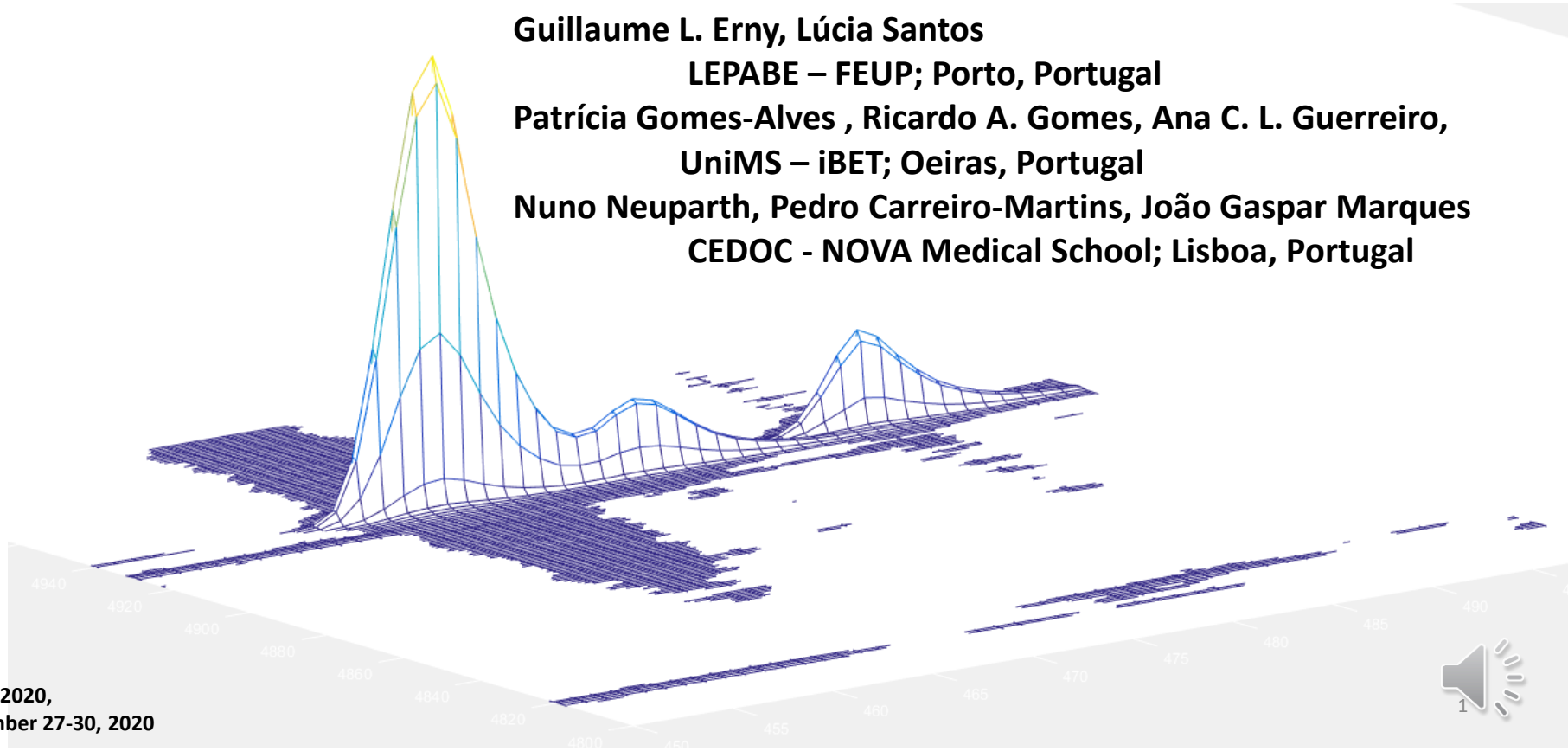
Lepabe
Laboratory for Process Engineering, Environment, Biotechnology and Energy

U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

COMPUTER-ASSISTED LC(CE)-HRMS TOOLS FOR UNTARGETED ANALYSIS: SOURCES OF ERRORS AND STRATEGIES TO OVERCOME THEM

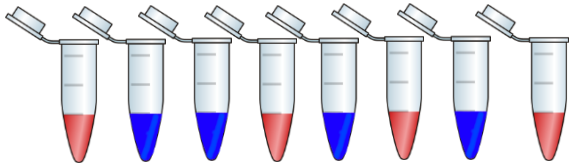
Guillaume L. Erny, Lúcia Santos
LEPABE – FEUP; Porto, Portugal
Patrícia Gomes-Alves , Ricardo A. Gomes, Ana C. L. Guerreiro,
UniMS – iBET; Oeiras, Portugal
Nuno Neuparth, Pedro Carreiro-Martins, João Gaspar Marques
CEDOC - NOVA Medical School; Lisboa, Portugal



e-MSB 2020,
September 27-30, 2020

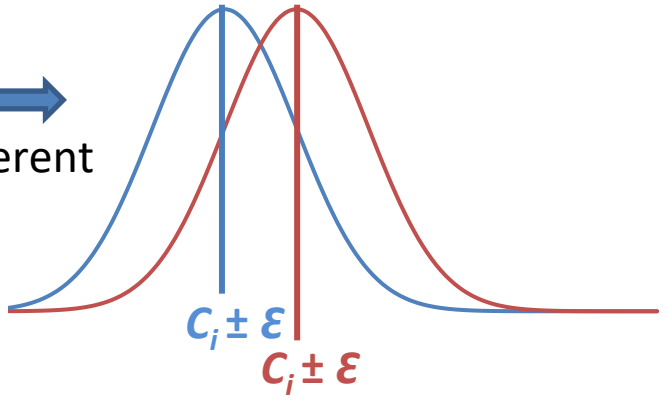
A Simple Goal: Finding markers specific to each

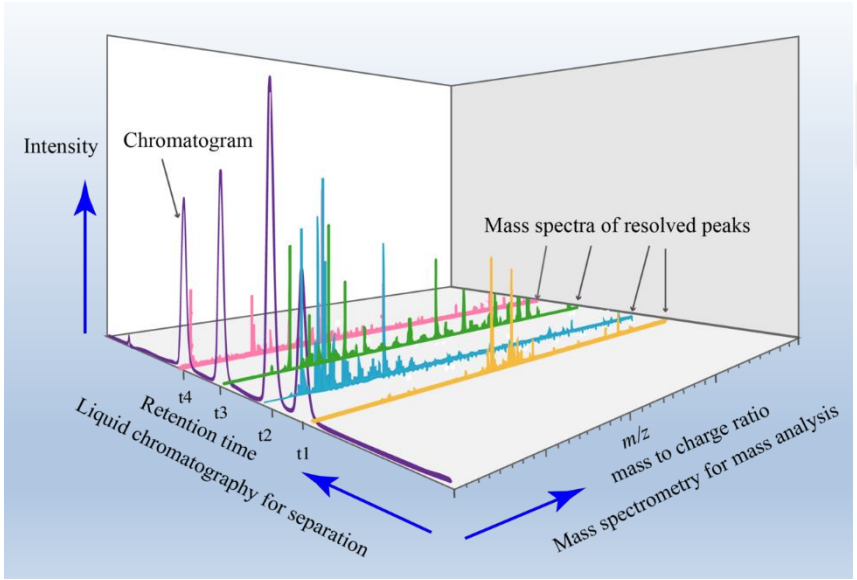
BIOLOGICAL STATE



Is there **any** components

That are **significantly** different between both state?





DATA PROCESSING

- Find the peaks
- Measured them
- Aligned them between dataset

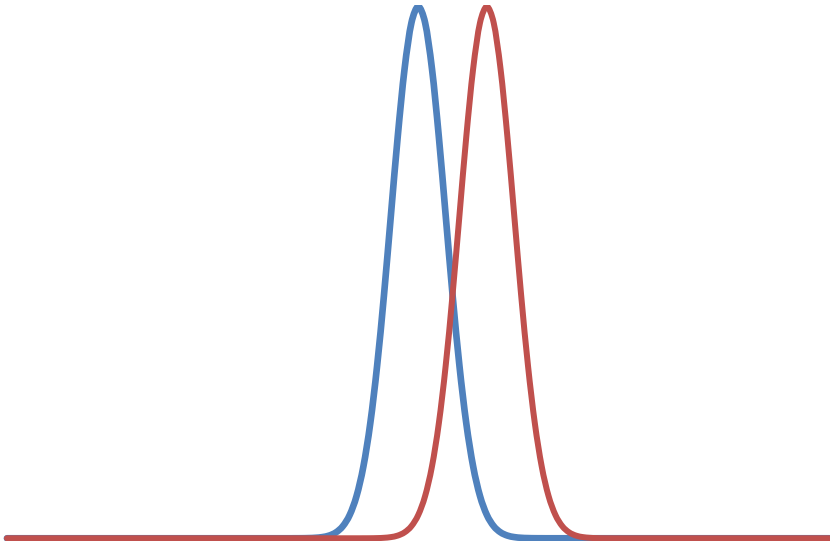
GARBAGE IN GARBAGE OUT…

1- Systematic error

Systematic (or determinate) error, causes the mean of a data set to differ **(constant shift)** from the *true* value.

2- Random error

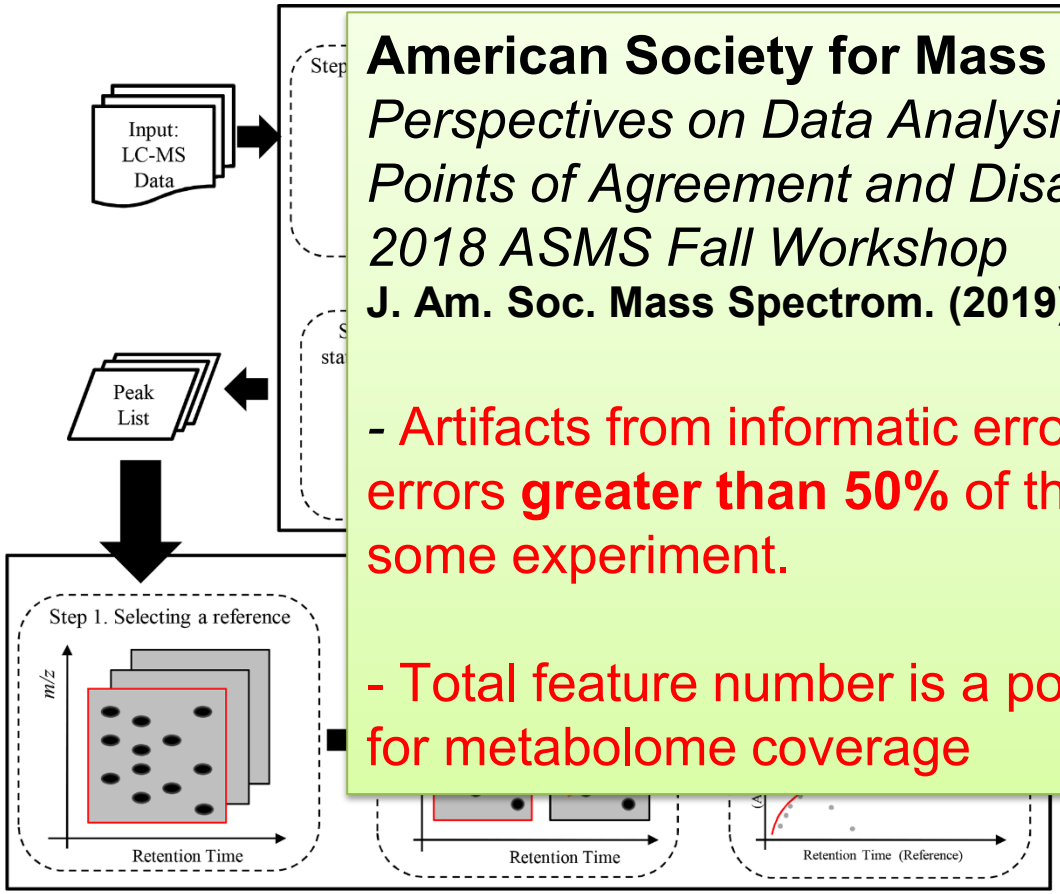
Random (or indeterminate) error, causes data to be **scattered more or less** symmetrically around a mean value.



3- Gross error

They usually occur only occasionally, are often large, and may cause a result to be either high or low. Gross errors lead to **outliers**, results that appear to differ **markedly** from all other data in a set of replicate measurements.

WHERE DOES ERRORS COME FROM? EXAMPLE WITH IMET-Q



American Society for Mass Spectrometry, 2019

Perspectives on Data Analysis in Metabolomics: Points of Agreement and Disagreement from the 2018 ASMS Fall Workshop

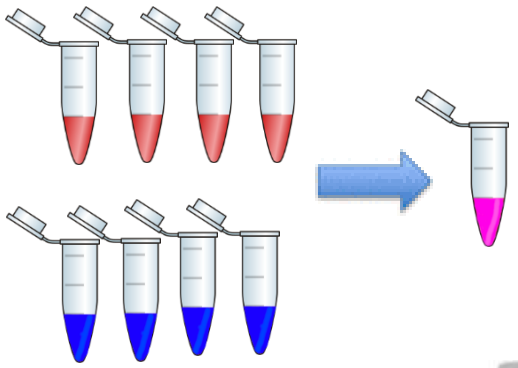
J. Am. Soc. Mass Spectrom. (2019) 30:2031–2036

- Artifacts from informatic errors and contaminants errors **greater than 50%** of the mined features in some experiment.

- Total feature number is a poor evaluation metric for metabolome coverage

Schematic depiction of iMet-Q workflow for peak detection and peak alignment, from: PLoS ONE 11(1):e0146112, licensed under CC BY 4.0

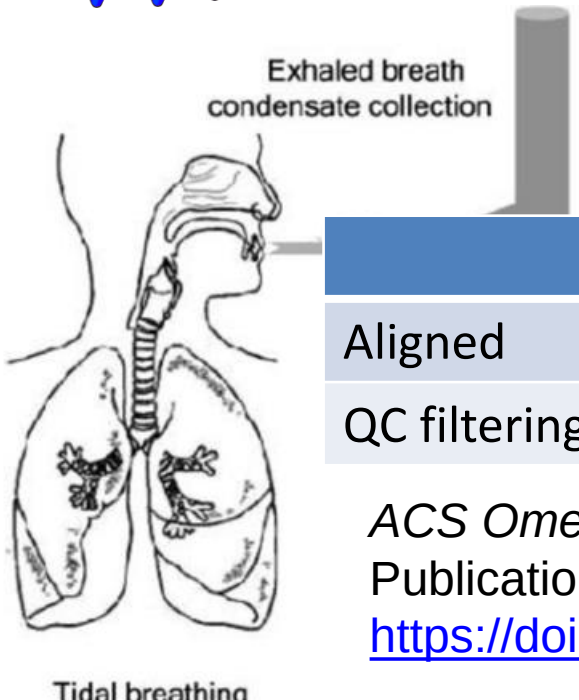
POOLED QC SAMPLES – THE MINIMAL REQUIREMENT?



- representative the entire collection of samples
- Injected at regular interval

✓ Validation of *features* based on frequency of detection and coefficient of variation of the pooled QC sample

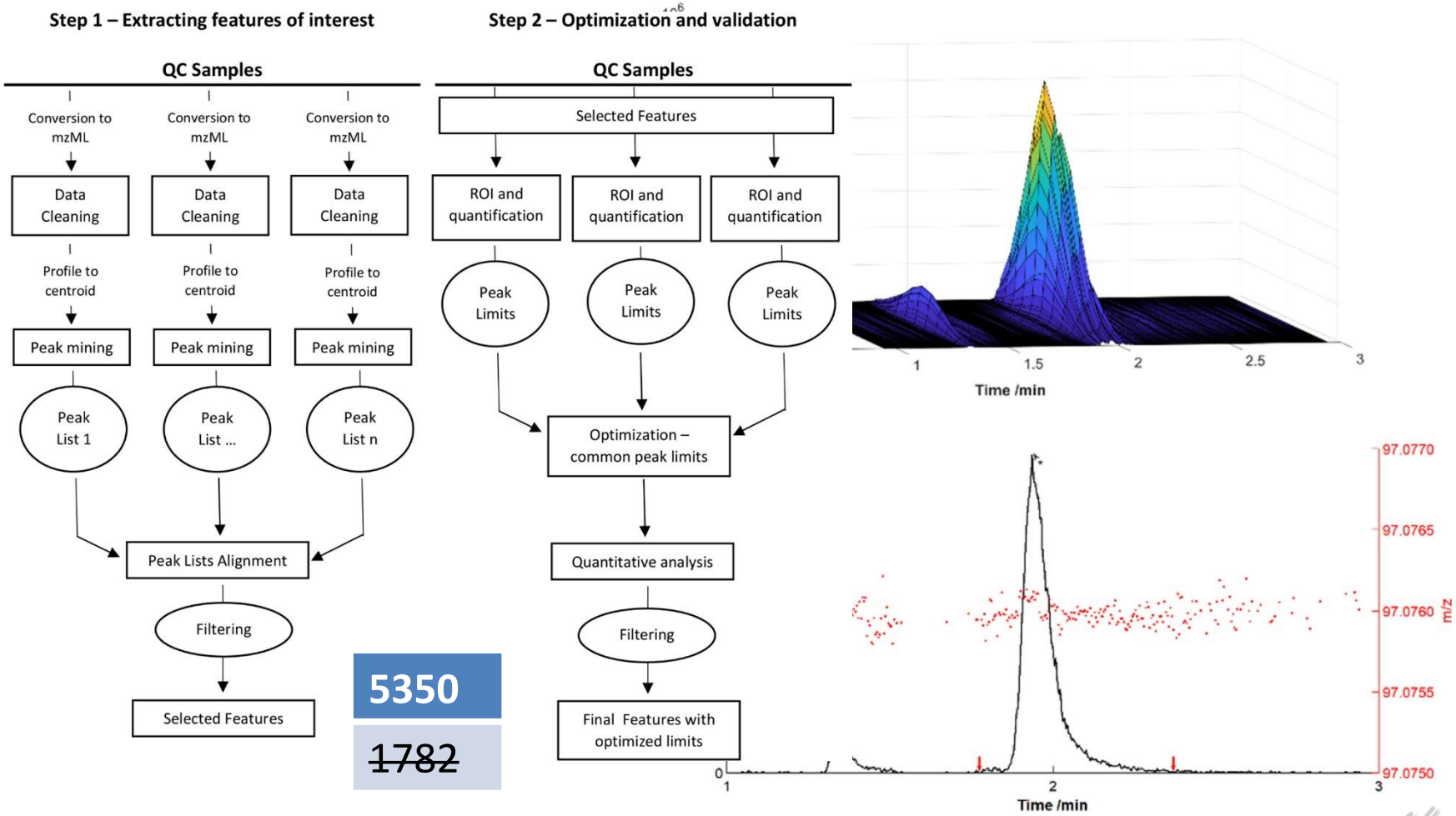
✓ Correction of instrumental variations



	msDial	xcMS	Finnee
Aligned	11560	7737	5350
QC filtering	1456	1587	1782

ACS Omega 2020, 5, 26, 16089–16098
Publication Date:June 23, 2020
https://doi.org/10.1021/acsomega.0c01610

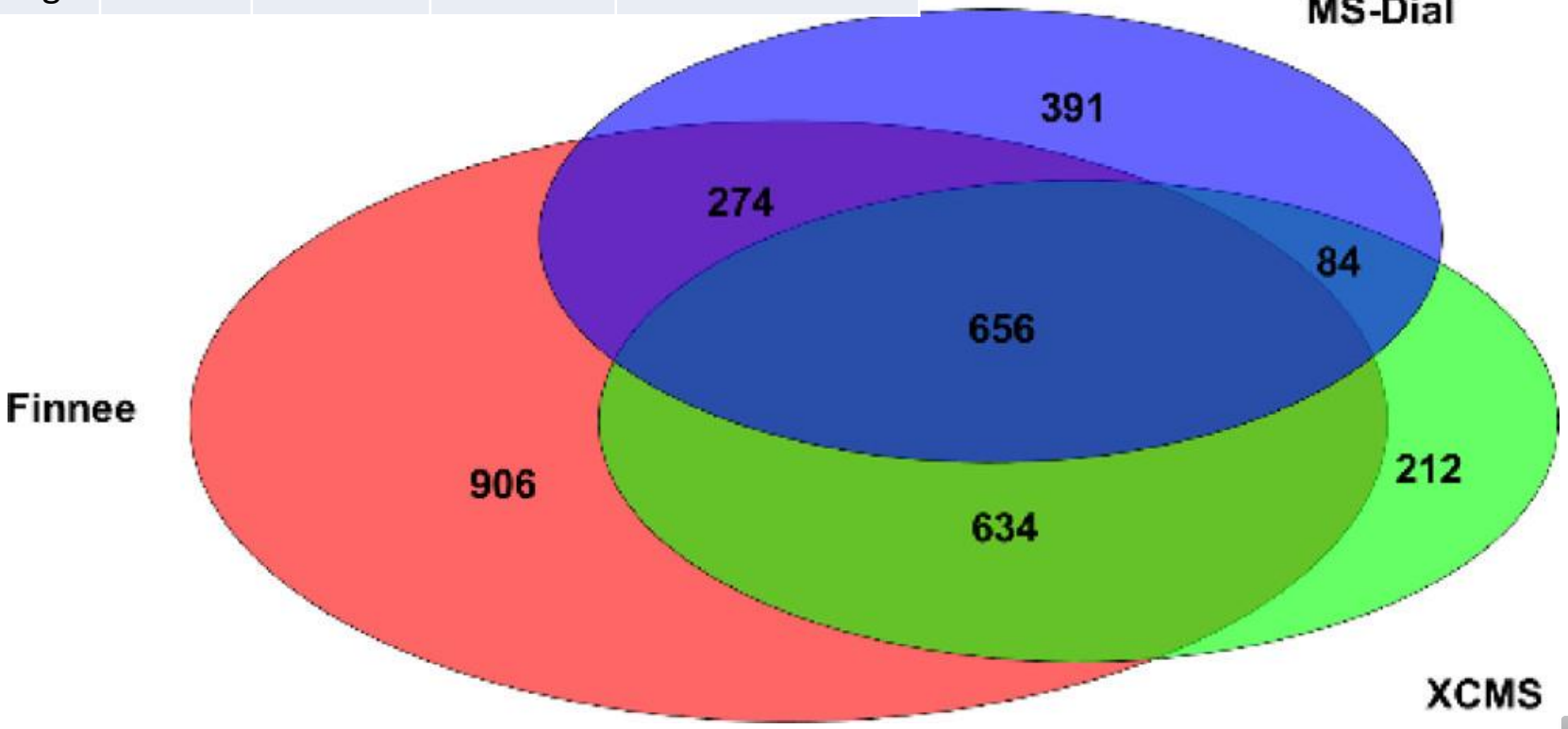
How to Reduce Gross Errors



ACS Omega 2020, 5, 26, 16089-16098

RESULTS (QC SAMPLES, FROM HEALTHY, COPD AND ATHMA PATIENTS)

	msDial	xcMS	Finnee	Finnee
Aligned	11560	7737	5350	5350
QC filtering	1456	1587	1782	**2491**



ACS Omega 2020, 5, 26, 16089-16098

CONCLUSIONS

- **The quality of your separation is paramount to reduce random and gross error**
  - ✓ Constant sampling rate
  - ✓ 10 points per peak minimal (LC & MS)
  - ✓ S/N > 10
  - ✓ Baseline resolution of peaks
- **Due to the high number of extracted ion profiles, gross errors are hard to find**
- **The number of transformation should be reduced, to decrease errors**

ACKNOWLEDGEMENTS

This work was financially supported by the projects:

(i) UID/ EQU/00511/2019 - Laboratory for Process Engineering, Environment, Biotechnology and Energy − LEPABE funded by national funds through FCT/MCTES (PIDDAC);

(ii) POCI-01-0145-FEDER-029702 funded by FEDER funds through COMPETE2020 − Programa Operacional Competitividade e Internacionalizaçã õ (POCI) and by national funds (PIDDAC) through FCT/ MCTES;

(iii) AstraZeneca − Projecto OLDER (CEDOC/ 2015/59);

(iv) iNOVA4Health - UID/Multi/04462/2013, financially supported by FCT/Ministerio da Educacão e Ciência, and co-funded by FEDER under the PT2020 Partnership Agreement.

ACKNOWLEDGEMENTS

THE TEAM.

Guillaume L. Erny,
Mónica S. F. Santos,
Lúcia Santos
***LEPABE – FEUP*; Porto, Portugal**

Patrícia Gomes-Alves ,
Ricardo A. Gomes,
Ana C. L. Guerreiro,
***UniMS – iBET*; Oeiras, Portugal**

Nuno Neuparth,
Pedro Carreiro-Martins,
João Gaspar Marques
***CEDOC* - NOVA Medical School; Lisboa, Portugal**

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR
REPÚBLICA PORTUGUESA
NORTE2020 PROGRAMA OPERACIONAL REGIONAL DO NORTE
COMPETE 2020
PORTUGAL 2020
UNIÃO EUROPEIA
Fundo Europeu de Desenvolvimento Regional

ABOUT FINNEE

Finnee is an open-source, **free to use** Matlab toolbox.
Anyone is welcome to test, exchange and **collaborate**!

Repository: https://github.com/glerny/Finnee2016
Blog: https://finneeblog.wordpress.com/