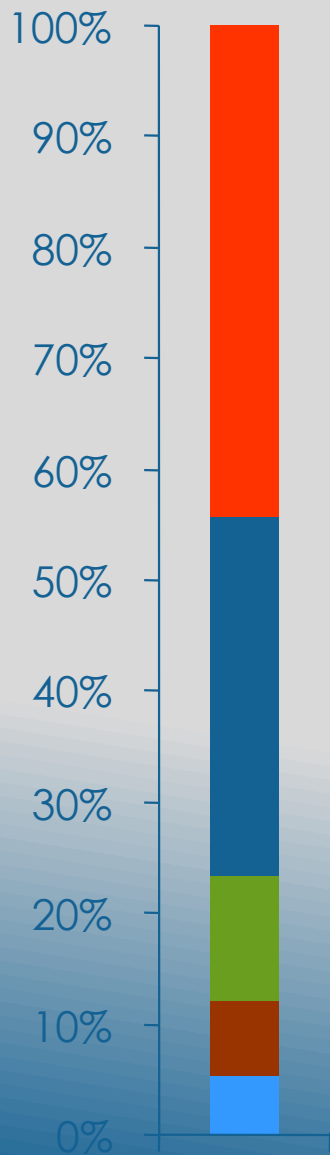


L'identification des criminels

*Les parades d'identification et les analyses
génétiques : éléments d'interprétation*

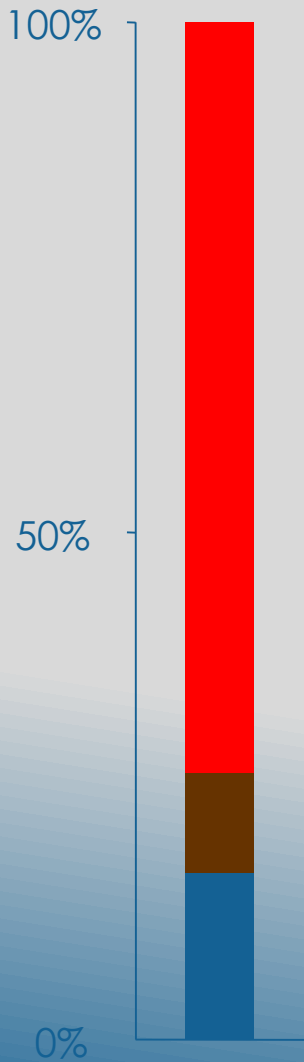
Les 5 conditions de succès de l'enquête de police judiciaire en France (N = 102 affaires)



Facteur le plus déterminant de l'identification du criminel	Affaires (%)
Témoignages	44,4%
Enquête de voisinage	32,3%
Investigations en cours d'enquête	11,3%
Premières constatations (empreintes, indices, traces, ADN, etc.)	6,8%
Dénonciations	5,3%

L'efficacité de la police scientifique dans la résolution des affaires

(arrestation d'un suspect dans 153 affaires traitées, Brodeur, 2005)



Type d'expertise scientifique	Affaires (%)
Autres éléments de l'enquête (filatures, témoignages, etc.)	74,0%
<i>Polygraphe</i>	<i>7,2%</i>
<i>Analyses ADN</i>	<i>2,6%</i>
Information (databases, renseignements)	4,6%
Analyses chimiques	3,3%
Analyses de sang	2,6%
Bertillonnage	2,0%
Analyses balistiques	2,0%
Autopsie	2,0%
Hypnose	0,0%
Analyse d'une scène de crime	0,0%



Type d'expertise scientifique

Affaires (%)

Polygraphe	7,2%
Analyses ADN	2,6%

Dans les 153 affaires traitées par Brodeur, ces deux éléments n'ont **jamais permis d'interpeller un suspect,...**

...mais seulement de **disculper** et de relâcher des suspects innocents...

L'importance « réelle » des témoignages

Dans **73 % des procès**, la seule preuve que pouvait fournir l'accusation ou la défense était un **témoignage oculaire**

(Delvin, 1973)

Témoignage **accusateur**

AVEC → **72%** de jugement de culpabilité

SANS → **18%** de jugement de culpabilité

(Loftus, 1979)

Les erreurs judiciaires

Les origines

A decorative graphic consisting of several parallel white lines of varying lengths, slanted upwards from left to right, located in the bottom right corner of the slide.

Quels facteurs peuvent expliquer les erreurs de justice ? *(Rattner, 1988: 205 cas d'erreurs judiciaires avérées)*

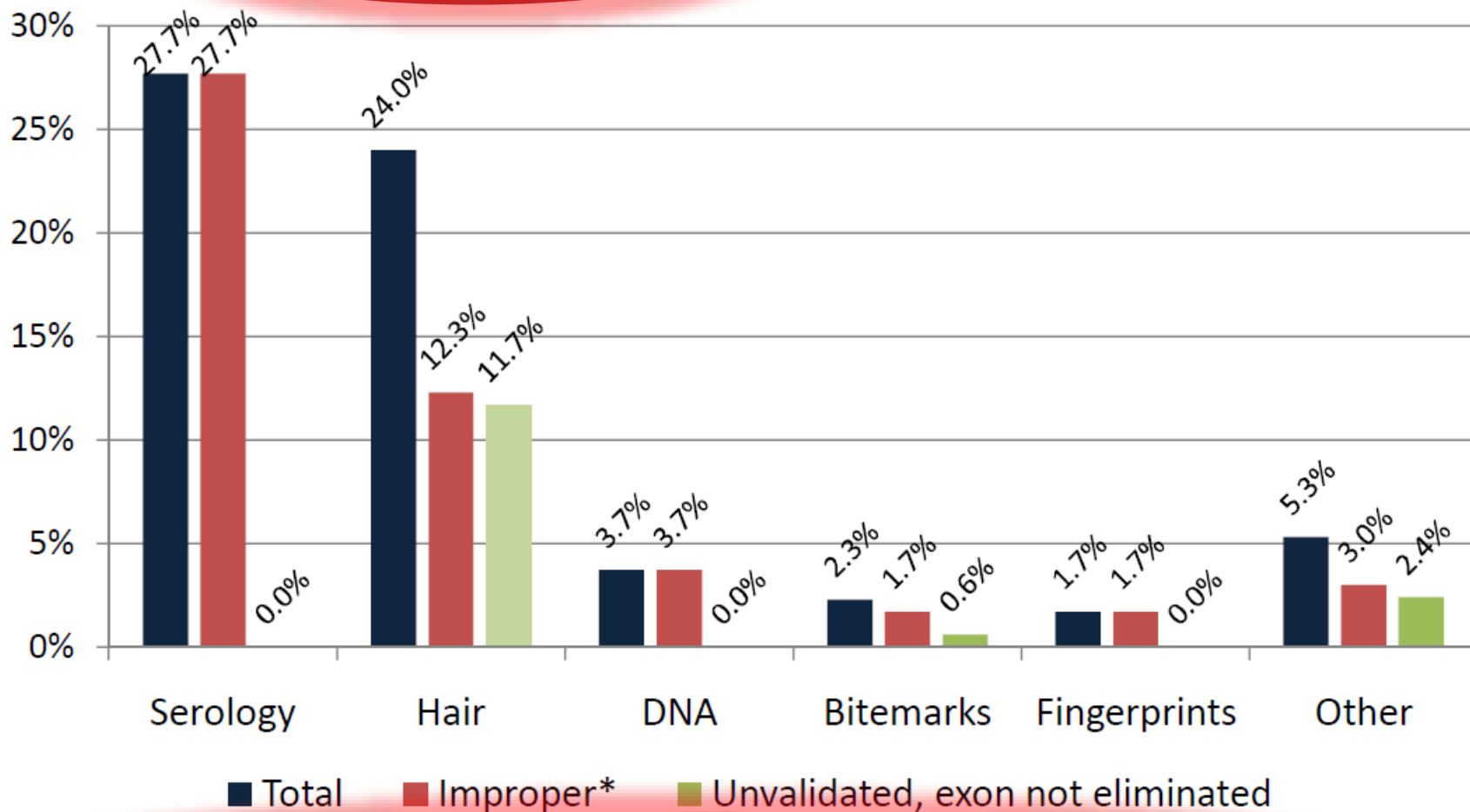


Les erreurs judiciaires

Les erreurs scientifiques existent et expliquent en partie les erreurs judiciaires

A decorative graphic consisting of several parallel white lines of varying lengths, slanted diagonally from the bottom right towards the top right, located in the lower right quadrant of the slide.

51% of 300 DNA Exonerations Involved Use of Improper/Unvalidated Forensic Science: Breakdown by Discipline



* **Improper category includes:** testimony or analysis which drew conclusions beyond the limits of science as known at that time; cases in which there was negligence in analysis, fabrications/alterations of reports and possible failures to conduct elimination testing or comparison; and withholding laboratory reports, analysis, data, or the very existence of evidence

Les séances d'identifications

(tapissages)

Quelle(s) conclusion(s) pour un tapissage ?

Le témoin (ou la victime) déclare...

« C'est le N°3 ! »



La probabilité de culpabilité (subjective) augmente

« Euuuhhh, je ne le vois pas... »



La probabilité de culpabilité diminue

DANGER si aucun autre élément de preuve ne vient confirmer la décision...

Le décalage entre la décision du témoin et la réalité

Il existe 6 issues possibles à une séance d'identification, dont **2 seulement sont**
CORRECTES (traduisent la réalité...)

Exactitude de la décision et conséquences judiciaires

Criminel présent dans l'alignement ?

(inconnu des enquêteurs)

Réponse du témoin	Oui <i>(arrestation du criminel)</i>	Non <i>(arrestation d'un suspect innocent)</i>
Il désigne le suspect	Identification correcte	Erreur de reconnaissance
Il désigne un distracteur	Erreur de reconnaissance	Erreur de reconnaissance
Il ne désigne Personne	Rejet incorrect	Rejet correct

Comment un témoin peut-il désigner un innocent ou un distracteur ?

(1) parce que la situation le pousse à le faire...

→ problème dans la formulation de la consigne donnée au témoin, «pression informelle» à la décision de la part des policiers, etc.

L'importance de la consigne donnée au témoin

Malpass et Devine (1981)

- Des étudiants jouent le rôle de témoins.
- Ils doivent identifier une personne qui a commis un délit lors de leur cour magistral.
- Les chercheurs utilisent des tapissages dans lesquels il n'y a pas le criminel (= > situation où les enquêteurs n'ont pas interpellé la bonne personne...)

Consigne **biaisée**

78% des témoins désignent quelqu'un

Consigne **non-biaisée**

33% choisissent quelqu'un

Comment un témoin peut-il désigner un innocent ou un distracteur ?

(1) parce que la situation le pousse à le faire...

(2) Parce que certains éléments
(« indices ») **amènent à penser qu'il est**
le suspect (seul à porter des menottes)

(3) Parce que le tapissage n'est pas fiable

Comment un témoin peut-il désigner un innocent ou un distracteur ?

- (1) parce que la situation le pousse à le faire...
- (2) Parce que certains éléments (« indices ») amènent à penser qu'il est le suspect (seul à porter des menottes)
- (3) Parce que le tapissage n'est pas fiable**

Les séances d'identifications

La fonction des tapissages

A decorative graphic consisting of several parallel white lines of varying lengths, slanted diagonally from the bottom right towards the top left, located in the lower right quadrant of the slide.

Lors d'une enquête, l'objectif principal est de mettre une identité sur le criminel

2 manières usuelles pour donner une identité :

- Faire **décrire** le criminel
- Procéder à une **séance d'identification**
(mais nécessite d'avoir un suspect)

Lors d'une enquête, l'objectif principal est de mettre une identité sur le criminel

Lorsque les enquêteurs reçoivent un témoin ou une victime, ils essaient très rapidement d'obtenir un signalement de l'auteur



donner une « identité » à l'auteur

Lors d'une enquête, l'objectif principal est
de mettre une identité sur le criminel

Au départ → 6,9 milliards de suspects potentiels

mais la description réduit fortement ce panel...



Lors d'une enquête, l'objectif principal est de mettre une identité sur le criminel

Cependant, l'incertitude reste élevée, la marge d'erreur est importante...

**« Difficile » de condamner quelqu'un
sur la seule base de la description**

***il faut réduire l'incertitude en obtenant d'autres
informations identitaires → GAIN D'INFORMATIONS***

Lors d'une enquête, l'objectif principal est de mettre une identité sur le criminel

Gain d'information = Réduction de la marge d'erreur



**Utiliser une parade
d'identification**

En résumé : à quoi sert un tapissage ?

Une description permet d'obtenir quelques éléments identitaires...

MAIS

à de rares exceptions près, elle ne permet pas de définir l'identité du criminel avec une marge d'erreur minimale.

En résumé : à quoi sert un tapissage ?

A préciser l'identité du criminel en en apprenant plus que l'on sait déjà...

Et que connaît-on déjà ?

La description du criminel...

Une affaire réelle bien connue...

L'affaire DSK / DIALLO



Une affaire réelle bien connue...

SUPREME COURT OF THE STATE OF NEW YORK
COUNTY OF NEW YORK: PART 51

THE PEOPLE OF THE STATE OF NEW YORK

-against-

DOMINIQUE STRAUSS-KAHN,

Defendant.

RECOMMENDATION
FOR DISMISSAL

Indictment No. 02526/2011

RECEIVED
CENTRAL CLERK'S OFFICE

AUG 22 2011

SUPREME COURT
CRIMINAL TERM
NEW YORK COUNTY

SUMMARY

The People of the State of New York move to dismiss the above-captioned indictment, which charges the defendant with sexually assaulting the complainant at a hotel in midtown Manhat-

HISTORY OF THE INVESTIGATION

A. Initial Investigation and Indictment

On May 14, 2011, the complainant, a housekeeper at the Sofitel Hotel on West 44th Street in Manhattan, reported to hotel security, and later to the NYPD, that she had been sexually assaulted by the defendant in his hotel suite. She first reported this incident to her immediate supervisor shortly after her interaction with the defendant, whose suite (Suite 2806) she had been assigned to clean. That supervisor summoned a more senior supervisor, to whom the complainant repeated her claim. The second supervisor notified hotel security and management personnel, who in turn notified the NYPD. NYPD uniformed police officers and detectives interviewed the complainant and had her transported to a local hospital for a medical examination later that afternoon.

In substance, the complainant reported to NYPD detectives, and later to prosecutors, that shortly after she entered the defendant's suite to perform her housekeeping duties, he emerged naked from the suite's bedroom, approached her, and grabbed her breasts without her consent. According to the complainant, the defendant closed the door to the suite, forced her into the bedroom, pushed her onto the bed, and attempted to forcibly insert his penis into her mouth, which caused his

Une affaire réelle bien connue...

The NYPD ascertained that the defendant was scheduled to depart on an Air France flight

at John F. Kennedy Airport that was headed for Europe. He was asked to disembark from that flight at approximately 4:45 p.m. by detectives assigned to the Port Authority Police Department, and was eventually taken into custody.

Les autorités savaient qui était le présumé « agresseur »

Une affaire réelle bien connue...

PROCEDURAL BACKGROUND

The defendant was taken into custody on May 14, 2011, and on the following day, was identified in a line-up by the complainant and arrested by the New York City Police Department (“NYPD”). The People filed a felony complaint on May 15, 2011, charging the defendant with the

Une affaire réelle bien connue...

PROCEDURAL BACKGROUND

The defendant was taken into custody on May 14, 2011

identified in a line-up by the complainant and

("NYPD"). The People filed

with the

**AUCUN GAIN
d'information possible...**

Une affaire réelle bien connue...

L'inconvénient est que **l'identification est l'élément qui emporte le plus la conviction des jurés**, quand bien même la parade n'apporte rien...

Elle devient alors un simple **« argument rhétorique »** qui permet de remporter la partie...

Si l'on ajoute à cela la **pression médiatique** déjà identifiée en 1932 par Lord Edwin Borchard, **DSK ne pouvait être que placé en accusation par le Grand Jury**

Les séances d'identifications

La fiabilité des tapissages

A decorative graphic consisting of several parallel white lines of varying lengths, slanted diagonally from the bottom right towards the top right, located in the lower right quadrant of the slide.

En apprendre plus...

Si **l'identité** du criminel est **encore inconnue**, alors on peut réaliser un **tapissage**

MAIS

Pour que les nouveaux éléments soient utiles, il faut qu'ils soient **VRAIMENT** nouveaux...

→ *Il ne faut pas que l'identification nous donne les mêmes éléments...*

Il faut « bloquer » la possibilité pour le témoin de réutiliser les informations déjà fournies dans la description

En apprendre plus...

Si l'on ne bloque pas les **anciennes** information, il y a risque de les faire passer pour de **nouvelles** informations

→ « recyclage »

→ **On induit la Cour en erreur...**

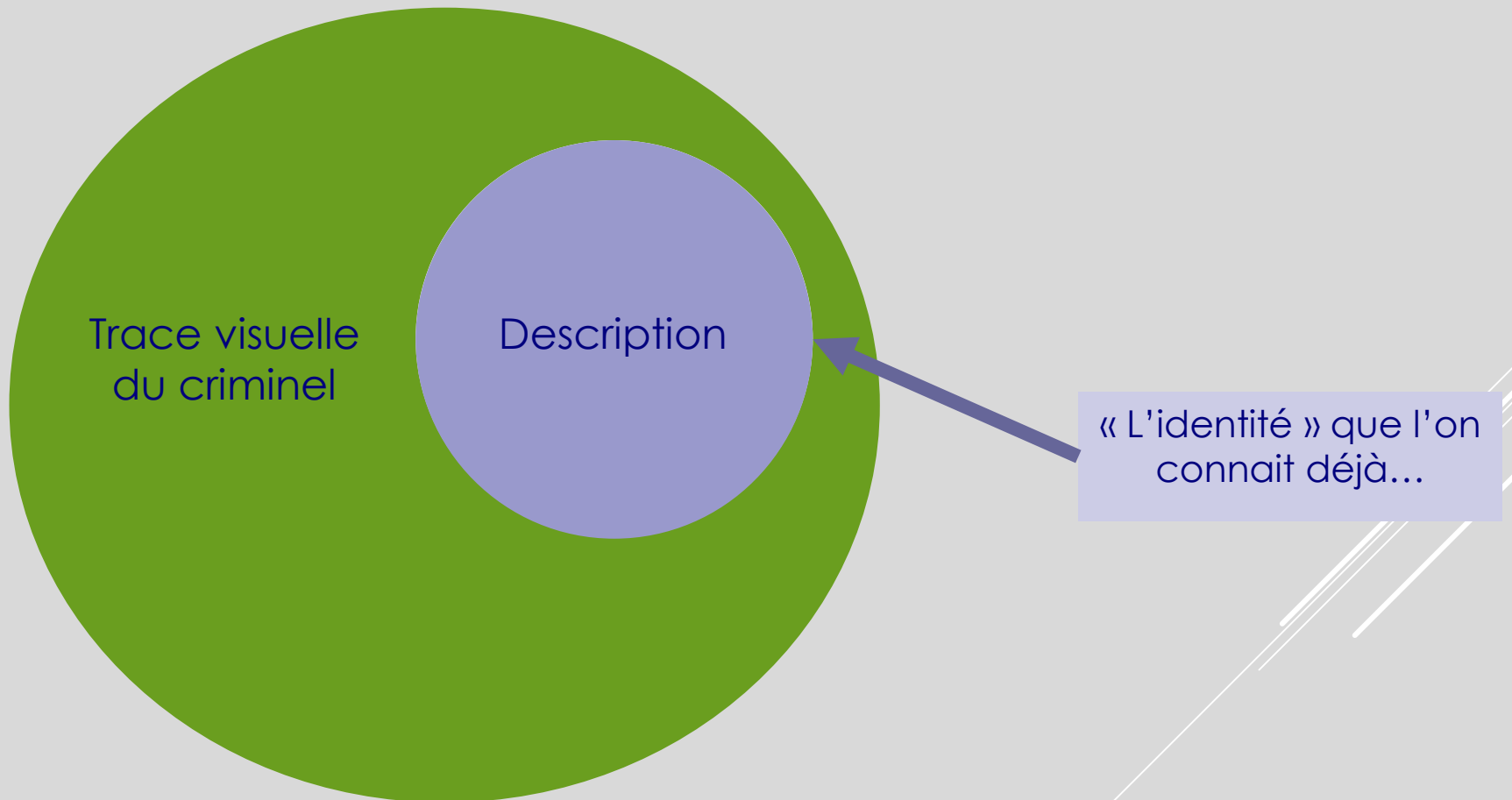


En apprendre plus...



La description obtenue par les enquêteurs provient exclusivement de la **trace visuelle** du criminel que le témoin conserve en mémoire

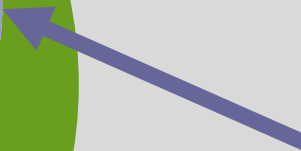
En apprendre plus...



« l'identité » optimale
que l'on voudrait bien
obtenir...



« L'identité » que l'on
connait déjà...



« l'identité » optimale
que l'on voudrait bien
obtenir...



Trace visuelle
du criminel

Description

« L'identité » que l'on
connait déjà...



« l'identité » optimale
que l'on voudrait bien
obtenir...

Description

Le risque est que le témoin « **recycle** » la
description, sans faire l'effort de rapporter
de nouvelles informations

Trace visuelle
du criminel

« L'identité » que l'on
connait déjà...

« l'identité » optimale
que l'on voudrait bien
obtenir...

Description

En « recyclant », il va donner l'impression
(fausse !) qu'il en sait plus que ce qu'il a
déjà dit, alors qu'il n'en sait pas plus...

Et tout le monde va le CROIRE ...

on

Comment être sûr qu'une désignation du suspect correspond bien à un **GAIN** **d'informations** ?

Empêcher les témoins de baser leur identification sur les mêmes informations déjà fournies dans la description

Les contraindre à en utiliser de nouvelles, pour autant qu'ils en aient encore



Comment être sûr qu'une désignation du suspect correspond bien à un **GAIN d'informations** ?

COMMENT FAIRE ?

Sélectionner des distracteurs qui **correspondent tous autant que le suspect à la DESCRIPTION** fournie par le témoin

Illustration du principe « En apprendre plus que l'on ne sait déjà... »

La description d'un criminel fournie par un témoin va apparaître

Ensuite...

La photo d'un (vrai !) tapissage va apparaître

Consigne :

le plus **RAPIDEMENT** possible, trouver la personne qui correspond **LE PLUS** à la description fournie.

1 8 6 4 9 5

« type maghrébin, 18-20 ans, mince, cheveux rasés »

« type maghrébin, 18-20 ans, mince, cheveux rasés »

1 8 6 4 9 5

« type maghrébin, 18-20 ans, mince, cheveux rasés »



« type maghrébin, 18-20 ans, mince, cheveux rasés »

Ce tapissage est

biaisé à l'encontre du suspect

puisque des personnes qui ne connaissent rien à

l'affaire *(vous !)* ont pu retrouver le suspect

Ce tapissage n'est pas
un élément de preuve :

*Il n'y a **aucun gain**
d'information possible*

Dilemme...

On est alors face à une **incertitude manifeste** si un témoin désigne le suspect dans une parade biaisée

2 solutions possibles, mais impossible de choisir entre les deux...

Soit le témoin s'est servi de **nouvelles** informations

→ Identification **valide**

→ **GAIN d'INFORMATIONS**

Soit le témoin s'est servi des **mêmes** informations (c-à-d la description)

→ Identification **non pertinente** puisque le suspect était le seul à correspondre à ces anciennes informations...

→ **AUCUN GAIN d'INFORMATIONS**

5

4

2

6

« jeune homme blond, pas très grand, assez maigre, cheveux courts »

« jeune homme blond, pas très grand, assez maigre, cheveux courts »

5

4

2

6

« jeune homme blond, pas très grand, assez maigre, cheveux courts »

5

4

2

6

« jeune homme blond, pas très grand, assez maigre, cheveux courts »

Existe-t-il d'autres biais ?

Dans le cas où les distracteurs correspondent autant que le suspect au signalement (c-à-d, **parade fiable**), peut-on tout de même retrouver le suspect ?

En d'autres termes, si la parade n'est pas biaisée, existe-t-il des « astuces » pour retrouver le suspect ?

Existe-t-il d'autres biais ?

Consigne

« retrouvez le suspect »

(mobilisation des « connaissances » rattachées au mot « suspect »)



10

16

15

14

13

« Retrouvez le suspect »

10

16

15

14

13

« Retrouvez le suspect »

10

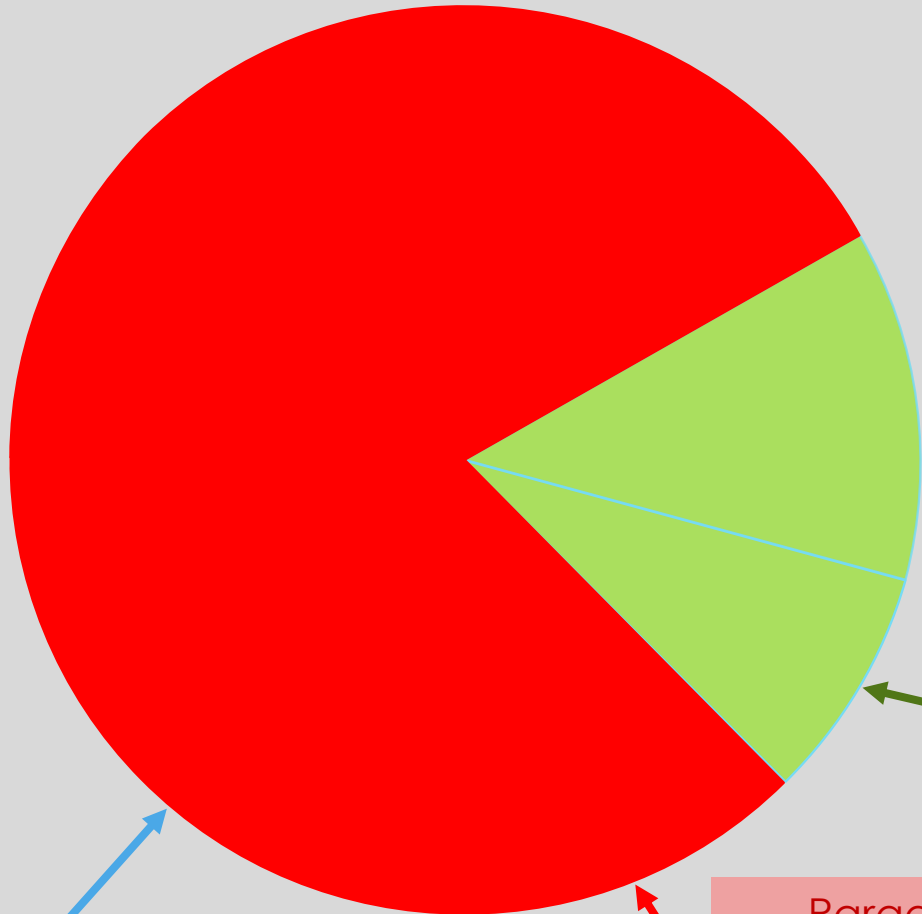
16

15

14

13

« Retrouvez le suspect »



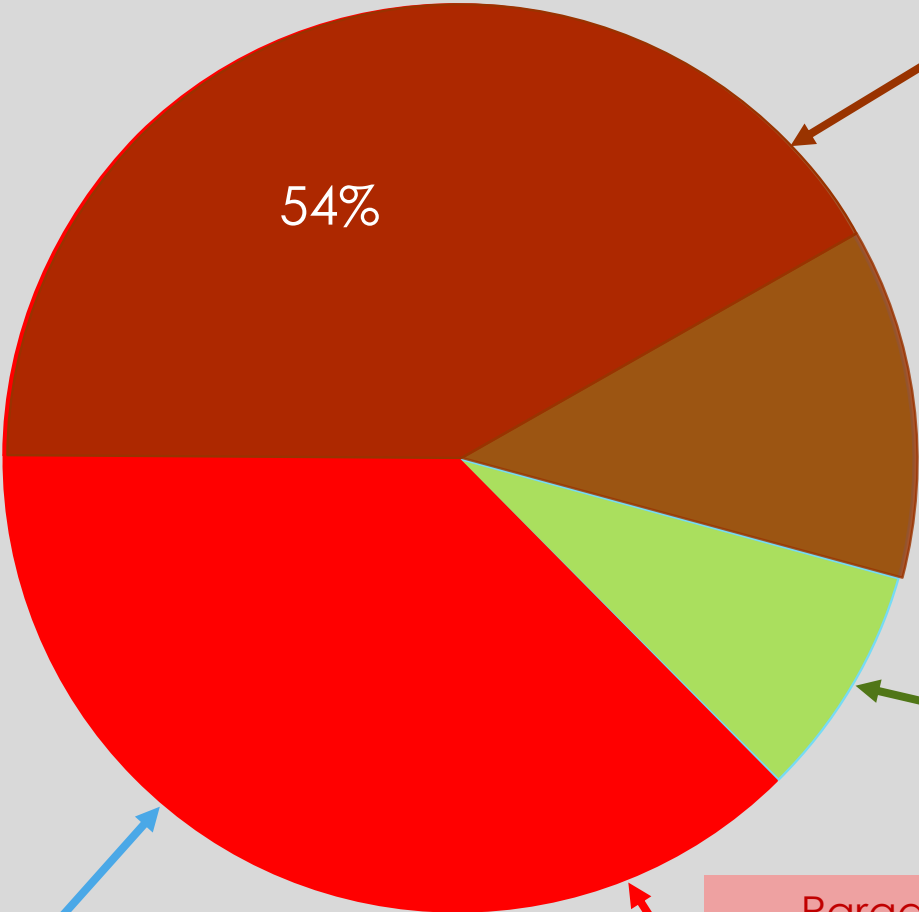
$N = 55$

Parades
biaisées
Test consigne
description

Parades non
biaisées

Parades biaisées

Test: Consigne
Suspect



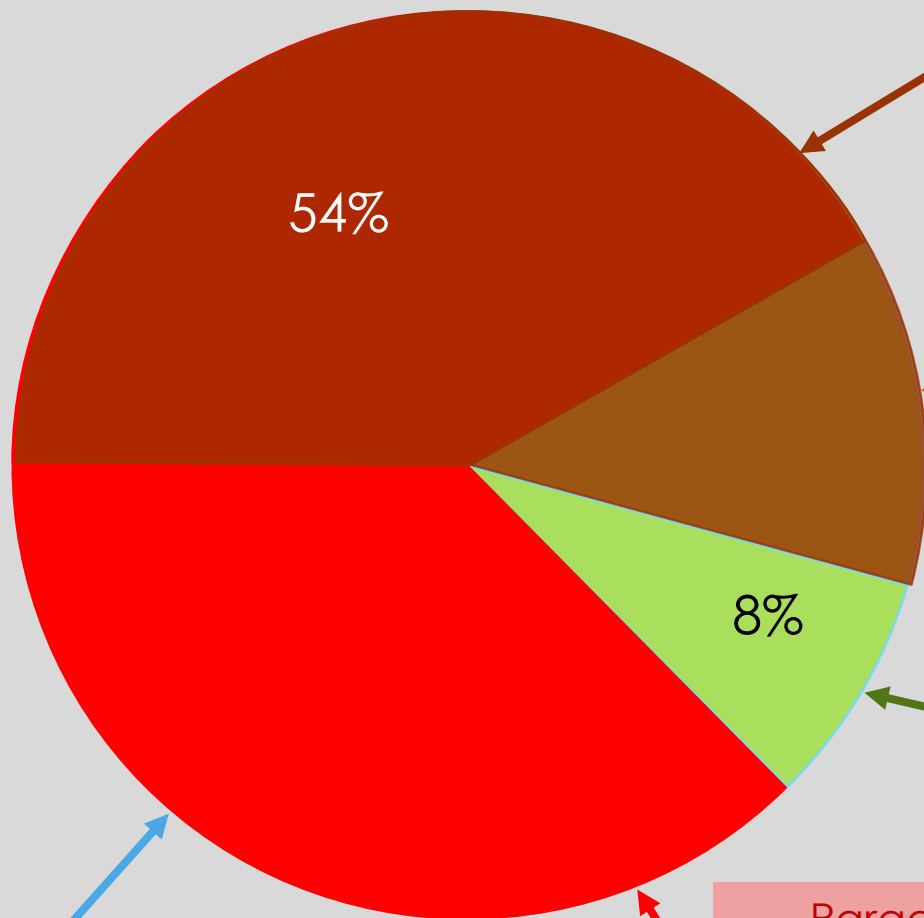
Parades non
biaisées

Parades
biaisées
Test consigne
description

$N = 55$

Parades biaisées

Test: Consigne
Suspect



Parades biaisées
avec la consigne
Suspect, mais pas
avec la consigne
Description

Parades non
biaisées

Parades
biaisées
Test consigne
description

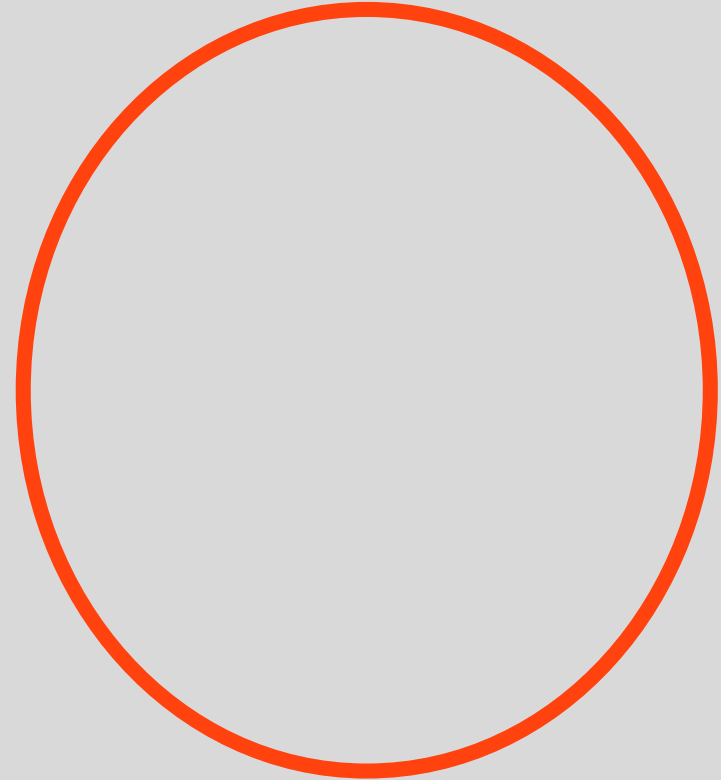
$N = 55$

Comment faire un tapissage fiable ?

Un exemple

A decorative graphic consisting of three parallel white lines of varying lengths, slanted upwards from left to right, located in the bottom right corner of the slide.

Homme, européen, 20-
25 ans, 1.70-75,
corpulence moyenne,
cheveux courts, bruns et
bouclés, avec les yeux
clairs



Homme, européen, 20-
25 ans, 1.70-75,
corpulence moyenne,
cheveux courts, bruns et
bouclés, avec les yeux
clairs

Les séances d'identifications

*Une autre source d'erreur :
Le mode de présentation*

Une autre source d'erreurs : le mode de présentation

La présentation simultanée de tous les membres permet au témoin de COMPARER les individus entre eux

PROBLEME...

Il peut alors choisir celui qui correspond **LE PLUS** à son souvenir, et non celui qui y correspond **EXACTEMENT**

Une autre source d'erreurs : le mode de présentation

La présentation simultanée de tous les membres permet au témoin de **COMPARER** les individus entre eux

PROBLEME...

Si le souvenir est **imprécis** (Wells, 1984), le témoin peut désigner quelqu'un à partir d'un **souvenir partiel**

POSSIBLE ERREURS D'IDENTIFICATION

Comment lutter contre le jugement relatif ?

Présentation SEQUENTIELLE des membres (Lindsay & Wells, 1985)

- Puisque l'on présente les membres un par un, **impossible de les comparer.**
- Une **décision** est **attendue** pour **chaque** membre.
- Dès que l'on a un **OUI**, on ne présente pas les membres restants

Performances de la présentation SEQ.

Méta-analyse de Steblay, Dysart, et Wells (2011)

Mesures	SIM	SEQ
Bonne reconnaissance ($S = C$)	58 %	32 %
Erreur de reconnaissance ($S \neq C$)	25 %	13 %
Rejet correct ($S \neq C$)	43 %	64 %
Désignation de distracteurs ($S = C$ ou $S \neq C$)	24 %	24 %
Rejet à tort ($S = C$)	27 %	41 %

Tout est affaire de point de vue...

La supériorité de la SEQ a conduit plusieurs états américains à l'adopter (CA, NJ, Car.Nor, Virg. Wisc.)

→ Acceptabilité sociale de la méthode.

MAIS...

Tout est affaire de point de vue...

Les plus conservateurs dénoncent un mode de présentation qui **défavorise l'accusation au détriment de la défense**, puisqu'il y a une diminution des bonnes reconnaissances avec la SEQ (S=C)

Pour Malpass et al. (2009) → **« la SEQ, c'est du marketing...! »**

Mais Lindsay et al., (2009) → **« Oui, mais il y aura toujours 50% d'erreurs de reconnaissances en moins avec la SEQ ! »**

Le dilemme de la Justice

Favoriser la défense (SEQ)

Ou

Favoriser l'accusation (SIM) ?

*Malpass (2006) introduit la
notion **d'Utilité Sociale***

(Bentham, 1789)

L'utilité sociale

Malpass (2006) :

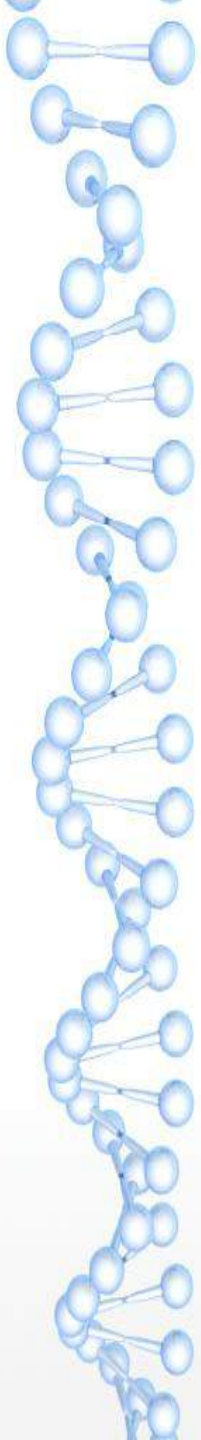
*« ne pas s'attacher à évaluer
les bénéfices respectifs des
modes de présentation, mais
en mesurer l'utilité pour la
société »*

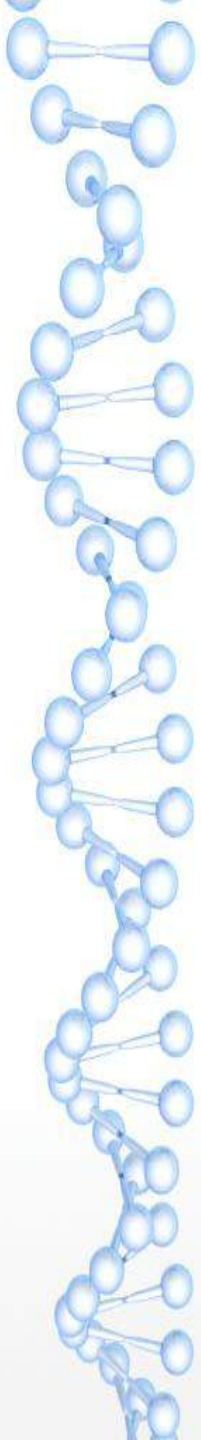
L'utilité sociale

Cela revient au dilemme suivant :

*Préférons-nous augmenter la probabilité de désignation au risque **d'augmenter** conjointement les **erreurs de reconnaissances**,*


*ou choisissons-nous d'éviter coûte que coûte d'accuser à tort et préférer **voir un criminel échapper à ses juges** ?*





L'examen des rapports d'analyse ADN

Éléments de réflexion et
d'interprétation...



N° 3121

ASSEMBLÉE NATIONALE

CONSTITUTION DU 4 OCTOBRE 1958

ONZIÈME LÉGISLATURE

Enregistré à la Présidence de l'Assemblée nationale
le 7 juin 2001

N° 364

SÉNAT

SESSION ORDINAIRE DE 2000-2001

Annexe au procès-verbal de la séance
du 7 juin 2001

**OFFICE PARLEMENTAIRE D'ÉVALUATION
DES CHOIX SCIENTIFIQUES ET TECHNOLOGIQUES**

RAPPORT

SUR

LA VALEUR SCIENTIFIQUE DE L'UTILISATION DES EMPREINTES GÉNÉTIQUES DANS LE DOMAINE JUDICIAIRE

par

M. Christian CABAL, Député

par M. Jean-Yves LE DÉAUT,
Premier Vice-Président de l'Office.

par M. Henri REVOL,
Président de l'Office.



Paragraphe 4.3. L'interprétation des résultats et l'utilisation des probabilités

On a précédemment indiqué qu'en cas de concordance entre une trace indiciaire inconnue et le profil génétique d'un suspect, l'expert a recours, pour évaluer la probabilité d'une coïncidence fortuite, à une base de données, établie au moyen d'une étude préliminaire de population, afin de connaître la fréquence d'apparition de chaque allèle dans la population générale.

Dans sa présentation, **l'expert se limite le plus souvent à la seule partie analytique de l'expertise** : il se contente de comparer les caractéristiques des traces retrouvées avec celles d'un suspect et de présenter la valeur de la fréquence de ces caractéristiques dans une population de référence. **Ainsi indiquera-t-il, par exemple, que la fréquence d'apparition d'une trace concordante avec le profil du suspect est de 1 personne sur 10 millions.**

Depuis plusieurs années, une abondante littérature scientifique estime que cette façon chiffrée de présenter la valeur de la preuve génétique doit être modifiée car elle ne permet pas véritablement aux juges de connaître la force probante du lien analytique établi entre la trace et le suspect.

Des procédures appropriées d'interprétation devraient donc, selon ces auteurs, être proposées et appliquées afin de faciliter la communication entre les mondes juridique et scientifique et d'éviter les pièges dus à l'intuition.

Sans entrer ici dans le détail des théories probabilistes, on s'en tiendra à quelques indications générales, renvoyant pour une présentation approfondie à la note que nous ont communiquée les professeurs Patrice MANGIN et Franco TARONI (cf. annexe n° 3).

Un consensus s'est établi depuis la fin des années 60 dans la communauté scientifique concernant une présentation de la preuve par l'ADN qui prémunisse contre ces erreurs d'interprétation. La solution réside dans l'application d'un modèle probabiliste, le théorème de Bayes et, plus spécifiquement, dans l'exploitation d'un **rapport de vraisemblance (« Likelihood Ratio ») intervenant dans le calcul Bayésien.**

Le modèle Bayésien permet de reconsidérer une mesure d'incertitude à propos de l'existence ou de la non-existence d'un fait sur la base d'une nouvelle information acquise. Cette approche est commune à tous les domaines scientifiques (notamment la médecine) où les données sont combinées avec des informations *a priori* afin de fournir des probabilités *a posteriori* sur l'existence ou la non-existence d'un fait particulier.

Cette chance *a posteriori* s'obtient en multipliant la chance *a priori* par le rapport de vraisemblance qui mesure la valeur de l'indice matériel.

L'expert n'a pas (à la différence du témoin) connaissance des circonstances particulières du cas et n'est donc pas en mesure d'évaluer correctement les chances *a priori* en faveur ou en défaveur des hypothèses posées par la Cour. Il ne peut s'exprimer que sur le rapport de vraisemblance.

Une évaluation complète doit nécessairement associer l'information analytique de l'expert et l'information, collectée pendant la phase d'enquête, sur les circonstances de l'affaire. L'approche Bayésienne clarifie les positions respectives du scientifique et du juge et définit leurs relations : l'expert se concentrera sur l'évaluation du rapport de vraisemblance et le juge se chargera de l'évaluation des chances *a priori* et *a posteriori*.

Dans une récente communication à la première conférence des utilisateurs de l'ADN organisée par Interpol à Lyon en novembre 1999, Raphaël COQUOZ, professeur à l'Institut de police scientifique et de criminologie de Lausanne, estimait que **les lacunes les plus importantes affectant le fonctionnement des laboratoires ADN concernaient le domaine de l'interprétation des preuves.** « *La capacité à exploiter les résultats dans un schéma rigoureux (tel que l'offre l'interprétation Bayésienne) est encore loin de l'optimal* ».

[...]

..., Raphaël COQUOZ, [...], estimait que **les lacunes les plus importantes affectant le fonctionnement des laboratoires ADN concernaient le domaine de l'interprétation des preuves.**

« La capacité à exploiter les résultats dans un schéma rigoureux [...] est encore loin de l'optimal ».

[...]



Quelques rappels et éléments de connaissances relatifs aux analyses ADN

Principes d'une analyse d'ADN

On s'intéresse au caractère **polymorphe** de certains secteurs des brins d'ADN (loci) → recherche des **variantes** : **les allèles**

Les méthodes actuelles reposent sur la **PCR** (*Polymérase Chain Reaction*) qui permet de dupliquer (amplifier) les éléments d'intérêt

1 cycle → 2 copies

10 cycles → 1024 copies

20 cycles → environ 1 million de copies

30 cycles → plus d'1 milliard de copies



Un cas pratique...

Crédit : Annexe 3 du rapport 3121 de l'Assemblée Nationale par M. le Député Cabal (2001) ; auteurs de l'Annexe 3 : F. Taroni et P. Mangin

Voici une conclusion typique d'un rapport d'expertise d'une analyse ADN:

« il n'y a pas de différence entre les profils génétiques établis sur le prélèvement et ceux établis chez le suspect. Une telle constellation de caractéristiques se retrouve chez environ 0,000001% (c'est-à-dire 1 chance sur 10 millions) de la population »

(note : 1 / 1 000 000 → max. technique d'identification RFLP ; 1 / 1 000 000 000 pour la méthode PCR)

Que peut-on conclure à partir de cette phrase ?



Un cas pratique...

Crédit : Annexe 3 du rapport 3121 de l'Assemblée Nationale par M. le Député Cabal (2001) ; auteurs de l'Annexe 3 : F. Taroni et P. Mangin

On peut logiquement émettre le raisonnement suivant :

La fréquence d'apparition de cette trace, concordante avec le profil du suspect, est de 1 sur 10 millions...

Donc la probabilité de trouver cette trace là si quelqu'un d'autre que le suspect l'a laissée est de 1 sur 10 millions...

Donc la probabilité que quelqu'un d'autre laisse cette trace est de 1 sur 10 millions...

Par conséquent, on peut être sûr à 99, 9999% que le suspect a laissé cette trace !



Un cas pratique...

Crédit : Annexe 3 du rapport 3121 de l'Assemblée Nationale par M. le Député Cabal (2001) ; auteurs de l'Annexe 3 : F. Taroni et P. Mangin

On peut logiquement émettre le raisonnement suivant :

La fréquence d'apparition de cette trace, concordante avec le profil du suspect, est de 1 sur 10 millions...

Donc la probabilité de trouver cette trace chez quelqu'un d'autre que le suspect est de 1 sur 10 millions...

Donc la probabilité que quelqu'un d'autre laisse cette trace est de 1 sur 10 millions...

**Conclusion à mettre
au profit de l'accusation**

Par conséquent, on peut être sûr à 99, 9999% que le suspect a laissé cette trace !



Mais quelle interprétation de cette valeur peut faire la **DEFENSE** ?

La fréquence d'apparition de cette trace, concordante avec le profil du suspect, est de 1 sur 10 millions

Donc la probabilité de trouver cette trace là si quelqu'un d'autre que le suspect l'a laissée est de 1 sur 10 millions...

Donc la probabilité que quelqu'un d'autre laisse cette trace est de 1 sur 10 millions...

Par conséquent, on peut être sûr à **99,9999%** que le suspect a laissé cette trace !



Mais quelle interprétation de cette valeur peut faire la **DEFENSE** ?

La fréquence d'apparition de cette trace, concordante avec le profil du suspect, est de 1 sur 10 millions

Donc la probabilité de trouver cette trace là si quelqu'un d'autre que le suspect l'a laissée est de 1 sur 10 millions...

Donc la probabilité que quelqu'un d'autre laisse cette trace est de 1 sur 10 millions...

Par conséquent, on peut être sûr à **99,9999%** que le suspect a laissé cette trace !

La fréquence d'apparition de cette trace, concordante avec le profil du suspect, est de 1 sur 10 millions

Un tel profil génétique se trouve chez environ 1 individu sur 10 millions

En considérant une population d'intérêt d'environ 20 000 000 d'individus en France qui peuvent avoir laissé cette trace, il y a environ 2 individus qui présentent le même profil génétique

Par conséquent, la probabilité que la trace ait été laissée par le suspect est de 1 sur 2

Nous avons donc **50% de chance de nous tromper**



Mais quelle interprétation de cette valeur peut faire la **DEFENSE** ?

La fréquence d'apparition de cette trace, concordante avec le profil du suspect, est de 1 sur 10 millions

Donc la probabilité de trouver cette trace là si quelqu'un d'autre que le suspect l'a laissée est de 1 sur 10 millions...

Donc la probabilité que quelqu'un d'autre laisse cette trace est de 1 sur 10 millions...

Par conséquent, on peut être sûr à **99,9999%** que le suspect a laissé cette trace !

La fréquence d'apparition de cette trace, concordante avec le profil du suspect, est de 1 sur 10 millions

Un tel profil génétique se trouve chez environ 1 individu sur 10 millions

En considérant une population d'intérêt d'environ 20 000 000 d'individus en France qui peuvent avoir laissé cette trace, il y a environ 2 individus qui présentent le même profil génétique

Par conséquent, la probabilité que la trace ait été laissée par le suspect est de 1 sur 2

Nous avons donc **50% de chance de nous tromper**

[...]

Des procédures appropriées d'interprétation devraient donc, selon ces auteurs, être proposées et appliquées afin de faciliter la communication entre les mondes juridique et scientifique et **d'éviter les pièges dus à l'intuition.**

[...]



Un autre exemple...

(State of Arizona vs Robert Wayne Johnson)

« Le pull-over de la victime a été examiné et le test ADN a montré que les traces s'accordaient avec le profil génétique du suspect. L'expert a calculé la possibilité d'une coïncidence génétique due au hasard dans la population, c'est-à-dire la probabilité que deux personnes non associées par des liens de parenté puissent avoir les mêmes caractéristiques génétiques, comme étant de **1 sur 312 millions** »

Un autre exemple...

(State of Arizona vs Robert Wayne Johnson)

« Le pull-over de la victime a été examiné et le test ADN [...], la probabilité que deux personnes non associées par des liens de parenté puissent avoir les mêmes caractéristiques génétiques, comme étant de 1 sur 312 millions »

Ainsi, la valeur de la probabilité présentée par l'expert devrait être, **si l'on veut respecter les lois des probabilités...**

$$P \text{ (existence d'une autre correspondance)} = 1 - (1 - f)^n$$

Probabilité qu'une autre personne sélectionnée de façon aléatoire dans la population générale (ou de référence) ne partage pas les caractéristiques du suspect



Un autre exemple...

(State of Arizona vs Robert Wayne Johnson)

« Le pull-over de la victime a été examiné et le test ADN [...], la probabilité que deux personnes non associées par des liens de parenté puissent avoir les mêmes caractéristiques génétiques, comme étant de 1 sur 312 millions »

Considérons que l'enquête de police ait permis de réduire la population potentielle à la taille d'une ville d'1 million d'habitants

→ **$n = 1\ 000\ 000$**

Probabilité pour qu'il existe dans la population concernée une autre personne qui possède les mêmes caractéristiques :

$$1 - [1 - (1 / 312\ 000\ 000)]^{1\ 000\ 000} = \mathbf{3,2 / 1\ 000}$$



Un autre exemple...

(State of Arizona vs Robert Wayne Johnson)

« Le pull-over de la victime a été examiné et le test ADN [...], la probabilité que deux personnes non associées par des liens de parenté puissent avoir les mêmes caractéristiques génétiques, comme étant de 1 sur 312 millions »

On passe de ...

1 chance sur 312 millions

à

3,2 chances sur 1 000



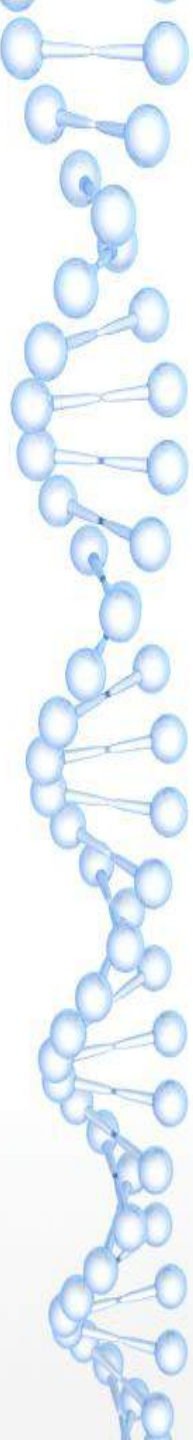
Un autre exemple...

(State of Arizona vs Robert Wayne Johnson)

« Le pull-over de la victime a été examiné et le test ADN [...], la probabilité que deux personnes non associées par des liens de parenté puissent avoir les mêmes caractéristiques génétiques, comme étant de 1 sur 312 millions »

L'expert a présenté à la Cour un raisonnement fallacieux que les mathématiciens appellent « **erreur de la probabilité d'une autre correspondance** »,

raisonnement au profit de l'accusation...



internet spécialisés et aux localisations téléphoniques de l'intéressé, Toufik Mekki aurait été, d'après elle, dans « l'impossibilité matérielle » de participer à l'agression.

Depuis leurs interpellations, les trois accusés se déclarent étrangers à ces faits.

L'accusation s'est appuyée sur un ensemble d'éléments suspects et un indice matériel, une trace ADN attribuée à Stéphane Oltra, relevée sur la scène de crime de Vimines. L'expert a mesuré à 1/45000 la probabilité qu'un autre puisse correspondre à ces caractéristiques ADN. « C'est plus que suffisant, accablant » selon l'avocat général. « À la probabilité de 1/7 milliards d'un profil complet, c'est une certitude. À celle-ci, d'un pro-

fil faible, c'est une éventualité qui n'est pas confortée par un autre élément de preuve », a protesté M^e Abad. L'avocat général a affirmé sa « certitude » de la culpabilité. « On ne peut pas forcer les pièces d'un puzzle », a insisté M^e Roux, en relevant « les incohérences, les erreurs et les interprétations » dans l'affaire.

M^e Abad a avoué sa crainte vis-à-vis d'un dossier « qui fait peur : frelaté, approximatif, incertain, sans preuve tangible, avec des apparences trompeuses, des fausses vérités et un doute extraordinaire. » Les accusés vont reprendre la parole ce matin, puis les jurés et la cour se retireront pour délibérer. Verdict dans la journée.

Frédéric CHIOLA



Mais alors, que peut-on conclure ? Qui a raison et qui a tort ? Et pourquoi ?

Il faut d'abord se poser la question :

« À quoi correspond cette valeur 1/312 000 000 ? »



fréquence relative d'apparition du profil génétique dans le population de référence !

(équivalent à un lancé de dés => *« au prochain lancé, j'ai 1 chance sur 6 d'obtenir un « 3 »... »*)

→ événements répétitifs et indépendants



Mais alors, que peut-on conclure ? Qui a raison et qui a tort ? Et pourquoi ?

Aucun intérêt pour établir la vérité judiciaire !

Mais elle présente 2 défauts majeurs :

(1) Crée des confusions

(2) Ne répond pas aux attentes de la Cour

→ *Un juge n'a que faire de savoir que cette trace est présente dans x % de la population... Ce qu'il veut savoir, c'est en quoi cette analyse accuse ou disculpe l'accusé !*



Mais alors, que peut-on conclure ? Qui a raison et qui a tort ? Et pourquoi ?

Tout le monde, l'accusation comme la défense, les enquêteurs comme les experts...

Parce que tout le monde commet des erreurs de raisonnement dans ce type de situation très contre intuitive !

Rien, ou pas grand-chose...



L'interprétation correcte d'un rapport d'analyse d'ADN...

Repose AVANT TOUT sur **une utilisation adaptée des probabilités...**

Mais les psychologues expérimentalistes ont montré que ***nous ne savons pas intuitivement exploiter*** correctement des données statistiques ou probabilistes...

On commet de nombreuses erreurs de raisonnement...

→ *c'est ce qui nous fait dire : « les probabilités, ça marche jamais !!!... »*



Biais de confirmation

La tâche des 4 cartes de Wason

- *Quatre cartes comportant un chiffre sur une face et une lettre sur l'autre, sont disposées à plat sur une table.*
- *Une seule face de chaque carte est visible.*
- *Les faces visibles sont les suivantes : **D, 7, 5, K.***
- *Quelle(s) carte(s) devez-vous retourner pour déterminer la ou les carte(s) qui ne respecte(nt) pas la règle suivante : **Si une carte a un D sur une face, alors elle porte un 5 sur l'autre face.***
- *Il ne faut pas retourner de carte inutilement, ni oublier d'en retourner une.*



Biais de confirmation

La tâche des 4 cartes de Wason

D

7

5

K

Si une carte a un D sur une face, alors elle porte un 5 sur l'autre face.



Biais de confirmation

La tâche des 4 cartes de Wason

Solution

retourner les deux cartes portant
la lettre **D** et le numéro **7**



Biais de confirmation

La tâche des 4 cartes de Wason

*Quatre personnes sont en train de boire dans un bar et vous disposez des informations suivantes : la **première** boit une boisson alcoolisée, la **seconde** a moins de 18 ans, la **troisième** a plus de 18 ans et la **dernière** boit une boisson sans alcool.*

Quelle(s) personne(s) devez-vous interroger sur leur âge ou sur le contenu de leur verre pour vous assurer que tous respectent bien la règle suivante : **Si une personne boit de l'alcool, elle doit avoir plus de 18 ans.**



Biais de confirmation

La tâche des 4 cartes de Wason

Quatre personnes sont en train de boire dans un bar et vous disposez des informations suivantes : la **première** boit une boisson alcoolisée, la **seconde** a moins de 18 ans, la **troisième** a plus de 18 ans et la **dernière** ne boit pas d'alcool.

Vous pouvez interroger sur leur âge ou sur la présence ou non d'alcool en levant la carte de la table. **l'habillage de la tâche nous perturbe ou pas...**

resp. la règle suivante : **Si une personne boit de l'alcool, elle doit avoir plus de 18 ans.**



Biais de confirmation

Comment en diminuer l'impact ?

En utilisant des **protocoles** précis pour prendre une décision :

- Ainsi la décision – **foncièrement probabiliste** – ne sera plus entachée du biais de confirmation...
- Mais comme elle reste probabiliste, **elle ne sera pas forcément juste**



Biais de confirmation

Comment en diminuer l'impact ?

En utilisant des **protocoles** précis pour prendre une décision :

Qui en utilise ?

- les médecins...
- les scientifiques...



Gérer les probabilités...

Comment gère-t-on les
informations statistiques et
probabilistes ?



Gérer les probabilités...

Evaluons Linda...

“Linda a 31 ans, parle couramment des langues étrangères, et est très brillante. Elle est diplômée en philosophie. Lorsqu’elle était étudiante, elle était très concernée par les problèmes de discriminations et de justice sociale, et elle a participé à plusieurs manifestations anti-nucléaire.”

Exercice : ranger les 8 propositions suivantes des moins probables aux plus probables



Gérer les probabilités...

- (a) Linda est professeur dans une école élémentaire
- (b) Linda travaille dans un magasin et prend des cours de yoga
- (c) Linda est très active dans un mouvement féministe
- (d) Linda est travailleur social dans un centre psychiatrique
- (e) Linda est membre de la ligue pour le droit des Femmes
- (f) Linda est libraire
- (g) Linda vend des assurances
- (h) Linda est libraire et très active dans un mouvement féministe



Gérer les probabilités...

- (a) Linda est professeur dans une école élémentaire
- (b) Linda travaille dans un magasin et prend des cours de yoga
- (c) Linda est très active dans un mouvement féministe
- (d) Linda est travailleur social dans un centre psychiatrique
- (e) Linda est membre de la ligue pour le droit des Femmes
- (f) Linda est libraire
- (g) Linda vend des assurances
- (h) Linda est libraire et très active dans un mouvement féministe



Gérer les probabilités...

Kahneman & Tversky (1982):

89% ont jugé que (h) était plus probable que (f), malgré qu'on ne puisse pas être (f) sans être (h)...

85% des étudiants du cours de science de décision de la Stanford Business School ont donné la même réponse



Gérer les probabilités...

Kahneman & Tversky (1982):

89% ont jugé que (h) était plus probable que (f), malgré qu'on ne puisse pas être f sans être h...

85% des étudiants du cours de science de décision de la Business School ont donné la même

Sophisme de conjonction



Négligence du taux de base dans les jugements probabilistes

Casscells et al. (1978)

Etudiants de la Medical Harvard School

*Si un test utilisé pour détecter une maladie dont la **prévalence est de 1/1000** a un taux de **faux positifs de 5%**.*

quel est le risque qu'une personne chez qui on trouve un résultat positif ait réellement la maladie,
en supposant que l'on ne sait rien sur ses symptômes?



Négligence du taux de base dans les jugements probabilistes



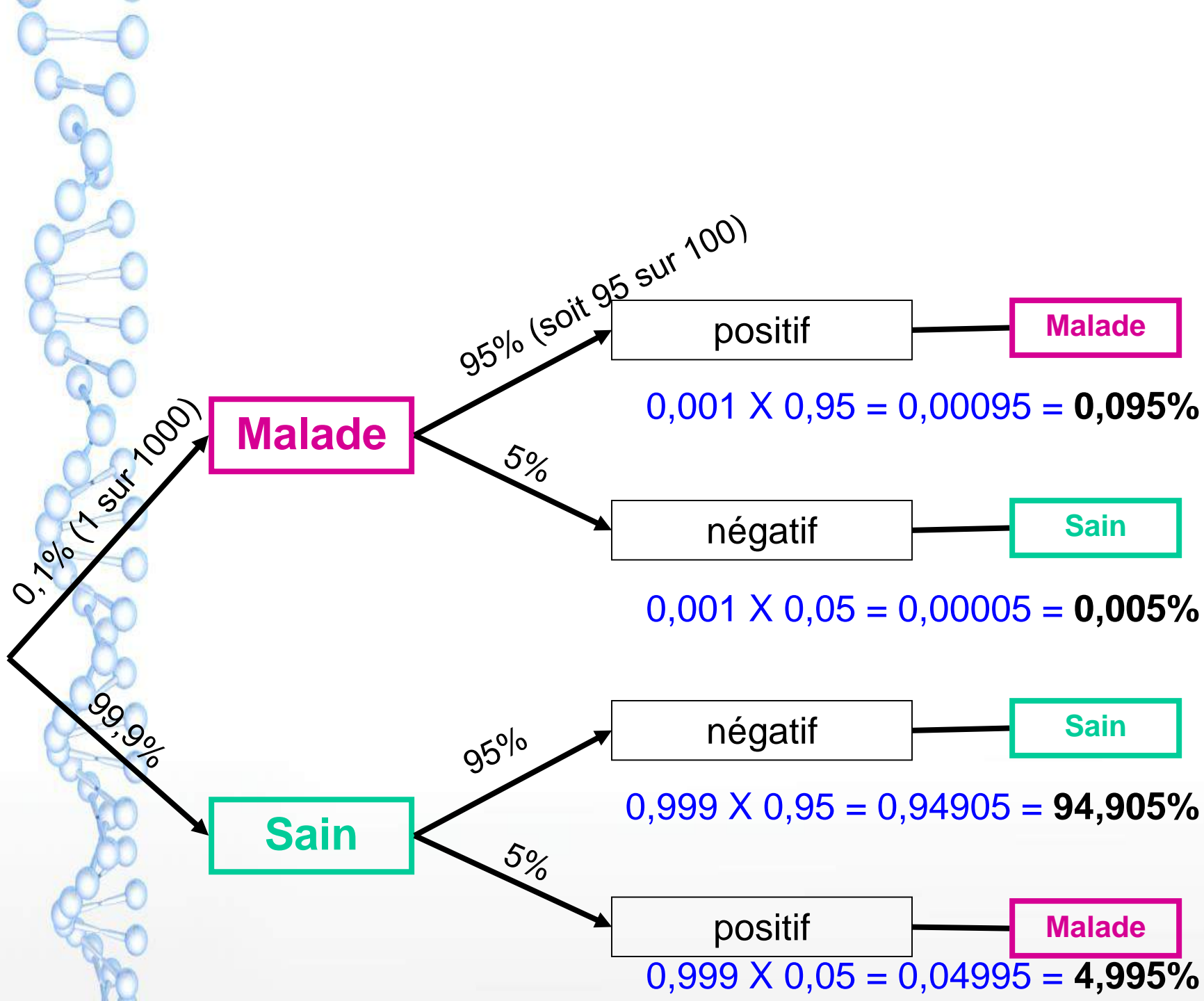
« bin, 2% » (soit $0.001 / 0.05 = 0.02$)

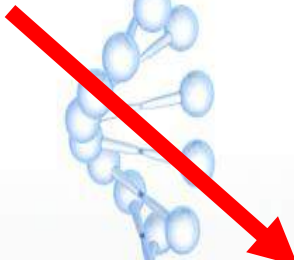
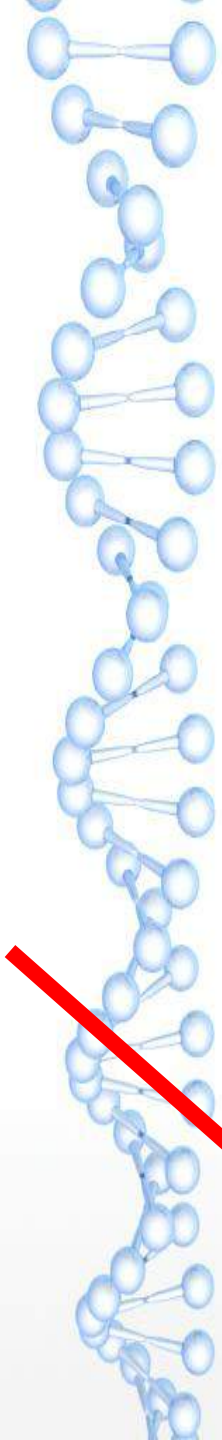
→ **18%** des étudiants



« bin, 95% »

→ **45%** des étudiants





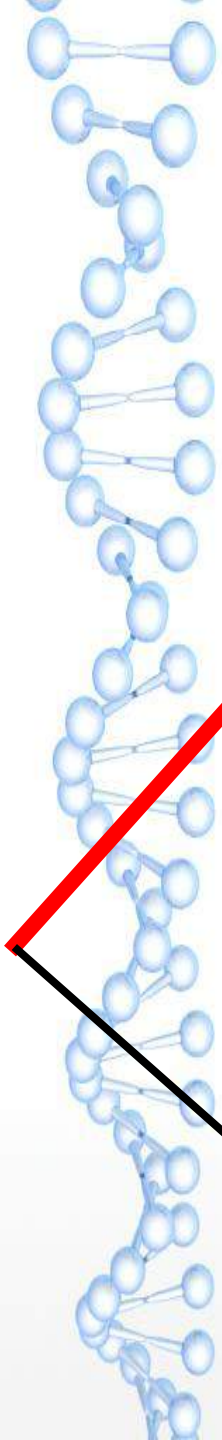
Sain



positif

FAUX POSITIF

Malade



Malade

positif

Malade

0,095%

$$\begin{aligned} &= 0,00095 / (0,00095 + 0,04995) \\ &= 0,00095 / 0,0509 \\ &= 0,018664 \end{aligned}$$

soit 1,87%

Sain

positif

Malade

4,995%

ATTENTION, ne pas confondre...

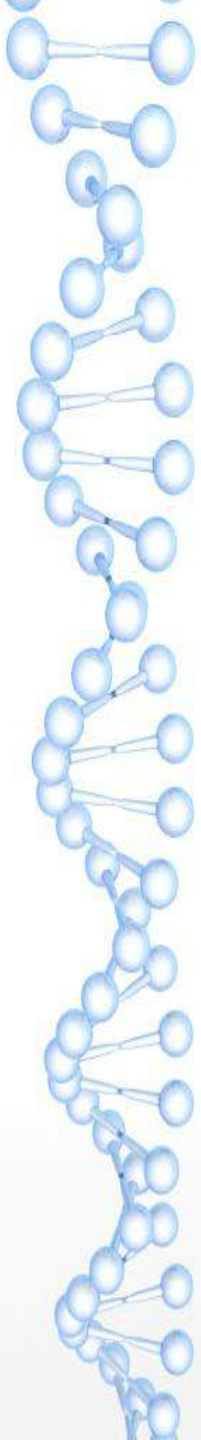
Probabilité (être malade | le test est POSITIF)

1,87%

\neq

Probabilité (le test est POSITIF | être malade)

95%



Mais alors, que faire, comment le faire et qui doit le faire ?



Mais alors, que faire, comment le faire et qui doit le faire ?

Utiliser la logique BAYESIENNE

- permet de gérer correctement les informations issues des rapports d'expertise et de l'ensemble des éléments de l'affaire
- permet de répartir les rôles de chacun (magistrats, experts, enquêteurs,...)
- Compatible avec l'expression de l'intime conviction
(le magistrat n'est pas « enfermé » dans un protocole rigide qui le « guide »)



Mais alors, que faire, comment le faire et qui doit le faire ?

Le théorème de Thomas BAYES,

révérend, pasteur et mathématicien Anglais (1702-1760)

*Ce théorème permet de **réviser** une mesure d'incertitude à propos de la véracité d'une hypothèse ou d'une proposition (« **c'est le criminel** ») suite à de nouvelles informations (ex. **ADN**)*



Mais alors, que faire, comment le faire et qui doit le faire ?

Le théorème de Thomas BAYES,

révérend, pasteur et mathématicien Anglais (1702-1760)

Nécessité d'exploiter un

Rapport de Vraisemblance (RV)

[...]

Un consensus s'est établi depuis la **fin des années 60** dans la communauté scientifique concernant une présentation de la preuve par l'ADN qui prémunisse contre ces erreurs d'interprétation.

La solution réside dans l'application d'un modèle probabiliste, le théorème de Bayes et, plus spécifiquement, dans l'exploitation d'un **rapport de vraisemblance** (« *Likelihood Ratio* ») intervenant dans le **calcul Bayésien**

[...]



Le rapport de vraisemblance (RV)

$$RV = \frac{p(\text{concordance des profils ADN} \mid \text{le suspect est le criminel})}{p(\text{concordance des profils ADN} \mid \text{le suspect est innocent})}$$

permet d'analyser les données au regard de **deux hypothèses fortement compatibles** avec les obligations policières d'instruction à charge et à décharge (Utilisé pour la première fois aux USA dans l'affaire People V Collins)



Le rapport de vraisemblance (RV)

$$RV = \frac{p(\text{concordance des profils ADN} \mid \text{le suspect est le criminel})}{p(\text{concordance des profils ADN} \mid \text{le suspect est innocent})}$$

RV > 1 → l'hypothèse de l'**accusation** semble la plus probable

RV < 1 →, l'hypothèse de la **défense** semble la plus probable

RV = 0 → preuve compatible avec les 2 hypothèses

Le calcul BAYESIEN

Prob. A priori

X

Rapport de
Vraisemblance
(RV)

=

Prob. A
posteriori

Cote en faveur de l'hypothèse de l'accusation au vu des éléments du dossier autres que l'ADN

→ Dépend exclusivement de la qualité du travail des ENQUETEURS

Cote fournie en faveur de l'hypothèse de l'accusation au regard de la seule analyse d'ADN

→ Fournie par l'EXPERT

C'est la **cote** en faveur de l'accusation calculée au vu de **TOUS les éléments** du dossier

→ Ce qui intéresse les **MAGISTRATS et JURES**

Probabilité <i>a priori</i> que le suspect soit le coupable au vu du dossier (témoignages, autres analyses...)			RV (fourni par l'expert)	Probabilité <i>a posteriori</i> que le suspect soit à l'origine de la trace , calculée par les magistrats et les jurés = (a priori X RV)/((a priori X RV)+1)	
<i>1 chance sur...</i>	<i>en fréquence</i>	<i>en %</i>		<i>en fréquence</i>	<i>en %</i>
1 000 000 000	0,000000001	0,0000001%	10 000 000	0,00990	0,990099%
100 000 000	0,000000010	0,000001%	10 000 000	0,09091	9,09091%
10 000 000	0,000000100	0,00001%	10 000 000	0,50000	50%
1 000 000	0,000001000	0,0001%	10 000 000	0,90909	90,9091%
100 000	0,000010000	0,001%	10 000 000	0,99010	99,0099%
10 000	0,000100000	0,01%	10 000 000	0,99900	99,9001%
1 000	0,001000000	0,1%	10 000 000	0,99990	99,9900%
100	0,010000000	1%	10 000 000	0,99999	99,9990%
10	0,100000000	10%	10 000 000	1,00000	100%
1	1,000000000	100%	10 000 000	1,00000	100%