

Using Mass Spectrometry to Trace Polyphenols in Human Urine

Crystal V. Carrillo¹, Jaewoo Choi², Rachel Hahn^{1,2}, Gerd Bobe^{1,2}, and Jan F. Stevens, Ph.D.^{2,3}

¹College of Agricultural Sciences, ²Linus Pauling Institute, ³College of Pharmacy,
Oregon State University

Abstract

Diet History Questionnaires (DHQ) can only provide general information about a person's nutritional intake. It is unable to cluster groups of people based on diets or lifestyles, nor identify specific foods that may play a role in disease prevention. The aim of this study was to validate the use of mass spectrometry as a method of assessing diet, in comparison to DHQs. Additionally, we used principal components analysis (PCA) to cluster subjects based on the concentrations of polyphenols. We compared values of polyphenols with values of DHQs. Sixty subjects' urine were part of an on-going clinical study of Xanthohumol and Prevention of DNA Damage at the Linus Pauling Institute. DHQs were from the National Institute of Health (NIH) II version. Urine was analyzed using mass spectrometry coupled with an Ultra High Performance Liquid Chromatography (UHPLC). An algorithm was created to compare both the intensities of polyphenols and food groups from DHQ. Principal component analysis was used to analyze polyphenols separately from DHQs. In comparing mass spectrometry results with DHQ, the strongest correlation value ($r > .60$, $P\text{-value} < 0.05$) was measured for secoisolariciresinol (SECO) and strawberry consumption. Using PCA, we found several polyphenols showed varying concentration between subjects; this validates mass spectrometry while showing the limited abilities of DHQs. Analyzing urine using mass spectrometry provided a more detailed analysis of diet when compared to the broad DHQs.

Introduction

History of DHQ and Diet Assessment Limitations

Diet seems to be a good indicator of a person's health or lifestyle. It can also be a means of learning what foods are correlated with lower incidences of certain chronic diseases. Currently Diet History Questionnaires (DHQs) are used to get an estimate of nutrient intake, yielding a 0.48 and 0.49 correlation coefficient correlation between estimated truth and DHQs for women and men respectively, (Subar et al. 2001). Although DHQs are better than other food frequency questionnaires, they have several limitations.

First, the questionnaire requires subjects to recall what they have consumed weeks or even a month prior to taking the survey. Subjects' interest ends rapidly, far before completing the survey.

Secondly, The length of a DHQ II from the National Institute of Health (NIH) is 35 pages, containing 153 questions, and it requires subjects to respond honestly. Finally, subjects give biased responses in addition to their lack of interest filling out the survey (Zamora-Ros et al. 2014). Although the subjects are told they will not be judged by responses, many are untruthful in their responses. All these factors leads to inaccurate responses, which yields incorrect results for nutrient intake (Zamora-Ros et al. 2014).

General Polyphenol Information

Polyphenols are secondary plant metabolites that are found in many foods such as fruits, vegetables, cereals and beverages like tea, coffee, and wines. There are more than 500 kinds of polyphenols, which include stilbene, flavonoids, phenolic acids, and lignans. There are a plethora of polyphenols with distinct molecular structures and MS/MS fragmentations that can be found in the urine. Polyphenols can only be found in foods that originated from plants, they are not

present in meat, poultry, and eggs (Manach et al. 2004). Researchers can use DHQs, but properly estimating polyphenol content is complex because many polyphenols are present in many different types of foods.

It is also important to consider bioavailability; bioavailability is influential on the metabolome because of the large molecular diversity in polyphenols. Polyphenols are usually excreted within 24-48 hours after consumption (Pérez-Jiménez et al. 2010). Recovery of polyphenols in urine is determined by bioavailability of polyphenols. For example anthocyanins are found at levels as 0.01% while isoflavones are at levels high as 43%. It is important to note that glycosylation of flavonoids and esterification of phenolic acids affects the absorption in the gut. Bioavailability should be taken into account while assessing how polyphenols in the urine are metabolized (Zamora-Ros et al. 2014).

An analysis of recent publications that have looked at total polyphenol intake concluded that it would be more beneficial to look at one phenol at a time. The author also add that biomarkers' limitations as diet assessors can be challenging because of the rapid absorption and elimination. Making observational studies unable to pin point specific biomarkers (Zamora-Ros et al. 2014). Studies show that phyto-estrogens act as inhibitors of carcinogenic and atherosclerotic processes (Mazur et al. 2000). There are many other studies connecting polyphenols to anti-inflammatory, anti-glycemic, and cancer chemopreventive effects (Stevens, 2014).

A controlled human study in Finland found that blackberries and strawberries had the highest amounts of Secoisolariciresinol (SECO) concentrations detected. The study measured urine and plasma concentrations before and after consumption of a strawberry meal. Subjects showed relatively high amounts of plant lignin SECO and elevated amount of enterolactone (ENL) in urinary excretion (Mazur et al. 2000). Many more publications have suggested that polyphenols

play a vital role in disease prevention. A systematic review explains that recent polyphenol studies, results were obtained with fixed servings and polyphenol intake was higher than normal diets (Zamora-Ros et al. 2014).

Mass spectrometry is a frequently used instrument; it is able to collect large amounts of lightweight molecules such as polyphenols. Mass spectrometry analysis on human urine has helped in discovering the health properties of polyphenols that have the potential to reduce chronic diseases (Manach et al. 2005).

Validating the Use of Polyphenols in the Urine Compared to Plasma

Polyphenols are harder to find in blood plasma than urine. Polyphenols have been found in higher concentrations in urine compared to plasma. In addition, the sample processing for urine is less complicated compared to plasma processing before analysis (Zamora-Ros, 2014). Previous studies have looked at mass spectrometry analysis of urine and blood, but phenols have not been found in the blood; instead they have been found bound to proteins in plasma, primarily albumins. The affinity of polyphenols to albumins depends on their chemical structure (Manach et al. 2005).

Direction of study

With increased availability of technology, nutrient assessment should be done with mass spectrometry. There is a need for a predictor of specific polyphenols that can lead to reduction of diseases that cannot be seen or determined using DHQs. DHQs are very limiting in what information they are able to communicate to researchers. The objective of this study, therefore, was to determine whether there is any correlation between the polyphenols found in human urine

and the DHQs food data output. We hypothesize that there should be a correlation between these two pieces of information considering both reflect what subjects have consumed.

Methods and Materials

The Subjects

Human subjects were used from an on-going clinical study called Xanthohumol and Prevention of DNA Damage. The study was approved by Oregon State University Institutional Review Board and followed the Human Research Protection Program regulations. The purpose of the study was to test the effects of Xanthohumol on oxidative DNA. To participate in the study, subjects had to stop consumption of high levels of flavonoids and xanthohumol in the normal diet (onions, teas including green/black tea and microbrew beers) for 2 weeks prior to study entry through conclusion of study (for the complete list refer to Xanthohumol and Prevention of DNA damage clinical study). Recruitment included subjects both male and female ages 18-50 years old. Advertisement included: OSU today, Craigslist, and flyers posted around Corvallis area. Each subject was in the study for a duration of 13 weeks. Diet restrictions included no onions, black tea, and microbrew brews for the complete list refer to Xanthohumol and Prevention of DNA Damage clinical study website (Clinicaltrials.gov). Subjects had a list of exclusion during the study, exclusions included; Engaging in vigorous exercise more than 6 hours per week and participation in another dietary study in the past 3 months, for complete list refer to Xanthohumol and Prevention DNA damage clinical study website.

There was also a screening that happened before the subject was able to further participate in the study. Smokers were excluded; also anyone that had too low or high LDL was excluded from study. After subjects were screened, they were given a list of foods/beverages to refrain from for the duration of the study (Steven et al. 2016).

Enrolled subjects visited the Clinical Research Center (CRC) six times. At the beginning and end of each treatment period (2-5), subjects also provided urine collected during the previous 24 hours. The screening visit lasted about 45 minutes and subsequent visits lasted 30 minutes or less.

The Design

24 hours urine samples of the 60 subjects were collected after 8-10 hours of fasting. Each subject was asked to fill out a dietary history questionnaire twice during the study, during visit 3 and 5. Each subject started the study at different times. Urine was collected and stored at -80°C.

A master list of polyphenols was created through PeakView and Phenol-Explorer (<http://www.phenol-explorer.eu>) databases. PeakView is a software program that found polyphenols at the appropriate molecular weight and retention times. Phenol-Explorer and METLIN (<https://metlin.scripps.edu/index.php>) databases were used to measure polyphenol intake accurately. Phenol-Explorer is a polyphenol repository of 500 polyphenols and food compositions. Phenol-Explorer contains the phenol name, class, subclass, molecular weight, chemical formula, and chemical structure. This database is able to show the composition of polyphenols in foods and beverages. There are 459 foods logged with 9 classes and 67 sub-classes of polyphenols (<http://www.phenol-explorer.eu>).

METLIN is a database repository of polyphenol/metabolite information and tandem mass spectrometry data used to help researchers positively identify metabolites. METLIN provides MS/MS metabolite data in both negative and positive electrospray ionization mode using four different collision energies. If there are multiple polyphenols that have the same molecular weight METLIN: Metabolite search can provide appropriate fragmentations for specific phenol

(<https://metlin.scripps.edu/index.php>). The majority of polyphenols were found in the negative ion mode, but both negative and positive ion modes were viewed.

Dietary History Questionnaire Software was used to get a rough estimate of what foods each subject consumed (Subar et al. 2001). Main categories looked at were whole grains, non-whole grains, dark green vegetables, orange vegetables, white potato, other starchy vegetables, tomatoes, other vegetables, citrus/melon/berry, soy, nuts, and legumes. Each category had at least 18 food items that fell within fruits and vegetable groups.

Once all the urine was collected from the subjects, urine was removed from -80° C, and allowed to thaw to make dilutions. Urine sample preparation consisted of diluting urine 6-fold with water 50 microliters of urine and 250 microliters of water. The tubes were vortexed for three seconds. Then 300 microliters of dilution was centrifuged at 1300 revolutions per minute for 5 minutes. Then 100 microliters of supernatant was aliquoted into mass spectrometry vials. This process is called normalization to specific gravity (Edmands et al. 2014). Urine samples were analyzed using liquid chromatography coupled to a high resolution accurate mass spectrometer called AB Sciex 5600 TripleTOF (AB Sciex, Maine) coupled to a Shimadzu Nexera Ultra High Performance Liquid Chromatography (UHPLC; Shimadzu, California).

A total of 133 urine sample were collected. There were 73 subjects that remained in the study through visit 3; then soon after 13 of those 73 did not complete the remainder of the study. All urine samples mass spectrometry data were uploaded to PeakView. Using PeakView and METLIN only, a library of 65 positively identified polyphenols was created to identify supposed polyphenols in the other subjects. Based on 65 polyphenols' intensities, only 10 polyphenols/metabolites were chosen to be looked at further. These 10 were chosen based on polyphenols that previous studies have looked at to have a better understanding of metabolism.

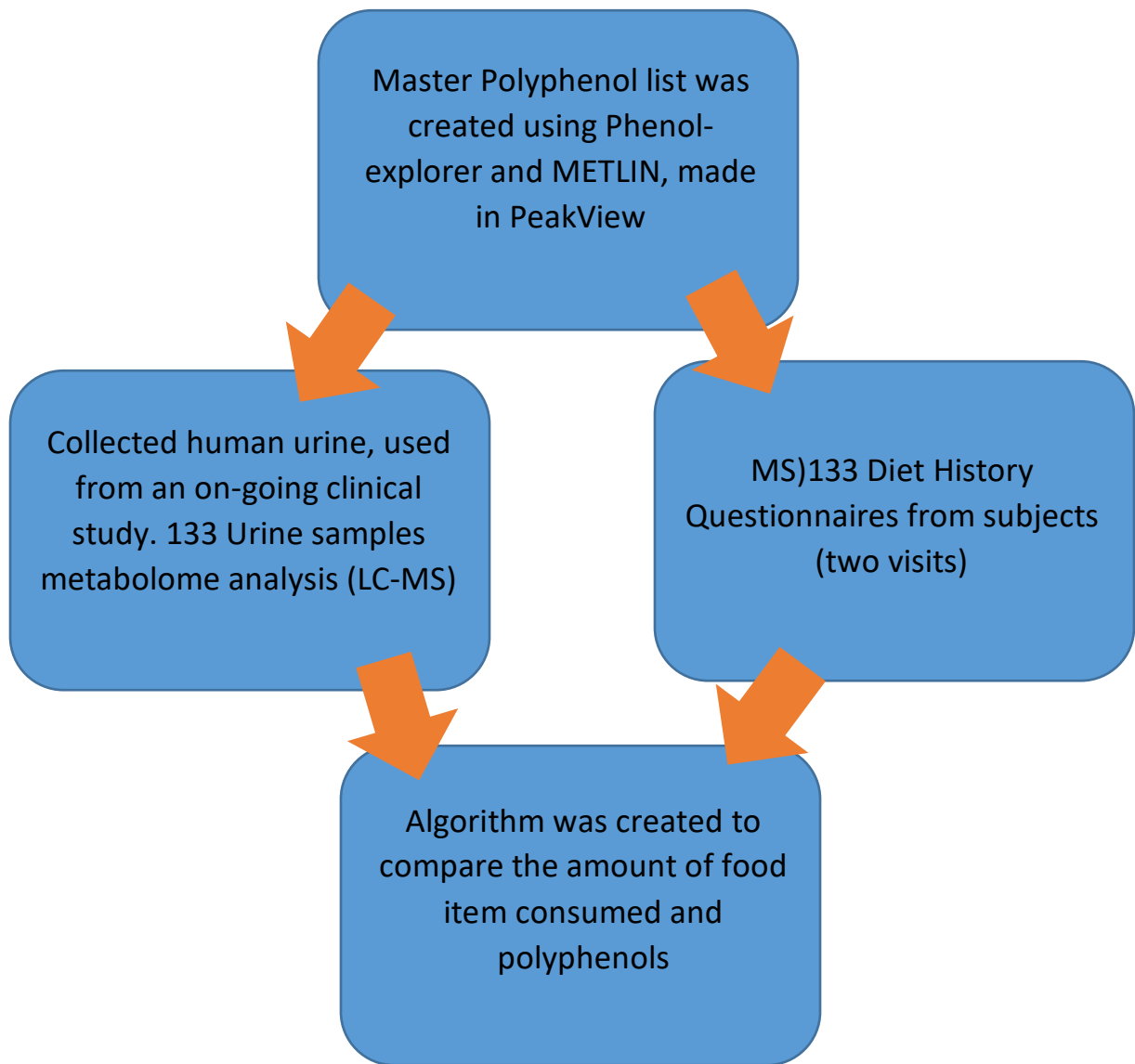


Figure 1 Method one compared only the intensities of polyphenols along with DHQs results.

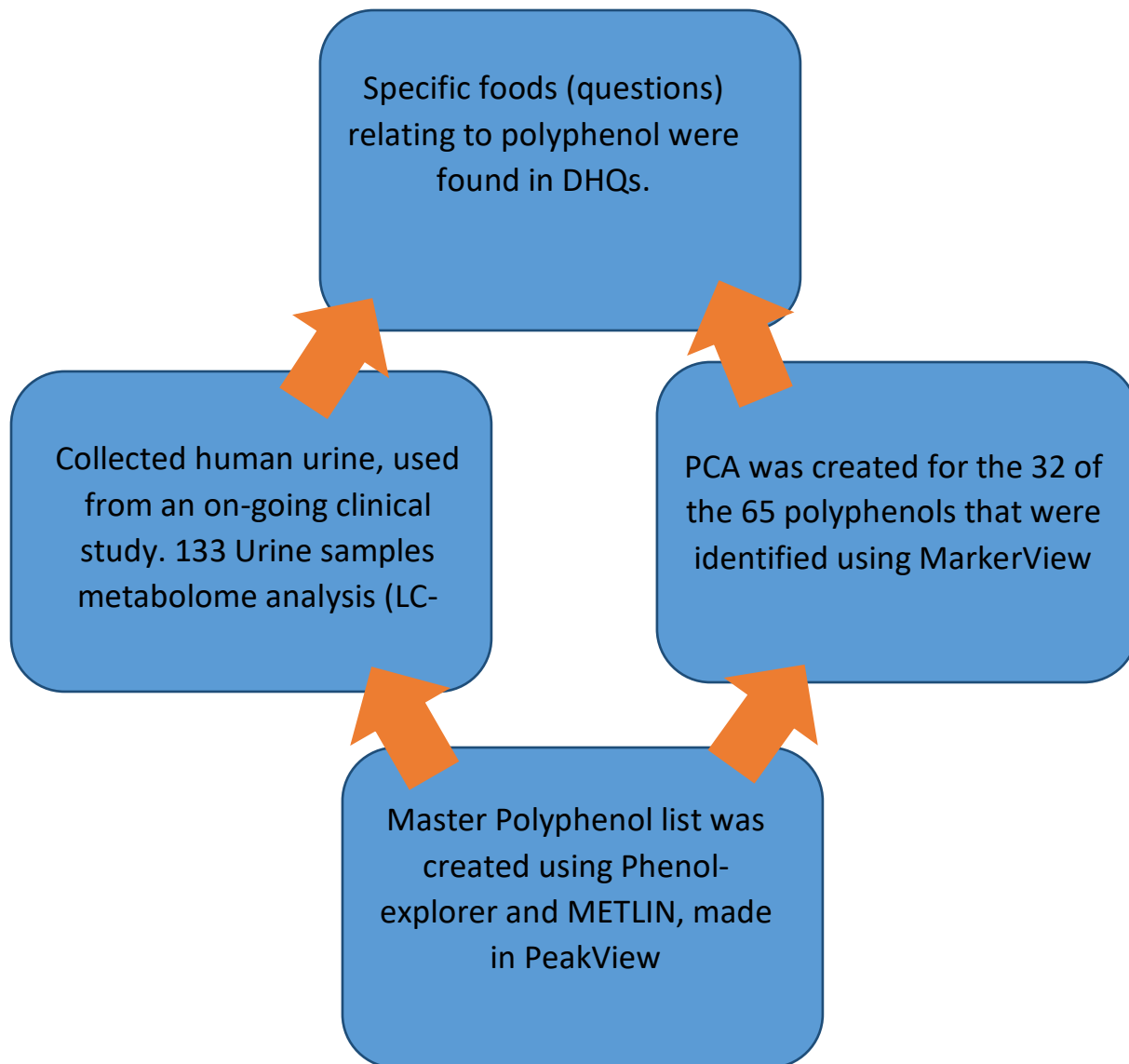


Figure 2 Shows a second method of PCA followed by looking at the DHQs

Sorting Through Raw Data

There were four excel files, which consisted of quantitative values for the DHQs' and mass spectrometry for both visit 3 and visit 5. In order to find any patterns between DHQs and polyphenol intensities, an algorithm was created comparing the two files and finding correlating coefficient values (Figure 1). DHQs only form general grouping of fruits and vegetables, such as dark green vegetables (broccoli, cilantro, kale, lettuce), which contains 39 other types of vegetables. This made it difficult to look at a single food that contained the specific polyphenol. The objective was to be able to predict the amount of phenols specific to a certain amount of food. To narrow down a specific food from a general dark green vegetable group would be challenging. Each selected question was given numerical values to the responses depending on how often and the amount consumed for each food. The values were computed for each subject and compared to intensities.

Different Approach

The second method followed the first step of method one; the master library list was used to find the polyphenols MarkerView had found from the mass spectrometry data. The second method used Principal Component Analysis (PCA) and only found 32 of the polyphenols compared to the 65 polyphenols found using the first method. MarkerView finds all molecular compounds in urine data, but only polyphenols that matched in retention time and molecular weight. The PCA preprocessing modes were set to logarithm and Pareto. PCA is very efficient at looking at multiple polyphenols and finding patterns or variances. It is important to note that PC-1 is an entirely new plane that represents maximum variance between variables this is the loading plot. It then graphs individual subjects on the score plot based in the amount of the found polyphenols

concentrations. If one subject had more of one polyphenol compared to another then it would be closer to the polyphenol with the higher concentration.

Results

Visit 5 data had the best results; it was assumed that visit 5 had better results because subjects were more cognizant of what was consumed over the previous 3 week period. From visit 5 data a phenol secoisolariciresinol found in strawberries showed a correlation coefficient of 0.6176 (Figure 3), P-value<0.05 (Table 1). The retention time was found at 13.36 seconds in the mass spectrometry urine samples. The parent compound is secoisolariciresinol and the metabolites are enterodiols and enterolactone (figure 3). The correlation coefficient values and p-values for enterodiols and enterolactone were not as strong as secoisolariciresinol (Table 1).

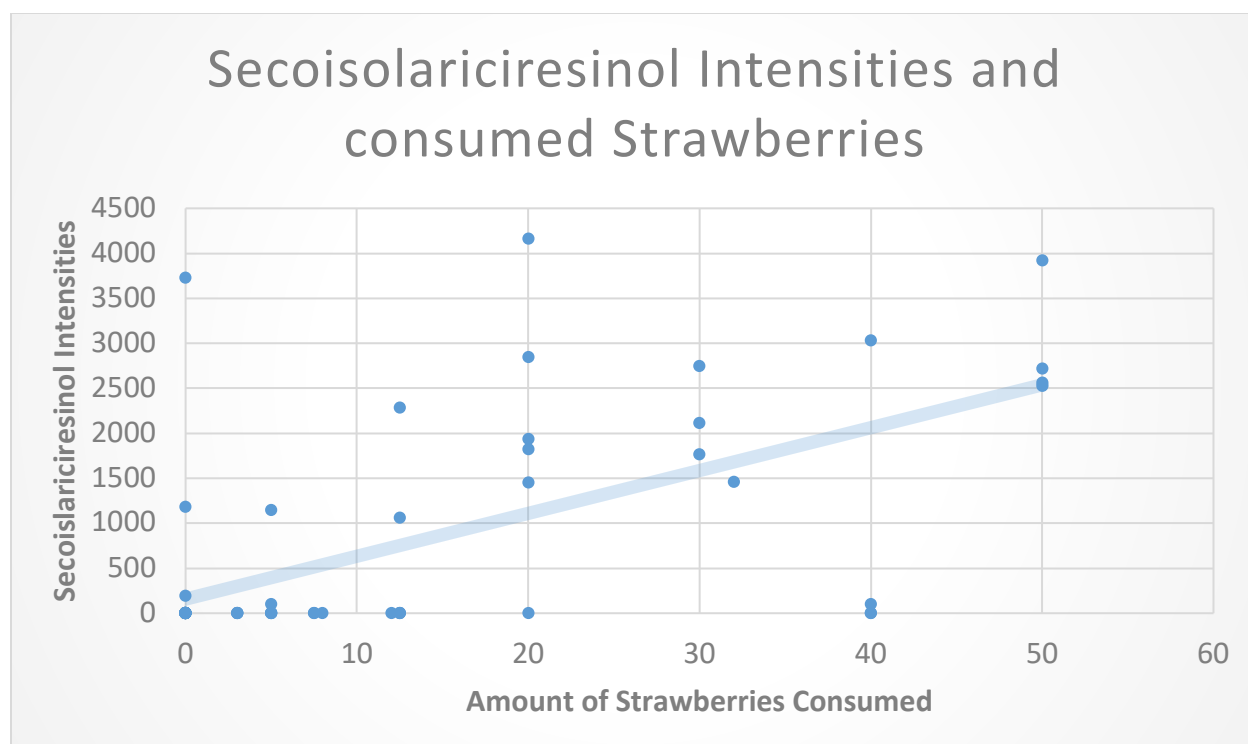


Figure 3 Secoisolariciresinol intensities increase with the increasing values of strawberry consumption.

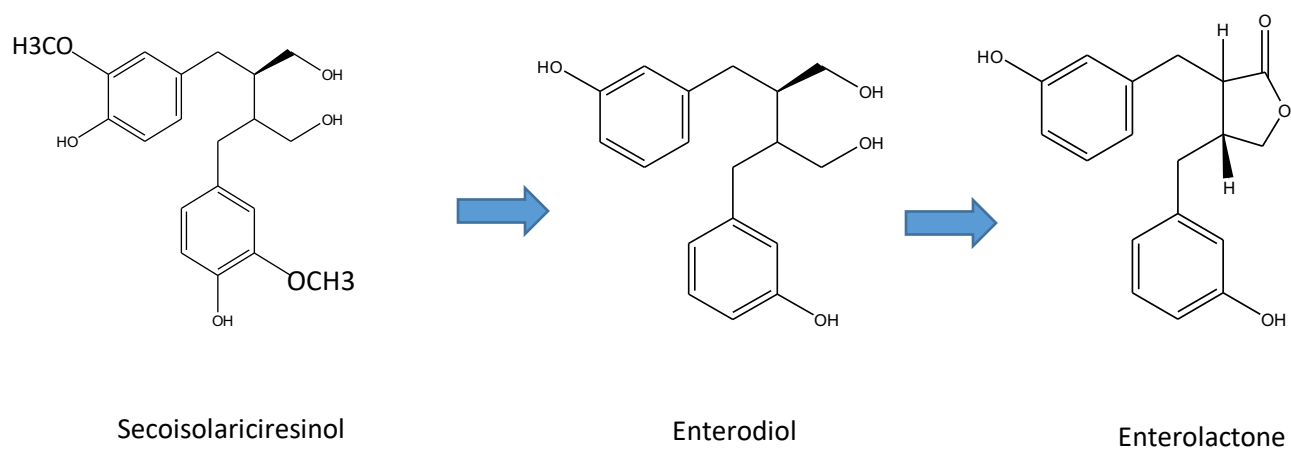


Figure 4 Metabolism of Secoisolariciresinol is followed by Enterdiol and Enterlactone. These metabolites are metabolized through the gut and may lead to different products depending on the microbiota.

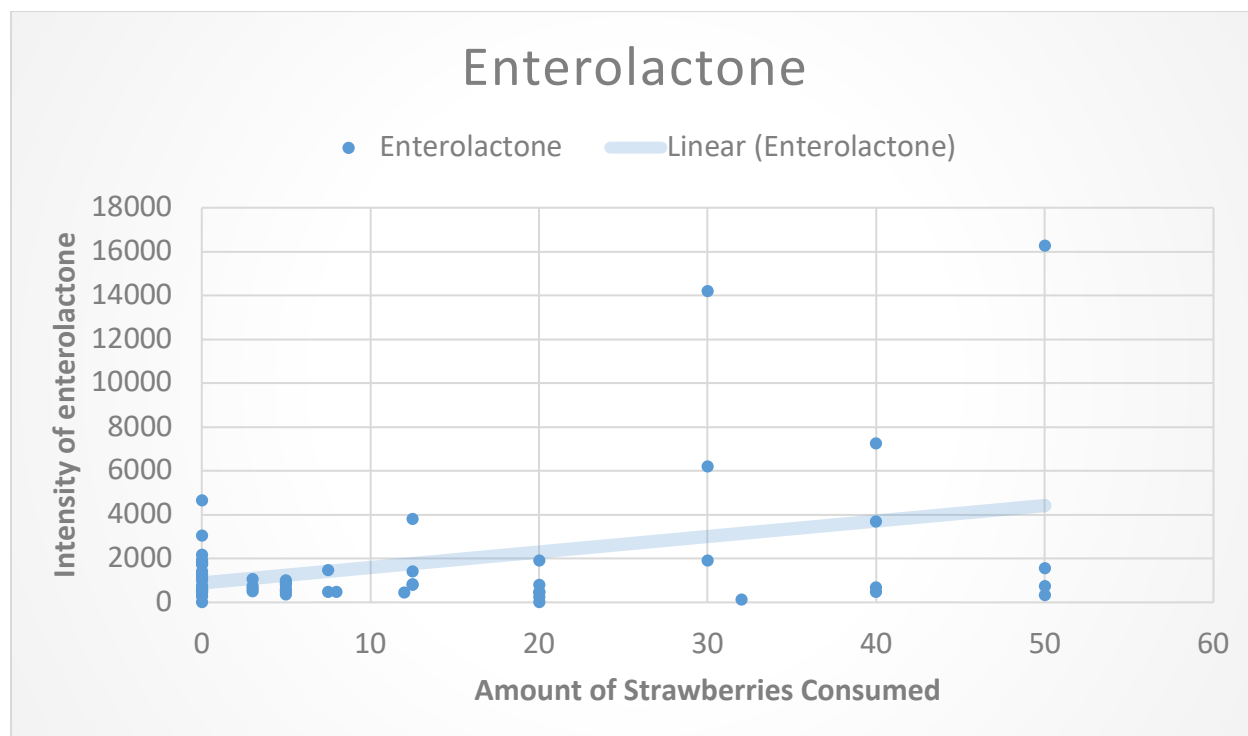


Figure 5 A graph was created comparing the amount of strawberries consumed and the intensities of Enterolactone. The metabolite Enterolactone shows similar patterns as Secoisolariciresinol.

Table 1 The p-values and correlation coefficients for enterolignans and ellagitannins polyphenols. The only polyphenols that showed a strong relationship between strawberries consumed were SECO and ENL.

Name	Correlation coefficient	p-value
Secoisolariciresinol diglucoside	-0.02688	0.6832
Secoisolariciresinol	0.61766	-3.47E5
Enterdiol	-0.08394	0.3478
Enterolactone	0.38201	0.01118
Ellagitannin	-0.14678	0.3282
Urolithin C	-0.04478	0.8358
Urolithin A	-0.05419	0.7236
Urolithin B	-0.13874	0.3840

PCA plots can be viewed looking at individual subject or by polyphenols. Both loading and score plots show the variance between PCA1 explains 55.8% of the variance while PCA2 only explains 8.6% (Figure 6).

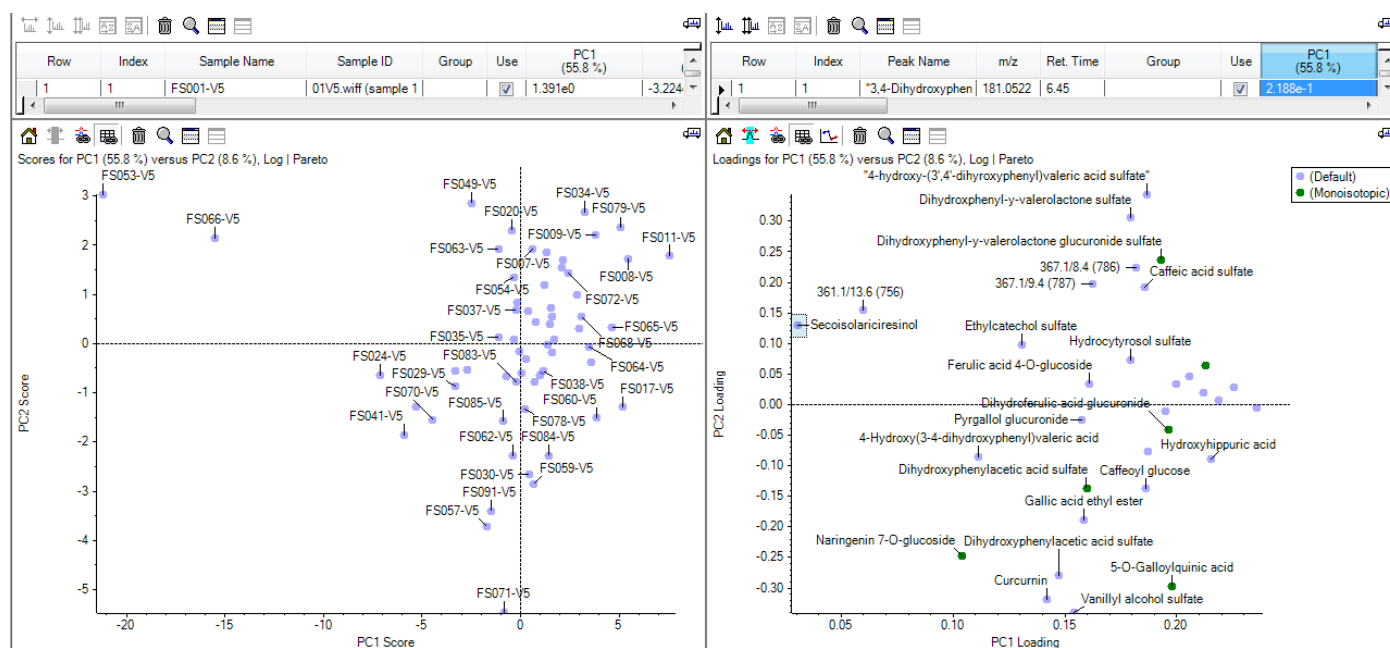


Figure 6 Subjects and polyphenols that are closer together are more similar to each other compared to others that are further away from each other.

Polyphenol Naringenin 7-O-glucuronide is a compound found mostly in grapefruit, tomatoes, and almonds. There were nine subjects with Naringenin 7-O gulcuronide exceptionally higher than the other subjects. MarkerView shows the peak area of the polyphenol for each subject and the numbered visit. When the amounts of polyphenols were graphed comparing the DHQs a few of the subjects showed that although the amount of naringenin was relatively high, the amount reported to DHQs were low (Figure 8).

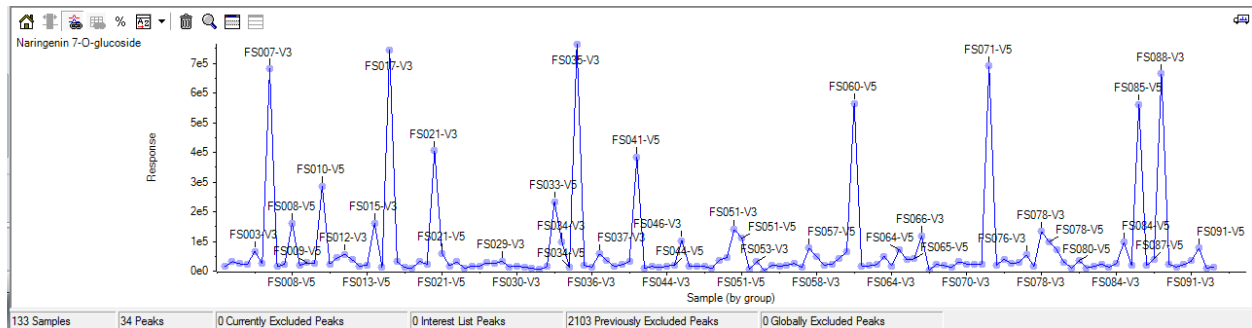


Figure 7 Naringenin-7-O glucuronide between subjects at different amounts. The subjects can be found on the x-axis and peak areas are shown on the y-axis. Naringenin 7-O-glucuronide can be found in grapefruit, tomato, and almonds

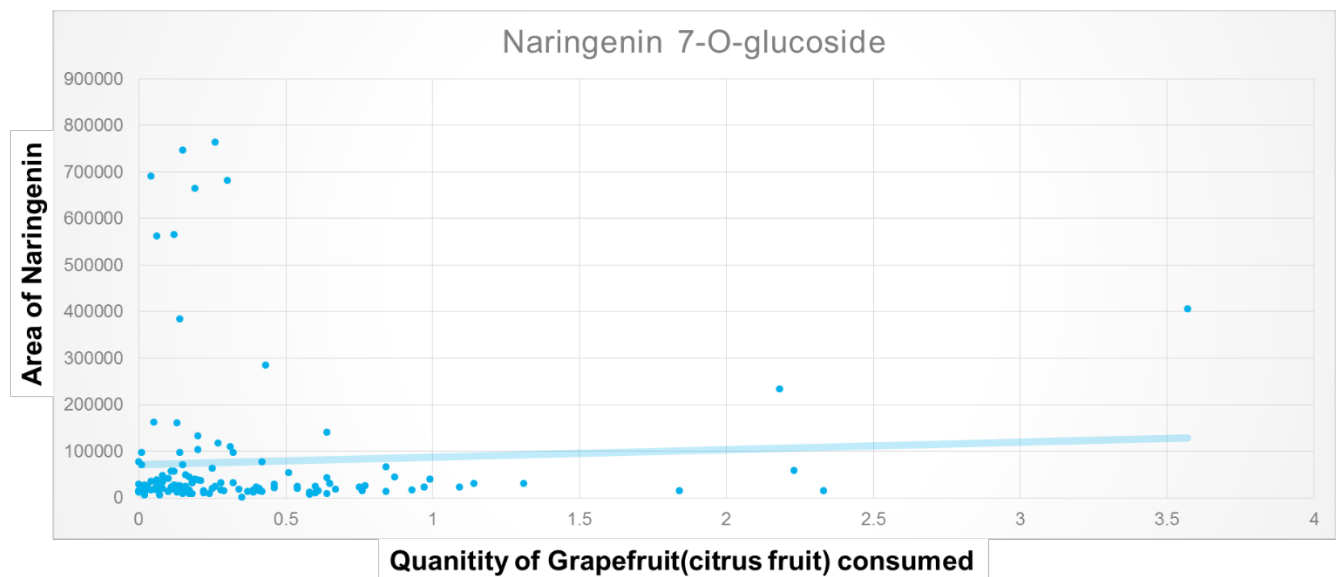


Figure 8 Amount reported by DHQs and amount of Naringenin for each subject were graphed. A majority of the cluster is focused on the lower left quadrant, this shows that both the amount of citrus fruits and naringenin were both low amount.

For Polyphenol, urolithin C, there were only 2 subjects (FS-060-V5 and FS092_V3). The area of the peaks of urolithin C is representative of its concentration (Figure 9). The peak area values were plotted against the amount of nuts a person consumed (Figure 10).

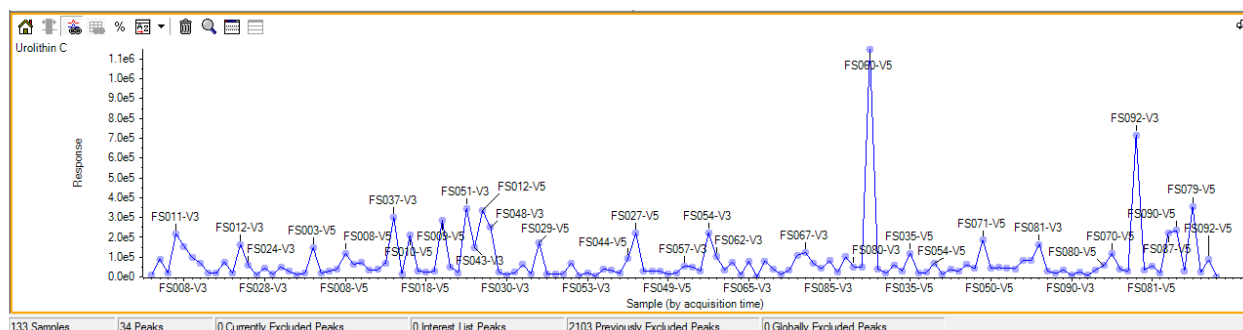


Figure 9 MarkerView shows the amount at which Urolithin was found within each subject.

Urolithin C is found in strawberries and pomegranate so either of these fruits can account for the higher concentrations in subjects.

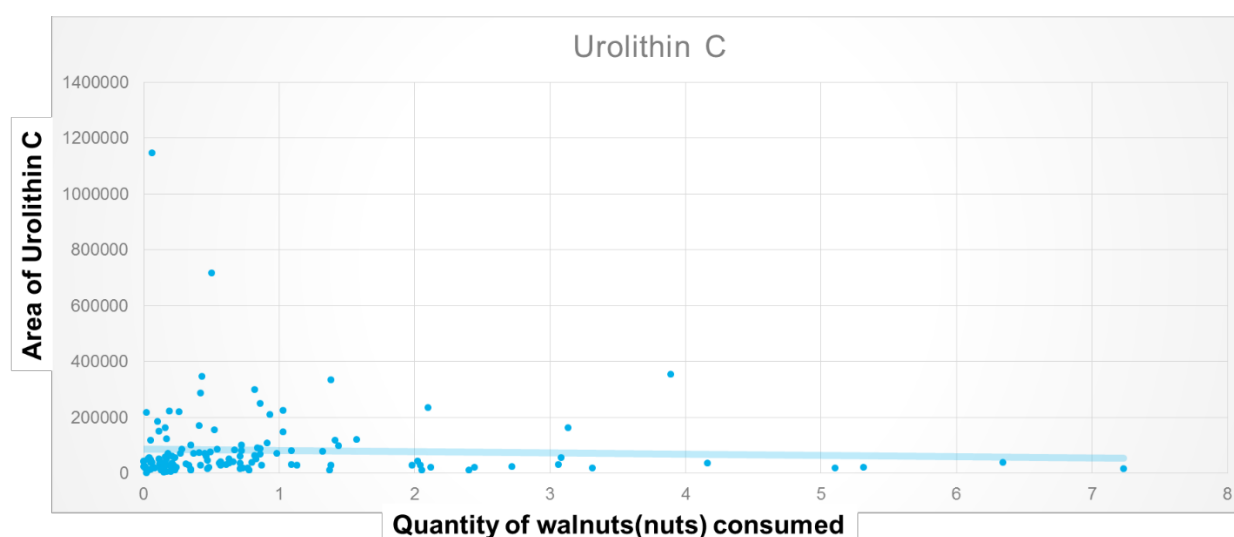


Figure 10 shows the quantity of walnuts (nuts) vs Urolithin C amounts. Urolithins are polyphenols that are well known in polyphenol studies. Urolithins are metabolites of gallic acids and ellagitannins.

The final polyphenol examined was dihydroxyphenyl valeric acid a compound found in chocolate. It can be seen between different subjects, this shows that each consumed a different amount to represent the varying peaks. There seems to be a positive correlation between chocolate and dihydroxyphenyl valeric acid(Graph 5).

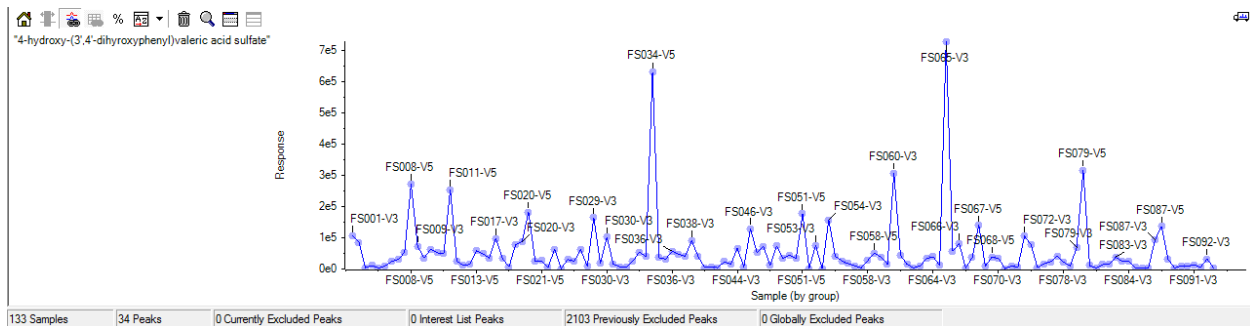


Figure 11: MarkerView shows the amounts of 4-hydroxy-(3',4'-dihydroxyphenyl) valeric acid, found in chocolate, between subjects. Subjects 34, 79, and 65 have the 3 highest concentrations in comparison to other subjects.

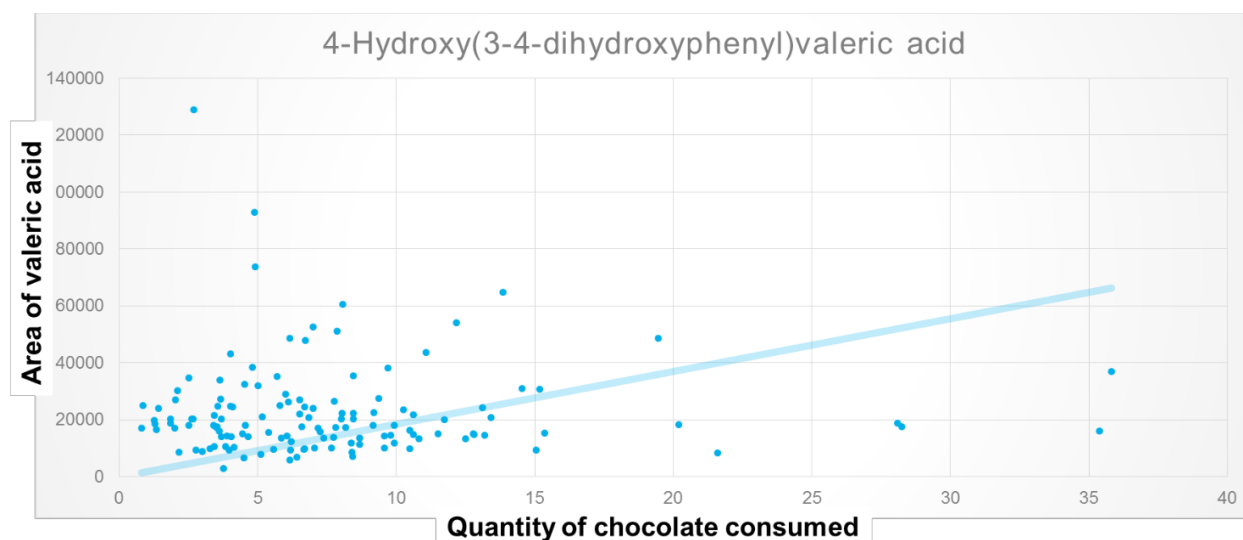


Figure 12 graphing the amount of chocolate consumed and the varying amounts of 4-hydroxy-(3',4'-dihydroxyphenyl) valeric acid. The data shows that there is not a strong correlation between chocolate consumed and valeric acid.

Discussion

Interpreting data

The correlation graph (Figure 3) shows a strong correlation between SECO and the number of strawberries consumed over the three-week period. If the graph were split into quadrants, anyone who occupied the top left corner would have reported not eating strawberries, but intensity of SECO would have reflected a large amount of strawberries consumed. The lower right quadrant shows subjects reported consuming 40 strawberries while SECO was not found in the urine.

Another biomarker that had a similar correlation was enterolactone (ENL), correlating at a 0.40 with a P-value<0.05. ENL is single product of SECO and matairesinol (MAT). The other lignans did not show any compelling values. Ellagitannins is another polyphenol found in strawberries. But a study by Pilar Truchado, found that urolithin A and urolithin A glucuronide expressed a high individual variation due to the number of gut microbiota converting ellagic acid into urolithin A (2011). Urolithin B was also observed to only show in 20% of total population,

which suggests that only a small portion of the population could have the microbiota needed to metabolize gallic acid into urolithin B. This could be due to the individual host and gut microbiota between people; there is a high variability of gut microbiota among individual (Truchado. 2011).

A principal component analysis from the 32 positively identified polyphenols shows SECO being the most dissimilar between subjects shown in the loading side.

Limitations

It's important to note the limitations of this study; there are several limitations that this study along with other have found when looking into polyphenols and other biomarkers. First, most polyphenols are excreted from the body 24 hours after consumption (Manach et al. 2005). DHQs were filled out according to what subject ate within a 3 week period, while urine was collected for only 24 hours before visits.

This study also looked at polyphenols and it was only the polyphenols that showed a correlation value. Also, out of the >500 polyphenols that were looked for, only 65 polyphenols were confidently identified due to the MS-MS data and retention times.

The different microbiota found in each individual is also a limitation, unfortunately not all the individuals contain the same gut microbiota that convert certain polyphenols to the same product.

Conclusion

Mass spectrometry provides advantages, used along with databases such as Phenol-Explorer and METLIN to better assess the benefits of incorporating certain polyphenols to prevent chronic diseases. Urine analysis can make a strong impact, narrowing down not only what foods but what polyphenols can lower certain diseases.

Acknowledgements

Special thanks to Fred Stevens, Jaewoo Choi, Sandra Uesegui, Rachel Hahn, and Gerd Bobe for aiding in proper lab techniques and my project. Multicultural Scholars Program for its mentorship and funding. Wanda Crannell and Kate Field for their guidance and advising.

References

- Delman DM, Fabian CJ, Kimler BF, Yeh H, Petroff BK. 2015. "Effects of Flaxseed Lignan Secoisolariciresinol Diglucoside on Preneoplastic Biomarkers of Cancer Progression in a Model of Simultaneous Breast and Ovarian Cancer Development." *Nutrition and Cancer Journal* 67(5):857-864.
- Edmands, William M. B., Pietro Ferrari, and Augustin Scalbert. 2014. "Normalization to Specific Gravity Prior to Analysis Improves Information Recovery from High Resolution Mass Spectrometry Metabolomic Profiles of Human Urine." *Analytical Chemistry* 86 (21): 10925–31. doi:10.1021/ac503190m.
- Manach, Claudine, Augustin Scalbert, Christine Morand, Christian Rémésy, and Liliana Jiménez. 2004. "Polyphenols: Food Sources and Bioavailability." *The American Journal of Clinical Nutrition* 79 (5): 727–47.
- Manach, Claudine, Gary Williamson, Christine Morand, Augustin Scalbert, and Christian Rémésy. 2005. "Bioavailability and Bioefficacy of Polyphenols in Humans. I. Review of 97 Bioavailability Studies." *The American Journal of Clinical Nutrition* 81 (1): 230S – 242S.
- Mazur, W. M., M. Uehara, K. Wähälä, and H. Adlercreutz. 2000. "Phyto-Oestrogen Content of Berries, and Plasma Concentrations and Urinary Excretion of Enterolactone after a Single Strawberry-Meal in Human Subjects." *British Journal of Nutrition* 83 (04): 381–87. doi:10.1017/S0007114500000489.
- Pérez-Jiménez, J., V. Neveu, F. Vos, and A. Scalbert. 2010. "Identification of the 100 Richest Dietary Sources of Polyphenols: An Application of the Phenol-Explorer Database." *European Journal of Clinical Nutrition* 64 (S3): S112–20. doi:10.1038/ejcn.2010.221.
- Rothwell J, Medina-Rejon A, Perez-Jimenez J, Neveu V, Knaze V, Slimani N, Scalbert A. 2014. "Effects of food processing on polyphenol contents: A systematic analysis using phenol-Explorer." *Molecular Nutrition & Food Research* 59(1): 160-170.
- Subar, A. F., F. E. Thompson, V. Kipnis, D. Midthune, P. Hurwitz, S. McNutt, A. McIntosh, and S. Rosenfeld. 2001. "Comparative Validation of the Block, Willett, and National Cancer Institute Food Frequency Questionnaires: The Eating at America's Table Study." *American Journal of Epidemiology* 154 (12): 1089–99.
- Stevens, Jan Fredrick., Maier, Claudia S. 2014. "Phytochemistry Reviews-The Chemistry of Gut Microbial Metabolism of Polyphenols."
- Stevens, Jan Fredrick., Xanthohumol and Prevention of DNA Damage. *Clinical Trials*. <https://clinicaltrials.gov/ct2/show/NCT02432651>
- Truchado P, Larrosa M, García-Conesa MT, Cerdá B, Vidal-Guevara ML, Tomás-Barberán FA, Espín JC. 2011. "Strawberry processing does not affect the production and urinary excretion of urolithins, ellagic acid metabolites, in humans." *Journal of Agricultural and Food Chemistry* 60 (23):5749-5754.
- Zamora-Ros, Raul, Marina Touillaud, Joseph A. Rothwell, Isabelle Romieu, and Augustin Scalbert. 2014. "Measuring Exposure to the Polyphenol Metabolome in Observational Epidemiologic Studies: Current Tools and Applications and Their Limits." *The American Journal of Clinical Nutrition*, July, ajcn.077743. doi:10.3945/ajcn.113.077743.

