Demonstration of the utility of biomarkers for dietary intake assessment; proline betaine as an example

Helena Gibbons[1], Charlotte JR Michielsen[1], Milena Rundle[2], Gary Frost[2], Breige A McNulty[1], Anne P Nugent[1], Janette Walton[3], Albert Flynn[3], Michael J Gibney[1] and Lorraine Brennan[1]*

[1]School of Agriculture and Food Science, Institute of Food and Health, University College Dublin, Dublin, Ireland

[2]Nutrition and Dietetic Research Group, Division of Endocrinology and Metabolism, Imperial College London, London W12 0NN, U.K.

[3]School of Food and Nutritional Sciences, University College Cork, Cork, Ireland

Keywords: citrus fruit, dietary assessment, dose-response, metabolomics, proline betaine.

Address correspondence to L Brennan, Institute of Food and Health, Belfield, University College Dublin, Dublin 4, Ireland. E-mail: lorraine.brennan@ucd.ie

Abbreviations: NANS, National Adult Nutrition Survey

Abstract (no more than 200 words)

Scope: There is a dearth of studies demonstrating the use of dietary biomarkers for determination of food intake. The objective of this study was to develop calibration curves for use in quantifying citrus intakes in an independent cohort.

Methods and results: Participants (n=50) from the NutriTech food-intake study consumed standardized breakfasts for three consecutive days over three consecutive weeks. Orange juice intake decreased over the weeks. Urine samples were analyzed by NMR-spectroscopy and proline betaine was quantified and normalized to osmolality. Calibration curves were developed and used to predict citrus intake in an independent cohort; the Irish National Adult Nutrition Survey (NANS) (n=565). Proline betaine displayed a dose-response relationship to orange juice intake in 24h and fasting samples ($p<0.001$). In a test set, predicted orange juice intakes displayed excellent agreement with true intake. There were significant associations between predicted intake measured in 24h and fasting samples and true intake($r=0.710$-0.919). Citrus intakes predicted for the NANS cohort demonstrated good agreement with self-reported intake and this agreement improved following normalization to osmolality.

Conclusion: The developed calibration curves successfully predicted citrus intakes in an independent cohort. Expansion of this approach to other foods will be important for the development of objective intake measurements.

1 Introduction

In an endeavor to overcome some of the methodological issues associated with current dietary assessment methods, dietary biomarkers are being utilized. Dietary biomarkers, for example urinary nitrogen a marker of protein intake, provide unbiased estimates of intake and can therefore be used to validate classical self-reporting approaches [1, 2]. More recently metabolomics has emerged as a valuable tool in the discovery of dietary biomarkers. A number of dietary biomarkers have been successfully identified including biomarkers of fish [3-5], red meat [6-8], cruciferous vegetables [9, 10], whole-grain cereals [11, 12] and coffee [13, 14]. However, the majority of these studies with the exception of alkylresorcinols, biomarkers of wholegrain intake, have not demonstrated dose-response relationships between dietary biomarkers and food intake. Furthermore, clear examples of how such biomarkers can be used for assessment of intake are lacking.

To date one of the most studied dietary biomarkers identified using a metabolomics approach is proline betaine [15, 16]. A number of acute and medium term interventions and cohort studies have identified proline betaine as a robust biomarker of citrus fruit intake [15-20]. Proline betaine was originally identified as a potential citrus fruit biomarker by Atkinson et al. [21]. Following this Heinzmann and colleagues preformed an acute intervention study [16]. In this acute study participants consumed a mixed-fruit meal and urine samples were collected and analyzed by ^{1}H Nuclear Magnetic Resonance (NMR) [16]. Multivariate analysis identified proline betaine as a potential biomarker of citrus fruit intake. Furthermore, proline betaine was assessed in a number of fruit and commercially available fruit juices and was found in higher concentrations in citrus fruit. The urinary excretion profile of proline betaine was measured following orange juice consumption in six participants. This biomarker was then confirmed using data from participants in the INTERMAP UK cohort, demonstrating a

high sensitivity and specificity for citrus fruit consumption (90.6 % and 86.3 %, respectively) and a significant correlation with citrus consumption ($R^2 = 0.80$) [16]. Furthermore, following consumption of 200 ml of orange juice as part of a standardized test breakfast Lloyd et al. identified proline betaine and a number of biotransformed products in postprandial urine [17]. Urinary proline betaine measurements also distinguished between low, medium and high habitual intakes of citrus foods (estimated by food frequency questionnaire (FFQ)) with sensitivities and specificities of 80.8 – 92.2 % and 74.2 – 94.1 %, respectively, for elevated proline betaine in high reporters of citrus fruit consumption [17]. In another study, urinary metabolomes were profiled for volunteers that had consumed an acute dose of orange or grapefruit juice, volunteers that had consumed orange juice regularly for one month as part of their habitual diet and also volunteers whom had reported high or low consumption of citrus products in a large cohort study [15]. Proline betaine was identified as a biomarker of citrus fruit intake in all research designs [15]. Considering that independent metabolomics studies with different population groups, different analytical methods and exposures consistently reported proline betaine as a marker of citrus fruit intake, the evidence base is therefore substantial to support its use. However, a clear demonstration of the utility of this biomarker in predicting citrus intake is lacking.

The objective of the present work was to develop calibration curves for use in quantifying citrus intakes in an independent cohort. This study investigated the dose-response relationship of proline betaine with orange juice intake in a controlled intervention study and subsequently developed calibration curves for use in quantifying citrus intakes.

2 Materials and methods

2.1 Study design and population; the NutriTech study

Ethical approval was received from London Brent Ethics Committee (reference number: 12/LO/0139). The NutriTech study is a randomized control trial comprised of two parts; the food intake study which aimed to investigate the use of metabolomic profiling as a method of independent food quantification and secondly the weight loss intervention that aimed to quantify the effect of diet on 'phenotypic flexibility' (adaptation of biological and physiological processes in the state of challenged homeostasis) (NCT01684917) (Supporting Information Figure S1). For the present study data from the NutriTech food intake study is included. Recruitment took place between June 2012 and April 2014. Participants attended the NIHR/Wellcome Trust Imperial Clinical Research Facility for three days over three consecutive weeks. Eligibility criteria included males and females of all ethnicities, aged between 18 and 65 y with a BMI of 18.5-35 kg/m^2 and free from any chronic medical condition. Participants (n = 50) were randomly assigned (using nested randomization based on sex, age and BMI) to one of five different diets; red meat, fish, poultry, processed meat or a supplement and vegetarian option (Supporting Information Table S1). An independent researcher, not linked to the study, performed the randomization by sealed envelopes. On each day of each intervention week participants consumed a standardized breakfast at 8 am and their test meals at midday (12 pm) and evening (7 pm). All test meals were designed to provide similar intakes of dietary energy and fiber but macronutrient composition varied over the intervention weeks with carbohydrate decreasing from week one to week three and protein and fat intake increasing from week one to week three (Supporting Information Figure S2). Leftovers were measured and recorded where appropriate. Urine and plasma samples were collected (Supporting Information Figure S3). On day one of each intervention week no biofluids were collected. A 24 h urine sample was collected on day two of each intervention

week. Participants began the 24 h urine collection at 8 am (or at the time of the first, fasting, morning urine void of day two). All voids throughout the day were collected in a single 5 L container. 24 h urine samples were kept chilled at 4°C until all urine had been collected and the final volume was recorded (at 8 am day three or at the time of the first, fasting, urine void of day three). On day three of each intervention week participants were only allowed void at the designated times: 0 h (11.55 am), 2 h (2 pm or 2 h after their midday meal) and 6 h (6 pm or 6 h after they eat their midday meal). Blood samples were also collected at 0 h, 2 h and 6 h. On day four of each intervention week, before participants left the clinical investigation unit, fasting blood and urine samples were collected. Participants returned to their normal dietary habits until returning to the study the subsequent week. The present study focused on the breakfast meal. Each diet received the same breakfast which included white bread, eggs, butter, yoghurt and orange juice. The amount of orange juice provided decreased from week one to week three. In weeks one, two and three the breakfast was designed for females to receive 250 g/d, 220 g/d and 50 g/d of orange juice respectively and males to receive 520 g/d, 450 g/d and 30 g/d respectively. Participants did not consume any other citrus fruits or juices during the three day intervention. The 24 h urine samples and the fasting urine samples are used in this present analysis.

2.2 Validation study; the NANS study

Dietary and urinary metabolomic data from the National Adult Nutrition Survey (NANS) was used to demonstrate further the utility of dietary biomarkers in predicting intake. Ethical approval was obtained from the University College Cork Clinical Research Ethics Committee of the Cork Teaching Hospitals (ECM 3 (p) 4 September 2008) and recruitment began in May 2008. NANS investigated habitual food and beverage consumption, lifestyle, health indicators and attitudes to food and health in a representative sample of 1500 adults aged 18 - 90 y in the

126 Republic of Ireland during 2008 - 2010 [22]. For the present study dietary and urinary
127 metabolomic data from 565 NANS participants are included in the analysis. The 565
128 participants were randomly selected from the main NANS database ensuring equal numbers
129 of men and women across the age range. Dietary data was collected, over four consecutive
130 days, using a four day semi-weighed food diary. Participants recorded detailed information on
131 the amount and type of all foods, drinks and nutritional supplements consumed over four
132 consecutive days in the food diary. Each of the 2552 food codes consumed during the survey
133 were assigned to one of 68 food groups. For the purpose of this analysis citrus containing
134 food groups (fruit squashes, cordials and fruit juice drinks, fruit juices and citrus fruit) were
135 combined to form the total citrus food group. Mean daily citrus intake (average citrus intake
136 based on the four days of recording) was computed for the total citrus food group. Under-
137 reporters of energy intake were identified as having a ratio of energy intake:BMR of < 1·1
138 [23]. During the data collection period, a 50 ml first void urine sample was also collected
139 from participants. All urine samples were centrifuged at 1800 x g for 10 min at 4ºC and stored
140 at -80ºC for analysis.
141
142 **2.3 Urine analysis and metabolite quantification**
143 Urine samples were prepared for ^{1}H NMR spectroscopy by the addition of 250 µL phosphate
144 buffer (0.2 mol KH_2PO_4/L, 0.8 mol K_2HPO_4/L) to 500 µL urine. After centrifugation at 5360
145 x g for 5 min at 4ºC, 10 µL sodium trimethylsilyl [2,2,3,3-$^{2}H_4$] proprionate (TSP) and 50 µL
146 deuterium oxide (D_2O) were added to 540 µL of the supernatant. Spectra were acquired on a
147 600-MHz Varian NMR spectrometer by using the first increment of a nuclear Overhauser
148 enhancement spectroscopy pulse sequence at 25ºC. Spectra were acquired with 16,384 data
149 points and 128 scans. Water suppression was achieved during the relaxation delay (2.5 s) and
150 the mixing time (100 ms). All ^{1}H NMR urine spectra were referenced to TSP at 0.0 parts per

million and processed manually with the Chenomx NMR Suite (version 7.7, Inc.; Edmonton, Canada) by using a line broadening of 0.2 Hz, followed by phase and baseline correction. A ^{1}H NMR spectrum was acquired for a proline betaine standard. This spectrum was added to the Chenomx Spectral Reference Library using the company's recommended spectral acquisition and formatting protocols. Proline betaine was identified and quantified by using the Chenomx Profiler (version 7.7). Osmolality was measured by using an Advanced Osmometer model 3D3 (Advanced Instruments, Norwood, MA). Aliquots of urine were tested for osmolality with the use of micro-osmometry by freezing point depression, with values reported as the number of solute particles, in moles, dissolved in a kilogram of urine. Metabolite concentrations were normalized to osmolality where appropriate, by dividing the metabolite concentration by the osmolality reading for the sample.

2.4 Statistical analyses

Paired sample t-tests were performed using IBM SPSS Statistics 20.0 to compare proline betaine concentrations between intervention week one and intervention week three in the 24 h and fasting urine samples. Ten participants were randomly selected from the 50 NutriTech participants and served as a test set, the remaining 40 participants served as a training set. Concentration curves were determined based on data from the training set and orange juice intakes were predicted in the test set based on the proline betaine concentrations in urine using curve-fitting software (WinCurveFit). Concentration curves were also determined using data from the NutriTech total population (n = 50) and citrus intakes were predicted in the NANS cohort. Bland and Altman plots were made via GraphPad Prism 6.0 to assess agreement between the predicted (based on proline betaine concentrations) citrus intake and actual (NutriTech intake) or recorded (food diary) citrus intake in the test set and the NANS

cohort [24]. The association between the actual intake and the predicted orange juice intake was also examined using Spearman's correlations.

3 Results

3.1 The NutriTech study population

Characteristics of the NutriTech participants (n = 50) are presented in Table 1. The training set (n = 40) comprised of 21 males and 19 females, with a mean age of 60 ± 4 y and a mean BMI of 28.5 ± 3.6 kg/m^2. The test set (n = 10) comprised of four males and six females, a mean age of 59 ± 5 y and a mean BMI of 29.2 ± 3.4 kg/m^2.

3.2 Proline betaine quantification

Proline betaine was quantified using Chenomx Profiler and concentrations were compared between intervention week one and intervention week three in the training set. Two samples were missing for the 24 h analysis and one sample was missing for the fasting sample, therefore data presented is based on 38 participants and 39 participants respectively. In both the 24 h urine samples and the fasting urine samples proline betaine concentrations significantly decreased from intervention week one to intervention week three ($p < 0.001$) in response to decreasing orange juice consumption (Table 2). Proline betaine also decreased significantly ($p < 0.001$) in both the 24 h urine and the fasting urine samples when normalized for osmolality (Supporting Information Table S2).

3.3 Development of calibration curves for prediction of orange juice intakes

Calibration curves were determined using proline betaine concentrations and actual orange juice intakes from the training set. This was completed for the 24 h and fasting urine, both

normalized and not normalized to osmolality. The calibration curve based on proline betaine concentrations from the training set 24 h urine sample not normalized to osmolality is presented in Figure 1 (Y=1.63E-03*X+1.31E-01). From this orange juice intake was predicted for the test set (n = 10) (Supporting Information Table S3).

Bland and Altman plots, used to assess the agreement between actual orange juice intakes and predicted orange juice intakes in the test set are presented in Figure 2. The 24 h urine sample had less than 10 % of the observations fall outside the 95 % limits of agreement (the dotted lines) (Figure 2A). Similar results are found with the fasting samples, as less than 10 % of the observations also fall outside the 95 % limits of agreement (Figure 2C). The mean difference (bias) between predicted and actual orange juice intake was small (43.1 and -18.1 g for the 24 h and fasting samples respectively). Overall these plots indicate good agreement between the predicted and actual orange juice intakes. 24 h urine samples and fasting urine samples normalized to osmolality also had less than 10 % of the observations fall outside the 95 % limits of agreement. The mean difference (bias) between predicted and actual orange juice intake was smaller for 24 h and fasting urine samples normalized to osmolality (9.8 and -4.1 g respectively) (Figure 2B, Figure 2D).

The association between actual orange juice intakes and predicted orange juice intakes was assessed using Spearman's correlations coefficient. Actual orange juice intake showed a significant association with predicted orange juice (Supporting Information Table S4). The spearman correlation was 0.712 ($p < 0.001$) and 0.710 ($p < 0.001$) for 24 h and fasting urine respectively, while proline betaine concentrations normalized to osmolality in the 24 h urine and the fasting urine samples had correlations of 0.859 and 0.919 ($p < 0.001$) respectively (Supporting Information Table S4).

3.4 Prediction of citrus intakes in an independent cohort

The calibration curve determined using NutriTech participant's (n = 50) fasting urine proline betaine concentrations was used to predict citrus intake for the NANS participants (n = 565). Bland and Altman plots were used to assess the agreement between participant's self-reported mean daily citrus intake and predicted citrus intakes from the participant's proline betaine concentrations in the fasting urine sample (normalized and not normalized to osmolality) (Figure 3A, Figure 3B). Mean daily citrus intake both normalized and not normalized for osmolality had <5 % of the observations fall outside the 95 % limits of agreement. The mean difference (bias) between recorded citrus intake and predicted citrus intake using proline betaine concentrations not normalized to osmolality was 21.6 g (Figure 3A). The mean difference (bias) between recorded citrus intake and predicted citrus intake using proline betaine concentrations normalized to osmolality was smaller (4.3 g) (Figure 3B). Disagreement between measurements was greatest for high predicted intakes. Twenty-two participants were predicted to have higher citrus intake compared to the self-reported data. Upon further investigation, seven participants were identified as under-reporters and three participants were supplement users. When data was normalized to osmolality the number of participants having predicted citrus intakes higher than recorded intakes was reduced (15 participants).

4 Discussion

The present study has made significant advancements in the dietary biomarker field. Primarily, the development of calibration curves successfully enabled proline betaine to be used to estimate citrus intakes in a large cross-sectional study. Furthermore, this was supported by demonstrating a dose-response relationship between proline betaine and orange juice intake. This approach, using dietary biomarkers to quantify food or beverage intake can

be developed and utilized in future studies, therefore aiding the translation of these biomarkers into practice.

Our study is an important demonstration of the successful use of dietary biomarkers. The study design enabled the examination of the dose-response relationship between the biomarker and actual food intake. Importantly, the orange juice was consumed as part of a mixed meal, which is more reflective of habitual dietary intake and demonstrates the sensitivity of proline betaine as it can still classify participant's intakes irrespective of other components of the diet. Demonstration of the use of the developed calibration curves to predict intake in the cross-sectional study was a valuable aspect of this study. Proline betaine concentrations successfully predicted citrus intakes. The ability to predict mean daily citrus intake is important for future use of proline betaine as a marker of habitual intake of citrus fruit. Although the calibration curves were built using orange juice the use of the NANS cohort demonstrated their ability to predict citrus intake which included both juices and fruit.

Both 24 h urine and fasting urine samples were examined in this study. Interestingly, the results indicate that the fasting samples performed well and once corrected for osmolality outperformed the 24 h samples. This is particularly important for nutritional epidemiology where many cohort studies have fasting samples collected and not 24 h urine samples. For future studies it also demonstrates the potential use of fasting samples once corrected for osmolality. Collection of a fasting sample is less burdensome on the volunteer and should enable easier compliance within studies. In the current study fasting samples were used to predict citrus intake in the free-living NANS population. Bland and Altman plots demonstrated good agreement between predicted and recorded intakes. The disagreement observed between methods in the NANS cohort was predominantly as a result of higher

predicted intakes compared to the self-reported intakes. Interestingly, when participants with self-reporting issues (under-reporters, overweight or obese participants) were taken into account, this disagreement accounted for less than 2 % of the total population. Agreement was further improved between predicted and recorded intakes when samples were normalized to osmolality. In both datasets in the current study normalization with osmolality improved the agreement between predicted and actual/reported intakes. A previous study also reported the importance of normalizing urine to osmolality for the detection of changes in metabolite profiles [25].

While there has been significant interest in using metabolomics to identify dietary biomarkers there has been a lack of studies demonstrating the use of such biomarkers in predicting intake. In a recent study a dose-response relationship between tartaric acid and grape intake was demonstrated [26]. Tartaric acid was subsequently quantified in participants (n=19) following four four-day dietary interventions which included 0 g/d, 50 g/d, 100 g/d, and 150 g/d of grapes in standardized diets in a randomized controlled trial. Predicted grape intake was found to be most accurate for 24 h urine samples compared to spot urine samples ($r^2 = 0.90$) [26]. In relation to citrus fruit biomarkers, Lloyd and colleagues demonstrated the potential quantitative relationship between proline betaine and citrus fruit consumption as urinary proline betaine levels differed among low, medium and high citrus consumers after an overnight fast [17]. However, estimations of consumption were based on self-reporting data from an FFQ and the dose-response of proline betaine with citrus intake was not investigated. Proline betaine has also been identified as a biomarker of citrus intake using three study designs; a short term intervention where an acute dose of orange/grapefruit juice was consumed, a medium term intervention where orange juice was consumed regularly for one month, and a cohort study where high or low consumers of citrus products were identified

from a 24 h recall [15]. The focus of this study however was on the discovery of biomarkers and therefore did not examine the dose-response. Furthermore, previous studies have shown that proline betaine has a relatively short half-life; however, this did not seem to impact on its ability to predict habitual dietary intake. An earlier study targeted four metabolites in 24 h urine samples following the consumption of controlled diets containing low red meat (60 g/d), medium red meat (120 g/d) and high red meat (420 g/d) [6]. Two metabolites demonstrated a dose-response relationship with meat intake, increasing as the amount of meat in the diet increased, however no further practical use of these biomarkers in quantifying red meat intake were demonstrated [6]. The current study used a well-controlled intervention study to develop calibration curves which enabled prediction of intake in a free-living cross-sectional cohort marks a very significant step forward in the field of dietary biomarkers.

There are a number of strengths associated with the present study. Primarily this study reveals how a dietary biomarker, discovered through a metabolomics based approach can be used to successfully predict food intakes in a large cross-sectional study. Thus clearly demonstrating the potential application of dietary biomarkers in dietary assessment. Furthermore, the NutriTech food intake study provides a successful strategy for dietary biomarker identification, enabling the assessment of the dose-response relationship between the biomarker and food source. However, it must be noted that although there was excellent agreement between predicted citrus intake and self-reported citrus intakes, further interventions with repeated measurements over time may be needed to assess the dose-response relationship for long term intakes. It is also important to acknowledge that this work reflects a food intake biomarker and addresses the significant issue of improving estimations of food intake. However, we did not assess if this results in improvements in nutrient intake data.

The present study represents an important advancement in biomarker research by demonstrating the utility of calibration curves to successfully quantify intakes of citrus food. This study illustrates a clear dose-response relationship between actual food intake and a dietary biomarker in a mixed meal setting. The results presented here are very promising for the field of dietary biomarkers; however more studies on dose-response relationships are essential for further progression in this area. This work will pave the way for further development of dietary biomarkers that can be used to predict unbiased intakes and that can be used to obtain more reliable risk estimates in diet-disease analyses.

Author contributions

H. G. conducted research, analyzed data and prepared the manuscript. C. J. R. M. assisted in the statistical analyses, M. R., G. F., B. A. M., A. P. N., J. W., A. F. and M. J. G. provided essential materials, L. B. designed research, conducted research, analyzed data and prepared the manuscript. All authors read and approved the final manuscript.

Acknowledgements/funding

We thank all volunteers for their commitment and patience during the study.

HG is supported by a research grant from NutriTech (project number: 2896511). LB acknowledges the following funding: The European Research Council ERC (647783) and SFI (JPI-HDHL/B3075, Foodball). Data used in part of this analysis were supported by joint funding from the Irish Department of Agriculture, Fisheries and Food and the Health Research Board under the Food for Health Research Initiative (2007–2012) (grant number 7FHRIUCC2).

Conflict of interest statement

The authors have declared no conflict of interest.

References

[1] Freedman, L. S., Kipnis, V., Schatzkin, A., Tasevska, N., et al., Can we use biomarkers in combination with self-reports to strengthen the analysis of nutritional epidemiologic studies? Epidemiol Perspect Innov. 2010, 7, 2.

[2] Bingham, S. A., Biomarkers in nutritional epidemiology. Public Health Nutr. 2002, 5, 821-827.

[3] Lloyd, A. J., Fave, G., Beckmann, M., Lin, W., et al., Use of mass spectrometry fingerprinting to identify urinary metabolites after consumption of specific foods. Am J Clin Nutr. 2011, 94, 981-991.

[4] Lloyd, A. J., Beckmann, M., Haldar, S., Seal, C., et al., Data-driven strategy for the discovery of potential urinary biomarkers of habitual dietary exposure. Am J Clin Nutr. 2013, 97, 377-389.

[5] Andersen, M. S., Rinnan, Å., Manach, C., Poulsen, S. K., et al., Untargeted Metabolomics as a Screening Tool for Estimating Compliance to a Dietary Pattern. J Proteome Res. 2014, 13, 1405-1418.

[6] Cross, A. J., Major, J. M., Sinha, R., Urinary Biomarkers of Meat Consumption. Cancer Epidemiol Biomarkers Prev. 2011, 20, 1107-1111.

[7] Stella, C., Beckwith-Hall, B., Cloarec, O., Holmes, E., et al., Susceptibility of human metabolic phenotypes to dietary modulation. J Proteome Res. 2006, 5, 2780-2788.

[8] Ross, A. B., Svelander, C., Undeland, I., Pinto, R., et al., Herring and Beef Meals Lead to Differences in Plasma 2-Aminoadipic Acid, β-Alanine, 4-Hydroxyproline, Cetoleic Acid, and Docosahexaenoic Acid Concentrations in Overweight Men. J Nutr. 2015, 145, 2456-2463.

[9] Edmands, W. M., Beckonert, O. P., Stella, C., Campbell, A., et al., Identification of human urinary biomarkers of cruciferous vegetable consumption by metabonomic profiling. J Proteome Res. 2011, 10, 4513-4521.

[10] Andersen, M. S., Reinbach, H. C., Rinnan, Å., Barri, T., et al., Discovery of exposure markers in urine for Brassica-containing meals served with different protein sources by UPLC-qTOF-MS untargeted metabolomics. Metabolomics. 2013, 9, 984-997.

[11] Ross, A. B., Bourgeois, A., Macharia, H. N. u., Kochhar, S., et al., Plasma alkylresorcinols as a biomarker of whole-grain food consumption in a large population: results from the WHOLEheart Intervention Study. Am J Clin Nutr. 2012, 95, 204-211.

[12] Landberg, R., Åman, P., Friberg, L. E., Vessby, B., et al., Dose response of whole-grain biomarkers: alkylresorcinols in human plasma and their metabolites in urine in relation to intake. Am J Clin Nutr. 2009, 89, 290-296.

[13] Heinzmann, S. S., Holmes, E., Kochhar, S., Nicholson, J. K., et al., 2-Furoylglycine as a Candidate Biomarker of Coffee Consumption. J Agric Food Chem. 2015, 63, 8615-8621.

[14] Lang, R., Wahl, A., Stark, T., Hofmann, T., Urinary N-methylpyridinium and trigonelline as candidate dietary biomarkers of coffee consumption. Mol Nutr Food Res. 2011, 55, 1613-1623.

[15] Pujos-Guillot, E., Hubert, J., Martin, J. F., Lyan, B., et al., Mass spectrometry-based metabolomics for the discovery of biomarkers of fruit and vegetable intake: citrus fruit as a case study. J Proteome Res. 2013, 12, 1645-1659.

[16] Heinzmann, S. S., Brown, I. J., Chan, Q., Bictash, M., et al., Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption. Am J Clin Nutr. 2010, 92, 436-443.

[17] Lloyd, A. J., Beckmann, M., Favé, G., Mathers, J. C., et al., Proline betaine and its biotransformation products in fasting urine samples are potential biomarkers of habitual citrus fruit consumption. Br J Nutr. 2011, 106, 812-824.

[18] May, D. H., Navarro, S. L., Ruczinski, I., Hogan, J., et al., Metabolomic profiling of urine: response to a randomised, controlled feeding study of select fruits and vegetables, and application to an observational study. Br J Nutr. 2013, 110, 1760-1770.

[19] Favé, G., Beckmann, M., Lloyd, A. J., Zhou, S., et al., Development and validation of a standardized protocol to monitor human dietary exposure by metabolite fingerprinting of urine samples. Metabolomics. 2011, 7, 469-484.

[20] Andersen, M. B., Kristensen, M., Manach, C., Pujos-Guillot, E., et al., Discovery and validation of urinary exposure markers for different plant foods by untargeted metabolomics. Anal Bioanal Chem. 2014, 406, 1829-1844.
[21] Atkinson, W., Downer, P., Lever, M., Chambers, S. T., et al., Effects of orange juice and proline betaine on glycine betaine and homocysteine in healthy male subjects. Eur J Nutr. 2007, 46, 446-452.
[22] Irish Universities Nutrition Alliance, 2011.
[23] Goldberg, G. R., Black, A. E., Jebb, S. A., Cole, T. J., et al., Critical evaluation of energy intake data using fundamental principles of energy physiology: 1. Derivation of cut-off limits to identify under-recording. Eur J Clin Nutr. 1991, 45, 569-581.
[24] Bland, J. M., Altman, D. G., Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet. 1986, 327, 307-310.
[25] Warrack, B. M., Hnatyshyn, S., Ott, K.-H., Reily, M. D., et al., Normalization strategies for metabonomic analysis of urine samples. J Chromatogr B. 2009, 877, 547-552.
[26] Garcia-Perez, I., Posma, J. M., Chambers, E. S., Nicholson, J. K., et al., An Analytical Pipeline for Quantitative Characterization of Dietary Intake: Application To Assess Grape Intake. J Agric Food Chem. 2016, 64, 2423-2431.

Table 1 NutriTech food intake study characteristics [a)]

Characteristics	**Training set (n = 40)**	**Test set (n = 10)**
Gender	21 (M) 19 (F)	4 (M) 6 (F)
Age (y)	60 ± 4	59 ± 5
BMI (kg/m^2)	28.5 ± 3.6	29.2 ± 3.4
Systolic blood pressure (mm Hg)	133.5 ± 15.9	127.8 ± 14.4
Diastolic blood pressure (mm Hg)	78.8 ± 10.6	74.2 ± 11.4

[a)]Data are Mean ± SD (all such vales). 10 participants were randomly selected from the 50 NutriTech food intake participants and served as the test set, the remaining 40 participants served as a training set.

Table 2 Mean proline betaine concentrations in the 24 h urine (n = 38) and fasting urine samples (n = 39) [a)]

mmol/L	**Week 1**	**Week 3**	***P*** [b)]
24 h proline betaine	0.74 ± 0.32	0.20 ± 0.09	<0.001
Fasting proline betaine	0.71 ± 0.34	0.20 ± 0.10	<0.001

[a)]Data are Mean ± SD.

[b)]Paired sample t-test was used to examine the differences between intervention week 1 and week 3.

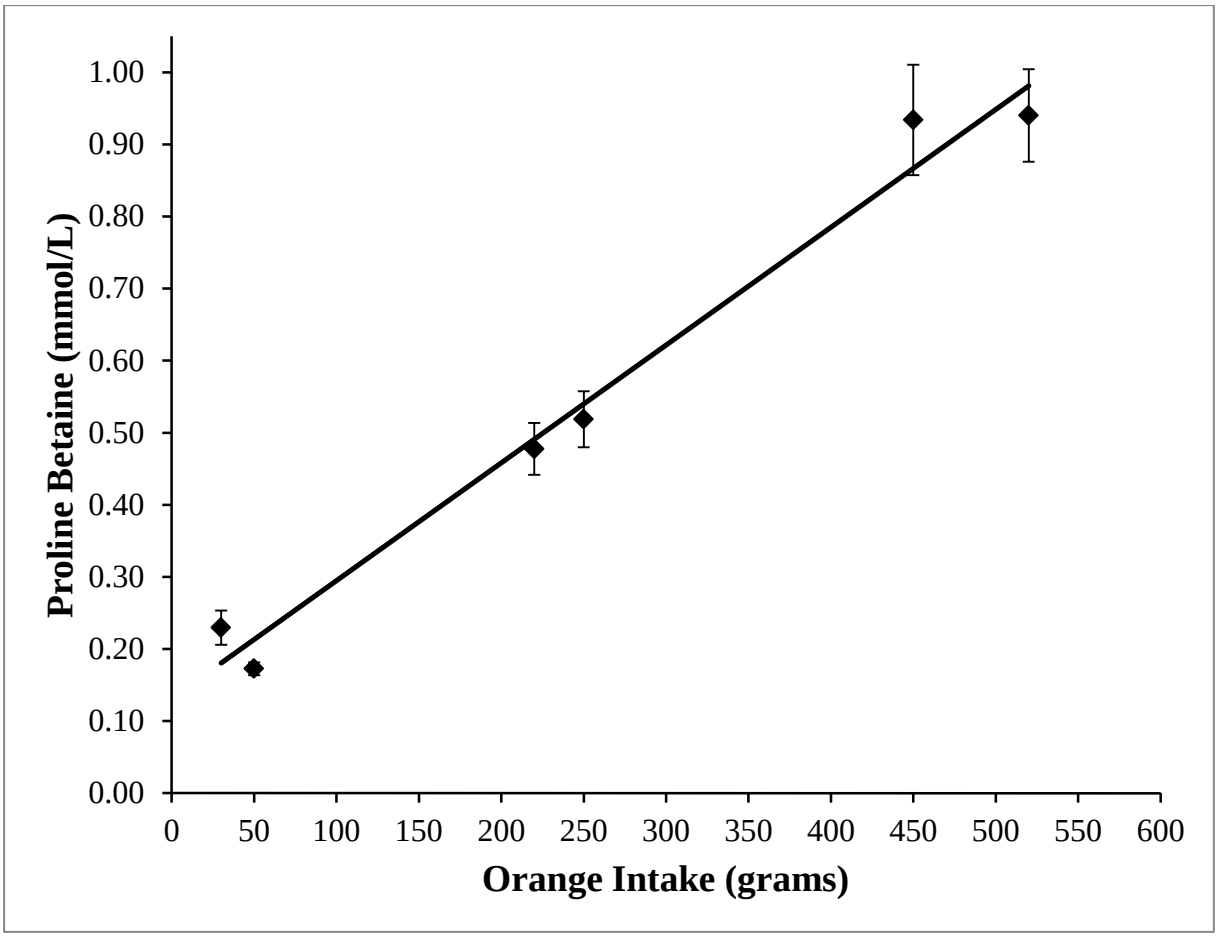

Figure 1. Calibration curve using the 24 h urine samples (mean and error bars (SEM) are presented). The x-axis; actual orange juice intake (grams) during the NutriTech food intake study, the y-axis; proline betaine concentrations measured in urine (mmol/L). Each point represents average proline betaine concentration for a particular orange juice intake. (Y=1.63E-03*X+1.31E-01)

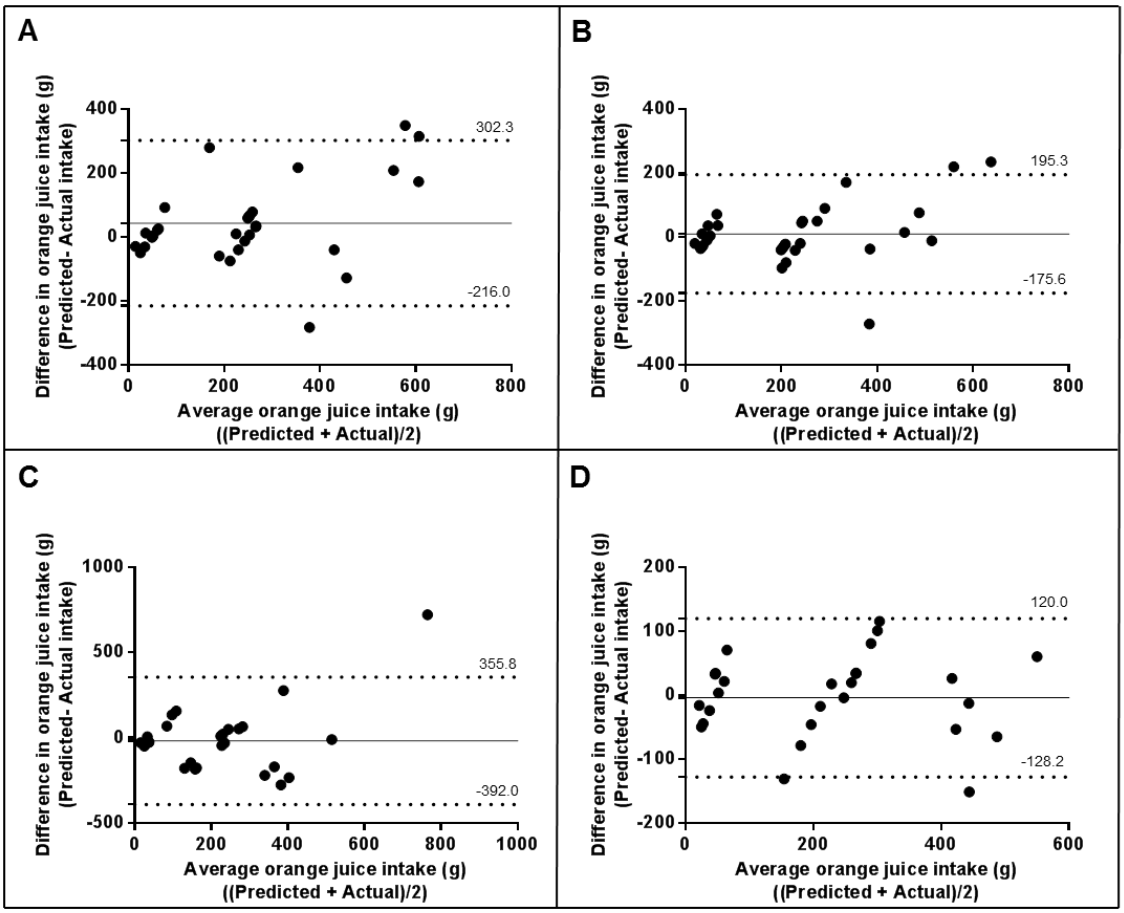

FIGURE 2

Figure 2. Bland and Altman plots for the test set **A:** Orange juice intakes were predicted from proline betaine concentrations measured in 24 h urine samples **B:** Orange juice intakes were predicted from proline betaine concentrations measured in 24 h urine samples normalized to osmolality **C:** Orange juice intakes were predicted from proline betaine concentrations measured in fasting urine samples **D:** Orange juice intakes were predicted from proline betaine concentrations measured in fasting urine samples normalized to osmolality, with mean difference and limits of agreement.

The solid line represents the mean difference and the dotted line represents the limits of agreement. 'Predicted' indicates the predicted orange juice intake based on urinary proline betaine concentrations. 'Actual' indicates the actual orange juice intakes according to the NutriTech study taking into account leftovers.

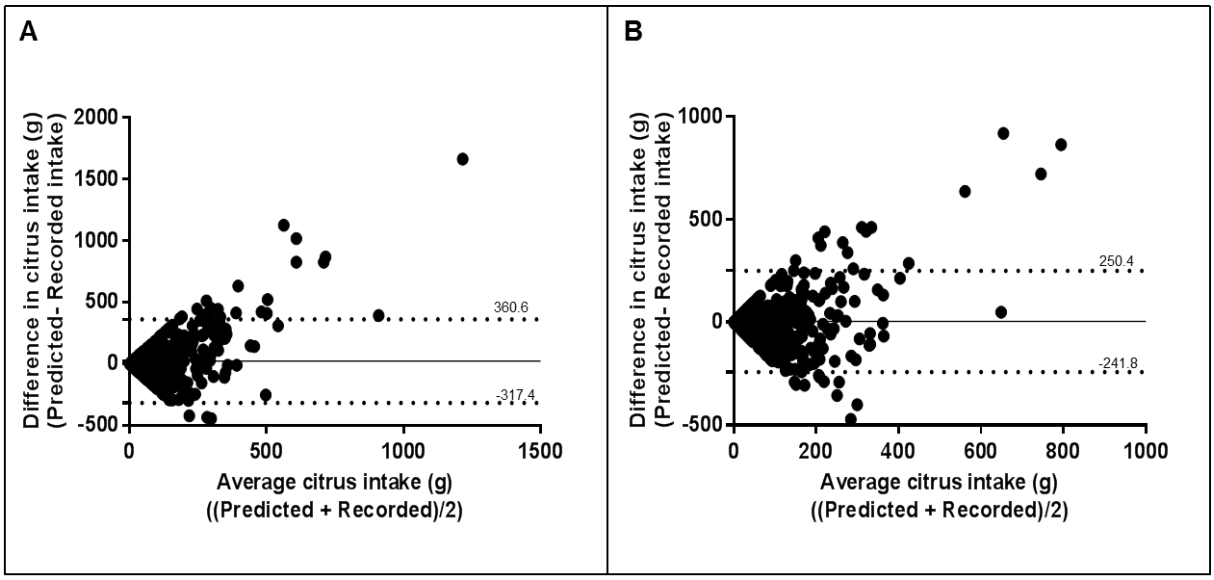

FIGURE 3

Figure 3. Bland and Altman plots for NANS **A:** Agreement between recorded mean daily citrus intake and citrus intakes predicted from proline betaine concentrations in a fasting sample **B:** Agreement between recorded mean daily citrus intake and citrus intakes predicted from proline betaine concentrations in a fasting sample normalized to osmolality, with mean difference and limits of agreement.

The solid line represents the mean difference and the dotted line represents the limits of agreement. 'Predicted' indicates the predicted citrus intake based on urinary proline betaine concentrations in a fasting sample. 'Recorded' indicates citrus intake recorded using the four-day food diary.