GC×GC-MS AND BAYESIAN TESTING IN FORENSICS: TOWARDS THE IDENTIFICATION OF SUSPECTS THROUGH THEIR ODOR

Isabelle RIVALS[1], Vincent CUZUEL[2], Guillaume COGNON[2], Roman Lecon
Didier THIEBAUT[3], Charles SAULEAU[2], Jérôme VIAL[3]

Équipe de Statistique Appliquée, UMRS 1158, ESPCI Paris, France
nstitut de Recherche Criminelle de la Gendarmerie Nationale, Cergy-Pontoise, France
UMR CBI 8231 - Laboratoire Sciences Analytiques, Bioanalytiques et Miniaturisation – ESPCI Paris – PSL Research University

- GCxGC 2018, Riva Del Garda, Italy

CONTEXT

Popularization of the echniques used by the police

Criminals are more attentive and cautious!

Human odor

USE OF TRAINED DOGS

- Sufficient for identification of a person
- Limited probative value in courts of justice

- Need for corroborative evidence by analytical tools:
  - **Support the information** provided by dogs
  - **Probative value** to evidence in courts of justice

OBJECTIVES

Develop a global strategy to characterize the olfactory ngerprints of individuals using **analytical and statistical tools**

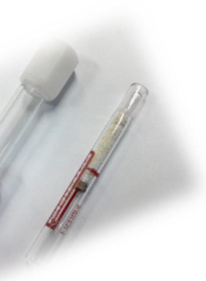
- Volatile compounds **at trace levels**: preconcentration step required
- **Complex mixtures**: multidimensional separation (GC×GC-MS)

Question to be answered

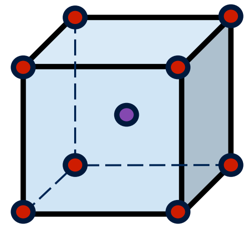
- Is the comparison of an “odor” reference chromatogram to a chromatogram obtained using an odor sample from a suspect (crime scene…) sufficient to prove that the odor belongs to the same person?

Global Strategy

Sampling/**Panel** → Separation and Detection → **Data Processin**



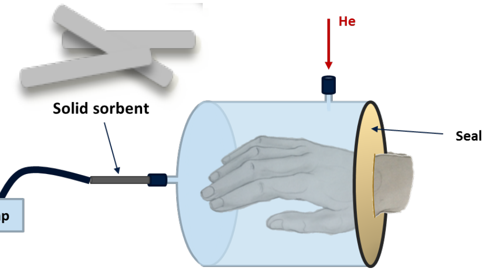
$$P(A|B) = \frac{P(B|A)\ P(A)}{P(B)}$$



PRECONCENTRATION AND ANALYSIS: PURGE AND TRAP - GCxGC

Direct sampling

Indirect sampling



Thermodesorption
coupled with
GC×GC-MS
VSP4000, Action Europe (Sausheim, France)

DB1MS-DB1701

2°C/min – modulation 8 s

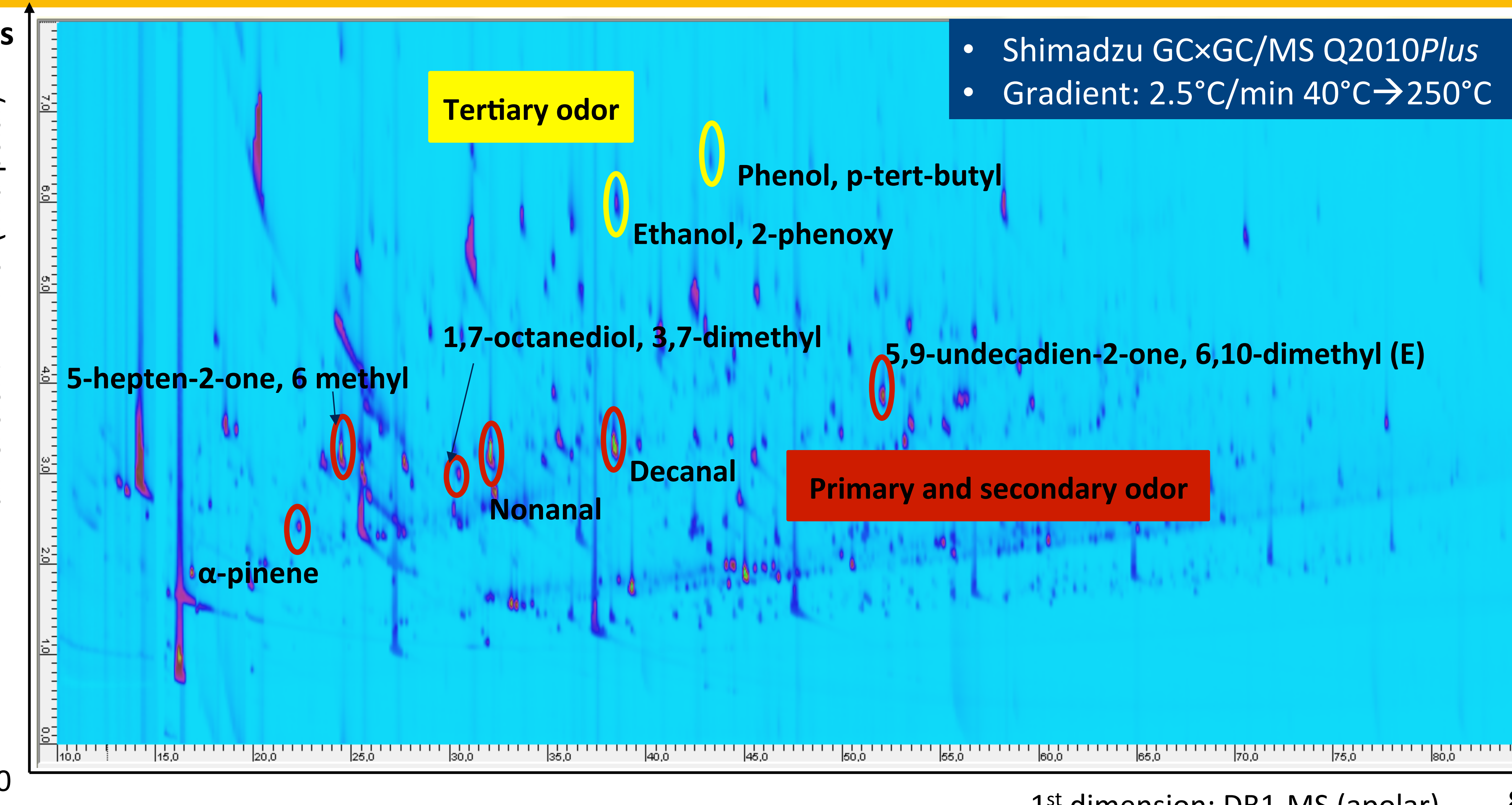
Desorption optimization DC

- **synthetic mixture** of huma odor (80 compounds [1])
- **full factorial design** 2^4

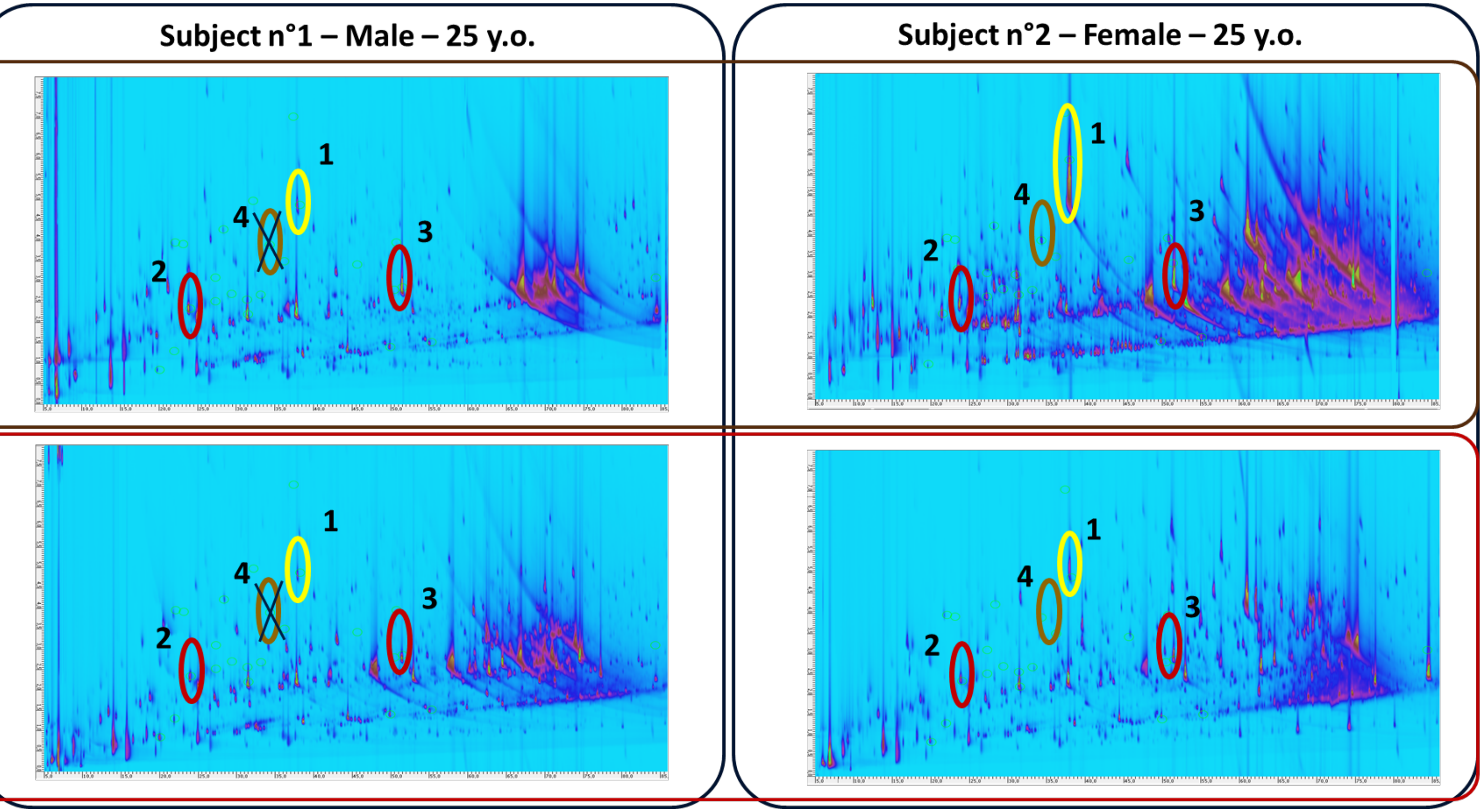
Sample temperature= 190°
Purge flow = 20 mL/min
Purge time= 20 min
Split = 0 mL/min

1- *Cuzuel et al.*, **A review: Origin, analytical characterization and use of human hands odor in forensics**, *2017, Journal of Forensic Sciences*

Chromatogram of a real sample



COMPARISON OF REAL SAMPLES?



- Complex sample
- **Comparison is n trivial**
- A lot of data to process

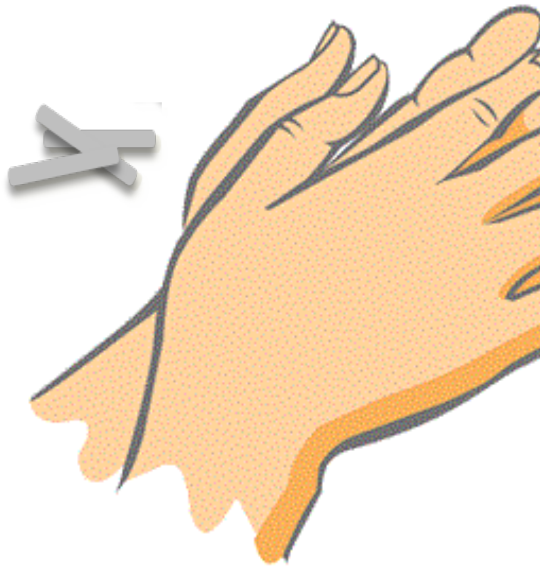
Need for an automated data processing to extract relevant information
Need for a panel of persons to evaluate the strategy

Chromatograms Of Real Samples: Panel

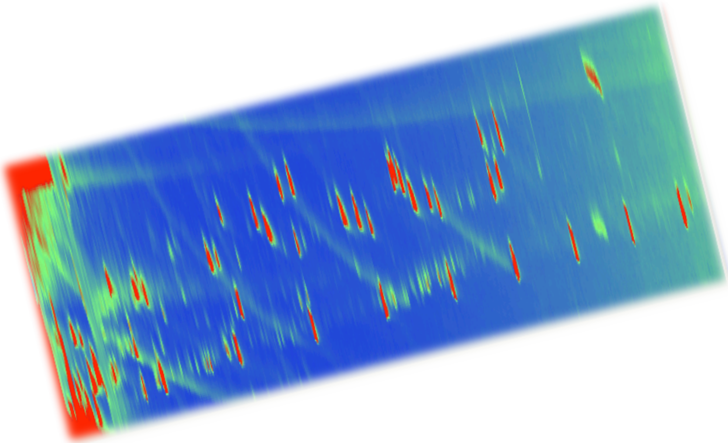
Panel of 119 persons

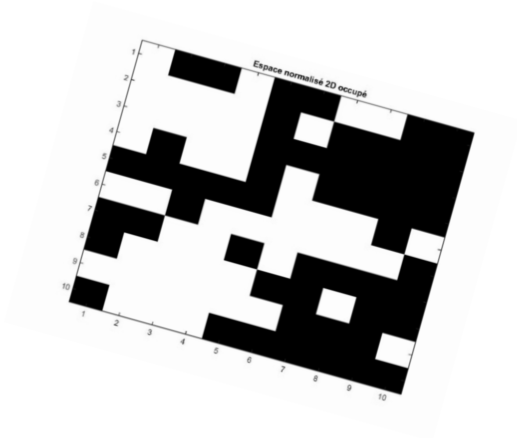
	gender		age (years)			phototype		
total	♂	♀	10-23	24-36	37-81	1	2	3
119	61	58	39	39	41	25	79	15

- Phototype 1 - skin is sun sensitive and does not burn
- Phototype 2 - intermediate skin
- Phototype 3 – well tanning skin
- 4 direct samplings of hands/person (Sorb-star®)
  - 15 minutes
  - Blank (sampling room)
- TD* - GC×GC-MS**
- 3 chromatograms/person

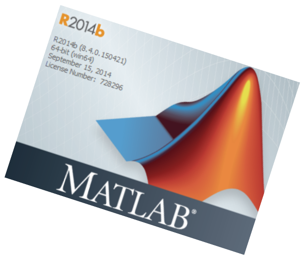


et al., **Sampling method development and optimization in view of human hand odor analysis by thermal desorption coupled with gas chromatography and mass spectrometry**, 2017, Anal. Bioanal. Chem.
l et al., **Human odor and forensics. Optimization of a comprehensive gas chromatography method based on orthogonality: how not to choose between criteria.**, 2017, Journal of Chromatography A





DATA PROCESSING / BAYESIAN APPROAC



CHROMATOGRAPHIC DATA PROCESSING WITH MATLAB

nversion of files	Pre-treatment	Detection of peaks	Transfer of data in the librairies	Treatment using statistics
Export in mpatible format (mzXML) **port** of data to Matlab	• Baseline correction • Selection of investigated zones	• Detection of local maxima • Extraction of associated informations (alkanes, ^{1}t, ^{2}t, LRI, MS spectrum, name)	• Import to NIST and **own** library (3 persons) **(>600** compounds**)** • **Library update** (known odor compounds*) • Identification and peak assignation	• Frequentist Approach • **Bayesian Approach**

hromatogram ⟺ 1 vector corresponding to 600 compounds peak intens

* Cuzuel et al., **Origin, analytical characterization and use of human odor in forensics**, 2017, *J. Forensic Sci.*

BAYESIAN APPROACH (*A POSTERIORI*)

: the two chromatograms are obtained from the same person

: the two chromatograms are NOT obtained from the same person

) represent the observed data (the two chromatograms), Bayes formula gives :

$$P(H_0 | D) = \frac{f(D | H_0)P(H_0)}{f(D | H_0)P(H_0) + f(D | H_1)P(H_1)}$$

otocole:

efinition of a **distance *d* between 2 chromatograms** (D ≡ d)

anel of chromatograms of individuals (119 persons sampled 4 times) splitted in **lependent calibration and test groups**

alibration group ➔ estimation of **distributions of *d*** for couples of chromatograms fron me person $f(d|H_0)$ and from different persons $f(d|H_1)$

est group ➔ estimation of **performance** (AUC, sensitivity, spécificity)

BAYESIAN APPROACH: CHOICE OF DISTANCE BETWEEN CHROM

Estimation of the statistical likelihood

Options :

d : distances between 600-vectors of intensities :

a) **euclidian distance**

b) **1 − Pearson correlation coefficient**

c) **1 − Spearman correlation coefficient**

• intensities **normalized** / **binarized** (b=c)

Calibration group (260 chromatograms / 75 persons)

341 couples of chromatograms for H_0 (same person)

33 329 couples de chromatograms for H_1 (différent persons)

➔ **histograms** of *d* values for H_0 and H_1

Ajustment of histograms using several **gaussian curves**

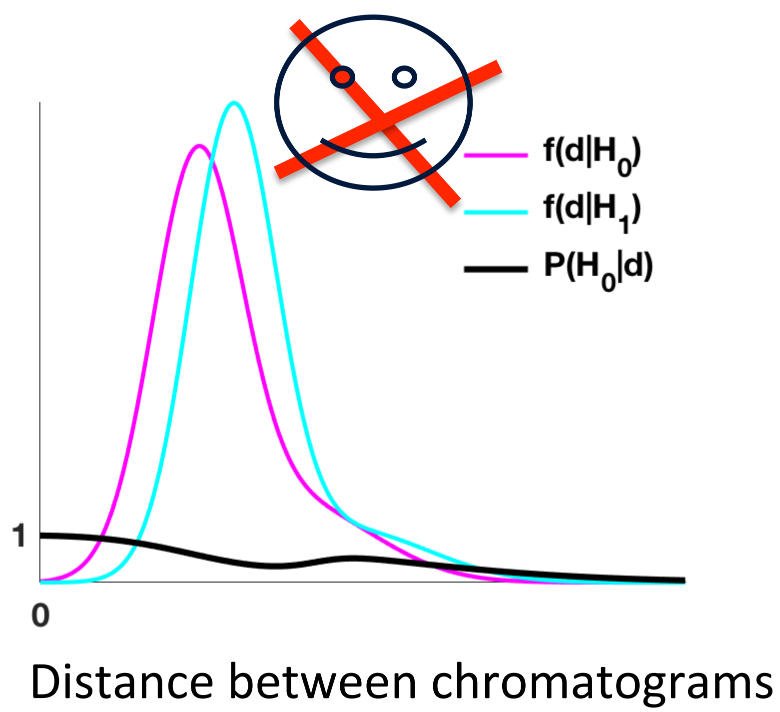
➔ $f(d|H_0)$ and $f(d|H_1)$

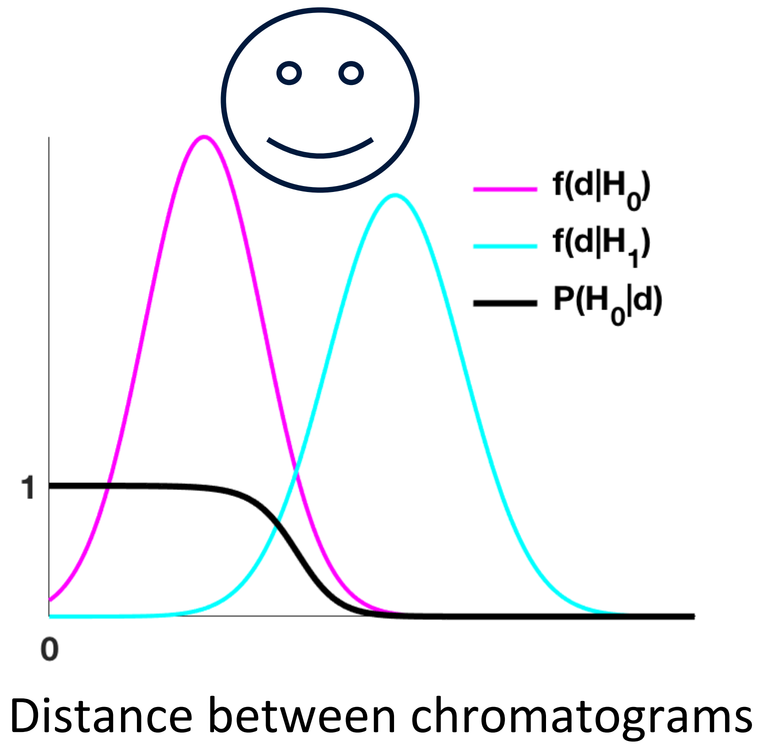
BAYESIAN APPROACH: EXPECTED RESULTS

- Probabilities *a priori* : $P(H_0) = P(H_1) =$ **0.5**

$$P(H_0\,|\,d) = \frac{f(d\,|\,H_0)P(H_0)}{f(d\,|\,H_0)P(H_0) + f(d\,|\,H_1)P(H}$$

- Fictitious examples of statistical likelihood

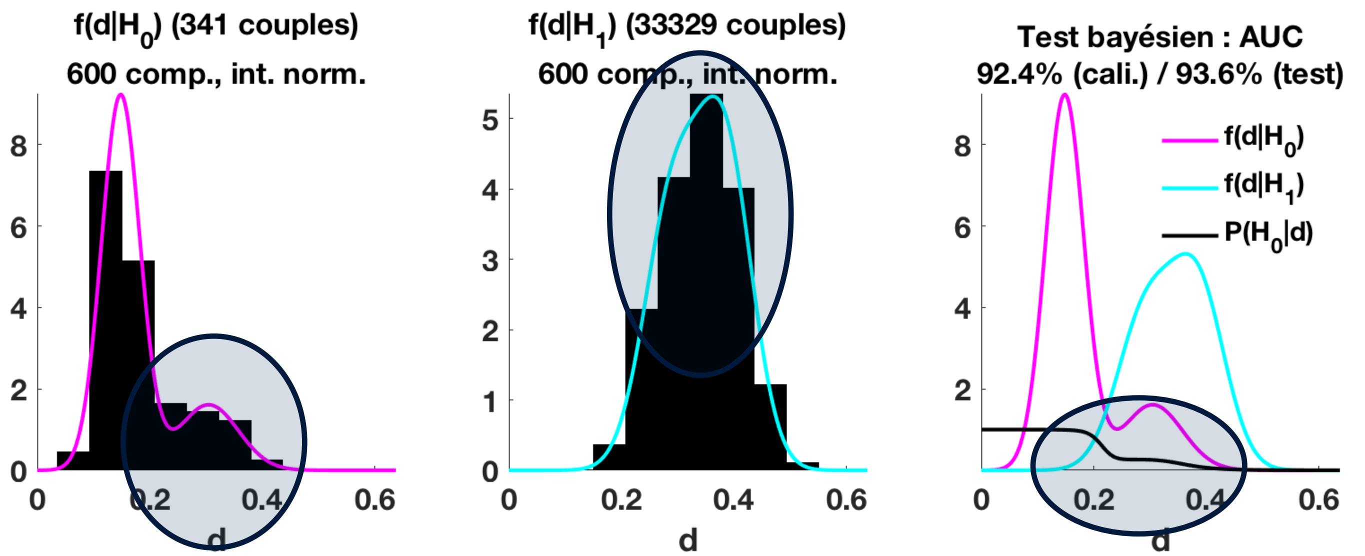




Bayesian Approach: Results Using 600 Compounds

(%AUC calibration / %AUC test)

intensities \ distance	euclidian	$1 - \rho_{Pearson}$	$1 - \rho_{Spearman}$
normalized	62.7% / 64.6%	74.6% / 74.7%	92.4% / 93.6%
binarized	88.4% / 91.6%	89.6% / 91.7%	

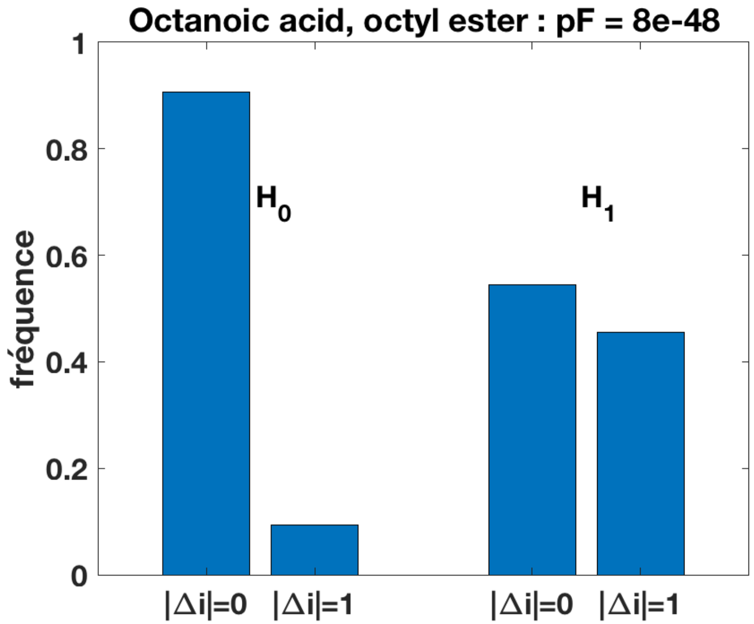


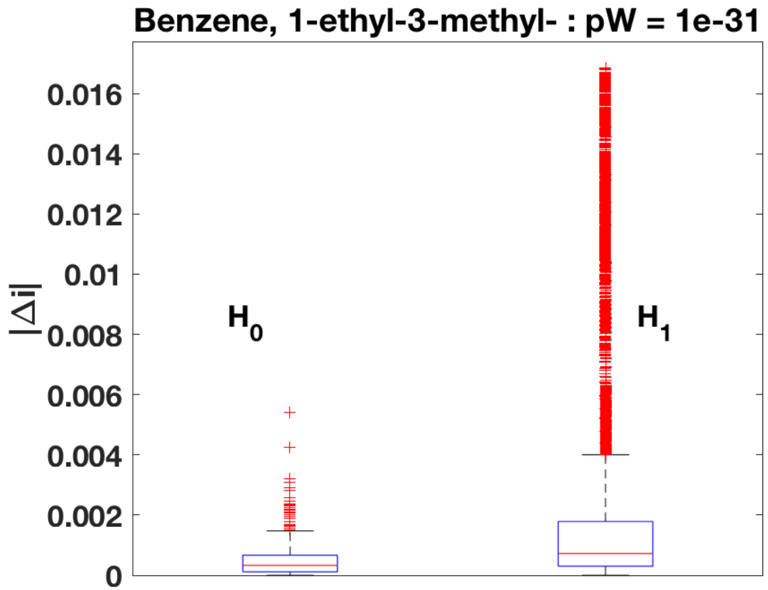
2 modes!

N.B. using the test group, there are 173 / 9 418 couples for H_0 / H_1 respectively

BAYESIAN APPROACH: DISCRIMINATING COMPOUNDS

- Discriminating compounds for H_0 and H_1 : those which **intensity differences** |Δi| are **significantly lower for H_0 than H_1**
- Quantification : ***p*-value** using **unilateral Fisher test** (binarized intensities) or **Wilcoxon** (normalized intensities) on |Δi|
- Examples :



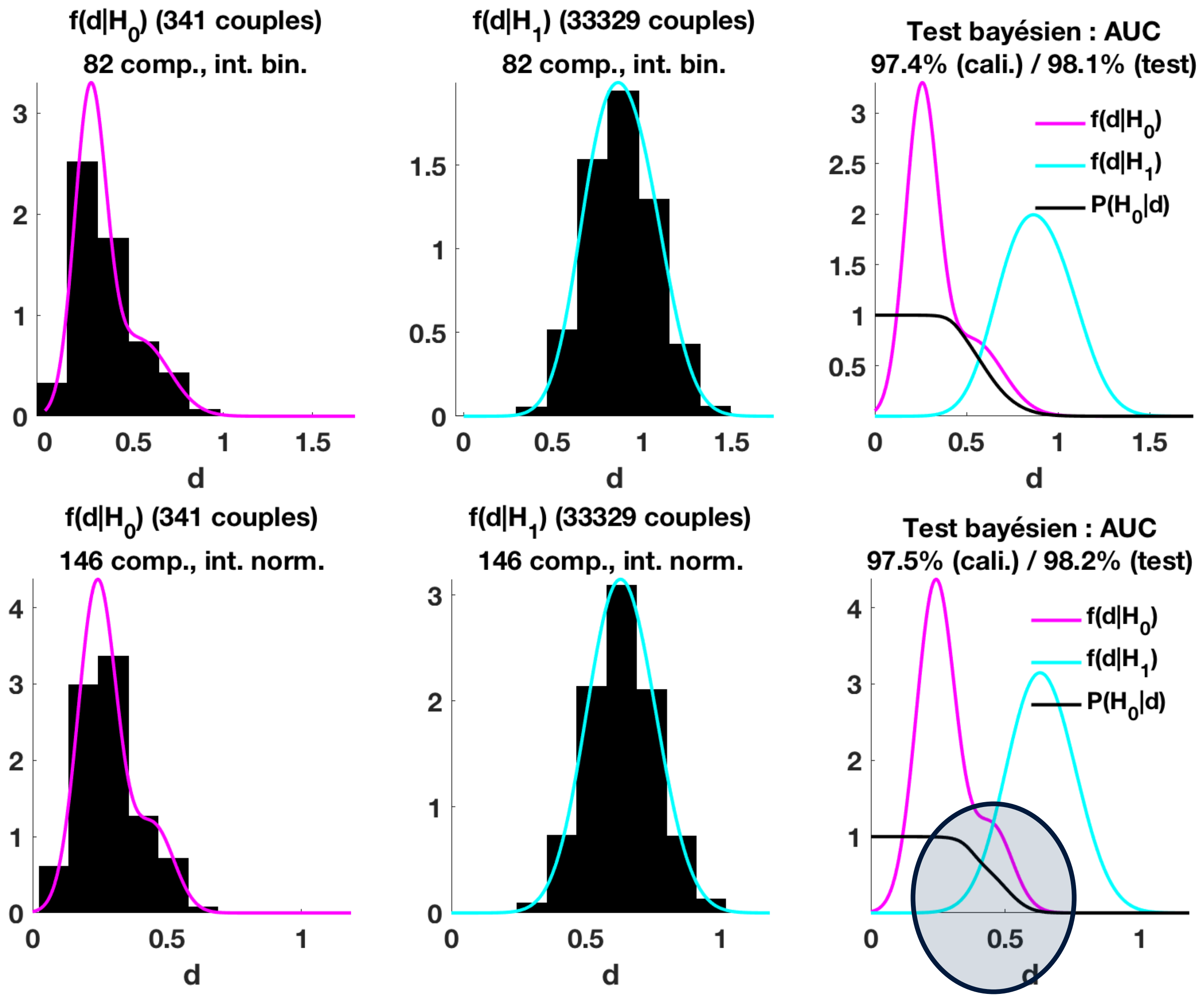


Bayesian Approach: Results using Discriminating

(%AUC calibration / %AUC test)

Binarized

Normalized



Bayesian Approach: Results using Discriminating Compounds

(%AUC calibration / %AUC test)

- **Threshold π value $-\log_{10}(p)$** of Fisher test (binarized intensities) or Wilcoxon (normalized intensities) : optimized value obtained using cross **validation** (K=3) on **calibration group**

distance / intensities	euclidian	$1-\rho_{Pearson}$	$1-\rho_{Spearman}$
normalized	π = 12 / 61 comp. **76.2% / 73.9%**	π = 13 / 54 comp. **78.1% / 75.2%**	π = 7 / 146 comp. **97.5% / 98.2%**
binarized	π = 18 / 82 comp. **93.1% / 94.8%**	π = 18 / 82 comp. **97.4% / 98.1%**	

Discussion

(%AUC calibration / %AUC test)

rformances

intensities	AUC	sensitivity	specificity	nb. compounds
binarized	97.4% / 98.1 %	89.4% / 90.0%	94.9% / 92.5%	82
normalized	97.5% / 98.2 %	89.1% / 85.9%	93.7% / 95.0%	146

Adequate distance ➔ **quantitative** exploitation of compounds intensities despite th alytical variability

Selection ➔ second modes of $f(d|H_0)$ et $f(d|H_1)$ are strongly decreased ➔ better result

Binarized : more **parsimonious (82/146 compounds to be used)**

57 common compounds for both classifiers

ota bene

same direct samples

no pollution by other odors

CONCLUSION AND PERSPECTIVES

- **Direct/non direct sampling procedures** for human (hand) odor analys
- **Comprehensive GC×GC-MS** method and data (ToF)
- Validation of procedures **in the field** with dog handlers
- Large Panel of individuals to test the model
- **Storage** of samples: standardized procedure
- Data processing in progress for real application
  - ✓ **Different samples (direct or not…) and sampling conditions**
  - ✓ Study of discriminating compounds
  - ✓ Normalization on discriminating compounds, more complex distance…
- **The final answer to the question must be YES or NO** not 98.2%

ACKNOWLEDGEMENTS

IRCGN

ESPCI PARIS

ESA

ACTION

ISIPCA
PARFUM • COSMÉTIQUE • ARÔME

phenomenex®
...breaking with tradition℠

analytics

îledeFrance

THANK YOU FOR YOUR ATTENTION!