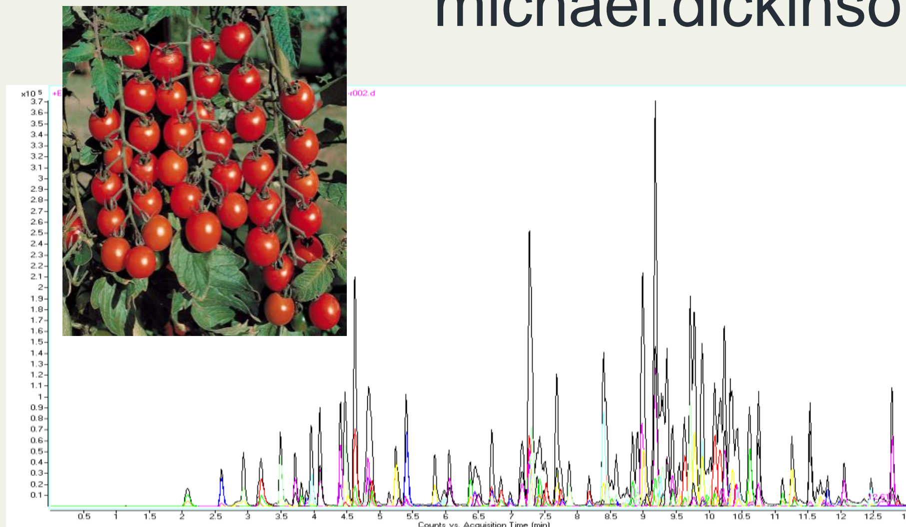
Validation of LC-Mass Spectrometry based methods and routine quality control

Mike Dickinson,
Fera Science Ltd.

michael.dickinson@fera.co.uk



International Symposium on Standardisation of Non Targeted Methods for Food Authentication. BfR, Berlin 28 – 29th November 2016

Outline

fera

- Why validate?
- How we validate non – targeted MS approaches
- What do we mean by quality control in this perspective?
- How we can use this QC effectively

fera

Why validate?

- Check or prove the validity or accuracy of…….
- Demonstrate or support the truth or value of……
- Make or declare legally valid……..

The profiling community in food authenticity should address:

- Raw data integrity – chemica
- Statistically sound experimer
  informatics and model creati
- Chain of custody implication
- Will our processes / systems

Ronald A. Fisher (1938)

"To call in the statistician after the experiment is done may be no more than asking him to perform a post-mortem examination: he may be able to say what the experiment died of"

* and this is why most claimed research findings are false

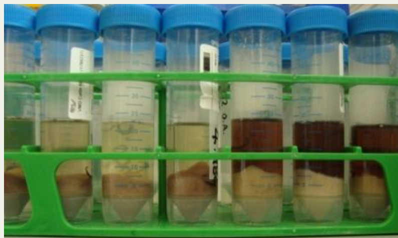
*Broadhurst, D. & Kell, D.B. (2006) *Metabolomics* **2**, 171-196

Pre experiment validation

Pre-experiment validation

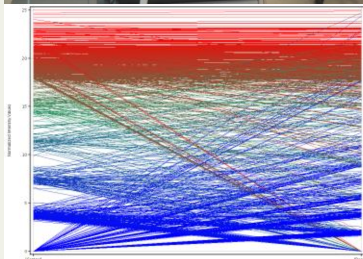
fera

1. How am I going to extract this sample taking into account matrix load on my system?



2. Will my extraction recover a range of metabolites efficiently and reproducibly?

3. How robust is my LC-MS system if I need to analyse multiple samples in one experimental run?



4. How do I handle the data to achieve a solution?

Pre-experiment validation – sample concentration?

More concentrated extract	Less concentrated extract
Advantages	Advantages
Potentially more metabolites	Less burden on LC-MS system
Greater sub sample can be taken = more representative	Lower ion suppression may give greater signal!
Potentially shorter sample prep	Potentially shorter sample prep

Using varying concentrations (and other sample extraction strategies), take some time to get to know your system

System = from sample on bench to final data output

Pre-experiment validation – knowing your system: LC-MS signal reproducibility

5 mg/ml matrix load

RSD = 14%

RSD = 9 %

RSD = 11%

Response (area)

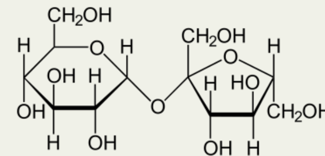
16,000,000
14,000,000
12,000,000
10,000,000
8,000,000
6,000,000
4,000,000
2,000,000
0

Inj no 1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59 61 63 65 67 69 71 73

proline

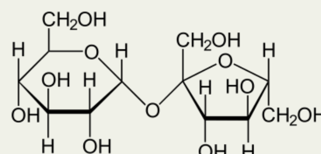
Istd (Tric)

Pre-experiment validation – knowing your system: LC-MS signal reproducibility



5 mg/ml matrix load

Pre-experiment validation – knowing your system: LC-MS signal reproducibility



1.25 mg/ml matrix load

Response (area)

500,000
450,000
400,000
350,000
300,000
250,000
200,000
150,000
100,000
50,000
0

RSD = 9 %

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42

Inj no

Dilution factor can be key to a successful MS profiling experiment !

Pre-experiment validation – knowing your system: repeatability

Day to day repeatability for 3 days (mean response over 7 vegetable extractions)

Valine, RSD 16%

Istd , RSD 11%

Trigonelline, RSD 8%

Proline, RSD 13%

Threonine, RSD 9%

Hydroxymaackiain, RSD 8%

Response (peak area)

40,000,000
35,000,000
30,000,000
25,000,000
20,000,000
15,000,000
10,000,000
5,000,000
0

0 1 2 3 4

Day

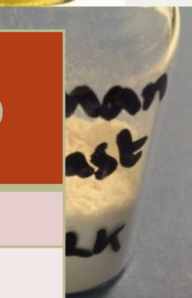
What significant variation are **you** adding to your dataset in large studies that require multiple consecutive batches?

Pre-experiment validation - Extraction recovery

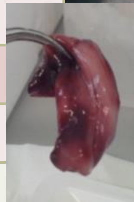
- Spiking of internal standards or "non-biological" compounds

Compound	m/z	RT	Mean % Recovery (n=5)	% CV (n=5)
No				
5- try				
Prog				
Test				
D3-				

Compound	m/z	RT	Mean % Recovery (n=5)	% CV (n=5)
Norleucine	131.0946	5.2	101	3
5-methyl- tryp	219.1128	12.0	99	2
Prog				
Test				
D3-				



Compound	m/z	RT	Mean % Recovery (n=5)	% CV (n=5)
Norleucine	131.0946	5.6	43	4
5-methyl-tryptophan	219.1128	11.6	67	3
D9-Progesterone	324.2881	16.7	70	2
D3-Testosterone	292.2354	16.1	96	2



NB. Dis-advanatges in using Istds in main profiling experiment

QC within experiment

Experimental QC

- Can be in many forms. E.g.

- 100 honey type A
- 100 honey type B
- 100 honey type C
- 100 honey type D

400 sample injections, > 2 days analysis time solid

Extract in random order over number of days, i.e. different batches.

Ideal to have

- Batch to batch variation check
- Pooled extract

QC's are technical replicates independent of batch or injection number

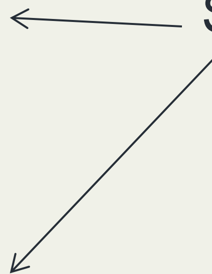
Example sequence set up:

Standards mix
Solvent Blank
Solvent Blank
Conditioner 1
Conditioner 2
Conditioner 3
Conditioner 4
Conditioner 5
Conditioner 6
Conditioner 7
Conditioner 8
Conditioner 9
Conditioner 10
IHR 1
Pooled QC 1
Sample 1
Sample 2
Sample 3
Sample 4
Sample 5
Sample 6
Pooled QC 2
Sample 7
Sample 8
Sample 9
Sample 10
Sample 11
Sample 12

Check system specs and then clear

Condition or “dirty” system to steady state

Batch to batch check,
Signal evaluation



Run all samples in random order – www.random.org

QCs allow signal assessment

PC 1 v PC2

PC2 : 7.67 % total variance

PC1 : 23.6 % total variance

10
5
0
-5
-10
-15

-8 -6 -4 -2 0 2 4 6 8 10 12

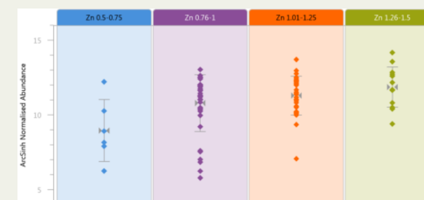
QA
Real
Spike1
Pool
Spike2

Begley P. *et al.* (2009) *Analytical Chemistry* **71,** 7038-7046

Signal assessment: checking reliability of the measurements

Formula	$C_4H_9NO_3$	$C_{11}H_{12}N_2O_2$	$C_{21}H_{30}O_2$	$C_{18}H_{32}O_2$	C8H7N	$C_8H_{10}N_4O_2$	$C_5H_4N_4O_3$	$C_9H_{11}NO_2$	$C_6H_{12}O_6$	C9H17N1O5
Cpd	**Threonine**	**Tryptophan**	**Progesterone**	**α - Linolenic acid**	**Indole**	**Caffeine**	**Uric acid**	**Phenylalanine**	**Glucose**	**Pantothenic acid (VitB5)**
M+H	120.06551	205.09714	315.23184	279.23184	118.06512	195.08764	167.02105	164.07169	179.0561	218.10338
RT	1.9	10.4	18.6	18.5	10.4	11.1	3.4	9.5	1.9	10
% RSD	**14**	**17**	**14**	**20**	**18**	**16**	**5**	**13**	**13**	**13**

E.g. you find a marker in the honey that distinguishes between A and B.
Largest mean fold change between groups: 3.5
T-test : P = 0.013



You check the %RSD of response across 65 QC's = 33%

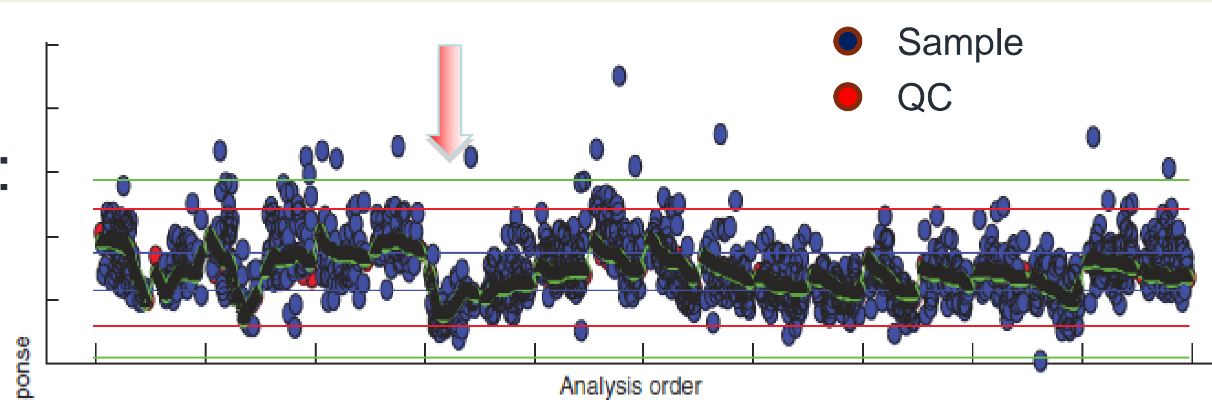
Many groups would now dismiss this potential marker:

Begley et al (2009) *Analytical Chemistry*, 81:7038 = < 30% RSD
Kirwan et al (2013) *Anal. and Bioanal. Chemistry,* 405:5147 = < 20%RSD

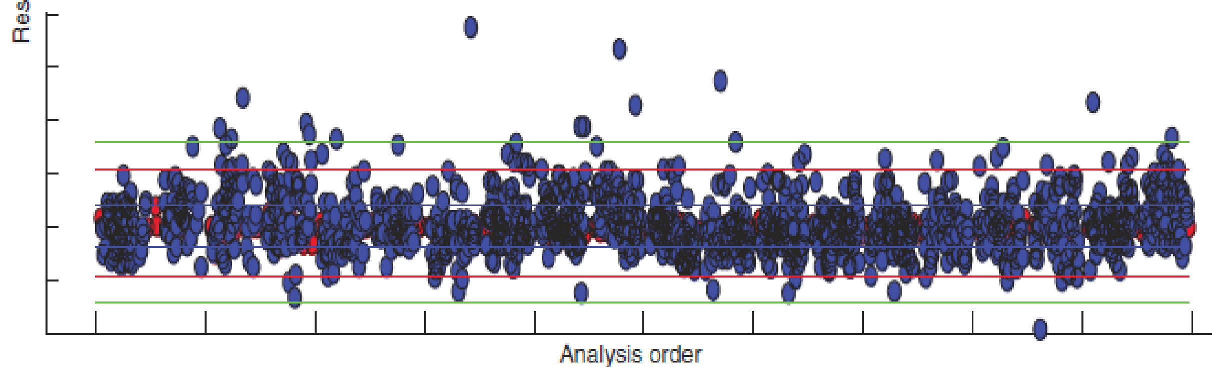
QCs allow signal correction

Instrument annual maintenance

Before:



After:



1060 | VOL.6 NO.7 | 2011 | NATURE PROTOCOLS

Dunn W. *et al.* (2011) *Nature Protocols* **6,** 1060-1083

Slide courtesy of Roy Goodacre, University of Manchester, UK.

QCs allow batch correction

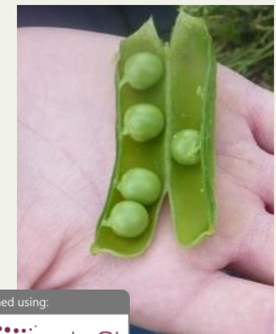
Raw Data –> PCA
Colour = Batch, Icon = Type

rpca$x[, 1] vs rpca$x[, 2]

This distance between two batches can be quantified using the "Bhattacharyya distance":

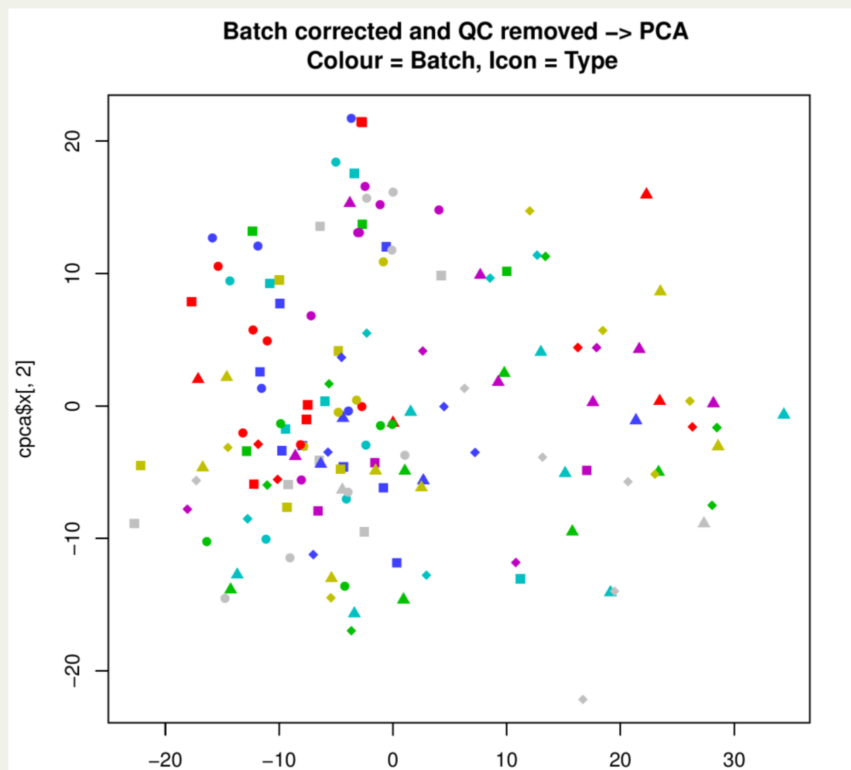
$$D_B = \frac{1}{8}(\mu_1 - \mu_2)^T \Sigma^{-1}(\mu_1 - \mu_2) + \frac{1}{2}\left(\frac{\det \Sigma}{\sqrt{\det \Sigma_1 \det \Sigma_2}}\right)$$

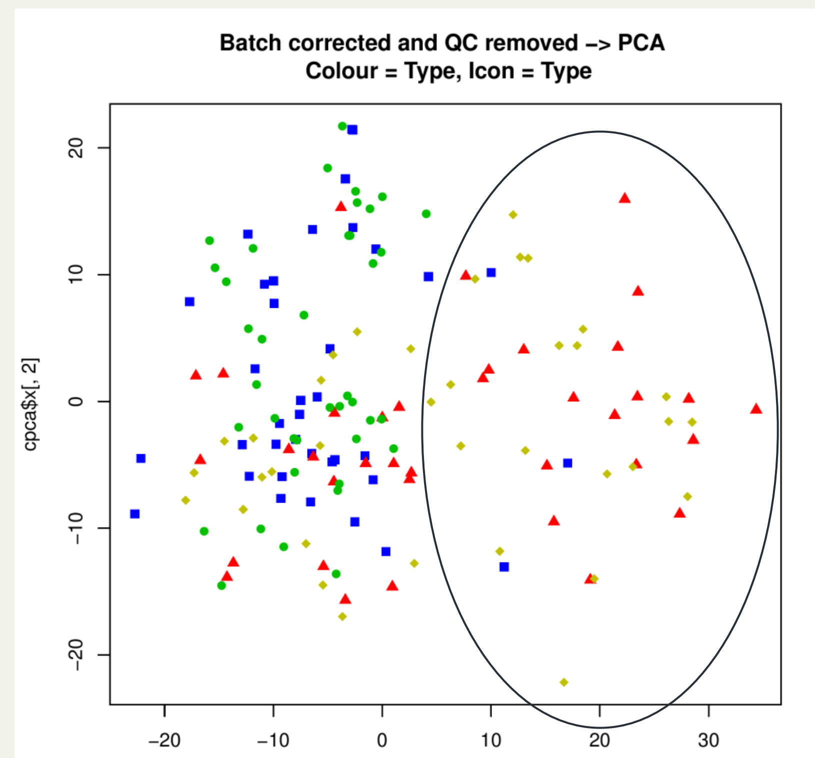
Wehrens R, et, al. *Metabolomics* (2016) 12:88



Rusilowicz, M., Dickinson M., Charlton A.J., O'Keefe S., and Wilson J (2015). Batch correction of liquid chromatography – mass spectrometry data without quality control samples.
Metabolomics 12:3

QCs allow batch correction





Corrected using median metabolite response from all QC's

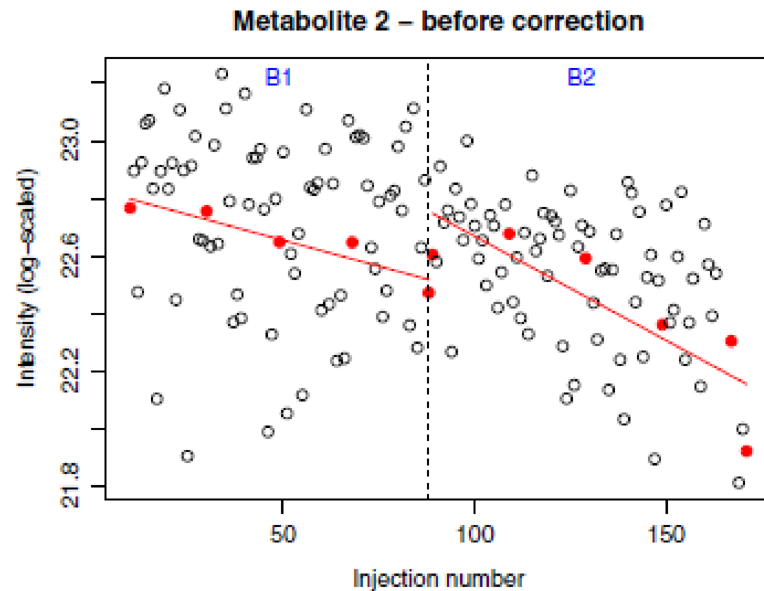
Metabolites with large numbers of zero values removed from dataset

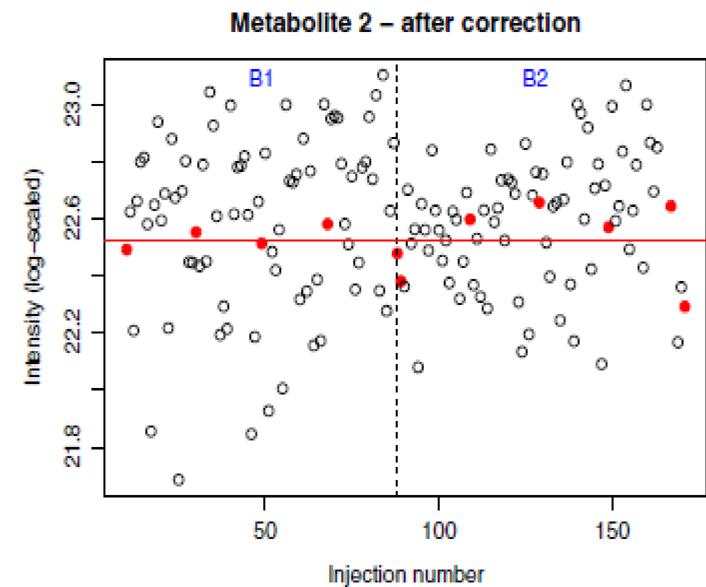
Allows us to start to make real sense from the data

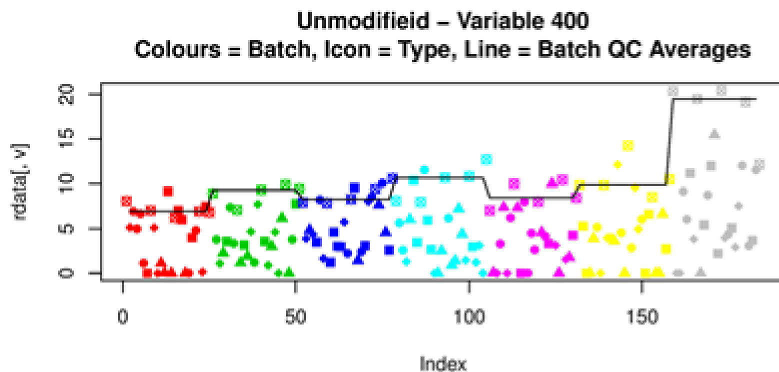
MetaboClust: interactive software for metabolomic time-series exploration and analysis. Rusilowivz, M., Dickinson M., Charlton A.J., O'Keefe S., and Wilson J (2016). *Chem. Intel. Lab. Sys* (Manuscript submitted)

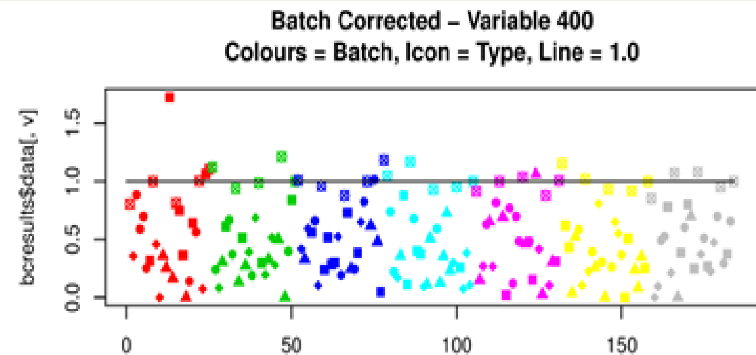
QCs allow within batch signal correction

Wehrens R, et, al. *Metabolomics* (2016) 12:88







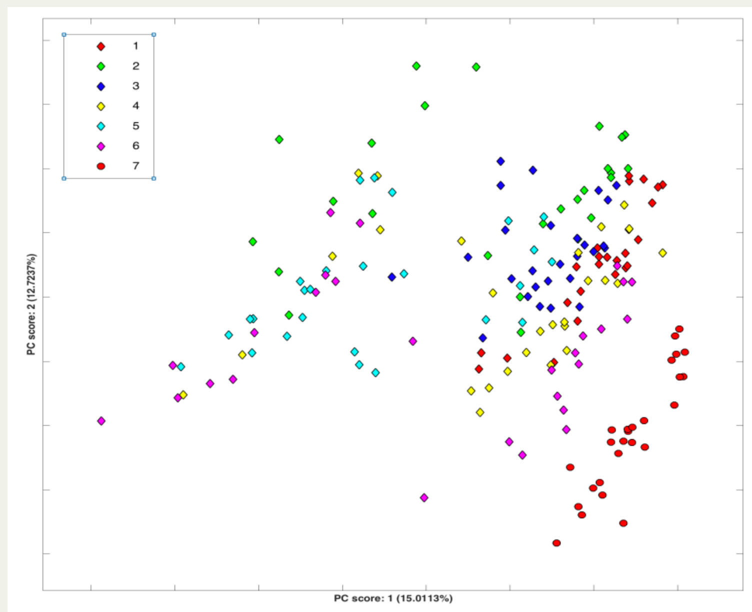


Rusilowicz, et, al. *Metabolomics* (2015) 12:3

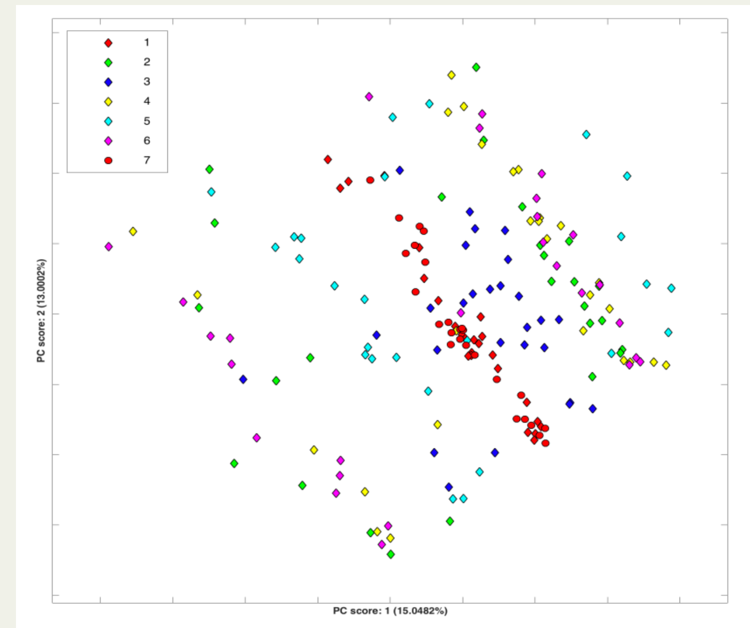
When QC correction doesn't work!

Why? Possible scenarios:

- Large random fluctuation changes in QC response profile



Uncorrected raw data showing obvious batch bias



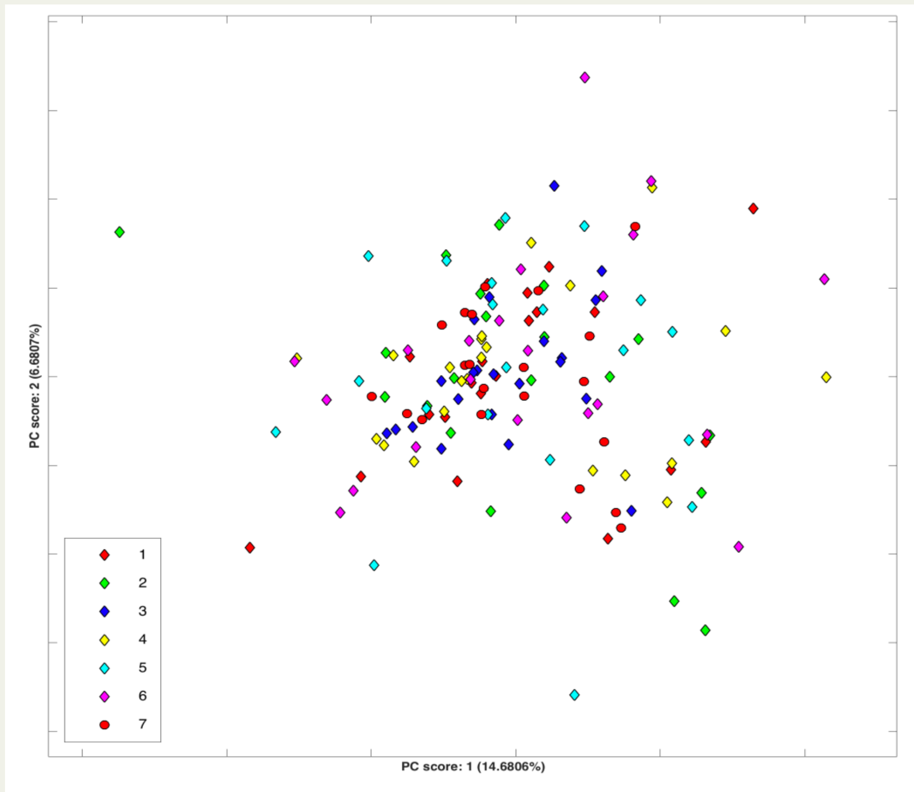
Corrected by QC – still showing batch bias

Corrected using median metabolite response from all QC's

Metabolites with large numbers of zero values removed from dataset

Rusilowicz, et, al. *Metabolomics* (2015) 12:3

When QC correction doesn't work!



- Corrected using a "moving median" metabolite response **from all data**
- The correction factor is different depending on a moving window across the data set.

$$C_{p,b,i} = median\left(X_{p,b,i-w} .. X_{p,b,i+w}\right)$$

- Metabolites with large numbers of zero values removed from dataset

Do we have an argument for reducing number of QC injections?

Rusilowicz, et, al. *Metabolomics* (2015) 12:3

fera

Summary

- Pre validation / knowing your system can save data analysis problems later
- Important to understand system performance during profiling experiment
- QC can be used to improve datasets post processing

Acknowledgements

- Dr James Donarski (Fera)
- Mr Mark Harrison (Fera)
- Mr Mark Parker (Fera)

- Professor Julie Wilson (University of York, UK)
- Dr Martin Rusilowicz (University of York, UK)
- Professor Roy Goodacre (University of Manchester, UK)
- Dr William Allwood (James Hutton Institute, UK)

Thank you for your attention

Importance of random order

N = 30

Day 1

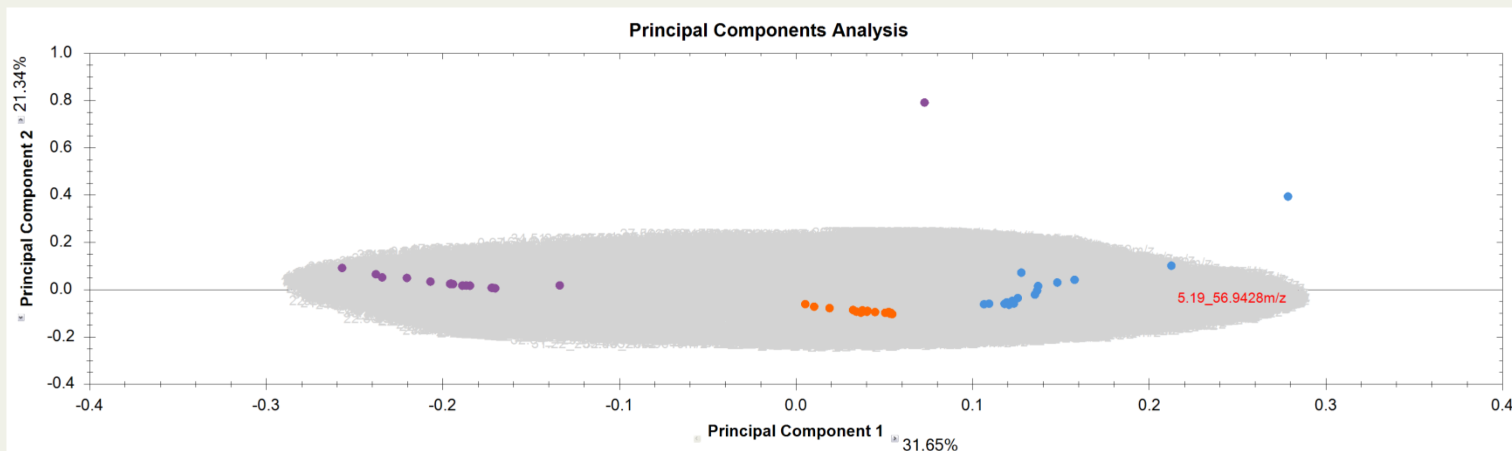
N = 30

Day 2

N = 30

Day 3

Profiling variety A vs B vs C



- **From 9,300 features (metabolites) detected, 2,500 potential markers with P<0.01, fold change > 3!!**
- **But… It's the same sample!**

Importance of random order and false detects correction

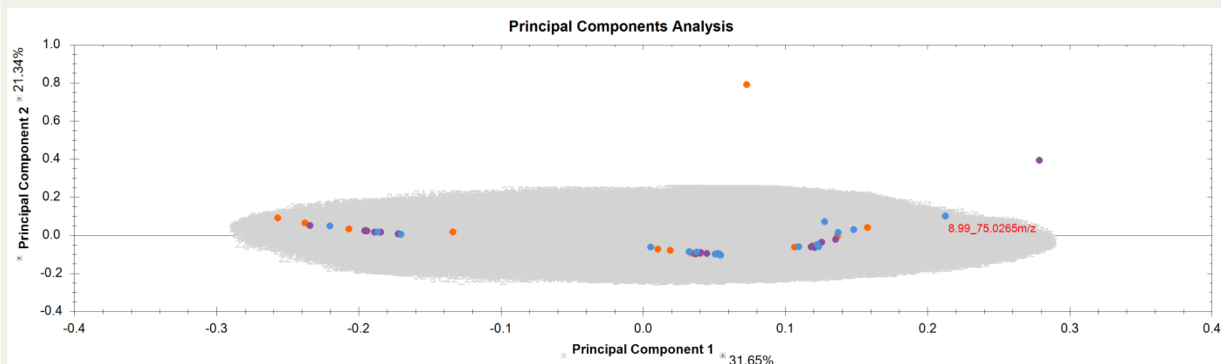
N = 30

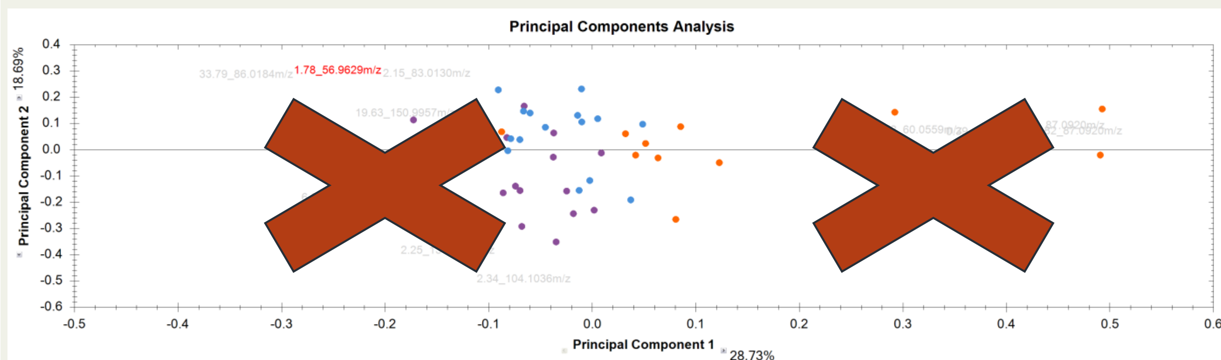
N = 30

N = 30

Extracted and analysed in random order

Profiling variety A vs B vs C





A filtered PCA!

$P < 0.01$

MF > 3

12 significant metabolites

After false detects removal = 0 metabolites