

Research Article

Urinary Metabolomic Markers of Protein Glycation, Oxidation, and Nitration in Early-Stage Decline in Metabolic, Vascular, and Renal Health

Jinit Masania,¹ Gernot Faustmann,^{2,3} Attia Anwar,¹ Hildegard Hafner-Giessauf,² Nasir Rajpoot,⁴ Johanna Grabher,² Kashif Rajpoot,⁵ Beate Tiran,⁶ Barbara Obermayer-Pietsch,⁷ Brigitte M. Winklhofer-Roob,³ Johannes M. Roob,^{2†} Naila Rabbani,¹ and Paul J. Thornalley^{1,8}

¹Warwick Medical School, Clinical Sciences Research Laboratories, University of Warwick, University Hospital, Coventry CV2 2DX, UK

²Clinical Division of Nephrology, Department of Internal Medicine, Medical University of Graz, 8036 Graz, Austria

³Human Nutrition & Metabolism Research and Training Center (HNMRC), Institute of Molecular Biosciences, Karl Franzens University of Graz, Universitätsplatz 2, 8010 Graz, Austria

⁴Department of Computer Sciences, University of Warwick, Coventry CV4 7AL, UK

⁵School of Computer Science, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

⁶Clinical Institute of Medical and Clinical Laboratory Diagnostics, Medical University of Graz, 8036 Graz, Austria

⁷Clinical Division of Endocrinology, Department of Internal Medicine, Medical University of Graz, 8036 Graz, Austria

⁸Diabetes Research Center, Qatar Biomedical Research Institute (QBRI), Hamad Bin Khalifa University, Qatar Foundation, P.O. Box 34110, Doha, Qatar

[†]Deceased

Correspondence should be addressed to Paul J. Thornalley; pthornalley@hbku.edu.qa

Received 23 May 2019; Revised 7 September 2019; Accepted 11 September 2019; Published 19 November 2019

Academic Editor: Marco Malaguti

Copyright © 2019 Jinit Masania et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The publication of this article was funded by Qatar National Library.

Glycation, oxidation, nitration, and crosslinking of proteins are implicated in the pathogenic mechanisms of type 2 diabetes, cardiovascular disease, and chronic kidney disease. Related modified amino acids formed by proteolysis are excreted in urine. We quantified urinary levels of these metabolites and branched-chain amino acids (BCAAs) in healthy subjects and assessed changes in early-stage decline in metabolic, vascular, and renal health and explored their diagnostic utility for a noninvasive health screen. We recruited 200 human subjects with early-stage health decline and healthy controls. Urinary amino acid metabolites were determined by stable isotopic dilution analysis liquid chromatography-tandem mass spectrometry. Machine learning was applied to optimise and validate algorithms to discriminate between study groups for potential diagnostic utility. Urinary analyte changes were as follows: impaired metabolic health—increased N_ε-carboxymethyl-lysine, glucosepane, glutamic semialdehyde, and pyrroline; impaired vascular health—increased glucosepane; and impaired renal health—increased BCAAs and decreased N_ε-(γ-glutamyl)lysine. Algorithms combining subject age, BMI, and BCAAs discriminated between healthy controls and impaired metabolic, vascular, and renal health study groups with accuracy of 84%, 72%, and 90%, respectively. In 2-step analysis, algorithms combining subject age, BMI, and urinary N_ε-fructosyl-lysine and valine discriminated between healthy controls and impaired health (any type), accuracy of 78%, and then between types of health impairment with accuracy of 69%–78% (*cf.* random selection 33%). From likelihood ratios, this provided small, moderate, and conclusive evidence of early-stage cardiovascular, metabolic, and renal disease with diagnostic odds ratios of 6 – 7, 26 – 28, and 34 – 79, respectively. We conclude that measurement of urinary glycated, oxidized, crosslinked, and branched-chain amino acids provides the basis for a noninvasive health screen for early-stage health decline in metabolic, vascular, and renal health.

1. Introduction

Diabetes mellitus, cardiovascular disease (CVD), and chronic kidney disease (CKD) are major noncommunicable chronic diseases in adults linked to premature death and loss of productive life in Westernised countries. Type 2 diabetes mellitus (T2DM) linked to development of insulin resistance and dysglycemia in prediabetes accounts for *ca.* 90% of cases of diabetes [1]. Spontaneous and potentially damaging modifications of proteins by reactive oxygen species (ROS) have been implicated in the pathogenesis of this disease development [2–4]. Protein glycation by glucose to form fructosamine adducts, particularly as assessed by glycated hemoglobin A1C, is a major clinical measure of glycemic control in diabetes and at lower levels is considered a diagnostic marker of prediabetes [5]. More recently, advanced glycation end products (AGEs), formed by the degradation of proteins glycated by glucose and by the direct reaction of proteins with reactive dicarbonyl compounds such as methylglyoxal, have been proposed as both mediators of health decline leading to T2DM, CVD, and CKD [6–8]. Examples of the latter are association of fructosamine-derived AGE, glucosepane, with development of T2DM [6] and association of increased formation of the methylglyoxal-derived AGE, hydroimidazolone MG-H1, with insulin resistance, risk of CVD, and development of CKD [9–11]. Oxidized and glycated proteins are targeted for cellular proteolysis which forms related oxidized and glycated amino acid metabolites, also called protein oxidation and glycation free adducts. These are released from cells and excreted in urine [12]. Oxidized and glycated amino acids are also absorbed from the intestinal tract after digestion of oxidized and glycated proteins of ingested food [13]. Pyrraline, an AGE formed only at the high temperatures of culinary processing of food, is absorbed and is a marker of the dietary AGEs [9]. Measurement of urinary fluxes of oxidized and glycated amino acids gives an estimate of total body exposure to these adducts—except for N_ε-fructosyl-lysine (FL) which may be further metabolised enzymatically [14].

Irreversible protein crosslinking increases with age and may contribute mechanistically to the related age-associated increased risk of T2DM, CVD, and CKD [1, 15]. Dityrosine is a major oxidative crosslink, particularly through enzymatic formation by dual oxidase (DUOX) [16], glucosepane is a major glycation-derived protein crosslink [17], and N_ε-(γ -glutamyl)lysine (GEEK) is a crosslink formed in proteins catalysed by transglutaminases [18]. Dityrosine and GEEK may also be absorbed from digested proteins in food [19, 20]. Trace level urinary fluxes of oxidized, glycated, and crosslinked amino acids reflect the flux of formation of glycated, oxidized, and crosslinked proteins—with also contributions from food. We hypothesized that early-stage changes in urinary fluxes of oxidized, glycated, crosslinked, and branched-chain amino acids may provide biomarkers supporting the early-stage diagnosis of impaired metabolic, vascular, and renal disease. Branched-chain amino acids (BCAAs) have been linked previously to the development of T2DM and CKD [21, 22].

In this study, we determined urinary fluxes of oxidized, glycated, crosslinked, and branched-chain amino acids in

healthy human subjects and subjects with early-stage decline in metabolic, vascular, and renal health, exploring potential diagnostic utility by data-driven machine-learning approaches. We based our sample collection and analysis on urine as a sample type favoured for clinical metabolomic applications because of the advantages of it being a readily available sample, ease of sample donation and collection, and less complex sample matrix than other body fluids such as serum or plasma. Oxidized, glycated, and crosslinked amino acids determined were as follows: FL—the major early-stage protein glycation adduct formed by glucose; MG-H1 and N_ε-carboxymethyl-lysine (CML)—the major AGEs; pyrraline—a marker of dietary AGE exposure and absorption [23, 24]; protein crosslinks—glucosepane, pentosidine, dityrosine, and GEEK; glutamic acid semialdehyde (GSA)—a “protein carbonyl” marker of oxidative damage; and 3-nitrotyrosine—a marker of protein nitration [25]. Protein modification analysis classes and associated processes reported thereby are summarised in Table 1.

2. Materials and Methods

2.1. Subject Study Groups and Sampling. A total of 200 study participants of the BIOCLAIMS cohort, recruited and investigated at the Medical University and Karl Franzens University of Graz, Austria, between May 2011 and November 2014, were investigated in this study. This cohort was conceived, subjects were recruited, and samples were collected as part of the EU FP7 BIOCLAIMS research project. The underlying hypothesis of this project was that maintenance of good health may be improved with biomarkers of good health or “health biomarkers” [26, 27]. We hypothesized that the range of flux of protein damage by glycation, oxidation, nitration, and crosslinking found in subjects of good health, singly or combinations thereof, may serve as health biomarkers and this could be tested by studying the changes in fluxes of related analytes in early decline in health—representing a challenge to health homeostasis. We selected early decline in metabolic, vascular, and renal health as examples of major health impact through further progression to type 2 diabetes, CVD, and renal failure. Measurement of urinary glycated, oxidized, nitrated, and crosslinked amino acids provide a surrogate measure of this. Changes in these analytes may thereby be used to diagnose early-stage decline of health. The subjects were assigned to one of the four groups.

2.1.1. Group 1: Healthy Controls (n = 55). The inclusion criteria for these subjects were as follows: BMI 18.5–29.9 kg/m², carotid artery intimal medial thickness (CIMT) \leq 75th percentile at the left plus right side, homeostatic model assessment of insulin resistance (HOMA-IR) index \leq 2.5 mU⁻¹ mM, A1C < 38 mmol/mol, (or either HOMA-IR or A1C was allowed above this threshold but not both) estimated glomerular filtrate rate (eGFR) > 60 ml/min/1.73 m² (deduced from increased serum creatinine by the Modification of Diet in Renal Disease Study equation [28]), and clinical chemistry tests within the normal range \pm 10%. The clinical chemistry tests included the

TABLE 1: Glycated, oxidized, and nitrated amino acid metabolites.

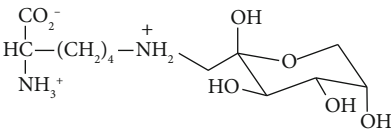
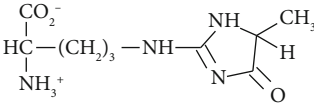
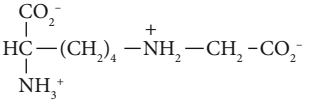
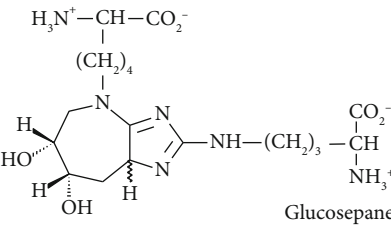
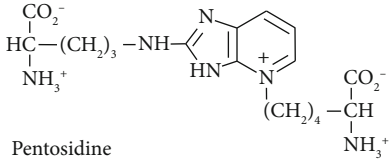
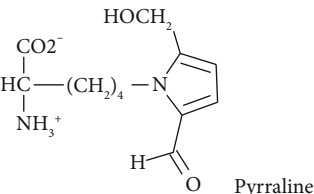
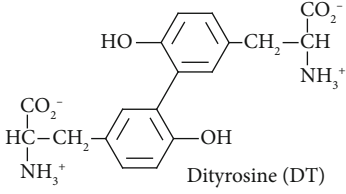
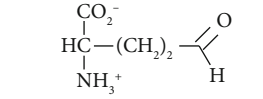
Metabolite class	Urinary metabolite	Comment
	 <p>N_ϵ-(1-Deoxy-D-fructos-1-yl)lysine (FL)</p>	Early-stage glycation adduct [25]. Formed from glucose nonenzymatically and reporting on exposure to increased glucose concentration. Repaired intracellularly by fructosamine 3-phosphokinase [14]. Free adduct absorbed after digestion of food proteins [69].
	 <p>MG-derived hydroimidazolone (MG-H1)</p>	A major quantitative arginine-derived AGE formed from methylglyoxal. Linked to increased fasting and postprandial glucose exposure, insulin resistance, and cardiovascular disease [9, 10, 47, 70]. Free adduct absorbed after digestion of food proteins [9].
	 <p>N_ϵ-Carboxymethyl-lysine (CML)</p>	A major quantitative lysine-derived AGE—particularly in food. Formed by the oxidative degradation of FL from and other sources. Free adduct absorbed after digestion of food proteins [71].
Glycation	 <p>Glucosepane</p>	Major quantitative crosslink formed in protein glycation [17].
	 <p>Pentosidine</p>	Low level pentose sugar-derived glycation crosslink and intense fluorophore. Considered to reflect pentose phosphate pathway activity [72].
	 <p>Pyrraline</p>	Glucose-derived AGE formed at high temperatures of culinary processing; originating only from food [23, 24].
Oxidation	 <p>Dityrosine (DT)</p>	Oxidative crosslink formed spontaneously in oxidative stress and enzymatically by dual oxidase (DUOX) [16, 25].
	 <p>Glutamic semialdehyde (GSA)</p>	Major “protein carbonyl” formed by the oxidative deguanidylation of arginine and oxidative ring-opening of proline [73].

TABLE 2: Mass spectrometric multiple reaction monitoring detection of protein glycation, oxidation, nitration, crosslinks, and branched-chain amino acids.

Analyte group	Analyte	R _t (min)	Parent ion (Da)	Ion (Da)	CE (eV)	Neutral fragment loss (es)	Internal standard and amount added
Glycation	FL	28.5	291.0	84.3	31	H ₂ CO ₂ , fructosylamine	[² H ₄]FL, 0.3 pmol
	CML	28.5	204.9	130.1	12	NH ₂ CH ₂ CO ₂ H	[¹³ C ₆]CML, 0.25 pmol
	MG-H1 [†]	11.6 & 12.5	229.2	114.3	14	NH ₂ CH(CO ₂ H)CH ₂ CH=CH ₂	[¹⁵ N ₂]MG-H1, 1.25 pmol
	Glucosepane	16.5	429.2	382.1	38	C ₂ H ₅ O	[¹³ C ₆]Glucosepane, 0.25 pmol
	Pyrraline	17.9	255.2	84.3	23	2-CHO-5-HOCH ₂ -pyrrole, H ₂ CO ₂	[¹³ C ₆ , ¹⁵ N ₂]Pyrraline, 1.00 pmol
Oxidative damage	Dityrosine	19.9	361.2	315.3	15	H ₂ CO ₂	[² H ₆]DT, 0.25 pmol
	GSA	32.2	114.0	68.0	15	H ₂ CO ₂	[² H ₃]AAA, 2.5 pmol
Nitration damage	3-NT	23.2	227.1	181.2	13	H ₂ CO ₂	[² H ₃]3-NT, 0.25 pmol
TG crosslink	GEEK	9.3	276.1	147.1	12	NH ₂ CH(CO ₂ H)CH ₂ CH=C=O	Nε-(γ-[¹³ C ₅]glutamyl) lysine, 2.5 pmol
BCAA	Leu	27.6	132.3	86.2	10	H ₂ CO ₂	[² H ₃]Leu, 250 pmol
	Ile	31.5	132.3	86.2	10	H ₂ CO ₂	[¹³ C ₆]Ile, 250 pmol
	Val	8.2	117.8	72.0	19	H ₂ CO ₂	[² H ₈]Val, 250 pmol

[†]For MG-H1, R_t values for the 2 epimers are given. LC-MS/MS was performed as described previously [25, 35] with chromatography using two Hypercarb™ (5 μm particle size, 0.2 × 50 mm and 0.2 × 250 mm) columns, column switching, and elution with 0.1% trifluoroacetic acid (TFA) in water and custom acetonitrile (MeCN) gradient. Different chromatography conditions were used for assay of GEEK: the column was Hypercarb™ (2 μm particle size, 0.2 × 150 mm) with isocratic elution at 0.2 ml/min with 3.75% MeCN and 0.1% TFA in water (solvent A) for 15 min. After each run the column was washed by elution with 50% tetrahydrofuran in 0.1% TFA in water for 20 min and reequilibrated by elution with solvent A at 0.4 ml/min for 15 min. Pentosidine was detected by in-line fluorimetry; excitation wavelength 320 nm, emission wavelength 365 nm [35].

analytes were quantified in the ultrafiltrate of second void urine after overnight fast by LC-MS/MS multiple reaction monitoring (MRM) as described previously [12, 35] with detection of additional analytes GEEK, leu, ile, and val. Urinary analytes were normalised to urinary creatinine because spot urine samples were used. Urine samples were filtered through 3 kDa pore size microspin filters (14,000 g, 4°C). Ultrafiltrate (5 μl) was mixed with a cocktail of stable isotopic standards (25 μl) and analysed by LC-MS/MS using an Acquity™ ultrahigh performance liquid chromatography-Xevo-TQS LC-MS/MS system (Waters, Manchester, U.K.). MRM detection conditions are summarised with chromatographic conditions, calibration, limit of detection, analytical recovery, and inter- and intrabatch coefficient of variance as given previously [35] (Table 2). GEEK was detected in a separate chromatography run to resolve it from isobaric glu-lys and lys-glu dipeptides [36]. Analyte amounts in test samples were deduced by interpolation of analyte/internal standard peak area ratios deduced from MRM mass chromatograms on calibration curves constructed by analysis of calibration standards.

Natural isotopic abundance analytical standards and stable isotope-substituted internal standards were purchased (Sigma-Aldrich, Poole, Dorset, U.K., and Cambridge Isotope Laboratories, Tewksbury, MA, USA) or synthesised in-house where unavailable commercially, as previously described [12, 37]. GSA was prepared from N-acetyl-L-ornithine using lysyl oxidase activity from eggshell membrane by modification of the method of Akagawa et al. [38]. Briefly, eggshell membranes (ESM) were isolated from 12 fresh hen eggs, washed thoroughly with distilled water, and cut into small

pieces (5 × 5 mm). Surface water was blotted from ESM with a filter paper. N_α-Acetyl-L-ornithine (43.6 mg, 0.25 mmol) was incubated with ESM (2.5 g) in 25 ml sodium phosphate buffer (20 mM, pH 9.0, 37°C) for 7 days with shaking. The eggshell membranes were removed by centrifugation (6000g, 10 min, 20°C), and the reaction mixture adjusted to pH 7.0. Acylase-1 (20 mg; grade I from porcine kidney, ≥2,000 units/mg protein, Sigma-Aldrich, Cat no. A3010) was added, and the reaction mixture was incubated at 25°C for 2.5 days. The reaction mixture was filtered (3 kDa) to remove acylase and the resulting GSA solution used without further purification. GSA was calibrated by derivatisation with 2-aminobenzaldehyde (OBA) to a dihydroquinazolinium adduct [39] for which the extinction coefficient of 2,800 M⁻¹ cm⁻¹ was assumed (as for related α-amino adipic semialdehyde adduct) [40]. GSA solution was incubated with 15 mM OBA in water at 20°C for 20 min and absorbance measured at 465 nm. The yield of GSA was 83%. The internal standard used for GSA was [²H₃]α-amino adipic acid ([²H₃]AAA; C/D/N Isotopes Inc., Pointe-Claire, Quebec, Canada) [41]. Stable isotopic N_ε-(γ-[¹³C₅]glutamyl) lysine ([¹³C₅]GEEK) was prepared by modification of the facile synthesis described previously [42]. L-[¹³C₅]Glutamic acid (5.0 mg, 34 μmol) and L-lysine (5.0 mg, 34 μmol) were suspended in pentan-1-ol (25 μl), methanol (2 μl), and water (2 μl) in a 0.3 ml reaction vial. The mixture was heated at 130°C for 5 h. The solvent was removed under reduced pressure, and the residues were lyophilised to dryness to give [¹³C₅]GEEK, 6.8 mg (yield: 71%) which was used without further purification.

2.3. Machine Learning. Support vector machine (SVM) learning methods were used to produce classification algorithms [43]. Only urinary analytes, age, BMI, and gender were included as features to limit analysis sample requirement to urine only. These were developed to discriminate between study groups in two different setups. In the first setup, SVMs were used as a supervised machine learning binary classifier that develops a model from the training data to discriminate between the healthy and a diseased class using a linear hyperplane in a high-dimensional space. Subsequently, the trained model is applied to unseen data to predict the study group class of the related subject. In the second setup, we developed a two-step algorithm. In step 1, the algorithm was trained to discriminate between healthy and diseased state where all disease impairments were considered as a disease class. In the second step, the algorithm was trained to identify one of the three disease impairments. For the three-class system explored (discriminating between impaired metabolic, vascular, and renal health study groups), the trained algorithm classifies the subject into one of three groups. For all classification experiments, we used a 2-fold cross-validation system where data is split into 50% each for training and testing and then this data split is alternated. This was repeated 100 times to test the stability and generalizability of the classification system. To assess the diagnostic characteristics, 8 performance metrics were computed: classification accuracy, sensitivity, specificity, positive likelihood rate, negative likelihood rate, positive predictive value, negative predictive value, and F-measure/score. The false positive rate is also described ($=1 - \text{specificity}$). The 95% CI are determined by Student's *t* distribution.

2.4. Statistical Analyses. Data are presented as mean \pm SD for parametric distributions and median (lower-upper quartile) for nonparametric distributions. For two groups, significance of the difference between means of parametric data was analysed by Student's *t*-test and medians of nonparametric data by the Mann-Whitney *U* test for independent samples; for more than two groups, significance of the difference between means of parametric data was analysed by ANOVA and medians of nonparametric data by the Kruskal-Wallis test for independent samples. Correlation analysis was performed by the Spearman rank correlation method. Data were analysed using SPSS, version 24.0. $P < 0.05$ was considered significant.

3. Results

3.1. Clinical Characteristics of Subjects Recruited. Characteristics of the subjects recruited with and without early-stage impairment of metabolic, vascular, and renal health are given (Table 3). Healthy control subjects without impaired metabolic, vascular, and renal health were younger and had lower BMI, A1C, and CIMT than study groups with impaired health. All health impairment study groups had increased plasma total cholesterol with respect to healthy controls, with increased LDL cholesterol in impaired metabolic and vascular health and decreased HDL cholesterol and increased triglycerides and systolic and diastolic blood pressure in impaired vascular and renal health. Urinary albumin and

total protein were increased only in impaired renal health, with respect to healthy controls. Changes in plasma LDL cholesterol and HDL cholesterol were not always accompanied by similar changes in plasma ApoB and ApoA1, respectively.

3.2. Urinary Amino Acid Metabolites. Urinary levels of glycosylated, oxidized, nitrated, crosslinked, and branched amino acid analytes for the 4 study groups are given in Table 4. Urinary fluxes of modified amino acids in healthy controls were in the order: FL>CML>MG-H1 \approx pyrraline>GSA>glucosepane>GEEK>DT>3-NT. For protein crosslinks, levels of urinary fluxes were in the order: glucosepane>GEEK>pentosidine>DT. For BCAAs, urinary excretion was in the order: val>leu>ile. With respect to healthy controls, changes found were as follows: impaired metabolic health—increased urinary excretion of CML, glucosepane, pyrraline, and GSA; impaired vascular health—increased glucosepane; and impaired renal health—increased BCAAs and decreased GEEK.

In correlation analysis, for healthy control subjects, there was no association of any urinary biomarkers with subject age. FL correlated positively with MG-H1 ($r = 0.84$) and 3-NT ($r = 0.62$), and MG-H1 correlated positively with 3-NT ($r = 0.72$). BCAAs correlated positively with each other ($r = 0.74 - 0.82$). There were also positive correlations of pyrraline with FL, CML, and MG-H1 ($r = 0.43$), suggesting significant contributions to urinary fluxes of these analytes from food—Table S1. In subjects with impaired metabolic health, surprisingly, there were no correlations of A1C or HOMA-IR with urinary glycation adducts. There were negative correlations of all BCAAs with A1C ($r = -0.38$ to -0.39) and positive correlations of BCAAs with eGFR ($r = 0.38 - 0.50$), GSA ($r = 0.57 - 0.61$), and GEEK ($r = 0.45 - 0.54$). There were positive correlations of pyrraline with CML, MG-H1, GSA, and GEEK ($r = 0.44 - 0.66$), suggesting these analytes had significant contributions from food—Table S2. In subjects with impaired vascular health, there were positive correlations of CIMT, A1C, and glucosepane with age ($r = 0.86$, $r = 0.51$, and $r = 0.37$, respectively) and negative correlation of eGFR with age ($r = -0.51$). Subjects with impaired vascular health, therefore, have age-related increase in CIMT thickness along with age-related early-stage decline in glucose tolerance and renal function and increased glucose-mediated protein crosslinking. There were positive correlations of pyrraline with CML ($r = 0.57$), MG-H1 ($r = 0.42$), GSA ($r = 0.49$), and DT ($r = 0.39$), suggesting these analytes had significant contributions from food—Table S3. In subjects with impaired renal health, there were positive correlations of CIMT ($r = 0.79$) and glucosepane with age ($r = 0.46$), suggesting that subjects with impaired renal function have increased CIMT and increased glucose-mediated protein crosslinking with age. There were positive correlations of pyrraline with CML ($r = 0.50$) and MG-H1 ($r = 0.55$), suggesting again these AGE free adducts had significant contributions from food—Table S4.

3.3. Machine Learning Analysis. We performed machine learning-based analysis on subject groups with and without early-stage decline in metabolic, vascular, and renal health.

TABLE 3: Clinical characteristics of subjects.

Variable	Healthy controls	Impaired metabolic health	Impaired vascular health	Impaired renal health	Significance (ANOVA/Kruskal-Wallis)
N	55	44	58	43	
Age (yrs)	33.3 (26.8 – 41.8)	57.5 (43.9 – 62.1)***	49.3 (34.1 – 59.0)***	57.2 (48.3 – 69.7)***,††	<0.001
Gender (M/F)	27/28	22/22	23/35	30/13	
BMI (kg/m ²)	23.1 ± 2.5	28.9 ± 4.5***	24.9 ± 3.5***,OOO	26.3 ± 3.6***,OO,†	<0.001
Smoker (current/ex/never)	13/11/31	8/13/23	9/7/42	2/17/24	
Alcohol intake (mg/day)	4,842 (1,778 – 9,481)	5,048 (1,056 – 10,428)	4,870 (1,851 – 11,658)	5,303 (1,035 – 14,532)	
FPG (mM)	4.79 ± 0.41	5.32 ± 0.58***	4.92 ± 0.54 ^{OOO}	5.12 ± 0.53***,†	<0.001
A1C (mmol/mol Hb)	34.6 ± 3.1	40.6 ± 2.3***	37.6 ± 3.0***,OOO	37.7 ± 4.4***,OOO	<0.001
Plasma insulin (IU/ml)	7.00 (4.90 – 9.10)	15.83 (12.49 – 19.00)***	7.55 (5.60 – 9.48) ^{OOO}	8.70 (6.20 – 13.65)***,OOO	<0.001
HOMA-IR (IU/ml/mM)	1.41 (1.11 – 1.96)	3.69 (2.93 – 4.34)***	1.66 (1.19 – 2.10) ^{OOO}	1.98 (1.47 – 3.23)*,OOO,†	<0.001
Plasma cystatin C (mg/l)	0.755 ± 0.088	0.810 ± 0.104**	0.773 ± 0.119	1.346 ± 0.224***,OOO	<0.001
eGFR (ml/min/1.73 m ²)	88.0 ± 12.5	79.6 ± 11.7***	84.5 ± 16.2	45.0 ± 8.8***,OOO	<0.001
CIMT (mm)	0.475 ± 0.065	0.572 ± 0.086***	0.643 ± 0.107***,OOO	0.605 ± 0.070***	<0.001
Total cholesterol (mmol/l)	4.69 ± 0.66	5.22 ± 1.00**	5.38 ± 0.84***	4.90 ± 0.93 ^{††}	<0.001
LDL cholesterol (mmol/l)	2.61 ± 0.60	3.11 ± 0.88**	3.17 ± 0.83***	2.86 ± 0.95	<0.01
Apolipoprotein-B (g/l)	0.794 ± 0.168	0.988 ± 0.220***	0.903 ± 0.203**, ^O	0.898 ± 0.229*	<0.001
HDL cholesterol (mmol/l)	1.75 ± 0.42	1.42 ± 0.46***	1.84 ± 0.48 ^{OOO}	1.42 ± 0.49***,†††	<0.001
Apolipoprotein-A1 (g/l)	1.60 ± 0.24	1.54 ± 0.27	1.69 ± 0.30 ^O	1.55 ± 0.28 [†]	<0.05
Triglycerides (mmol/l)	0.72 ± 0.27	1.51 ± 0.92***	0.81 ± 0.38 ^{OOO}	1.37 ± 1.09***,††	<0.001
Systolic blood pressure (mmHg)	115 ± 12	126 ± 20**	115 ± 12 ^{OO}	123 ± 14** ^{††}	<0.001
Diastolic blood pressure (mmHg)	75 ± 7	82 ± 10***	77 ± 9 ^{OO}	80 ± 9**	<0.001
Urinary albumin (mg/24h)	7.0 (5.1 – 9.2)	7.5 (5.4 – 12.8)	6.7 (4.6 – 9.3)	26.5 (6.6 – 129.4)***,OO,†††	<0.001
Urinary total protein (mg/24h)	99 (76 – 120)	89 (64 – 111)	89 (72 – 115)	121 (83 – 246)***,OO,††	<0.01

Data are mean ± SD or median (lower-upper quartile). Significance: *, **, and ***; P < 0.05, P < 0.01, and P < 0.001 with respect to healthy controls; o, oo, and ooo; P < 0.05, P < 0.01, and P < 0.001 with respect to impaired metabolic health; and †, ††, and †††; P < 0.05, P < 0.01, and P < 0.001 with respect to impaired vascular health.

TABLE 4: Urinary amino acid biomarkers.

Variable	Healthy controls	Impaired metabolic health	Impaired vascular health	Impaired renal health	Significance (ANOVA/Kruskal-Wallis)
FL (nmol/mg creatinine)	26.5 (17.3 – 39.4)	21.5 (17.1 – 33.9)	23.2 (16.7 – 38.5)	23.9 (19.2 – 29.2)	
CML (nmol/mg creatinine)	15.6 (13.7 – 20.6)	19.6 (15.4 – 25.8)*	16.2 (13.1 – 22.5) ^o	18.3 (14.1 – 21.1)	
MG-HI (nmol/mg creatinine)	9.75 (5.93 – 15.47)	7.79 (6.53 – 13.68)	8.72 (5.23 – 14.79)	9.68 (7.29 – 11.87)	
Glucosepane (nmol/mg creatinine)	2.84 (2.41 – 3.36)	3.17 (2.64 – 3.95)*	3.19 (2.59 – 4.58)**	2.84 (2.44 – 4.21)	<0.05
Pentosidine (nmol/mg creatinine)	0.258 (0.207 – 0.287)	0.266 (0.228 – 0.306)	0.270 (0.212 – 0.315)	0.237 (0.172 – 0.286)	
Pyrraline (nmol/mg creatinine)	9.11 (5.69 – 13.67)	12.5 (7.2 – 16.8)*	11.2 (6.3 – 15.6)	7.2 (5.6 – 10.9)	<0.05
DT (nmol/mg creatinine)	0.053 (0.044 – 0.061)	0.051 (0.043 – 0.061)	0.055 (0.046 – 0.067)	0.052 (0.044 – 0.059)	
GSA (nmol/mg creatinine)	7.45 (5.51 – 8.96)	9.36 (7.03 – 12.16)**	7.74 (5.92 – 9.40) ^o	8.08 (6.42 – 9.91)	<0.05
3-NT (nmol/mg creatinine)	0.012 (0.007 – 0.026)	0.011 (0.005 – 0.025)	0.012 (0.006 – 0.022)	0.012 (0.008 – 0.017)	
GEEK (nmol/mg creatinine)	0.422 (0.202 – 0.927)	0.359 (0.222 – 0.619)	0.403 (0.250 – 0.647)	0.277 (0.185 – 0.471)*	
Leu (nmol/mg creatinine)	22.9 (20.6 – 29.0)	24.4 (19.7 – 32.0)	22.8 (19.0 – 27.9)	16.5 (13.9 – 21.0)***, ^{ooo} , ^{†††}	<0.001
Ile (nmol/mg creatinine)	11.2 (9.2 – 13.1)	11.6 (9.1 – 15.0)	10.8 (8.6 – 13.0)	7.2 (5.8 – 9.0)***, ^{ooo} , ^{†††}	<0.001
Val (nmol/mg creatinine)	32.7 (29.8 – 37.7)	36.0 (28.6 – 45.1)	29.9 (25.8 – 37.1) ^o	19.6 (15.4 – 26.3)***, ^{ooo} , ^{†††}	<0.001
BCAA (nmol/mg creatinine)	67.2 (60.3 – 80.8)	73.3 (57.0 – 97.7)	64.6 (56.2 – 77.7)	44.2 (35.9 – 55.0)***, ^{ooo} , [†]	<0.001

Data are median (lower-upper quartile) with N values as given in Table 2. Significance: *, **, and ***, $P < 0.05$, $P < 0.01$, and $P < 0.001$ with respect to healthy controls; o, oo, and ooo, $P < 0.05$, $P < 0.01$, and $P < 0.001$ with respect to impaired metabolic health; and †, ††, and †††, $P < 0.05$, $P < 0.01$, and $P < 0.001$ with respect to impaired vascular health.

TABLE 5

(a) Algorithm outcome to detect health impairment comparing against good health

Algorithm Features	Set 1			Set 2		
	Health impairment	Metabolic	Age, BMI, FL, and val Vascular	Renal	Metabolic	Age, BMI, and leu Vascular
Accuracy (%)	83.5 (83.0 – 84.1)	70.5 (69.8 – 71.2)	89.9 (89.4 – 90.4)	84.0 (83.4 – 84.5)	71.6 (70.9 – 72.2)	85.4 (84.9 – 85.9)
Sensitivity (%)	79.1 (78.1 – 89.1)	69.7 (68.6 – 70.7)	87.8 (86.8 – 88.8)	79.9 (78.9 – 81.0)	67.7 (66.5 – 69.0)	80.6 (79.5 – 81.7)
Specificity (%)	87.1 (86.1 – 88.1)	71.4 (70.1 – 72.7)	91.6 (90.9 – 92.4)	87.2 (86.3 – 88.1)	75.6 (74.4 – 76.9)	89.2 (88.3 – 90.1)
Positive likelihood ratio	7.95 (7.24 – 8.65)	2.77 (2.59 – 2.96)	13.2 (12.2 – 14.2)	7.73 (7.09 – 8.36)	3.17 (2.97 – 3.37)	9.56 (8.74 – 10.4)
Negative likelihood ratio	0.24 (0.23 – 0.25)	0.43 (0.41 – 0.44)	0.13 (0.12 – 0.14)	0.23 (0.22 – 0.24)	0.43 (0.41 – 0.44)	0.22 (0.20 – 0.23)
Positive predictive value (%)	83.9 (82.9 – 84.9)	72.6 (71.7 – 73.5)	89.6 (88.8 – 90.4)	83.9 (83.0 – 84.8)	75.2 (74.3 – 76.0)	86.1 (85.2 – 87.1)
Negative predictive value (%)	84.2 (83.6 – 84.8)	69.4 (68.6 – 70.2)	90.9 (90.2 – 91.6)	84.7 (84.1 – 85.4)	69.4 (68.6 – 70.2)	85.6 (85.2 – 86.5)
F-measure	0.81 (0.80 – 0.82)	0.71 (0.70 – 0.72)	0.88 (0.87 – 0.89)	0.82 (0.81 – 0.82)	0.71 (0.70 – 0.72)	0.83 (0.82 – 0.83)
Diagnostic odds ratio	25.6 (23.1 – 28.1)	5.7 (5.3 – 6.1)	78.5 (70.0 – 87.0)	27.1 (24.6 – 29.6)	6.5 (6.1 – 6.9)	34.3 (30.9 – 37.7)

Diagnostic performance data are reported as mean (95% CI) of 100 times repeated 2-fold validation experiments.

(b) Algorithm outcome with 2-fold validation to detect health impairment by type using the SVM algorithm comparing against good health

Algorithm Features	Set 3		
	Health impairment	Metabolic	Age, BMI, and BCAA Vascular
Accuracy (%)	84.2 (83.6 – 84.7)	71.8 (71.1 – 72.4)	88.3 (87.8 – 88.9)
Sensitivity (%)	80.1 (79.2 – 81.0)	68.7 (67.6 – 69.8)	85.2 (84.0 – 86.5)
Specificity (%)	87.4 (86.5 – 88.3)	75.0 (73.9 – 76.1)	90.8 (90.0 – 91.6)
Positive likelihood ratio	8.04 (7.35 – 8.73)	3.13 (2.90 – 3.35)	11.6 (10.7 – 12.5)
Negative likelihood ratio	0.23 (0.22 – 0.24)	0.42 (0.40 – 0.43)	0.16 (0.15 – 0.17)
Positive predictive value (%)	84.2 (83.3 – 85.1)	74.8 (74.0 – 75.6)	88.5 (87.6 – 89.3)
Negative predictive value (%)	84.8 (84.3 – 85.4)	69.8 (69.1 – 70.5)	89.3 (88.5 – 90.1)
F-measure	0.82 (0.81 – 0.82)	0.71 (0.71 – 0.72)	0.86 (0.86 – 0.87)
Diagnostic odds ratio	27.9 (25.2 – 30.6)	6.6 (6.1 – 7.1)	56.8 (51.1 – 62.5)

Diagnostic performance data are reported as mean (95% CI) of 100 times repeated 2-fold validation experiments.

Several methods for subject assignment to one of the above four groups were tested with the SVM method producing superior diagnostic performance [43]. Features included in algorithm training were as follows: age, gender, BMI, alcohol intake, smoking status (current, former smoker, and never smoked), and urinary metabolites—FL, CML, MG-H1, GSP, pentosidine, pyrraline, DT, GSA, 3-NT, GEEK, and leu, ile, and val or total BCAA. Selection of optimal subset of features was data-driven to those combinations that minimise the group assignment error rate.

In the initial computations, we trained algorithms for each health impairment and healthy subjects. Few features were required for optimum diagnosis outcome. We found three combinations of features, Sets 1 – 3, which gave similar diagnostic performance: Set 1 features—age, BMI, and urinary excretion of FL and valine; Set 2 features—age, BMI, and urinary excretion of leu; and Set 3 features—age, BMI, and urinary excretion of val (Table 5). Diagnostic accuracies were 71 – 90%, sensitivities 68 – 88%, and specificities 71 – 92%. The highest accuracies for impaired metabolic, vascular, and renal health were 84%, 72%, and 90%, respectively, with

lowest false positive rates of 13%, 24%, and 8%. The highest positive likelihood ratios for impaired metabolic, vascular, and renal health were 8.0, 3.2, and 13.2, respectively, indicating that this biomarker set gave moderate evidence of impaired metabolic health, limited evidence of impaired vascular health, and convincing evidence of impaired renal health [44].

As a second approach, we performed a two-step algorithm analysis where step 1 distinguished between good health and impaired health of any kind investigated and then step 2 distinguished between the three different types of impaired health—metabolic, vascular, or renal (Figure 1). In step 1, outcome of the algorithm computations indicated that again few features were required for this: age, BMI, and urinary excretion of FL and valine. For step 1, the accuracy was 78% (random selection = 0.50) and sensitivity/specificity was 82/77% (Table 6). The positive likelihood ratio was 3.7, indicating the test provides moderate evidence of health impairment. In step 2, a 3-class algorithm was produced to distinguish between metabolic, vascular, or renal health impairment. Outcome of the algorithm computations

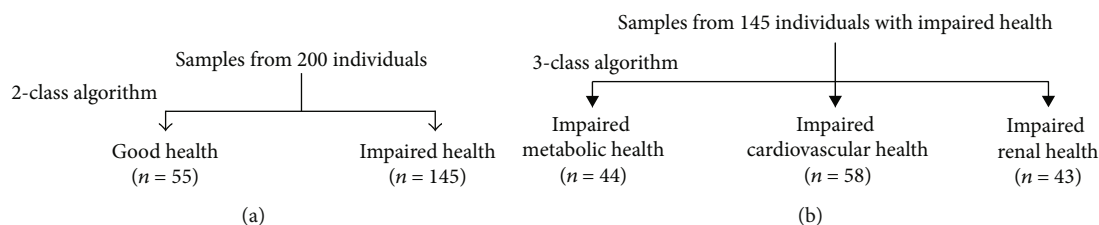


FIGURE 1: Training and validation of a multiclass algorithm for detection and discrimination of (a) good health versus impaired health disease and (b) impaired metabolic, vascular, and renal health.

TABLE 6: Algorithm outcome with 2-fold validation to detect health impairment using the SVM algorithm.

Algorithm	Step 1 – 2-class (good vs. impaired health)		Step 2 – 3 class algorithm (impaired metabolic, vascular, or renal health)		
	Age, BMI, FL, and val		Age, BMI, FL, and val		
Health impairment	All		Metabolic	Vascular	Renal
Accuracy (%)	78.2 (77.7 – 78.7)		73.4 (72.9 – 74.0)	68.7 (68.1 – 69.3)	78.1 (75.5 – 78.7)
Sensitivity (%)	81.5 (80.4 – 82.6)		54.9 (53.1 – 56.8)	58.6 (57.2 – 60.0)	67.4 (65.9 – 68.7)
Specificity (%)	77.0 (76.3 – 77.7)		81.5 (80.4 – 82.6)	75.4 (74.2 – 76.6)	82.6 (81.7 – 83.5)
Positive likelihood ratio	3.67 (3.57 – 3.77)		3.43 (3.20 – 3.66)	2.64 (2.51 – 2.77)	4.59 (4.19 – 4.98)
Negative likelihood ratio	0.24 (0.22 – 0.25)		0.55 (0.53 – 0.57)	0.55 (0.53 – 0.56)	0.39 (0.38 – 0.41)
Positive predictive value (%)	57.7 (57.0 – 58.3)		57.8 (56.5 – 59.1)	62.3 (61.3 – 63.2)	63.0 (61.9 – 64.2)
Negative predictive value (%)	91.8 (91.4 – 92.2)		80.9 (80.4 – 81.4)	73.5 (72.9 – 74.0)	85.9 (85.2 – 86.5)
F-measure	0.67 (0.66 – 0.68)		0.55 (0.54 – 0.56)	0.60 (0.59 – 0.61)	0.64 (0.63 – 0.65)
Diagnostic odds ratio	14.7 (14.0 – 15.4)		5.4 (5.0 – 5.8)	4.3 (4.1 – 4.5)	9.8 (8.9 – 10.7)

Diagnostic performance data are reported as mean (95% CI) of 100 times repeated 2-fold validation experiments.

indicated that features required for this are as follows: age, BMI, and urinary excretion of val. The accuracy for step 2 was 69–78% (random selection = 0.33). The positive likelihood ratios were in the range of 2.6 – 4.6, indicating a small increase in likelihood of health impairment with a positive outcome.

4. Discussion

This study produced quantitative estimates of glycosylated, oxidized, nitrated, crosslinked, and branched-chain amino acids in healthy subjects and changes in early-stage decline of metabolic, vascular, and renal health. The levels found suggested that for amino acid crosslinks, glucosepane had the highest urinary excretion, and for oxidized amino acids, GSA had the high urinary excretion. The urinary excretion of 3-NT was relatively low. This study also revealed the potential diagnostic utility of urinary BCAAs and FL. Positive likelihood ratios suggested that the use of urinary excretion of BCAA and FL as features in diagnostic algorithms gave small, moderate, and conclusive evidence of increased likelihood of early-stage cardiovascular, metabolic, and renal disease, respectively.

We chose urine analysis as the basis for our health screen. The ease of sample collection and analysis sees urine analysis as a current focus of development in quantitative clinical metabolomics. Compared to plasma or serum, urine has the disadvantage of being less well-buffered than plasma or serum and metabolite levels influenced by the time since previous bladder voiding and food consumption. In the current method, urine ultrafiltrate is mixed with stable isotopic stan-

dards in 0.1% trifluoroacetic acid which stabilizes analytes from risk of degradation at neutral and high pH; all are stable at low pH and ambient temperature. A further advantage is that plasma concentrations of glycosylated, oxidized, and nitrated amino acids are highly dependent on glomerular filtration rate due to related renal clearance whereas urinary concentrations of glycosylated, oxidized, and nitrated amino acids are not—being related rather to flux of protein glycation, oxidation, and nitration [11, 45]. For these reasons, we chose urine analysis as a preferred sample analysis protocol for health screening related to markers of protein glycation, oxidation, and nitration—although in machine learning analysis, only FL and BCAAs were informative features for subject group classification.

We defined impaired metabolic health herein by reference to detection of impaired glucose tolerance by increased A1C and to insulin resistance represented by the HOMA-IR. This is associated with exposure to increased glucose concentrations as judged by continuous glucose monitoring studies [46]. It was surprising that urinary excretion of glucose-derived FL was not increased in impaired metabolic health, *cf.* the marked increase in urinary excretion in patients with diabetes [47]. This may be due to efficient repair of FL free adduct by fructosamine 3-phosphokinase at levels formed in prediabetes [48] and a relatively large and variable contribution from food which may have increased data dispersion and thereby precluding detection of differences between study groups. Increased glucose exposure was rather reflected in increased urinary excretion of the glucose-dependent

glycation-derived crosslink, glucosepane. Glucosepane is formed from the degradation of FL and is not repaired enzymatically, potentially providing greater sensitivity for detection of dysglycemia. In analysis of protein glycation and oxidation adducts in plasma protein in a prospective study, we recently found that levels of glucosepane were the best predictor of development of T2DM [6]. Increased urinary excretion of GSA and CML—formed oxidatively from FL [25]—in early-stage impaired metabolic health may be indicative of oxidative stress. There was also increased urinary excretion of pyrroline which may indicate increased food consumption in this study group. The positive correlations of CML and GSA with pyrroline suggest that urinary increases of these metabolites were linked to increased food consumption. Surprisingly, there was no increase in urinary BCAAs in impaired metabolic health. Increased plasma levels of BCAA were found previously in overweight and obese subjects and correlated positively with insulin resistance [49, 50].

Vascular health impairment was associated with increased urinary excretion of glucosepane. Both glucosepane and CIMT correlated positively with age. Glucosepane is a major protein crosslink and may impact negatively on vascular contractility and function [25].

Renal health impairment was associated with decreased urinary excretion of individual and total BCAAs and also decreased urinary GEEK. The latter is a protein crosslink formed enzymatically by transglutaminases. Increased transglutaminase activity was associated with development of CKD [51]. Our finding is consistent with decreased removal of the GEEK crosslinks produced by increased transglutaminase activity in early-stage decline in renal function. A decrease of urinary excretion of BCAAs in chronic renal insufficiency was found previously—reviewed in [22], linked to both change in tubular transport and decreased interorgan exchange from impaired output by peripheral tissues [52].

In the machine learning-based analysis, we found that combinations of age, BMI with FL, and val, leu, or total BCAA was able to discriminate between healthy controls and subjects with early-stage metabolic, vascular, or renal health decline. Declining metabolic health, referenced to A1C and HOMA-IR, is associated with increased risk of developing type 2 diabetes [53]. Declining vascular health, referenced to increased CIMT, reflects asymptomatic atherosclerosis and increased risk of myocardial infarction and stroke [54]. Declining renal health, referenced to increased serum creatinine and decreased eGFR, is a measure of increased risk of renal disease linked to premature mortality and progression to renal failure [55]. The dominant involvement of subject age is expected as a nonmodifiable risk factor for these health impairments and BMI through association with insulin resistance of impaired metabolic health.

In subjects with impaired metabolic health, BMI correlated positively with urinary excretion of MG-H1, which is maintained when normalised to pyrroline, i.e., corrected for dietary AGE intake. MG-H1 is the major AGE formed from the reactive dicarbonyl metabolite, methylglyoxal. This suggests a link of impaired metabolic health with increased MG exposure, or dicarbonyl stress, for which independent

evidence has emerged recently [56]. The failure of urinary excretion of CML and MG-H1 to emerge as diagnostic features for diagnosis of early decline in metabolic and vascular health may be due to a large and variable contribution to urinary excretion of these analytes from digested food, compared to increases of endogenous formation.

Potential utility of measurement of glycated, oxidized, nitrated, and crosslinked amino acids and BCAAs in urine is suggested by the ease of urine sample collection and that damaged amino acids and disturbance in the levels of BCAAs may provide biomarkers of early-stage health decline preceding the development of chronic disease of high prevalence, morbidity, and mortality—diabetes, CVD, and CKD. A urine screening test for early-stage health decline would be a valuable clinical diagnostic asset—particularly when diagnostic assessment for all three early-stage health declines could be made from one analytical run. Screening for prediabetes in overweight and obese adults is considered to be cost effective when linked to subsequent implementation of lifestyle interventions to prevent T2DM [57]. There is profound limited awareness of early-stage CKD with only 9% of people with stage 3 CKD aware of their health impairment [58]—screening for CKD by serum creatinine and deduced estimated GFR (eGFR) is considered not cost effective other than for patients with T2DM and thereby increased risk of CKD [59]. In this initial study, however, we conclude that although there are changes in urinary glycated, oxidized, crosslinked, and branched-chain amino acids in these early-stage health declines, their measurement provides some but limited clinical diagnostic classifications for health screening.

An alternative noninvasive measure of insulin resistance is the ^{13}C -glucose test which requires ingestion of a stable isotopically labelled ^{13}C -glucose drink and collection of breath samples over the subsequent 6 h for measurement of exhaled $[\text{C}^{13}]/[\text{C}^{12}]\text{CO}_2$ ratio. Our test had similar accuracy and DOR values for detection of insulin resistance without requirement for administration of stable isotope and sample collections over 6 h [60]. An alternative noninvasive method for detection of impaired glucose tolerance is measurement of skin autofluorescence but this is compared poorly, with DOR of 2.7 [61], *cf.* DOR of 25.6 – 27.1 for methods herein. Alternative biomarkers for increased CIMT as an indicator of impaired vascular health are red cell distribution width and serum podocalyxin with DOR values of 3.9 and 3.2, respectively [62, 63]. Our algorithms fared better with values of 5.7 – 6.6. An alternative noninvasive biomarker of impaired renal health is salivary creatinine for which the DOR value for the same classification of CKD as applied herein was 17.8 [64], *cf.* DOR values herein of 34.3 – 78.5. In the diagnostic algorithm for detection of any early-stage health decline studied (metabolic, vascular, and renal health), the negative predictive value of 92% and DOR of 14.7 (Table 6) could provide the basis for screening for the absence of early-stage health decline, after further independent validation.

Potential implementation of this urinary health screening test would be facilitated with further studies to ease access, increase sample analysis throughput, and further validation on one or more independent clinical subject cohorts. Ease

of access to the test would be improved through urinary sample donation at primary health care centers. For this, further studies of sample stability during storage and shipment at ambient temperature for urinary analytes of N_ε-fructosyl-lysine (FL) and BCAAs would be required. Studies on sample storage and preanalytical processing validation for unfocused metabolomic analysis of urine were recently reviewed [65]. For analyte quantitation, sample throughput may be increased with FL, BCAAs, and creatinine analysed in one run by stable isotopic dilution analysis LC-MS/MS [66] or alternatively by ¹H nuclear magnetic resonance [67]. We estimate minimum sample analysis time would be ca. 10 – 20 min and 7 min, for these methods, respectively. For further independent validation of diagnostic outcomes, the assessment of good health versus health impairment (any type) had sensitivity of 81.5%. From statistical considerations based on this and absolute precision of ±5% [68], a validation study requires a minimum of 464 subjects for a 1:1 ratio of cases and controls.

We studied early-stage health decline as part of the rationale of the BIOCLAIMS research project and the related concept of “health biomarkers”—see above. There is also a consensus view that if health decline can be detected in the early stages then progression to frank disease may be prevented. This is particularly important for CVD where a first clinical disease event may be fatal and for diabetes and renal failure where there are difficult-to-treat debilitating and life-threatening complications. Also, screening of early-stage health decline could be made more cost effective if assessments of early-stage development of multiple disease are made concurrently.

5. Conclusions

- (i) We investigated the urinary excretion of oxidized, glycosylated, nitrated, crosslinked, and branched-chain amino acids and their association with early-stage decline in metabolic, vascular, and renal health
- (ii) We found characteristic changes in early-stage health decline where algorithm features of age, body mass index (BMI), fructosyl-lysine (FL), and branched-chain amino acids (BCAAs) combined to give small, moderate, and conclusive evidence of increased likelihood of early-stage cardiovascular, metabolic, and renal disease, respectively
- (iii) After further validation, urinary measurement of FL and BCAAs could help improve ease of access to diagnosis of impaired metabolic health in prediabetes and asymptomatic arterial stenosis and chronic kidney disease where remedial lifestyle and therapeutic intervention may prevent progression to advanced disease and premature mortality

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflict of interest in relation to this work.

Authors' Contributions

Naila R and PJT designed the study. JMR, HHG, BWR, GF, and JG recruited and investigated the study subjects. GF, JG, BT, and BOP generated clinical and laboratory data. BWR, JMR, and GF assigned the study subjects to clinically meaningful contrasting groups and characterized the study subjects. Nasir R, KR, JR, BWR, Naila R, and PJT analysed and interpreted the analytical data, Nasir R, BWR, and PJT drafted the report, and all authors reviewed and revised it critically and approved the final version to be published. Naila R and PJT had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Prof. Johannes M. Roob deceased during the final preparation of this manuscript.

Acknowledgments

This work was carried out with financial support of the European Union Framework Programme 7 FP7 2007-2013 under grant agreement no. 244995 (BIOCLAIMS Project) and the Austrian Federal Ministry of Science, Research and Economy to Karl-Franzens University of Graz and Medical University of Graz. The authors thank Petra Kieslinger, Daniela Berger, Verena Schaberl, Agnes Schriebl, Theopisti Maimari, Natalie Walter, Cornelia Missbrenner, Nicole Hacker, and Verena Zachhuber for excellent technical assistance.

Supplementary Materials

The supplementary materials include 4 tables. Supplementary Table 1: correlation analysis of clinical characteristics and urinary analytes: healthy controls. Supplementary Table 2: correlation analysis of clinical characteristics and urinary analytes: impaired metabolic health. Supplementary Table 3: correlation analysis of clinical characteristics and urinary analytes: impaired vascular health. Supplementary Table 4: correlation analysis of clinical characteristics and urinary analytes: impaired renal health. (*Supplementary Materials*)

References

- [1] S. S. Lim, T. Vos, A. D. Flaxman et al., “A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010,” *The Lancet*, vol. 380, no. 9859, pp. 2224–2260, 2012.
- [2] A. Ceriello and E. Motz, “Is oxidative stress the pathogenic mechanism underlying insulin resistance, diabetes, and cardiovascular disease? The common soil hypothesis revisited,” *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 24, no. 5, pp. 816–823, 2004.
- [3] E. Schleicher and U. Friess, “Oxidative stress, AGE, and atherosclerosis,” *Kidney International*, vol. 72, article S0085253815525477, pp. S17–S26, 2007.

- [4] Z. A. Massy, P. Stenvinkel, and T. B. Drueke, "Progress in uremic toxin research: the role of oxidative stress in chronic kidney disease," *Seminars in Dialysis*, vol. 22, no. 4, pp. 405–408, 2009.
- [5] American Diabetes Association, "2. Classification and diagnosis of diabetes: standards of medical care in diabetes—2019," *Diabetes Care*, vol. 41, Supplement 1, pp. S13–S28, 2019.
- [6] C.-J. Chiu, N. Rabbani, S. Rowan et al., "Studies of advanced glycation end products and oxidation biomarkers for type 2 diabetes," *BioFactors*, vol. 44, no. 3, pp. 281–288, 2018.
- [7] N. Rabbani and P. J. Thornalley, "Advanced glycation end products in the pathogenesis of chronic kidney disease," *Kidney International*, vol. 93, no. 4, pp. 803–813, 2018.
- [8] N. M. J. Hanssen, L. Engelen, I. Ferreira et al., "Plasma levels of advanced glycation endproducts N^ε-(carboxymethyl)lysine, N^ε-(carboxyethyl)lysine, and pentosidine are not independently associated with cardiovascular disease in individuals with or without type 2 diabetes: the Hoorn and CODAM studies," *Journal of Clinical Endocrinology & Metabolism*, vol. 98, no. 8, pp. E1369–E1373, 2013.
- [9] M. Xue, M. O. Weickert, S. Qureshi et al., "Improved glycemic control and vascular function in overweight and obese subjects by glyoxalase 1 inducer formulation," *Diabetes*, vol. 65, no. 8, pp. 2282–2294, 2016.
- [10] B. K. Kilhovd, A. Juutilainen, S. Lehto et al., "Increased serum levels of methylglyoxal-derived hydroimidazolone-AGE are associated with increased cardiovascular disease mortality in nondiabetic women," *Atherosclerosis*, vol. 205, no. 2, pp. 590–594, 2009.
- [11] S. Agalou, N. Ahmed, R. Babaei-Jadidi, A. Dawnay, and P. J. Thornalley, "Profound mishandling of protein glycation degradation products in uremia and dialysis," *Journal of the American Society of Nephrology*, vol. 16, no. 5, pp. 1471–1485, 2005.
- [12] P. J. Thornalley, S. Battah, N. Ahmed et al., "Quantitative screening of advanced glycation endproducts in cellular and extracellular proteins by tandem mass spectrometry," *Biochemical Journal*, vol. 375, no. 3, pp. 581–592, 2003.
- [13] N. Ahmed, P. J. Thornalley, R. Lüthen et al., "Processing of protein glycation, oxidation and nitrosation adducts in the liver and the effect of cirrhosis," *Journal of Hepatology*, vol. 41, no. 6, pp. 913–919, 2004.
- [14] G. Delpierre, D. Vertommen, D. Communi, M. H. Rider, and E. Van Schaftingen, "Identification of fructosamine residues deglycated by fructosamine-3-kinase in human hemoglobin," *Journal of Biological Chemistry*, vol. 279, no. 26, pp. 27613–27620, 2004.
- [15] R. J. W. Truscott, K. L. Schey, and M. G. Friedrich, "Old proteins in man: a field in its infancy," *Trends in Biochemical Sciences*, vol. 41, no. 8, pp. 654–664, 2016.
- [16] W. A. Edens, L. Sharling, G. J. Cheng et al., "Tyrosine cross-linking of extracellular matrix is catalyzed by Duox, a multidomain oxidase/peroxidase with homology to the phagocyte oxidase subunit gp91^{phox}," *Journal of Cell Biology*, vol. 154, no. 4, pp. 879–892, 2001.
- [17] K. M. Biemel, D. A. Friedl, and M. O. Lederer, "Identification and quantification of major Maillard cross-links in human serum albumin and lens protein - evidence for glucosepane as the dominant compound," *Journal of Biological Chemistry*, vol. 277, no. 28, pp. 24907–24915, 2002.
- [18] L. Lorand and R. M. Graham, "Transglutaminases: crosslinking enzymes with pleiotropic functions," *Nature Reviews Molecular Cell Biology*, vol. 4, no. 2, pp. 140–156, 2003.
- [19] K. Yasumoto and F. Suzuki, "Aspartyl- and glutamyl-lysine crosslinks formation and their nutritional availability," *Journal of Nutritional Science and Vitaminology*, vol. 36, no. 4, Supplement I, pp. S71–S77, 1990.
- [20] Y. Y. Ding, X. Tang, X. R. Cheng et al., "Effects of dietary oxidized tyrosine products on insulin secretion via the thyroid hormone T3-regulated TRβ1–Akt–mTOR pathway in the pancreas," *RSC Advances*, vol. 7, no. 86, pp. 54610–54625, 2017.
- [21] C. J. Lynch and S. H. Adams, "Branched-chain amino acids in metabolic signalling and insulin resistance," *Nature Reviews Endocrinology*, vol. 10, no. 12, pp. 723–736, 2014.
- [22] N. J. M. Cano, D. Fouque, and X. M. Leverve, "Application of branched-chain amino acids in human pathological states: renal failure," *The Journal of Nutrition*, vol. 136, no. 1, pp. 299S–307S, 2006.
- [23] A. Foerster and T. Henle, "Glycation in food and metabolic transit of dietary AGEs (advanced glycation end-products): studies on the urinary excretion of pyrroline," *Biochemical Society Transactions*, vol. 31, no. 6, pp. 1383–1385, 2003.
- [24] C. Hohmann, K. Liehr, C. Henning et al., "Detection of free advanced glycation end products in vivo during hemodialysis," *Journal of Agricultural and Food Chemistry*, vol. 65, no. 4, pp. 930–937, 2017.
- [25] P. J. Thornalley and N. Rabbani, "Detection of oxidized and glycated proteins in clinical samples using mass spectrometry — a user's perspective," *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. 1840, no. 2, pp. 818–829, 2014.
- [26] B. van Ommen, J. Keijer, S. G. Heil, and J. Kaput, "Challenging homeostasis to define biomarkers for nutrition related health," *Molecular Nutrition & Food Research*, vol. 53, no. 7, pp. 795–804, 2009.
- [27] C. Picó, F. Serra, A. M. Rodríguez, J. Keijer, and A. Palou, "Biomarkers of nutrition and health: new tools for new approaches," *Nutrients*, vol. 11, no. 5, article 1092, 2019.
- [28] A. S. Levey, J. P. Bosch, J. B. Lewis, T. Greene, N. Rogers, and D. Roth, "A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation," *Annals of Internal Medicine*, vol. 130, no. 6, pp. 461–470, 1999.
- [29] D. R. Matthews, J. P. Hosker, A. S. Rudenski, B. A. Naylor, D. F. Treacher, and R. C. Turner, "Homeostasis model assessment: insulin resistance and β-cell function from fasting plasma glucose and insulin concentrations in man," *Diabetologia*, vol. 28, no. 7, pp. 412–419, 1985.
- [30] American-Diabetes-Association, "Standards of medical care in diabetes-2019," *Diabetes Care*, vol. 39, Supplement 1, pp. S1–S193, 2019.
- [31] W. A. Riley, R. W. Barnes, M. G. Bond, G. Evans, L. E. Chambless, and G. Heiss, "High-resolution B-mode ultrasound reading methods in the atherosclerosis risk in communities (ARIC) cohort," *Journal of Neuroimaging*, vol. 1, no. 4, pp. 168–172, 1991.
- [32] J. H. Stein, C. E. Korcarz, R. T. Hurst et al., "Use of carotid ultrasound to identify subclinical vascular disease and evaluate cardiovascular disease risk: a consensus statement from the American Society of Echocardiography Carotid Intima-Media Thickness Task Force endorsed by the Society for Vascular Medicine," *Journal of the American Society of Echocardiography*, vol. 21, no. 2, pp. 93–111, 2008.

- [33] N. Denarie, J. Gariépy, G. Chironi et al., "Distribution of ultrasonographically-assessed dimensions of common carotid arteries in healthy adults of both sexes," *Atherosclerosis*, vol. 148, no. 2, pp. 297–302, 2000.
- [34] K. J. Knecht, J. A. Dunn, K. F. McFarland et al., "Effect of diabetes and aging on carboxymethyllysine levels in human urine," *Diabetes*, vol. 40, no. 2, pp. 190–196, 1991.
- [35] N. Rabbani, F. Shaheen, A. Anwar, J. Masania, and P. J. Thornalley, "Assay of methylglyoxal-derived protein and nucleotide AGEs," *Biochemical Society Transactions*, vol. 42, no. 2, pp. 511–517, 2014.
- [36] C. Schafer, M. Schott, F. Brandl, S. Neidhart, and R. Carle, "Identification and quantification of ϵ -(γ -glutamyl)lysine in digests of enzymatically cross-linked leguminous proteins by high-performance liquid chromatography–electrospray ionization mass spectrometry (HPLC-ESI-MS)," *Journal of Agricultural and Food Chemistry*, vol. 53, no. 8, pp. 2830–2837, 2005.
- [37] M. O. Lederer and H. P. Bühler, "Cross-linking of proteins by Maillard processes—characterization and detection of a lysine-arginine cross-link derived from d-glucose," *Bioorganic & Medicinal Chemistry*, vol. 7, no. 6, pp. 1081–1088, 1999.
- [38] M. Akagawa, D. Sasaki, Y. Ishii et al., "New method for the quantitative determination of major protein carbonyls, α -amino adipic and γ -glutamic semialdehydes: investigation of the formation mechanism and chemical nature in vitro and in vivo," *Chemical Research in Toxicology*, vol. 19, no. 8, pp. 1059–1065, 2006.
- [39] H. J. Vogel and B. D. Davis, "Glutamic γ -semialdehyde and Δ^1 -pyrroline-5-carboxylic acid, intermediates in the biosynthesis of proline^{1,2}," *Journal of the American Chemical Society*, vol. 74, no. 1, pp. 109–112, 1952.
- [40] L. Naranjo, E. M. de Valmaseda, O. Bañuelos et al., "Conversion of pipercolic acid into lysine in *Penicillium chrysogenum* requires pipercolate oxidase and saccharopine reductase: characterization of the lys7 gene encoding saccharopine reductase," *Journal of Bacteriology*, vol. 183, no. 24, pp. 7165–7172, 2001.
- [41] K. Sadilkova, J. S. M. Gospe, and S. H. Hahn, "Simultaneous determination of alpha-amino adipic semialdehyde, piperideine-6-carboxylate and pipercolic acid by LC-MS/MS for pyridoxine-dependent seizures and folinic acid-responsive seizures," *Journal of Neuroscience Methods*, vol. 184, no. 1, pp. 136–141, 2009.
- [42] P.-A. H. Finot and R. Bertholet, "Preparation of ϵ -(γ -glutamyl)-lysine," US Patent 4,052,372, 1977.
- [43] P. Sajda, "Machine learning for detection and diagnosis of disease," *Annual Review of Biomedical Engineering*, vol. 8, no. 1, pp. 537–565, 2006.
- [44] D. A. Grimes and K. F. Schulz, "Refining clinical diagnosis with likelihood ratios," *The Lancet*, vol. 365, no. 9469, pp. 1500–1505, 2005.
- [45] N. Rabbani, K. Sebekova, K. Sebekova Jr., A. Heidland, and P. J. Thornalley, "Accumulation of free adduct glycation, oxidation, and nitration products follows acute loss of renal function," *Kidney International*, vol. 72, no. 9, pp. 1113–1121, 2007.
- [46] C. L. Chan, L. Pyle, L. Newnes, K. J. Nadeau, P. S. Zeitler, and M. M. Kelsey, "Continuous glucose monitoring and its relationship to hemoglobin A1c and oral glucose tolerance testing in obese and prediabetic youth," *The Journal of Clinical Endocrinology & Metabolism*, vol. 100, no. 3, pp. 902–910, 2015.
- [47] N. Ahmed, R. Babaei-Jadidi, S. K. Howell, P. J. Beisswenger, and P. J. Thornalley, "Degradation products of proteins damaged by glycation, oxidation and nitration in clinical type 1 diabetes," *Diabetologia*, vol. 48, no. 8, pp. 1590–1603, 2005.
- [48] B. S. Szweggold, S. Howell, and P. J. Beisswenger, "Human fructosamine-3-kinase. Purification, sequencing, substrate specificity, and evidence of activity in vivo," *Diabetes*, vol. 50, no. 9, pp. 2139–2147, 2001.
- [49] C. B. Newgard, J. An, J. R. Bain et al., "A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance," *Cell Metabolism*, vol. 9, no. 4, pp. 311–326, 2009.
- [50] S. Cheng, E. P. Rhee, M. G. Larson et al., "Metabolite profiling identifies pathways associated with metabolic risk in humans," *Circulation*, vol. 125, no. 18, pp. 2222–2231, 2012.
- [51] J. R. Schelling, "Tissue transglutaminase inhibition as treatment for diabetic glomerular scarring: it's good to be glueless," *Kidney International*, vol. 76, no. 4, pp. 363–365, 2009.
- [52] G. Garibotto, E. Paoletti, F. Fiorini et al., "Peripheral metabolism of branched-chain keto acids in patients with chronic renal failure," *Mineral and Electrolyte Metabolism*, vol. 19, no. 1, pp. 25–31, 1993.
- [53] J. P. Crandall, W. C. Knowler, S. E. Kahn et al., "The prevention of type 2 diabetes," *Nature Clinical Practice Endocrinology & Metabolism*, vol. 4, no. 7, pp. 382–393, 2008.
- [54] M. W. Lorenz, H. S. Markus, M. L. Bots, M. Rosvall, and M. Sitzer, "Prediction of clinical cardiovascular events with carotid intima-media thickness: a systematic review and meta-analysis," *Circulation*, vol. 115, no. 4, pp. 459–467, 2007.
- [55] P. Sharma, K. McCullough, G. Scotland et al., "Does stage-3 chronic kidney disease matter?: a systematic literature review," *British Journal of General Practice*, vol. 60, no. 575, pp. e266–e276, 2010.
- [56] N. Rabbani and P. J. Thornalley, "Dicarbonyl stress in cell and tissue dysfunction contributing to ageing and disease," *Biochemical and Biophysical Research Communications*, vol. 458, no. 2, pp. 221–226, 2015.
- [57] T. J. Hoerger, K. A. Hicks, S. W. Sorensen et al., "Cost-effectiveness of screening for pre-diabetes among overweight and obese U.S. adults," *Diabetes Care*, vol. 30, no. 11, pp. 2874–2879, 2007.
- [58] M. Kurella Tamura, S. Anand, S. Li et al., "Comparison of CKD awareness in a screening population using the modification of diet in renal disease (MDRD) study and CKD epidemiology collaboration (CKD-EPI) equations," *American Journal of Kidney Diseases*, vol. 57, no. 3, Supplement 2, pp. S17–S23, 2011.
- [59] B. Manns, B. Hemmelgarn, M. Tonelli et al., "Population based screening for chronic kidney disease: cost effectiveness study," *BMJ*, vol. 341, no. 1, article c5869, 2010.
- [60] K. Tanaka, T. Matsuura, D. Shindo et al., "Noninvasive assessment of insulin resistance in the liver using the fasting ¹³C-glucose breath test," *Translational Research*, vol. 162, no. 3, pp. 191–200, 2013.
- [61] P. R. Baginsky, M. Jolicoeur, L. Hurlbert, and J. Revels, "Skin autofluorescence: a tool for identification of prediabetes (preDM) and diabetes (DM)," *Diabetes*, vol. 60, Supplement 1, pp. A328–A329, 2011.
- [62] D. Ren, J. Wang, H. Li, Y. Li, and Z. Li, "Red blood cell distribution width and carotid intima-media thickness in patients

- with metabolic syndrome,” *BMC Cardiovascular Disorders*, vol. 17, no. 1, article 44, 2017.
- [63] M. Shoji, M. Takemoto, K. Kobayashi et al., “Serum podocalyxin levels correlate with carotid intima media thickness, implicating its role as a novel biomarker for atherosclerosis,” *Scientific Reports*, vol. 8, no. 1, p. 245, 2018.
- [64] R. Renda, “Can salivary creatinine and urea levels be used to diagnose chronic kidney disease in children as accurately as serum creatinine and urea levels? A case-control study,” *Renal Failure*, vol. 39, no. 1, pp. 452–457, 2017.
- [65] V. L. Stevens, E. Hoover, Y. Wang, and K. A. Zanetti, “Pre-analytical factors that affect metabolite stability in human urine, plasma, and serum: a review,” *Metabolites*, vol. 9, no. 8, p. 156, 2019.
- [66] N. Takahashi, G. Boysen, F. Li, Y. Li, and J. A. Swenberg, “Tandem mass spectrometry measurements of creatinine in mouse plasma and urine for determining glomerular filtration rate,” *Kidney International*, vol. 71, no. 3, pp. 266–271, 2006.
- [67] T. Tynkkynen, Q. Wang, J. Ekholm et al., “Proof of concept for quantitative urine NMR metabolomics pipeline for large-scale epidemiology and genetics,” *International Journal of Epidemiology*, vol. 48, no. 3, pp. 978–993, 2019.
- [68] J. Xia, D. L. Broadhurst, M. Wilson, and D. S. Wishart, “Translational biomarker discovery in clinical metabolomics: an introductory tutorial,” *Metabolomics*, vol. 9, no. 2, pp. 280–299, 2012.
- [69] H. F. Erbersdobler and V. Faist, “Metabolic transit of *Amadori* products,” *Nahrung/Food*, vol. 45, no. 3, pp. 177–181, 2001.
- [70] N. Ahmed, R. Babaei-Jadidi, S. K. Howell, P. J. Thornalley, and P. J. Beisswenger, “Glycated and oxidized protein degradation products are indicators of fasting and postprandial hyperglycemia in diabetes,” *Diabetes Care*, vol. 28, no. 10, pp. 2465–2471, 2005.
- [71] R. Liardon, D. de Weck-Gaudard, G. Philipossian, and P. A. Finot, “Identification of N.epsilon.-carboxymethyllysine: a new Maillard reaction product in rat urine,” *Journal of Agricultural and Food Chemistry*, vol. 35, no. 3, pp. 427–431, 1987.
- [72] F. Wang, Y. Zhao, Y. Niu et al., “Activated glucose-6-phosphate dehydrogenase is associated with insulin resistance by upregulating pentose and pentosidine in diet-induced obesity of rats,” *Hormone and Metabolic Research*, vol. 44, no. 13, pp. 938–942, 2012.
- [73] J. R. Requena, C. C. Chao, R. L. Levine, and E. R. Stadtman, “Glutamic and amino adipic semialdehydes are the main carbonyl products of metal-catalyzed oxidation of proteins,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, pp. 68–74, 2001.
- [74] D. Tsikas, “Analytical methods for 3-nitrotyrosine quantification in biological samples: the unique role of tandem mass spectrometry,” *Amino Acids*, vol. 42, no. 1, pp. 45–63, 2012.