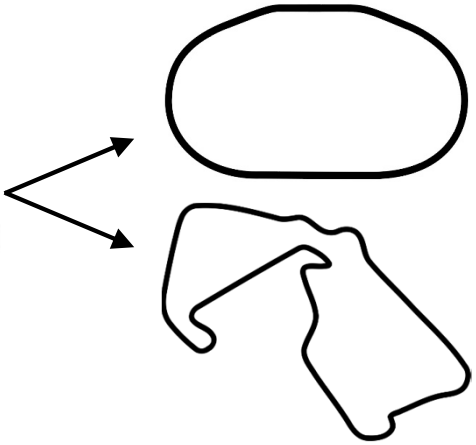
QC Benchmarker: a streamlined web application to comprehensively evaluate instrument performance and direct troubleshooting

Benjamin A. Neely
Paul E. Anderson
W. Clay Davis
Magnus Palmblad

Identification of certain commercial equipment, instruments, software or materials does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products identified are necessarily the best available for the purpose.

System Suitability – how you test is important

Race Car



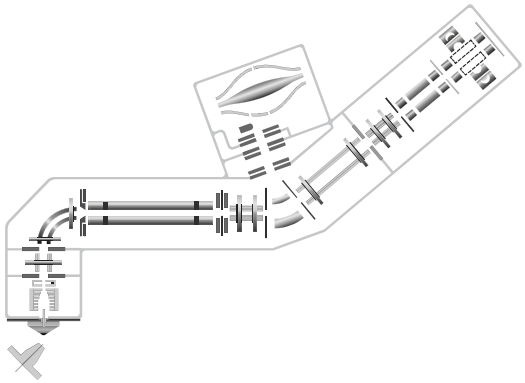
Metrics

Track speed
High-speed corners

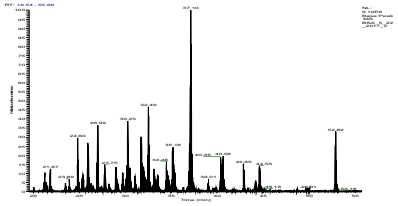
Track speed
Low-speed corners
High-speed corners
Braking
etc.

?

Modern Mass Spectrometer



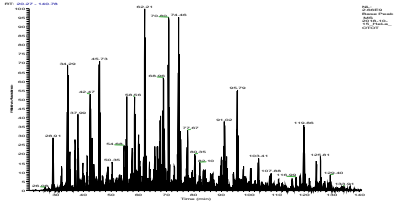
BSA Digest



Metrics

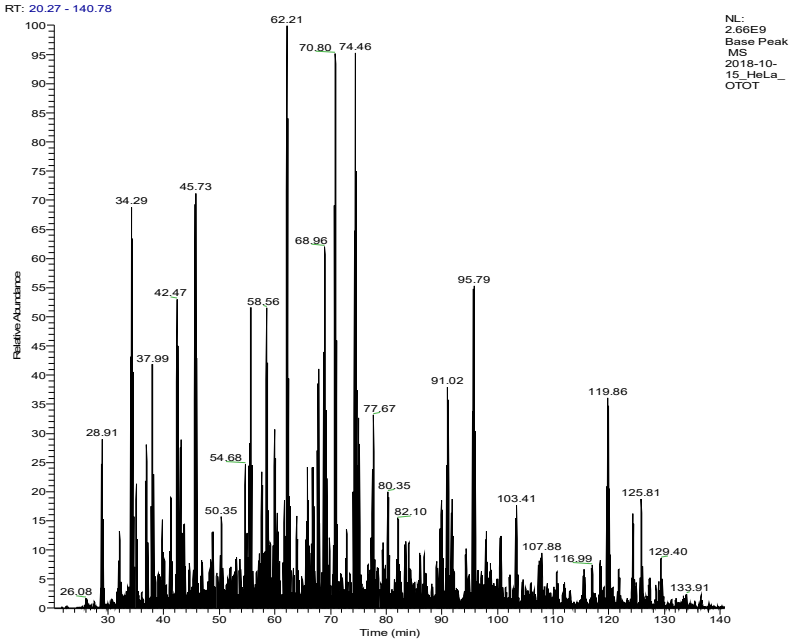
22k MS/MS
3552 PSMs
103 Peptide groups
77% coverage of BSA

HeLa Digest



75k MS/MS
28,469 PSMs
21,689 Peptide groups
4140 Protein groups

Proteomics - System suitability sample and metrics



ID-based Metrics:

- Detection of specific peptides
- Rates of peptide identification
- Coverage of known proteins

ID-free Metrics:

- Peak width/shape
- TIC
- Mass error
- Charge state distribution
- …
- etc.

Complex Digest:	Basic metrics:
HeLa	Ex.
Yeast	75k MS/MS
E. coli	28,469 PSMs
C. elegans	21,689 Peptide groups
Arabidopsis	4140 Protein groups

- Determine whether system is working
  … from LC to MS/MS
- Communicate quality across time and labs

QC – the current toolset

ID-Free + ID-based

Panorama RawTools QuiC™ powered by Pulsar QCQuan

QCLOUD pmartR PTXQC MS stats QC QCMAP

SIMPATIQCO Server

NIST MSQC Pipeline

Attention: Support for NIST MSQC Pipeline is discontinued

Tools derived/inspired by MSQC:
MassQC
Metriculator
QuaMeter

Specific application QC tools

- TMT - TVT: Triple Knockout Proteome Standard Visualization Tool

ID-Free only

- spRoCoP - Statistical Process Control in Proteomics
- Raw Beans
- QC-ART

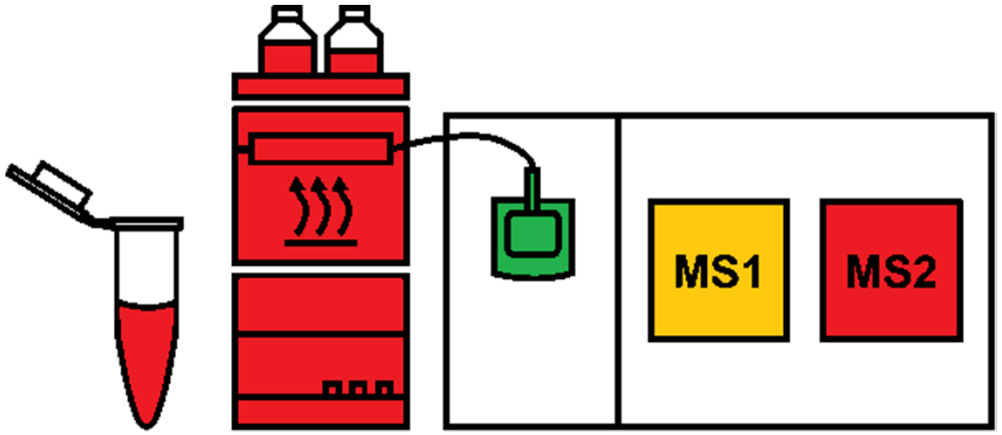
- Local and/or web-based versions
- Can work on any lab's QC sample
- Some require the presence of predefined *m/z*
- Can be used for quality controls within experimental runs

** this is not an exhaustive list, please excuse any omissions.*

Yet another tool!!!

Focus on a weekly QC sample run (known sample, known conditions) for system suitability, so that metrics/reports can be made **without user specified parameters**.

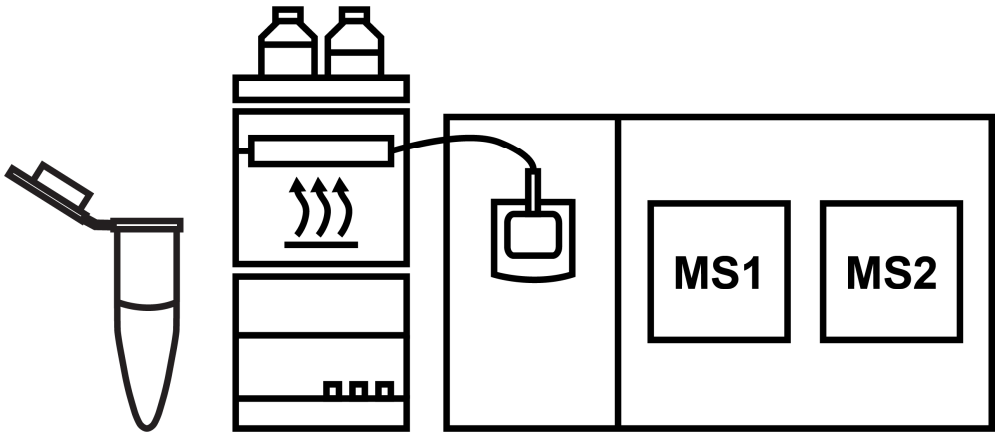
Primary User



TIC
Peak Widths
PSMs
?

Universal material and methods makes quality comparable and more useful

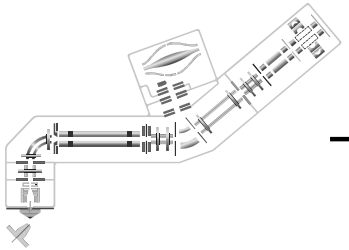
Core Lab



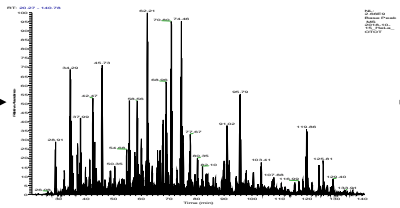
Report pre-run performance to users

Report instrument performance to future users

QC Benchmarker – Goals and Rationale



Commercial Digest
(HeLa, Yeast, *E. coli*)

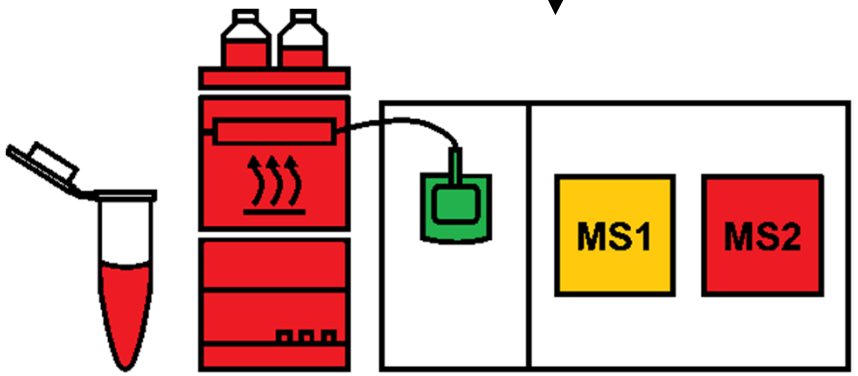


raw

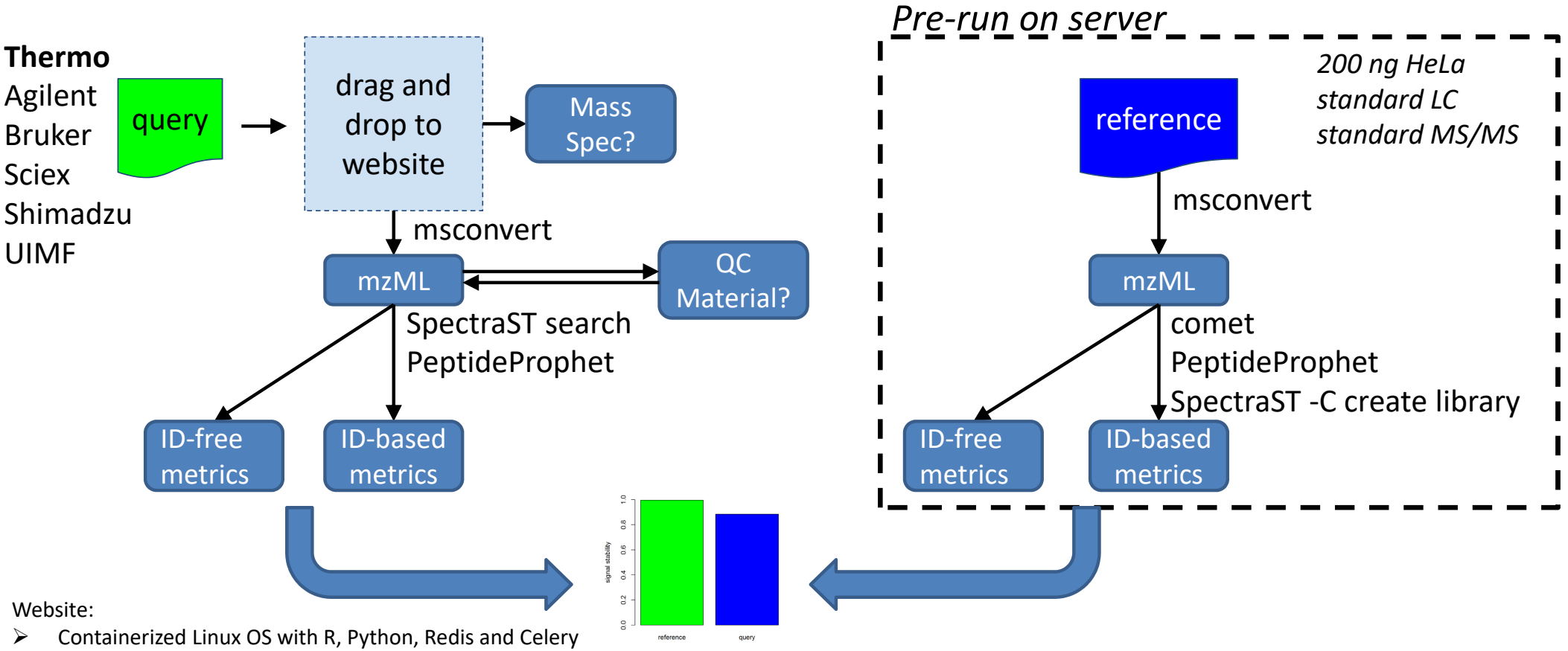
drag and drop to website

no parameter required

1) Use commercial digest standards and standard methods
2) Upload raw file to web-app
3) Return high-level summary (good, marginal, bad) with in-depth reports for troubleshooting



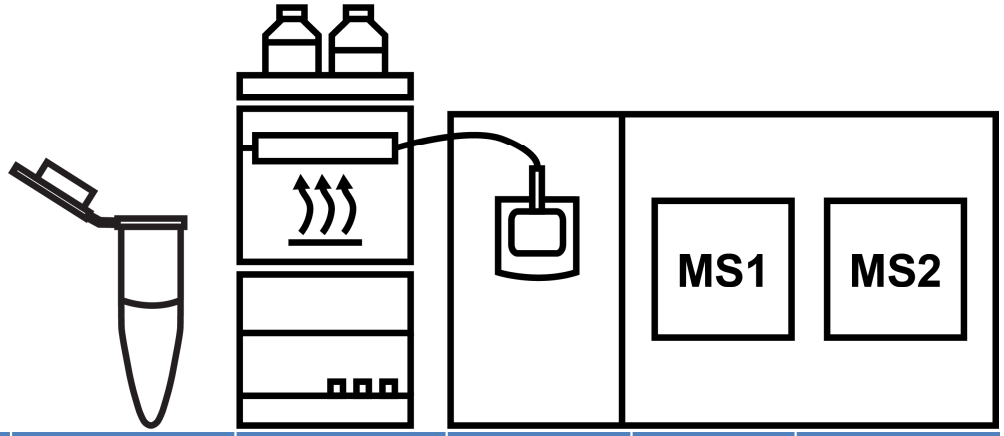
QC Benchmarker Implementation – Focus on automation and speed



Website:

- Containerized Linux OS with R, Python, Redis and Celery
- TPP (wine) used for conversion and identifications
- Existing R packages: mzR, MSnbase, pepXMLTab

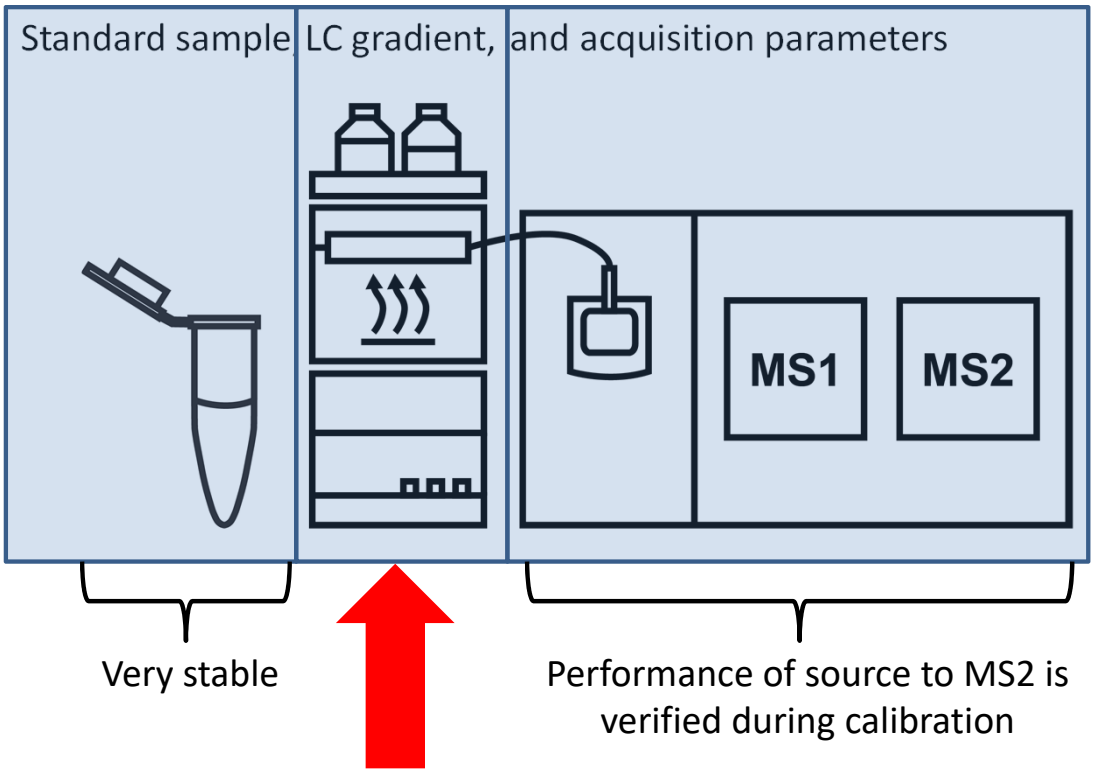
ID-free + ID-based metrics



	Sample	LC	Source	MS1	MS2
ID-free	amount loaded on column	Peak-widths, peak capacity	stability	TIC	Total signal
ID-based	% spectra matched	pH of mobile phases	z-dist of PSMs	mass error	mass error; # PSMs

**there are more metrics than these*

Development of QC Benchmarker



Problems

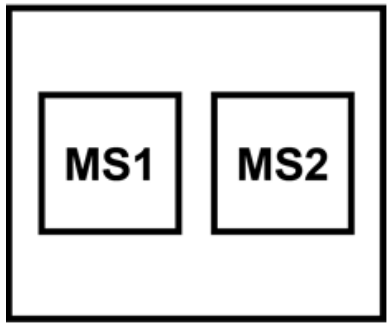
Pumps, column, mobile phases, system contaminants, etc.

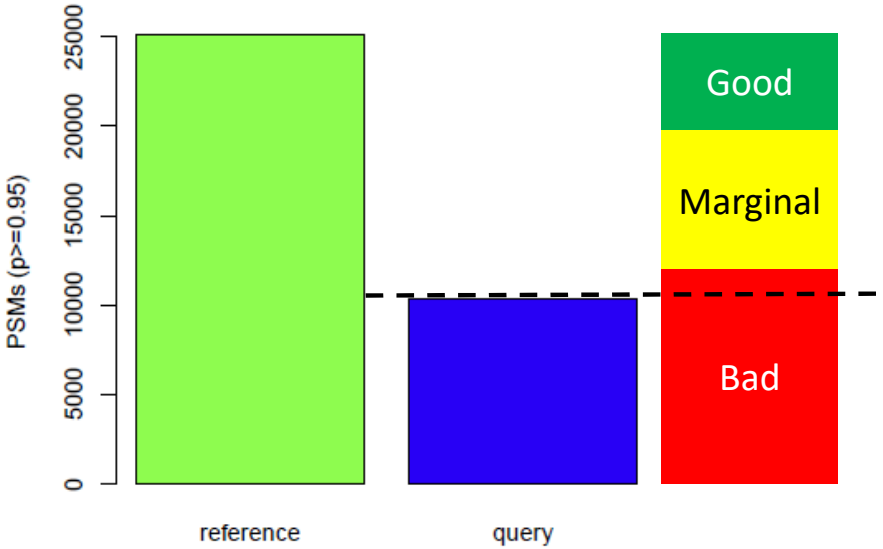
The development of data used for QC Benchmarker testing required using a standard sample, gradient and MS/MS. For now this has focused on thermo instruments and a 2 hour 200ng HeLa injection.

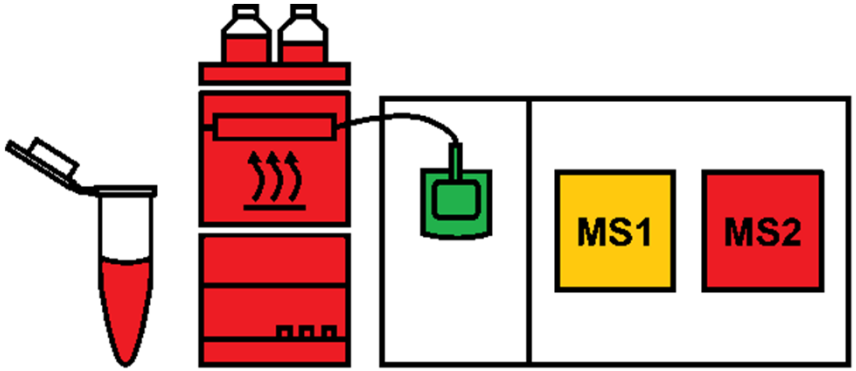
Generated ~70 runs specifically looking at:

- Mobile phase pH (0.1% formic acid to no formic acid)
- Different amounts of HeLa on column (50 to 1000 ng)
- Sub-optimal collision energy
- Attempted to degrade HeLa
- Runs before and after MS cleaning
- Attempted to decrease transfer efficiency
- Mass error (couldn't get over 8 ppm)
- Running on a different platform (QE classic v Lumos)
- 15 runs across a year

Example of a performance threshold – PSMs

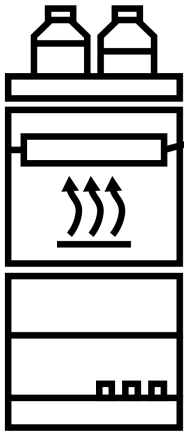




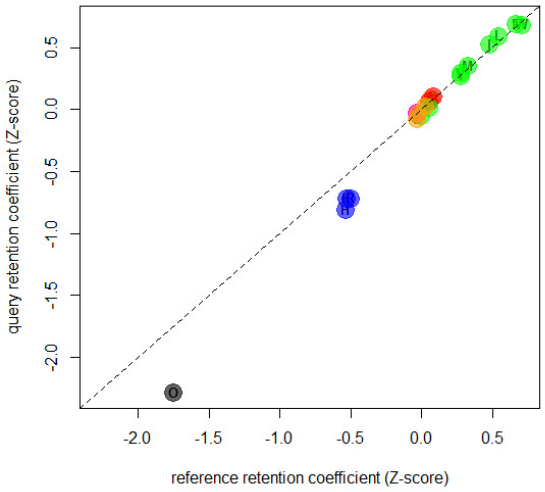


- Thresholds are based on instrument and type of acquisition
- Since each instrument is unique, qualitative thresholds may not be applicable, but metrics are still valid

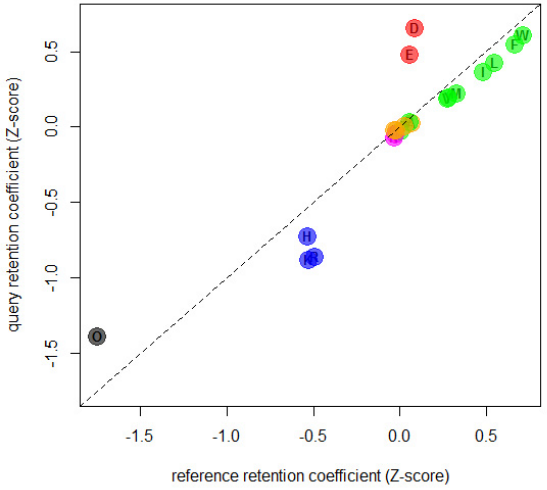
Machine Learning-based prediction of mobile phase pH and pump performance



acceptable



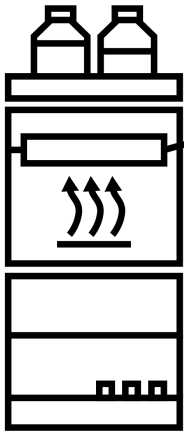
pH shift

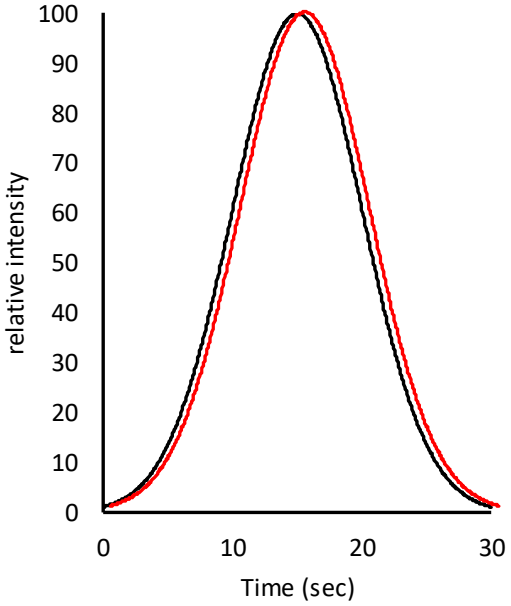


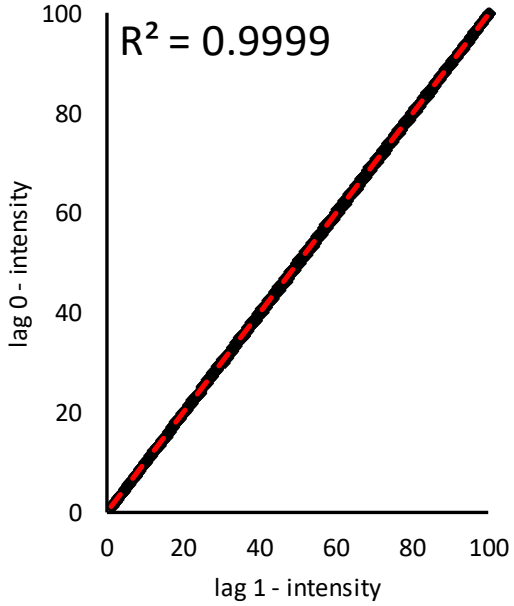
- Decreased correlation may indicate mobile phase pH is off
- A deviation of the intercept indicates the pumps aren't performing correctly

Palmblad et al., 2002 doi: 10.1021/ac0256890

Autocorrelation-based peak width determination

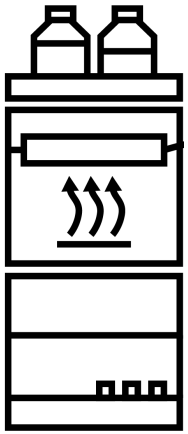


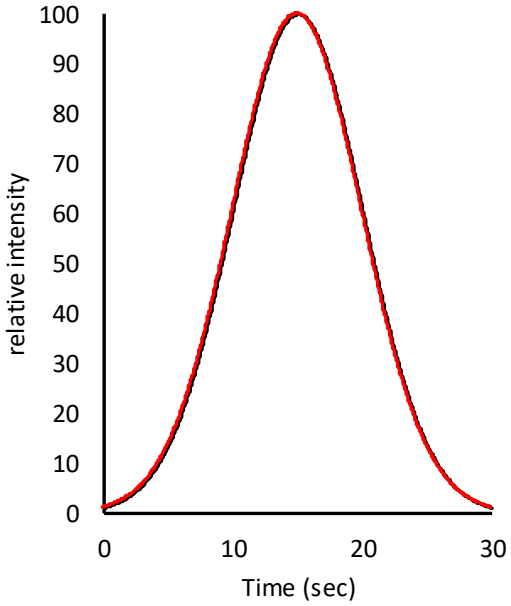




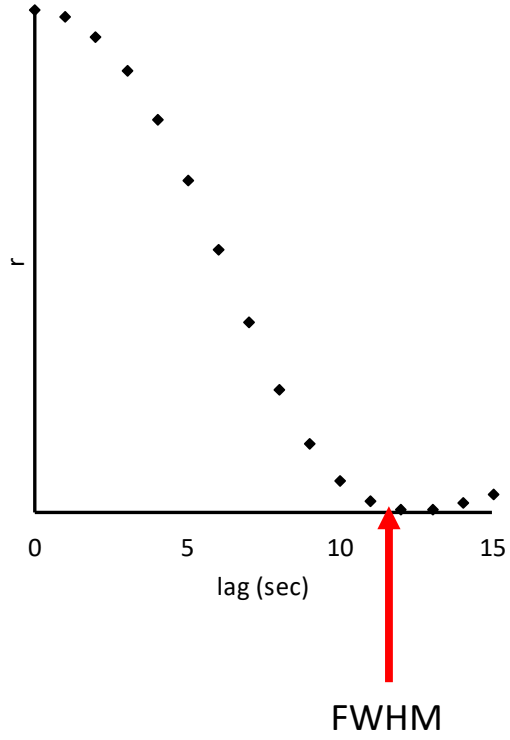
Shifted all points by 50 ms

Autocorrelation-based peak width determination



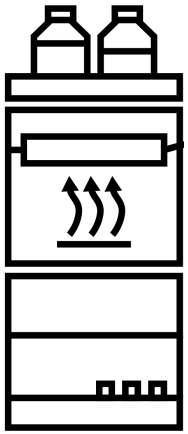


30 second peak at base
FWHM = 11.7 sec

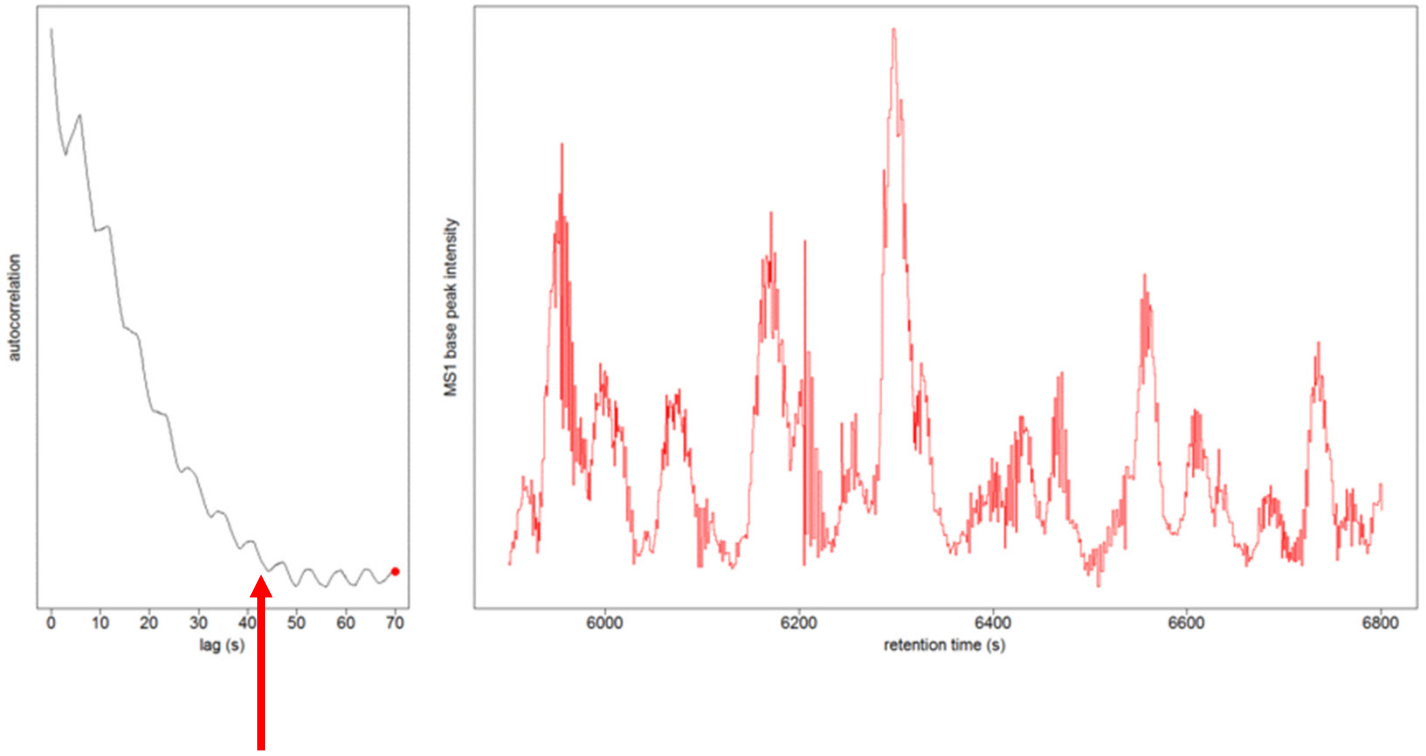


FWHM

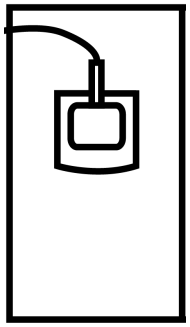
Autocorrelation-based peak width determination

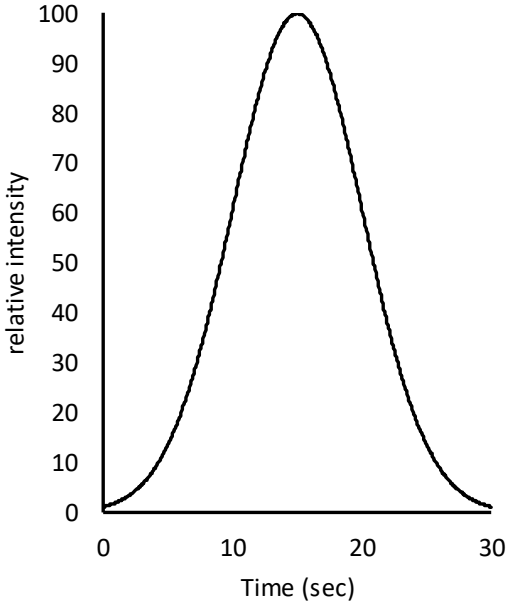


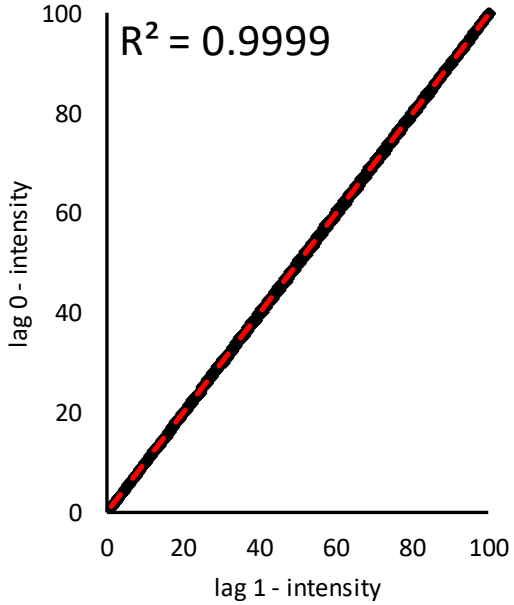
- Perform across the whole gradient



Autocorrelation-based measure of spray stability

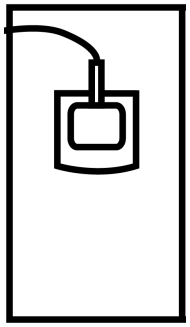


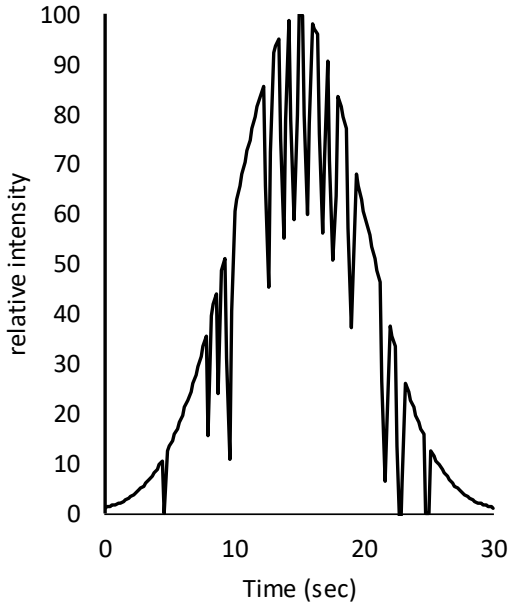


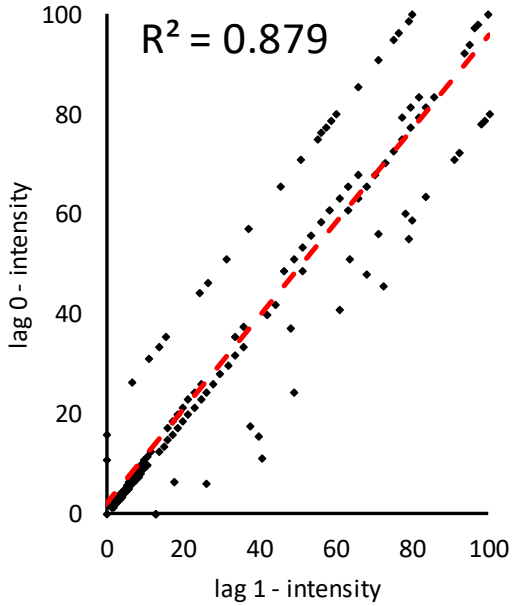


Shifted all points by 50 ms

Autocorrelation-based measure of spray stability

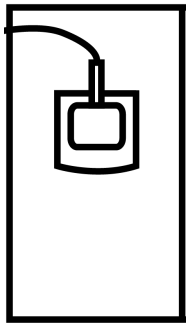


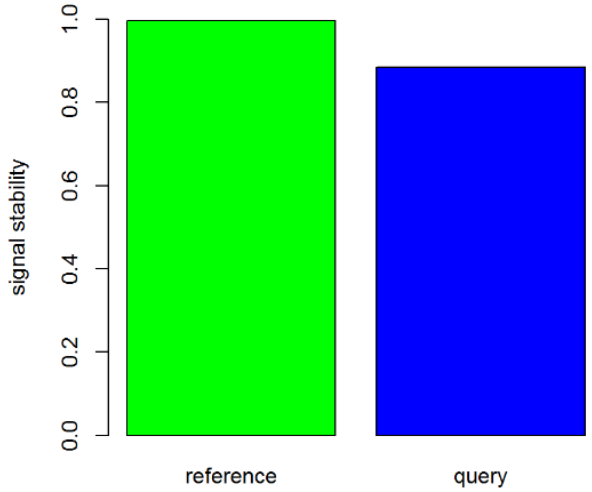




Shifted all points by 50 ms

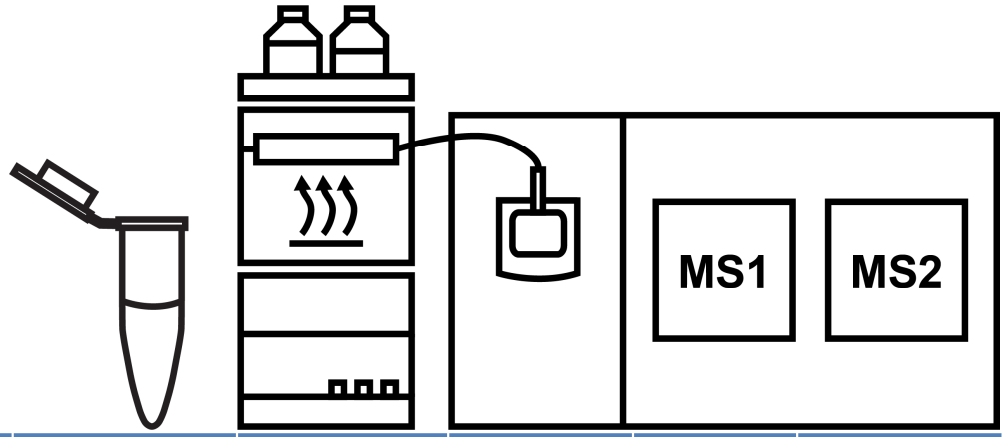
Autocorrelation-based measure of spray stability





- The lag-1 autocorrelation generates a summary of peak stability across the run
- Still working to determine what constitutes good, marginal or bad

ID-free + ID-based metrics



	Sample	LC	Source	MS1	MS2
ID-free	amount loaded on column	peak widths, peak capacity	stability	TIC	Total signal
ID-based	% spectra matched	pH of mobile phases	*z*-dist of PSMs	mass error	mass error; # PSMs

**there are more metrics than these*

msqc.live

Documentation | About | Disclaimer | Privacy Concerns

QC Benchmarker BETA

Save this url for future reference: http://34.74.220.180/index/qc_benchmarker_lbpqsk56

Drop raw file here

By uploading a file you acknowledge reading the disclaimer

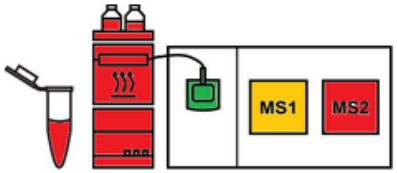
msqc.live

Documentation | About | Disclaimer | Privacy Concerns

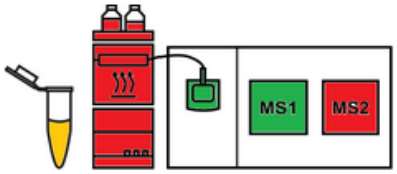
QC Benchmarker BETA

Save this url for future reference: http://104.196.70.201/index/qc_benchmarker_lbpqsk56

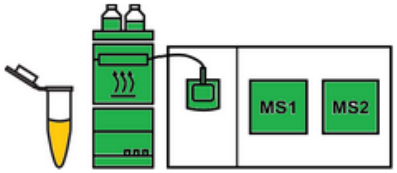
2018-10-18_HeLa_OTOT



2018_06_27_HeLa_OTOT



2018-05-04_HeLa_OTOT



** click on colored modules for full reports underneath*

Ongoing development (v1)

- Summary statistics and figures related to more “classic” quality measures
- Test on non-Thermo platforms
- Generate data and test for Water’s *E. coli* and Promega Yeast
  - please recommend any commercially available digest
- Establish relevant thresholds per platform
- Allow for shorter gradients or different loading amounts
- Model column heater effects and more pH data points

Future development (v2)

- Allow user to specify thresholds
- Allow user to import specific reference file (ex. experiment specific conditions and QC)
- Locally evaluate temporal variations on large data sets
- Evaluate mass spec-based metabolomics data

Questions