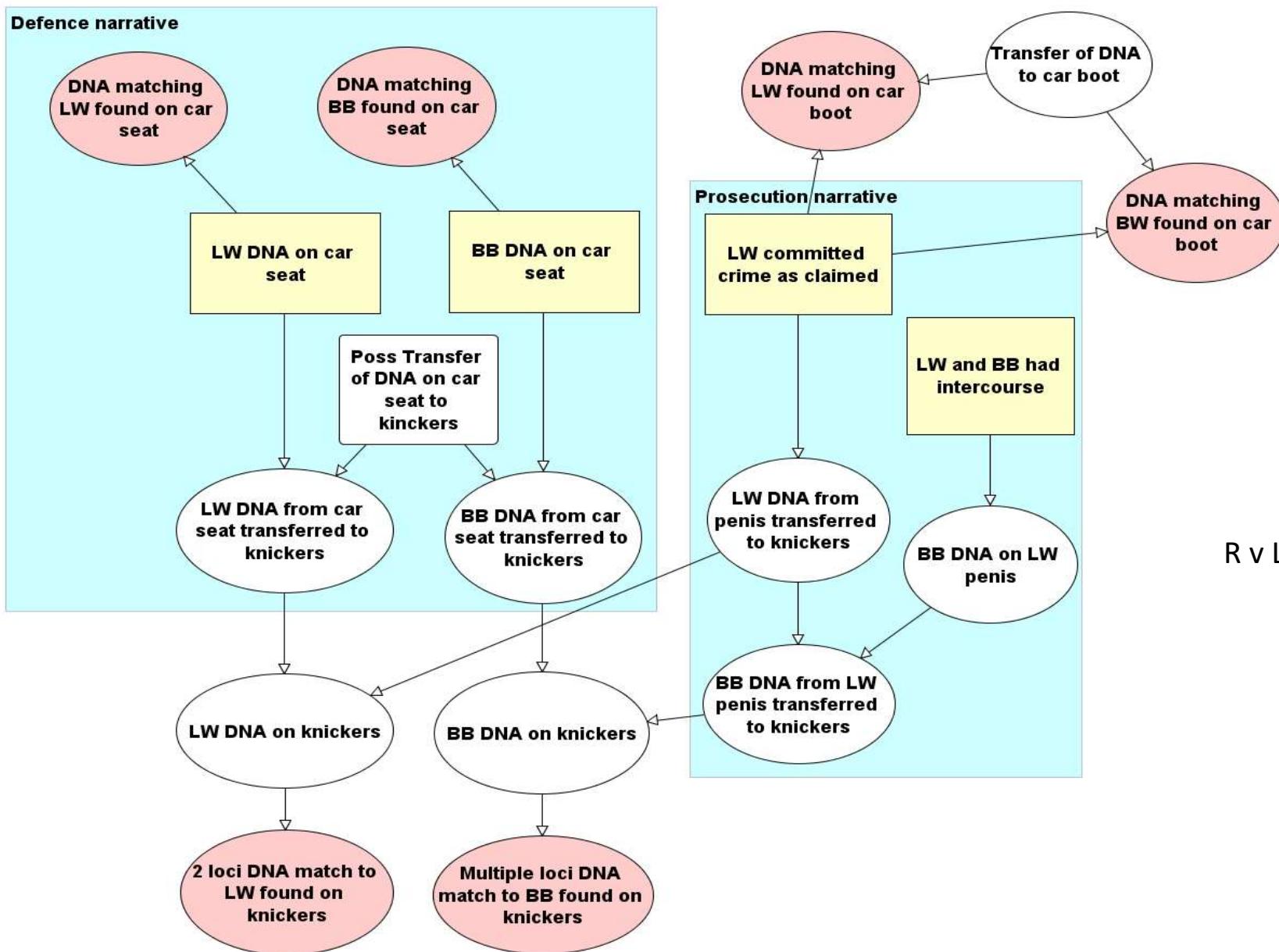
a) Over-simplistic prosecution argument



b) More realistic argument



## R v LW Case



# The case of the Kandinsky painting and Bayesian networks

**Support The Guardian**  
[Contribute →](#) [Subscribe →](#)

**The Guardian**

[Sign in](#)

[News](#) [Opinion](#) [Sport](#) [Culture](#) [Lifestyle](#) [≡](#)

[World](#) ► [Europe](#) [US](#) [Americas](#) [Asia](#) [Australia](#) [Middle East](#) [Africa](#) [Inequality](#) [Cities](#) [More](#)

**Netherlands**

## Dutch art panel's ruling against Jewish family criticised as 'step back'

Committee backed museum in battle over Kandinsky painting obtained before Nazis invaded



gaat naar families die recht zouden hebben oproofkunst in ruil voor een deel van de opbrengst.

Palmer vroeg de Britse wiskunde-hoogleraar Norman Fenton hoe groot de kans is dat alle vijf aannames in het advies over de Kandinsky juist zijn. „Fenton schatte die kans op minder dan 3 procent”, zegt Palmer in een reactie op het advies. „En al die aannames moeten juist zijn, wil het advies overeind blijven.”

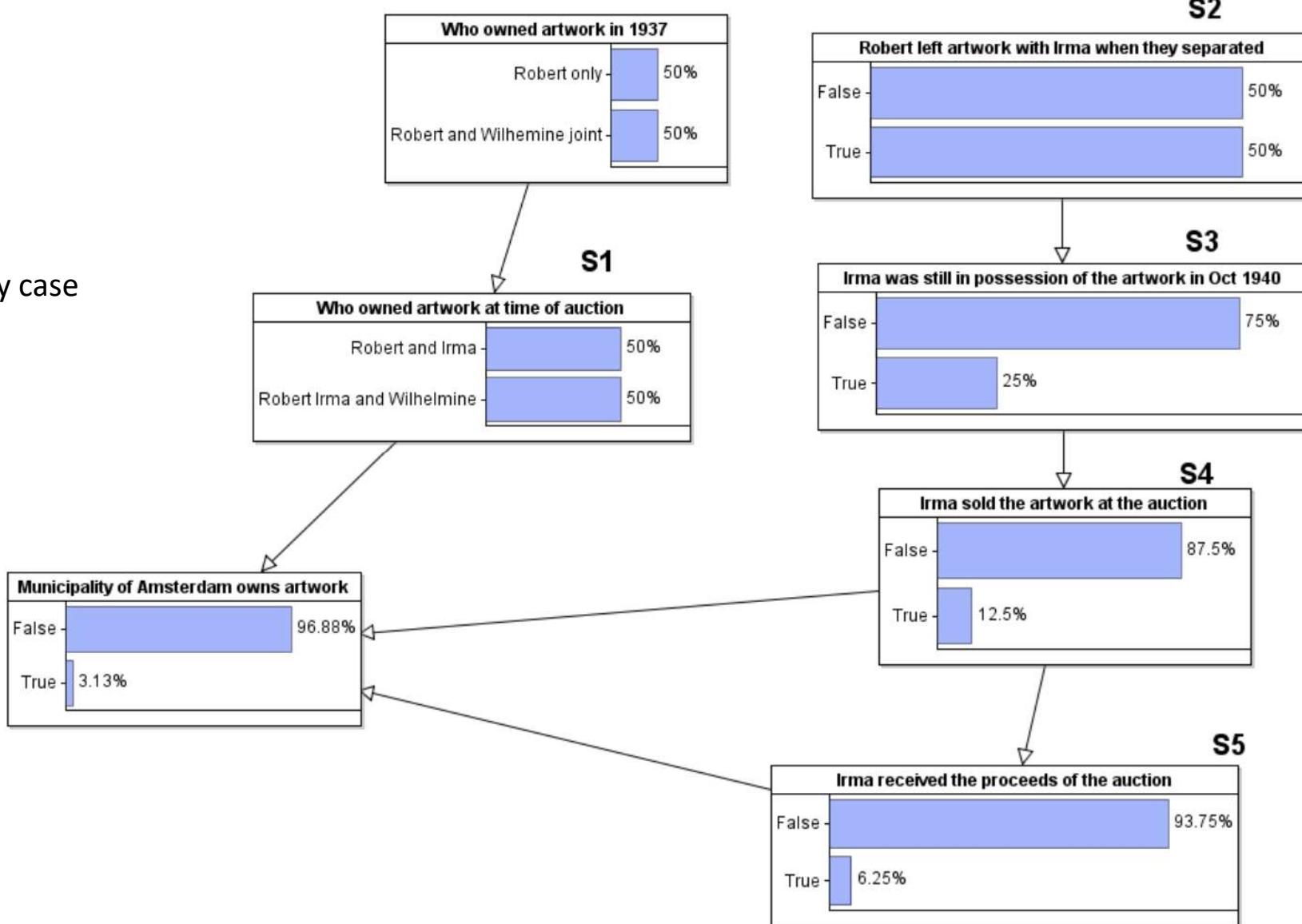
Als de bestuursrechter het advies nietig zou verklaren, kunnen de eisers zich opnieuw tot de Restitutiecommissie wenden.

▲ A detail from Painting with Houses (1909) by Wassily Kandinsky. Photograph: World History Archive/Alamy

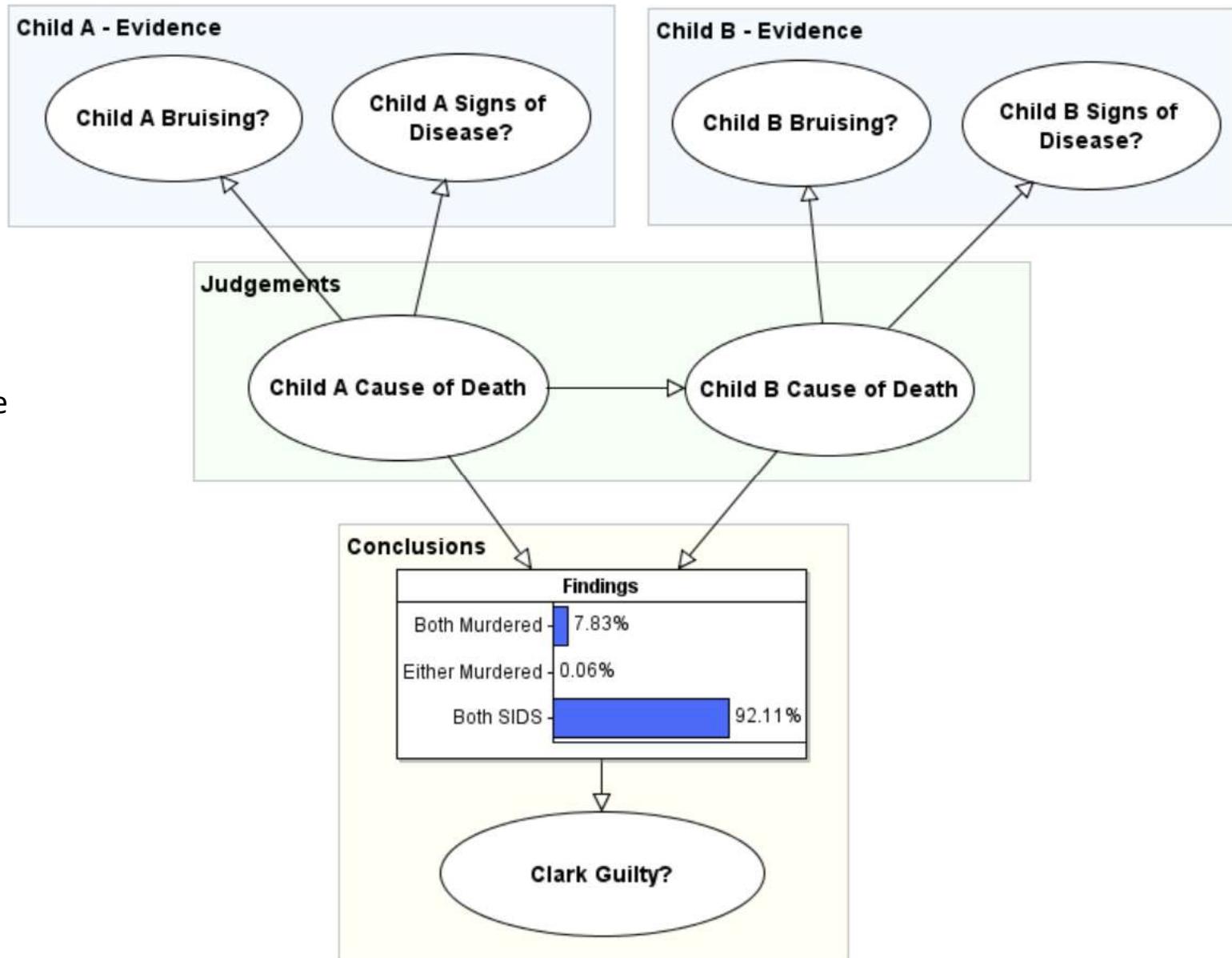
**Daniel Boffey in Brussels**

Wed 5 Dec 2018 05.00 GMT

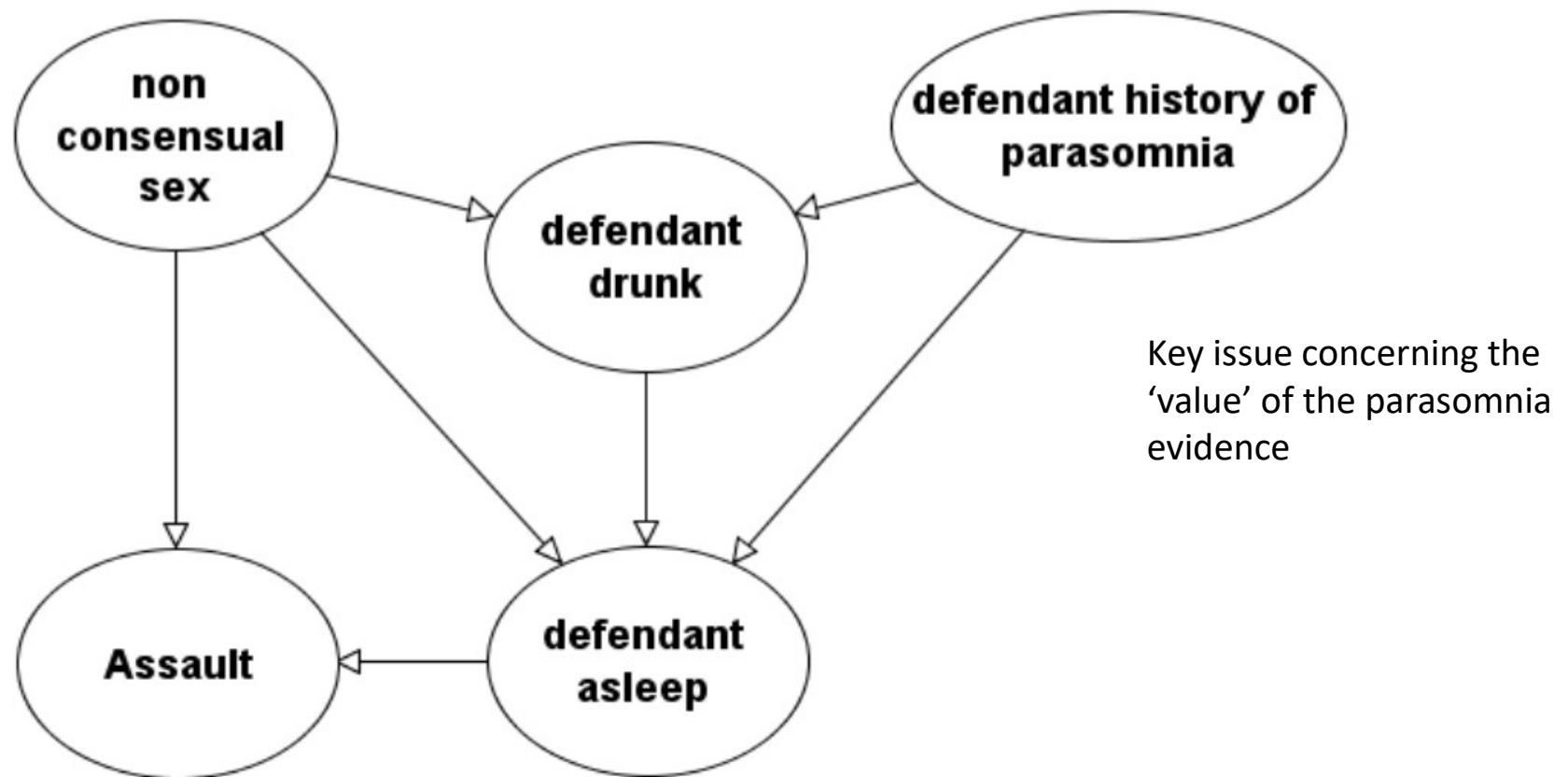
## Kandinsky case



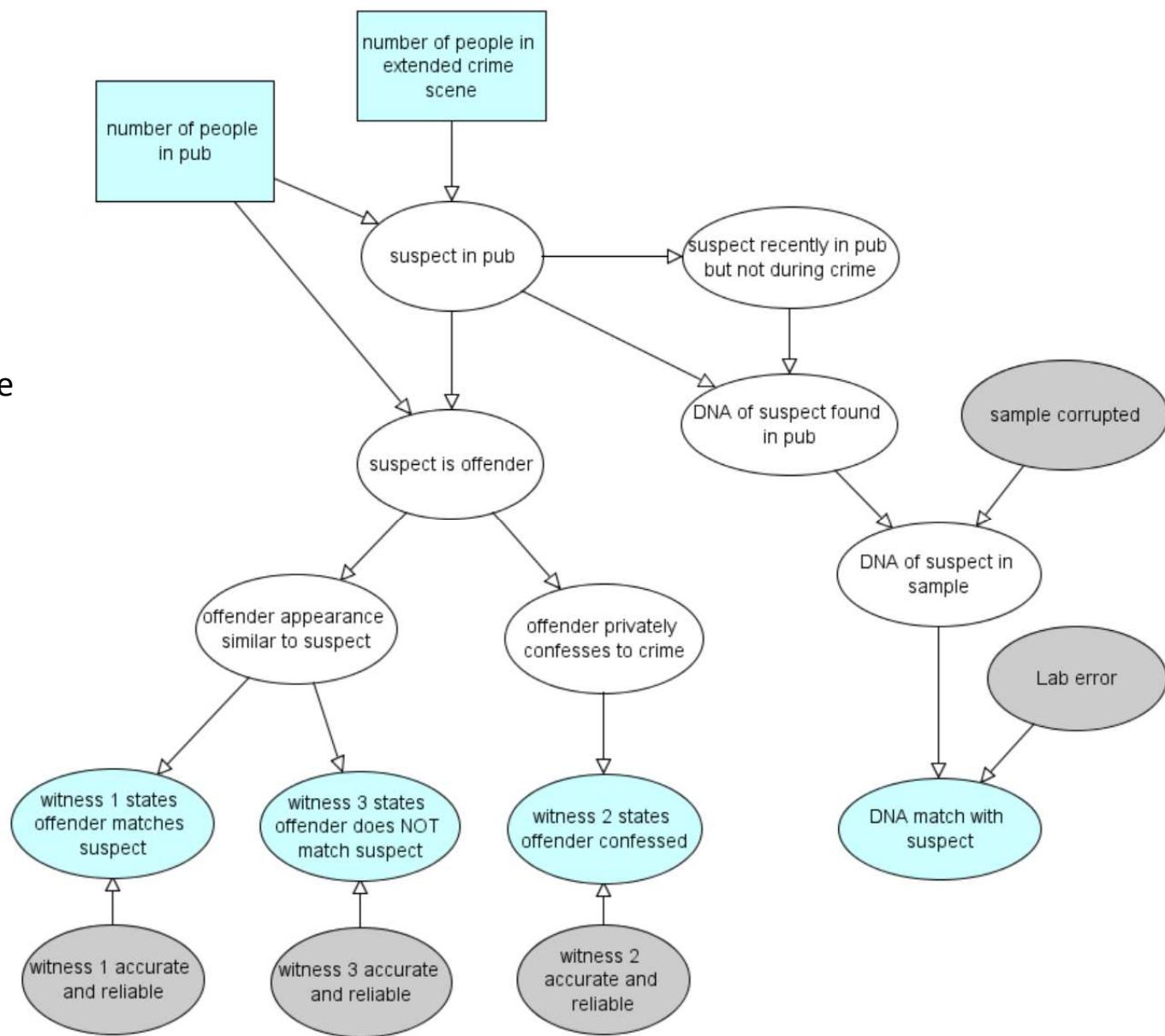
Sally Clark case



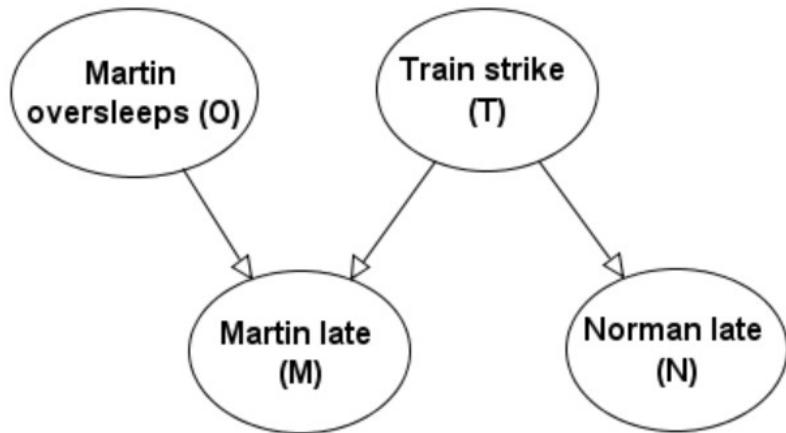
# Recent case: parasomnia



## Example: Handling conflicting evidence



# Do not attempt a formulaic calculation from first principles



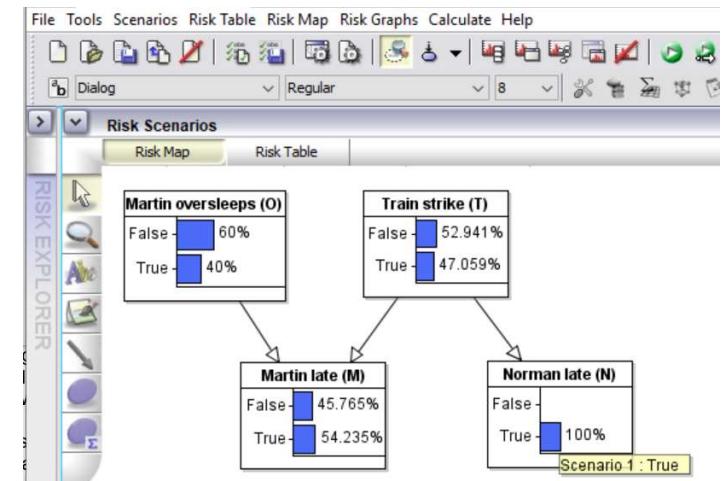
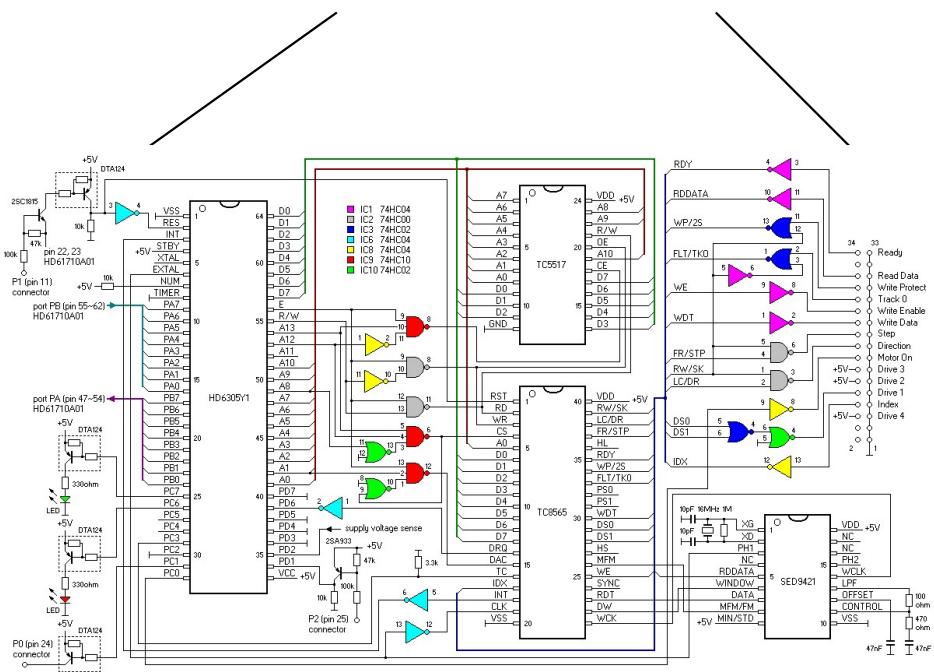
$$\begin{aligned}
 P(M = \text{True}) &= \sum_{O,T} P(M = \text{True} | O, T)P(O)P(T) \\
 &= P(M = \text{True} | O = \text{True}, T = \text{True})P(O = \text{True})P(T = \text{True}) \\
 &\quad + P(M = \text{True} | O = \text{True}, T = \text{False})P(O = \text{True})P(T = \text{False}) \\
 &\quad + P(M = \text{True} | O = \text{False}, T = \text{True})P(O = \text{False})P(T = \text{True}) \\
 &\quad + P(M = \text{True} | O = \text{False}, T = \text{False})P(O = \text{False})P(T = \text{False}) \\
 &= 0.8 \times 0.4 \times 0.1 + 0.6 \times 0.4 \times 0.9 + 0.6 \times 0.6 \times 0.1 + 0.3 \times 0.6 \times 0.9 \\
 &= 0.446
 \end{aligned}$$

$$\begin{aligned}
 P(M = \text{True} | N = \text{True}) &= \sum_{O,T} P(M = \text{True} | O, T)P(O)P(T | N = \text{True})
 \end{aligned}$$

Using the preceding results for  $P(T = \text{True} | N = \text{True})$  and  $P(T = \text{False} | N = \text{True})$  along with we get

$$\begin{aligned}
 &\sum_{O,T} P(M = \text{True} | O, T)P(O)P(T | N = \text{True}) \\
 &= P(M = \text{True} | O = \text{True}, T = \text{True})P(O = \text{True})P(T = \text{True} | N = \text{True}) \\
 &\quad + P(M = \text{True} | O = \text{True}, T = \text{False})P(O = \text{True})P(T = \text{False} | N = \text{True}) \\
 &\quad + P(M = \text{True} | O = \text{False}, T = \text{True})P(O = \text{False})P(T = \text{True} | N = \text{True}) \\
 &\quad + P(M = \text{True} | O = \text{False}, T = \text{False})P(O = \text{False})P(T = \text{False} | N = \text{True}) \\
 &= 0.8 \times 0.4 \times 0.47059 + 0.6 \times 0.4 \times 0.52941 + 0.6 \times 0.6 \times 0.47059 + 0.3 \times 0.6 \times 0.52941 \\
 &= 0.542353
 \end{aligned}$$

# Hence the Calculator Analogy



$$\begin{aligned}
 P(M = \text{True}) &= \sum_{O,T} P(M = \text{True} | O, T)P(O)P(T) \\
 &= P(M = \text{True} | O = \text{True}, T = \text{True})P(O = \text{True})P(T = \text{True}) \\
 &\quad + P(M = \text{True} | O = \text{True}, T = \text{False})P(O = \text{True})P(T = \text{False}) \\
 &\quad + P(M = \text{True} | O = \text{False}, T = \text{True})P(O = \text{False})P(T = \text{True}) \\
 &\quad + P(M = \text{True} | O = \text{False}, T = \text{False})P(O = \text{False})P(T = \text{False}) \\
 &= 0.8 \times 0.4 \times 0.1 + 0.6 \times 0.4 \times 0.9 + 0.6 \times 0.6 \times 0.1 + 0.3 \times 0.6 \times 0.9 \\
 &= 0.446
 \end{aligned}$$

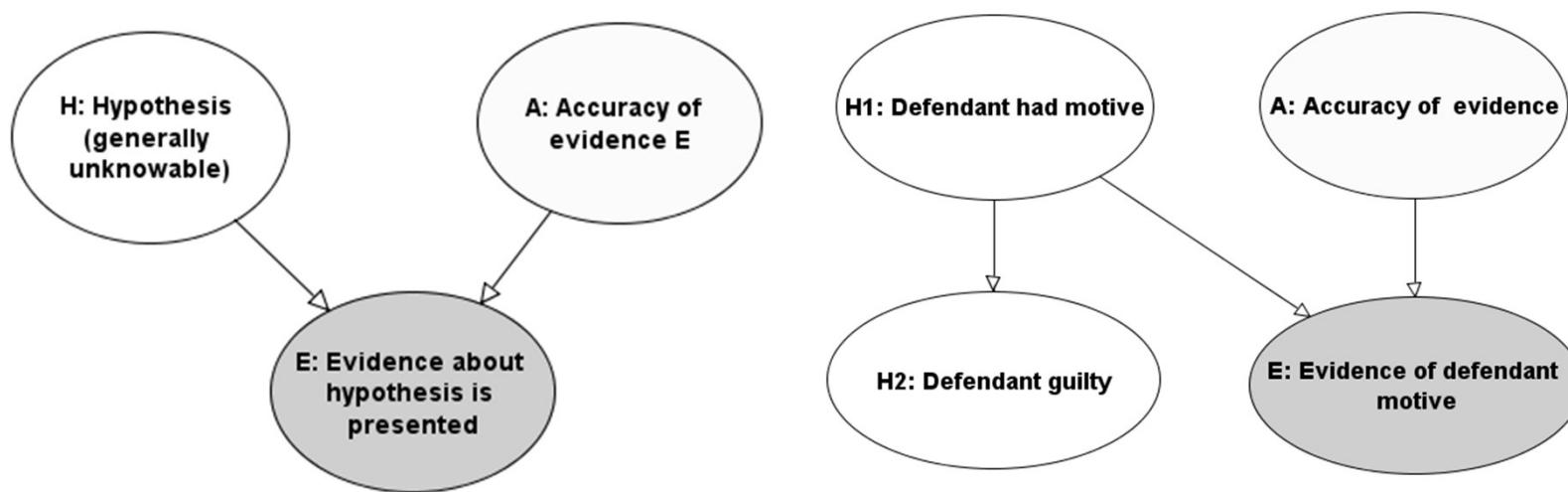
$$\begin{aligned}
 P(M = \text{True} | N = \text{True}) &= \sum_{O,T} P(M = \text{True} | O, T)P(O)P(T | N = \text{True}) \\
 &= P(M = \text{True} | O = \text{True}, T = \text{True})P(O = \text{True})P(T = \text{True} | N = \text{True})
 \end{aligned}$$

Using the preceding results for  $P(T = \text{True} | N = \text{True})$  and  $P(T = \text{False} | N = \text{True})$  along with we get

$$\begin{aligned}
 \sum_{O,T} P(M = \text{True} | O, T)P(O)P(T | N = \text{True}) &= P(M = \text{True} | O = \text{True}, T = \text{True})P(O = \text{True})P(T = \text{True} | N = \text{True}) \\
 &\quad + P(M = \text{True} | O = \text{True}, T = \text{False})P(O = \text{True})P(T = \text{False} | N = \text{True}) \\
 &\quad + P(M = \text{True} | O = \text{False}, T = \text{True})P(O = \text{False})P(T = \text{True} | N = \text{True}) \\
 &\quad + P(M = \text{True} | O = \text{False}, T = \text{False})P(O = \text{False})P(T = \text{False} | N = \text{True}) \\
 &= 0.8 \times 0.4 \times 0.47059 + 0.6 \times 0.4 \times 0.52941 + 0.6 \times 0.6 \times 0.47059 + 0.3 \times 0.6 \times 0.52941 \\
 &= 0.542353
 \end{aligned}$$

# Methods to make building legal BN arguments easier

'idioms' for common argument fragments (accuracy of evidence, motive/opportunity, alibi evidence)



Fenton, N. E., D. Lagnado and M. Neil (2013). "A General Structure for Legal Arguments Using Bayesian Networks." Cognitive Science 37, 61-102

Likelihood ratio: the worst part  
(DNA mixture evidence)

# Likelihood Ratio (The WORST part)

Even if the hypotheses are exhaustive a very high LR tells us nothing about the probability of the prosecution hypothesis being true

This is a little problematic for single profile DNA evidence....

But catastrophic (and not widely understood that it is or why) for mixed profile DNA evidence

# Special problems of ‘low template DNA’

Tiny amounts of DNA can now be ‘analysed’ – very few cells of unspecified tissue

Generally there will be multiple contributors found (more than 2 alleles at loci)

Difficult to determine how many contributors and to isolate a single contributor

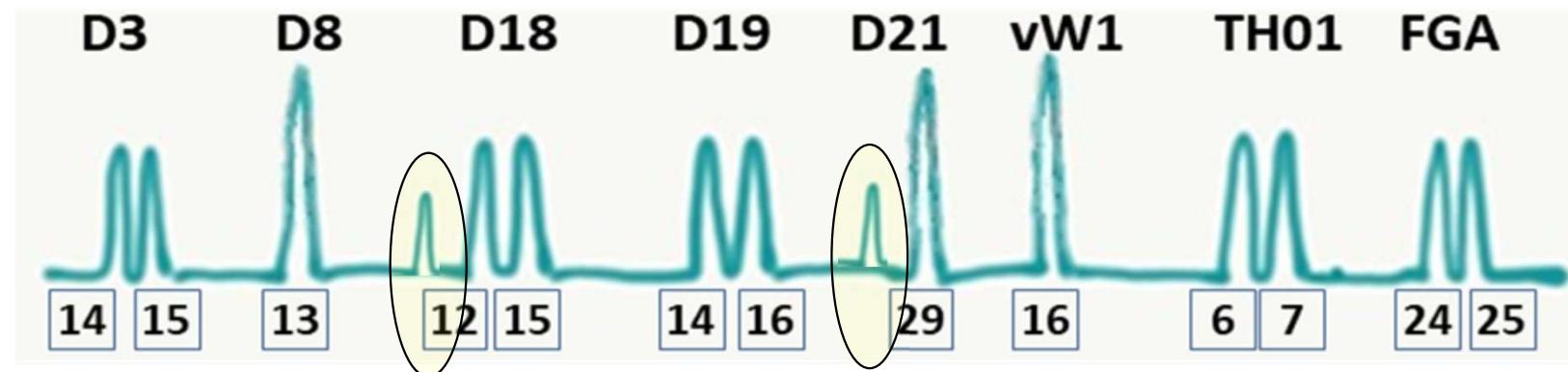
Some contributors’ DNA may be there via secondary, tertiary etc transfer – so

‘finding Fred’s DNA at crime scene’ does not mean Fred was certainly there.

‘Background DNA’ easily introduced

‘Stochastic effects’ (dropin, dropout, stutter) in testing process mean ‘peaks’ below or above threshold can be wrongly classified

‘Probabilistic genotyping software’ can only partially address the problems



# Likelihood Ratio: in mixed DNA evidence

Prosecution hypothesis: “*The suspect and one unknown unrelated person are included in the DNA*”

Defence hypothesis: “*The DNA comes from two unknown unrelated persons both unrelated to the defendant*”

***These hypotheses are NOT mutually exclusive and exhaustive***

***Cannot be certain about the number of people in the mixture and the extent to which they may be related to the suspect***

# Mixed DNA evidence

DNA sample from  
crimes scene

Locus	D3S1358	vWA	D16S539
Alleles found	14, 15, 16	16, 17, 18	9,10,11, 12
Suspect alleles	15, 16	17, 17	9, 11

DNA from suspect

We have a ‘match’.

Suppose  $H$  is the prosecution hypothesis: “the suspect DNA in included in the mixture”

$$LR = \frac{\text{Probability of } E \text{ given } H}{\text{Probability of } E \text{ given not } H} = \frac{1/1000}{1/1000000} = 1000$$

Note that this is typically a very small number

...but this is much smaller still

These are estimates based on many assumptions

As in the single profile case, if we find a ‘match’ on 14 loci then the LR is approximately  $10^{14}$  ....but unlike the single profile case, this does *not* provide overwhelming support for  $H$

# The special problems of mixed DNA evidence

DNA sample from  
crimes scene

Locus	D3S1358	vWA	D16S539	....
Alleles found	14, 15, 16	16, 17, 18	9,10,11, 12	...

DNA from suspect

Suspect alleles	15, 16	17, 17	9, 11	...
-----------------	--------	--------	-------	-----

Possible DNA  
mixture  
profiles that  
match the  
crime scene  
sample

Contributor 1	14, 16	16, 18	10, 12	...	
Contributor 2	15, 16	17, 17	9, 11	...	✓
Contributor 1	14, 16	16, 17	9,10	...	✗
Contributor 2	14, 15	17, 18	11,12	...	✗
Contributor 1	14, 16	16, 17	9,12	...	✗
Contributor 2	15, 15	18, 18	10, 11	...	✗
Contributor 1	14, 15	16, 16	9,12	...	✗
Contributor 2	16, 16	17, 18	10, 11	...	✗

Suspect's DNA profile  
included in this mixture.

All of these possibilities  
*exclude* the suspect's  
DNA profile. They are  
just as likely as any that  
include it.

There are many more  
such combinations

••••••••••••

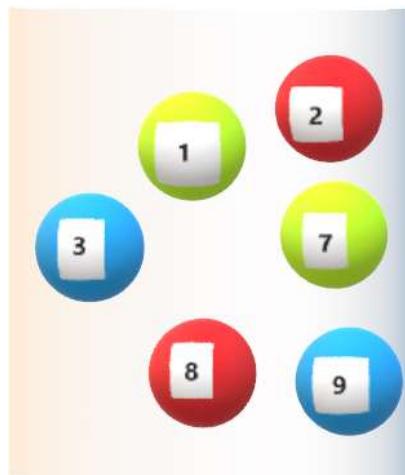
A high LR for a mixed profile ‘match’ is completely different to a high LR for a single profile match

Suppose we have a LR of over 10 billion for a single profile match (e.g. match on every locus). Then, even if our prior belief that it comes from the suspect is  $1/(7 \text{ billion})$  there is a greater than 50% posterior probability.

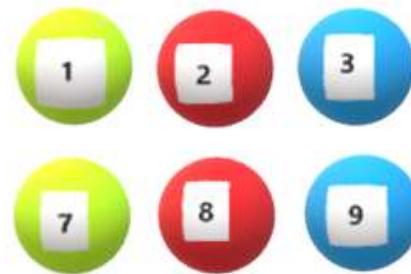
However, if the suspect’s alleles are included on every locus of a mixed profile then – while this will certainly lead to an astronomically high LR – the posterior probability that the person is a contributor is theoretically very low.

Imagine everybody carries 3 coloured balls (yellow, red, blue) with numbers 0-9.

Two people throw their balls into a box



Fred's 'profile'

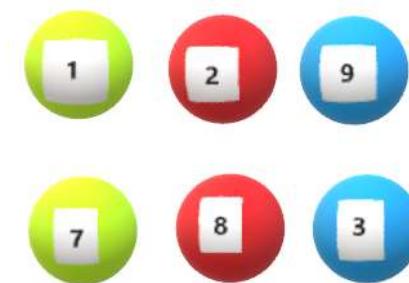


H: "Fred's profile is in the mixture"

LR = 166. So we are "166 times more likely ..."

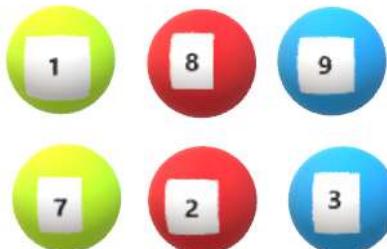


"Fred's profile  
NOT in this  
mixture"

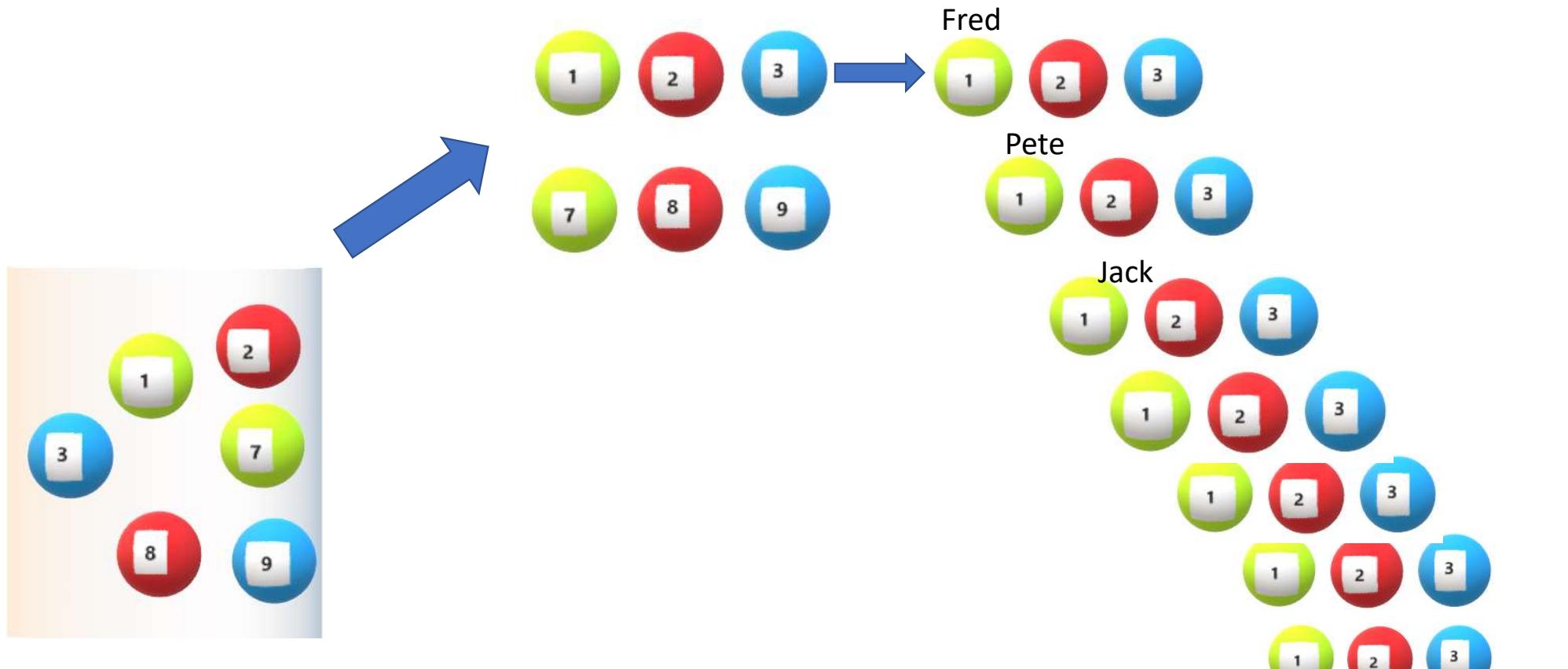


"Fred's profile  
NOT in this  
mixture"

"Fred's profile  
NOT in this  
mixture"



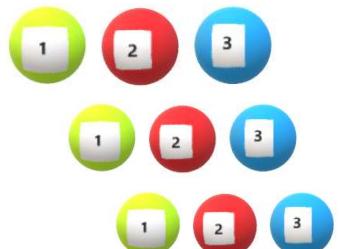
***So the posterior probability of H is only 1/4***



..and H is wrongly assumed to be “Fred is a contributor to the mixture”

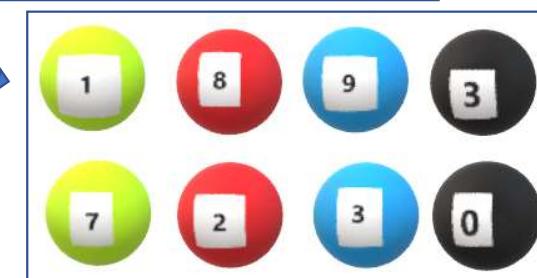
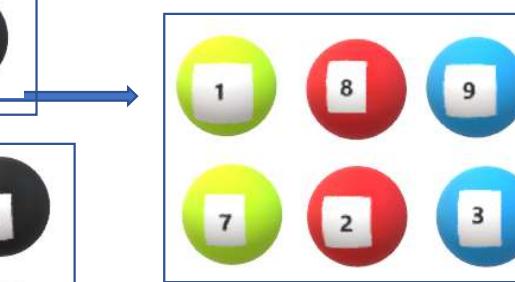
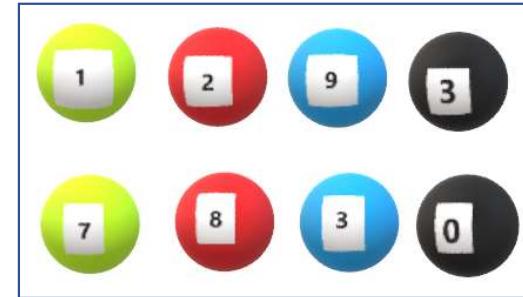
There is only a  $\frac{1}{4}$  probability his profile is in mixture and even if it is he is no more likely to be THE contributor with that profile than many others

So saying “*The evidence that Fred is a contributor is 166 times more likely than he is not*” is both misleading and WRONG



Now suppose everybody carries 4 coloured balls (yellow, red, blue, black)

Again 2 throw in their balls



Only 1 out of 8 possible combinations of contributors include Fred's profile... **But LR is 714**

...For 24 coloured balls

LR is approximately  $6 \times 10^{16}$ - an astronomically high number.

But the posterior probability of H given E is incredibly low, namely  $\frac{1}{2^{n-1}}$  (approx one in 8.4 million) because the suspect's profile is in exactly one of the  $2^{n-1}$  possible pairs of profiles that result in the mixture.

**Analogy:** For 2-person mixed DNA profile with 24-loci 'identified' "Evidence is many billions of times more likely if the suspect is a contributor than if he is not"

....but (in absence of other evidence) ***it is almost certain that he cannot be a contributor***

## But things are even more uncertain...

For low template mixtures, ‘matches’ may be reported for very few loci

Common cause of ‘stochastic effects’ are not taken into account

We can never be sure of the number of different contributors

The hypotheses are not exhaustive for two reasons (the ‘unrelated’ issue + the ‘number of contributors’ issue

Expert reports: “*It is a billion time more likely to obtain the DNA results if the suspect is a contributor than if he is not*”

Expert **should** report “*If we could be certain that there were exactly 2 contributors to this mixture, then it is a billion times more likely to obtain the DNA results if the suspect is a contributor than if two people unrelated to the suspect were the contributors....However, in the absence of any other evidence it is still very unlikely that anybody matching the suspect’s DNA profile is even in the mixture at all*”

<https://arxiv.org/abs/2106.05328>



Cornell University

the

arXiv.org > stat > arXiv:2106.05328

Search...  
Help | Advanced

**Statistics > Applications**

[Submitted on 9 Jun 2021]

# Calculating the Likelihood Ratio for Multiple Pieces of Evidence

Norman Fenton, Martin Neil

When presenting forensic evidence, such as a DNA match, experts often use the Likelihood ratio (LR) to explain the impact of evidence . The LR measures the probative value of the evidence with respect to a single hypothesis such as 'DNA comes from the suspect', and is defined as the probability of the evidence if the hypothesis is true divided by the probability of the evidence if the hypothesis is false. The LR is a valid measure of probative value because, by Bayes Theorem, the higher the LR is, the more our belief in the probability the hypothesis is true increases after observing the evidence. The LR is popular because it measures the probative value of evidence without having to make any explicit assumptions about the prior probability of the hypothesis. However, whereas the LR can in principle be easily calculated for a distinct single piece of evidence that relates directly to a specific hypothesis, in most realistic situations 'the evidence' is made up of multiple dependent components that impact multiple different hypotheses. In such situations the LR cannot be calculated . However, once the multiple pieces of evidence and hypotheses are modelled as a causal Bayesian network (BN), any relevant LR can be automatically derived using any BN software application.

Comments: 27 pages, 12 figures

Subjects: **Applications (stat.AP)**

Cite as: [arXiv:2106.05328 \[stat.AP\]](#)  
(or [arXiv:2106.05328v1 \[stat.AP\]](#) for this version)

# Sample output from mixed profile DNA analysis

Locus	Alleles in sample	Suspect alleles
D3S1358	14, 15, 16	15, 15
vWA	15, 16, 17, 19	17, 19
D16S539	10, 11, 12	12, 12
D8S1179	11, 12, 13, 14	11, 12
D18S51	12, 14, 17, 25	13, 25
...	...	...
D5S818	7, 10, 11, 12, 13	10, 11
D13S317	8, 9, 10, 12, 13	9, 10
D7S820	8, 8.3, 9, 9.3, 10, 12	10, 11
...	...	...
D12S391	15, 16, 18, 19	18, 18
D2S1338	16, 17, 19, 20, 24, 25	23, 25

'voids' – but the  
probabilistic genotyping  
takes care of these

## RESULTS OF ANALYSIS

Assumed number of contributors: 4 *This is a very strong assumption*

Likelihood Ratio: It is  $4 \times 10^6$  times more likely to obtain the DNA results if suspect is a contributor than if he is not a contributor

Summary conclusion: Strong support for inclusion of suspect

Suggests exhaustive hypotheses but this is NOT done in the calculations

*Not when all uncertainty is considered*

## **What the expert should have said**

“If we could be certain that there were exactly 4 contributors to this mixture, and if we totally ignore every type of possible error that could be introduced during the DNA collection and analysis process, then it is  $4 \times 10^6$  times more likely to obtain the DNA results if the suspect is a contributor than if four people unrelated to the suspect were the contributors.

However, if there were a different number contributors then it is possible that the suspect could be definitively excluded.

Also, because the suspect’s alleles were not detected at several loci, it is possible that there are people related to the suspect who are more likely to have been contributors.

Even if we are sure of the number of contributors and can exclude the possibility of people related to the suspect being contributors then, in the absence of any other evidence linking the suspect to the DNA sample, it is still very unlikely that he is a contributor.”

But there are priors that can be used and which can destroy the defence arguments against the LR

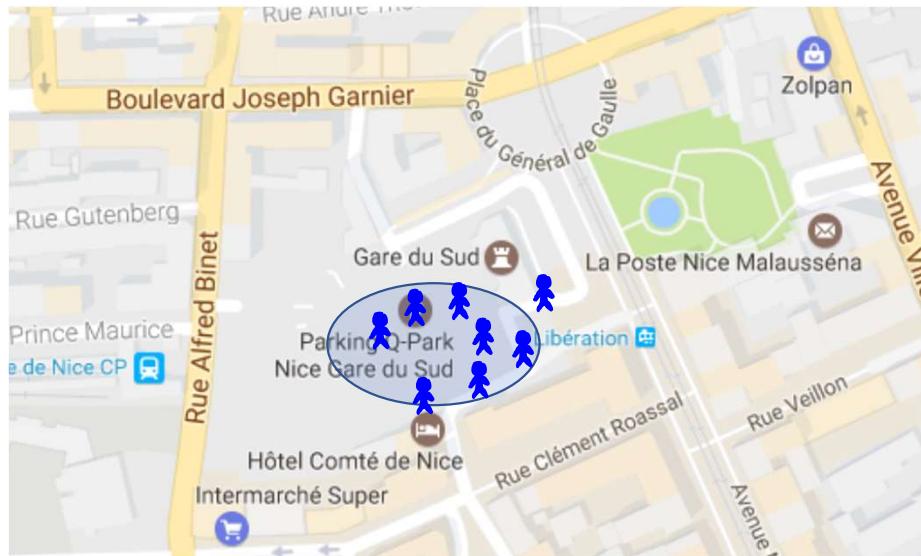
Like it or not people form their own subjective priors

The ‘opportunity prior’ work provides an objective prior for a large class of crimes

# Opportunity prior: Crime Scene and Crime Time

Assume a crime has taken place, it was committed by one person against one other person (e.g. murder, assault, robbery). Then:

- The ***crime scene*** (CS): smallest physical area within which it is certain the crime happened.
- The ***crime time*** (CT): smallest time interval ( $t, t'$ ) between which it is certain the crime took place.



Imagine we can observe people who are in the CS at any time during CT

## Number of people at Crime Scene during Crime Time: ( $n$ )

- We generally do not know who was at CS during CT. But it is possible to estimate number of people  $n$  (other than the victim) who were.
- By definition the criminal is one of these  $n$  people.

If the suspect was at CS during CT then the truly fair prior probability of guilt is  $1/n$

*As close to ‘innocent until proven guilty’ as possible*

We also handle the case where suspect was NOT at CS during CT by using notion of **extended CS**

# Crushing the ‘anybody in the world’ fallacy

- Suppose only TWO people Fred and Joe were at the CS during CT
- $P(\text{Fred is guilty}) = \frac{1}{2}$
- Suppose only TWO people Fred and an unknown other were at the CS during CT.
- The ‘other’ can be anybody in the world. So what is  $P(\text{Fred is guilty})$ ?
- Fallacy is to assume  $n$  is different (e.g. much higher) in this case

# The opportunity prior and the LR

Suppose that a mixed DNA profile of blood is found at a crime scene. The prosecution claims the suspect's DNA is in the mixture with a very high LR. As we have seen - in the absence of other evidence - this is essentially meaningless. It tell us nothing about the probability the suspect's DNA is actually in the mixture.

However, if the suspect was known to have been one of at most 100 people who were at the crime scene, then even taking account of all of the potential errors in the LR calculation, the opportunity prior of 1/100 generally leads to a high posterior probability that the suspect's DNA was in the mixture.

# Law, Probability & Risk

[Issues](#) [Advance articles](#) [Submit ▾](#) [Purchase](#) [Alerts](#) [About ▾](#)

## Article Navigation

### The opportunity prior: a proof-based prior for criminal cases

Norman Fenton ✉, David Lagnado, Christian Dahlman, Martin Neil

*Law, Probability and Risk*, Volume 18, Issue 4, December 2019, Pages 237–253, <https://doi.org/10.1093/lpr/mgz007>

Published: 20 May 2019 Article history ▾

“ Cite Permissions Share ▾

#### Abstract

One of the greatest challenges to the use of probabilistic reasoning in the assessment of criminal evidence is the ‘problem of the prior’, i.e. the difficulty in establishing an acceptable prior probability of guilt. Even strong supporters of a Bayesian approach have often preferred to ignore priors and focus on the likelihood ratio (LR) of the evidence. But to calculate if the probability of guilt, given the evidence reaches the probability required for conviction (the standard of proof), the LR has to be combined with a prior. In this article, we propose a solution to the ‘problem of the prior’: the defendant shall be treated as a member of the set of ‘possible perpetrators’ defined as the people who had the same or better opportunity as the defendant to commit the crime. For this purpose, we introduce the concept of an ‘extended crime scene’. The number of people who had the same or better opportunity as the defendant is the number of people who were just as close or closer to the crime scene, in time and space. We demonstrate how the opportunity prior is incorporated into a generic Bayesian network model that allows us to integrate other evidence about the case.

<https://doi.org/10.1093/lpr/mgz007>

### **Scenario 1**

A murder is committed in a crowded pub. The victim is known to have put up a fight and DNA from one other person is found on the victim. The DNA profile matches Jimmy and, according to the DNA expert analyst:

*"This DNA evidence very strongly supports the hypothesis that it was Jimmy's DNA on the victim. In fact, it is over a billion times more likely under this hypothesis than under the hypothesis that it came from a person other than Jimmy"*

On a scale of 0 to 100 how sure are you that:

- the DNA came from Jimmy?
- Jimmy was the murderer?

### **Scenario 2**

A murder is committed in a crowded pub. The victim is known to have put up a fight and DNA from at least two other people is found on the victim. A DNA profile that matches Jimmy's is in the DNA mixture and, according to the DNA expert analyst :

*"This DNA evidence very strongly supports the hypothesis that Jimmy was one of the people whose DNA was on the victim. In fact it is over a billion times more likely under this hypothesis than under the hypothesis that Jimmy was not a contributor"*

On a scale of 0 to 100 how sure are you that:

- Jimmy's DNA was on the victim?
- Jimmy was the murderer?

Conclusions and way  
forward

# Misplaced optimism?

“I assert that we now have a technology that is ready for use, not just by the scholars of evidence, but by trial lawyers.”

Edwards, W. (1991). "Influence Diagrams, Bayesian Imperialism, and the Collins case: an appeal to reason." Cardozo Law Review 13: 1025-107

# Summary

Correct probability reasoning is central to far more cases than people imagine – but errors of reasoning continue to plague the system

The LR is a standard measure used by forensic scientists to measure probative value of evidence

By Bayes, the LR is a valid measure of probative value of evidence only when comparing two mutually exclusive and exhaustive hypotheses

In practice forensic scientists do NOT compare exhaustive hypotheses, hence their presentation of LR results is misleading

For realistic cases the LR calculation requires a Bayesian network model

The LR for low template DNA mixture evidence is especially misleading – its probative value is generally massively overestimated

# For more information

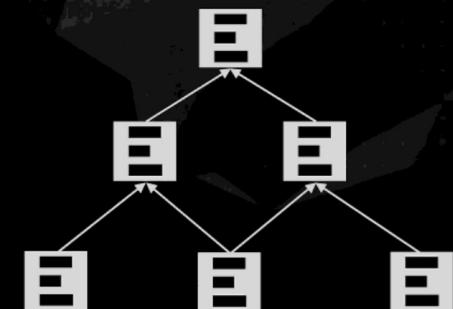
[eecs.qmul.ac.uk/~norman](http://eecs.qmul.ac.uk/~norman)  
[twitter.com/profnfenton](https://twitter.com/profnfenton)  
[probabilityandlaw.blogspot.com/](http://probabilityandlaw.blogspot.com/)  
[agenarisk.com](http://agenarisk.com)

Copyrighted Material

Risk Assessment and  
Decision Analysis with  
Bayesian Networks

SECOND EDITION

NORMAN FENTON  
MARTIN NEIL



CRC Press  
Taylor & Francis Group  
A CHAPMAN & HALL BOOK  
Copyrighted Material

AGENARISK