When is Enough Actually Enough-How Does the Number of Replicates Influence the Quality of Non-Targeted Analysis Results?

Karen E. Butler,[1,2] Erica Bakota,[3] Christine M. Fisher,[2] Brian Ng,[2] Ann M. Knolhoff[2]

[1]Joint Institute for Food Safety and Applied Nutrition, University of Maryland, 5825 University Research Ct, College Park, MD 20740; [2]Center for Food Safety and Applied Nutrition, Food and Drug Administration, 5001 Campus Dr, College Park, MD 20740; [3]Kansas City Human and Animal Food Laboratory, Food and Drug Administration, 10749 W. 84th Terrace, Lenexa, KS 66214

Introduction

- Non-targeted analysis (NTA) is incredibly useful for the detection and identification of unknown compounds.
- Thousands of compounds can be detected within a single food sample with liquid chromatography coupled to high resolution mass spectrometry (LC/HR-MS); however, ensuring compounds are reliably detected and extracted from each data file is a challenge.
- This study investigates the impact of extraction and injection replicates on the quality of the data output.

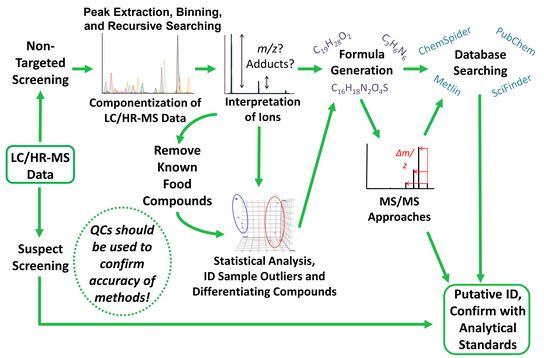
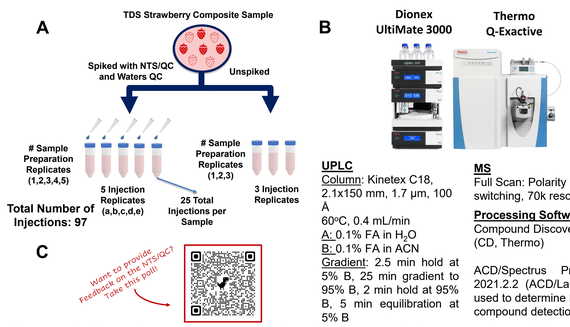


Figure 1. Data Analysis Workflow. Figure modified from Knolhoff, A. M. & Croley, T. R. *J. Chrom. A*. **2016**, *1428*, 86-96.

Materials and Methods

Analytical Strategy for Non-Targeted Analysis of Food Samples



UPLC
Column: Kinetex C18, 2.1x150 mm, 1.7 µm, 100 Å
60°C, 0.4 mL/min
A: 0.1% FA in H_2O
B: 0.1% FA in ACN
Gradient: 2.5 min hold at 5% B, 25 min gradient to 95% B, 2 min hold at 95% B, 5 min equilibration at 5% B

MS
Full Scan: Polarity switching, 70k resolution

Processing Software
Compound Discoverer 3.3 (CD, Thermo)

ACD/Spectrus Processor 2021.2.2 (ACD/Labs) was used to determine NTS/QC compound detection.

Figure 2. Experimental design (**A**) and instrumental parameters (**B**) used to analyze a pooled strawberry sample collected as part of the FDA's Total Diet Study. Each sample was spiked with a previously developed quality control standard mixture (86 compounds covering a broad range of chemical properties: NTS/QC, Knolhoff, A.M. et al, *Anal. Chem*. **2021**, 93(3), 1596-1603) and the LCMS QC Reference Standard (Waters: 9 compounds). More details for method parameters can be found here: Knolhoff, A.M. et al, *Anal. Chem*. **2021**, 93(3), 1596-1603. Please provide feedback on the NTS/QC using the QR code listed in part **C**; this will be used to make a standard like this more widely available and useful to a variety of research disciplines.

Results and Discussion

Analysis of Extraction and Injection Replicates

- A comparable number of molecular features was extracted when processing the same number of injection replicates from different extraction replicates.
- Low %RSDs were observed between the same number of injection replicates processed, with the lowest observed for duplicate replicates of duplicate preparations.

Table 1. Molecular Features Extracted from Data

Replicate Type	Replicates Processed Together	Features Found	Average	%RSD
Single injection of triplicate preparations	1a, 2a, 3a	5379	5318	1.06
	1b, 2b, 3b	5268	5318	1.06
	1c, 2c, 3c	5307	5318	1.06
Duplicate injections of duplicate preparations	1a, 1b, 2a, 2b	5788	5809	0.51
	1a, 1b, 3a, 3b	5843	5809	0.51
	3a, 3b, 2a, 2b	5796	5809	0.51
Triplicate injections of individual preparations	1a, 1b, 1c	5413	5371	1.32
	2a, 2b, 2c	5348	5371	1.32
	3a, 3b, 3c	5262	5371	1.32
	4a, 4b, 4c	5383	5371	1.32
	5a, 5b, 5c	5448	5371	1.32

1,2,3,4,5: extraction replicates
a,b,c: injection replicates

- A similar number of QC compounds were extracted from submissions of subsets of both extraction and injection replicates

Table 2. QC Compounds Extracted from Data

Replicate Type	Replicates Processed Together	NTS/QC Compounds Found (Detectable: 82)	Waters QC Compounds Found (Detectable: 9)
Single injection of triplicate preparations	1a, 2a, 3a	71	9
	1b, 2b, 3b	68	9
	1c, 2c, 3c	70	9
Duplicate injections of duplicate preparations	1a, 1b, 2a, 2b	73	9
	1a, 1b, 3a, 3b	70	9
	3a, 3b, 2a, 2b	70	9
Triplicate injections of individual preparations	1a, 1b, 1c	70	9
	2a, 2b, 2c	70	9
	3a, 3b, 3c	71	9
	4a, 4b, 4c	68	9
	5a, 5b, 5c	68	9

1,2,3,4,5: extraction replicates
a,b,c: injection replicates

Injection & Extraction Replicates- Why Do They Help?

- **Increased Confidence:** In sample preparation and feature extraction quality.
- **Recursive Searching:** Allows for molecular features not initially extracted from an injection to be re-examined and potentially extracted in subsequent processing, which can increase the number of extracted features when submitting multiple injections in a single analysis (**Tables 1, 2, and Figures 3 and 4).**
- **Chemometrics:** Allows for statistical determinations, such as %RSD, to be used to evaluate data quality.

Assessing Extraction of QC and Overall Molecular Features from the Data

- **9/9** Waters QC compounds were extracted in the 1st replicate data file
- **69/82** NTS/QC features were extracted in 1st replicate data file
- An increase in reps facilitated extraction of **5** more NTS/QC features (**74** total)
- **8** NTS/QC features not extracted

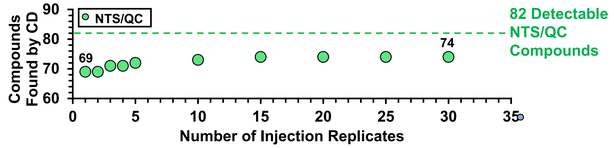


Figure 3. Standard compounds extracted from data created by injecting the same sample 30 times. Data files were processed together as follows: Rep 1; 1+2; 1+2+3; …; Reps 1 through 30.

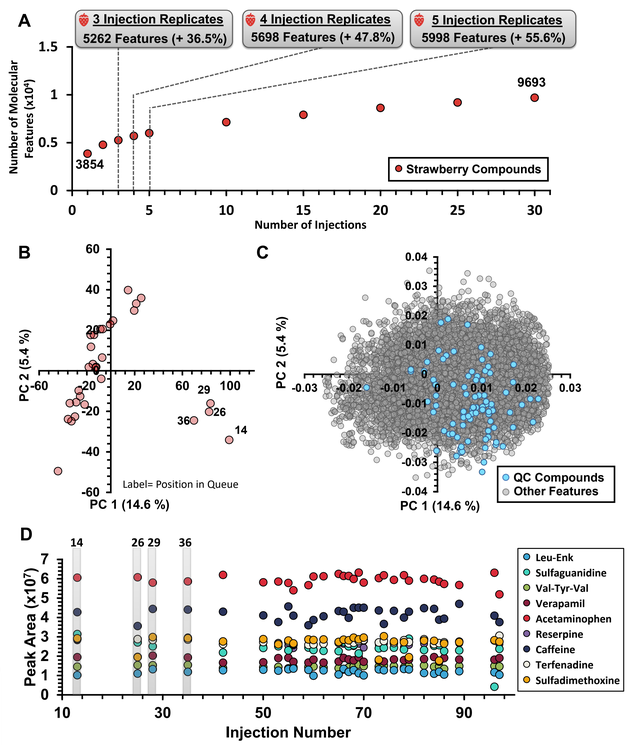


Figure 4. (A) As in Figure 3, processing increasing numbers of injection replicates together facilitated an increase in the total number of molecular features extracted from the data. The PCA and loadings plot for the 30 injection replicates are shown in parts **B** and **C**, respectively. Peak areas for the Waters QC compounds for each injection are shown in part **D**.

Time Considerations- Where is the Time Going?

- Sample preparation: ~5.5 hrs. per sample batch
- Data collection: 35 mins per file
- Data processing: ~3-11 hrs., depending on the number of files processed together (A detailed breakdown of the times for individual processing nodes for an example submission is shown in **Figure 5**).

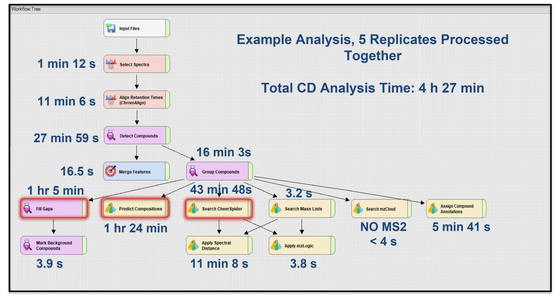


Figure 5. CD 3.3 workflow used for analysis and processing time for each node with 5 replicates processed together.

Conclusion

- Analyzing replicates is important for targeted as well as non-targeted analyses to determine variability and reproducibility. Analyzing one more replicate, even if it's injecting the same sample, can increase the number of features extracted from the data by an average of ~12% per injection replicate.
- Preliminary results demonstrate that triplicate extractions, triplicate injections, and duplicate injection replicates of duplicate preparations resulted in comparable feature extraction from the data.
- The number of replicates that should be used is situational and may depend on a number of factors such as available sample amount, time, and analysis cost.
- Future work will focus on implementing this NTA method and strategy on matrices in other areas of the AOAC foods triangle (**Figure 6**).

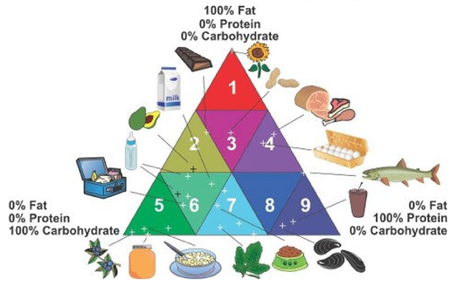


Figure 6. AOAC foods triangle, which classifies foods based on their fat, protein, and carbohydrate content. (*https://www.nist.gov)*).