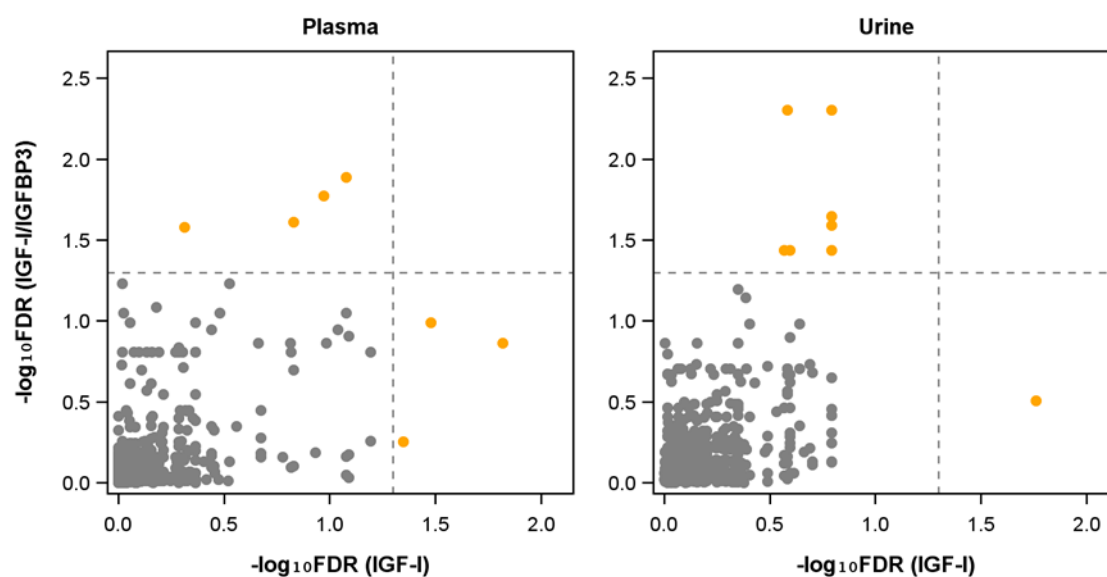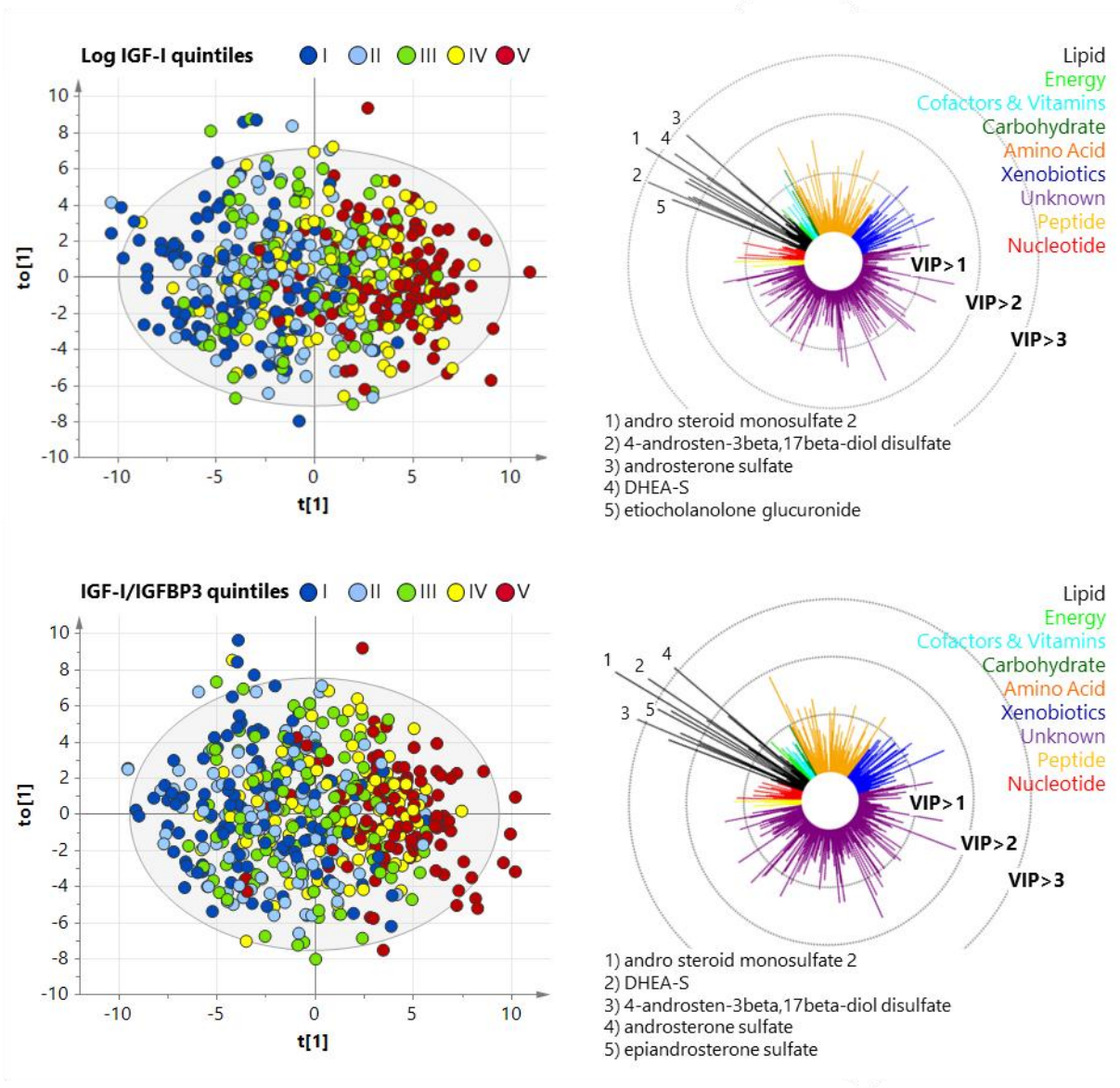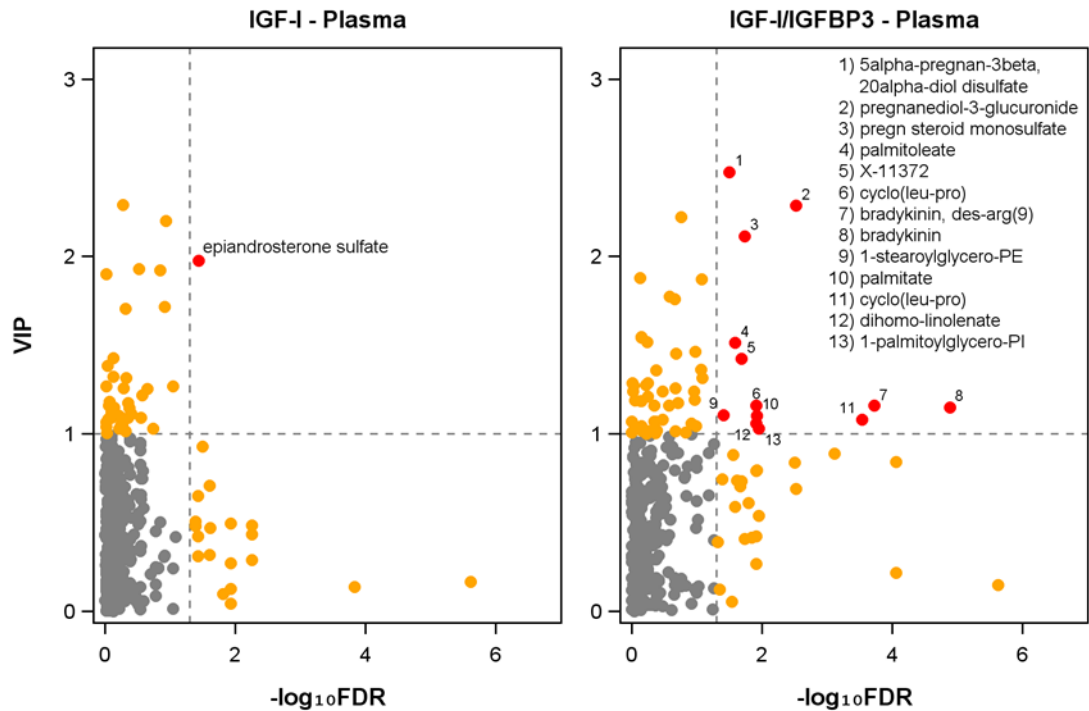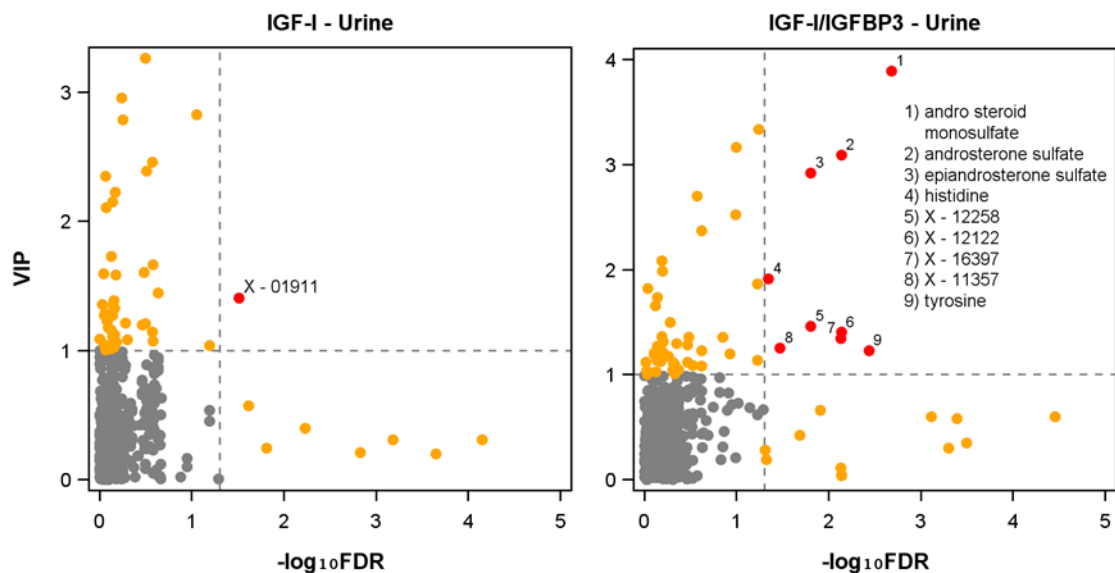## Supplemental figures

Supplemental Figure 1. Corrected p-values from univariate linear regression analysis of metabolites associated with IGF-I and IGF/IGFBP3 ratio among women in plasma (left panel) and urine (right panel). The dotted line denotes a false discovery rate (FDR) of 0.05, displaying the significance threshold. Orange marked metabolites are related either to IGF-I or the IGF-I/IGFBP-3 ratio. Regression models are adjusted for age, smoking, alcohol consumption, physical activity, waist circumference, liver diseases, LDL cholesterol and hypertension.

Supplemental Figure 2. OPLS analyses for modelling IGF-I and IGF-I/IGFBP-3 ratio based on **urine metabolites** among women. Left side: OPLS score plots for log IGF-I (top line) and IGF-I/IGFBP-3 ratio (bottom line) showing the predictive component t[1] and the first orthogonal component $t_0[1]$. Right side: Corresponding variable influence on projection (VIP) values of the predictive component. The five metabolites with the highest VIP were mentioned.

Supplemental Figure 3. Corrected p-values from univariate linear regression analysis of **plasma metabolites** associated with IGF-I (left side) and IGF/IGFBP3 ratio (right side) against the variable influence on projection (VIP) values based on orthogonal partial least squares (OPLS) regression analyses among women.


Supplemental Figure 4. Corrected p-values from univariate linear regression analysis of **urine metabolites** associated with IGF-I (left side) and IGF/IGFBP3 ratio (right side) against the variable influence on projection (VIP) values based on orthogonal partial least squares (OPLS) regression analyses among women.
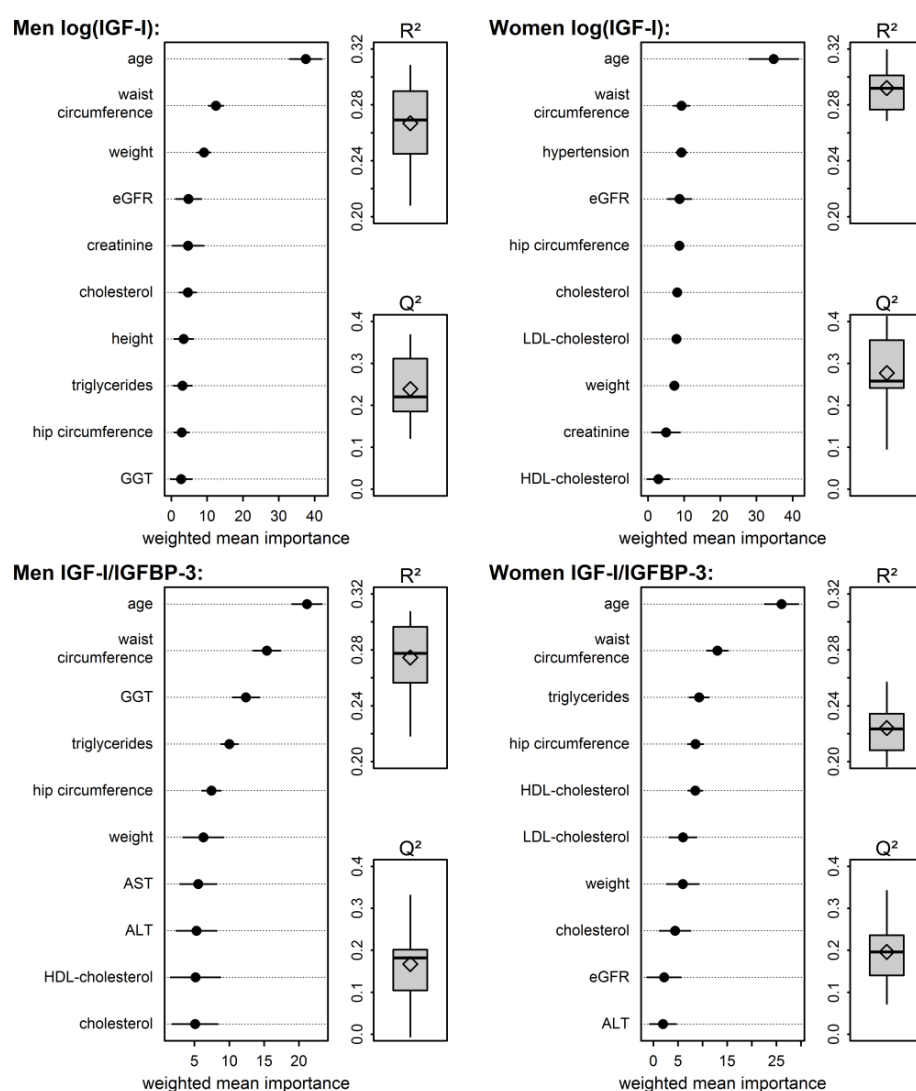
Supplemental Figure 5. Final results from random forest regression analyses predicting log(IGF-I) (*upper panel*) and the IGF-I/IGFBP-3 ratio (*lower panel*) in men (*left panel*) and women (*right panel*) in a two-stage cross validation. For each hormone and sex the dotchart shows the mean and standard deviation of a weighted (root mean square error) rank for the ten most important variables across the 10 outer loops. Additionally, explained (R^2) and predicted variance (Q^2) are displayed as boxplots. eGFR = estimated glomerular filtration rate; GGT = gamma-glutamyl transpeptidase; AST = aspartate amino transferase; ALT = alanine amino transferase.
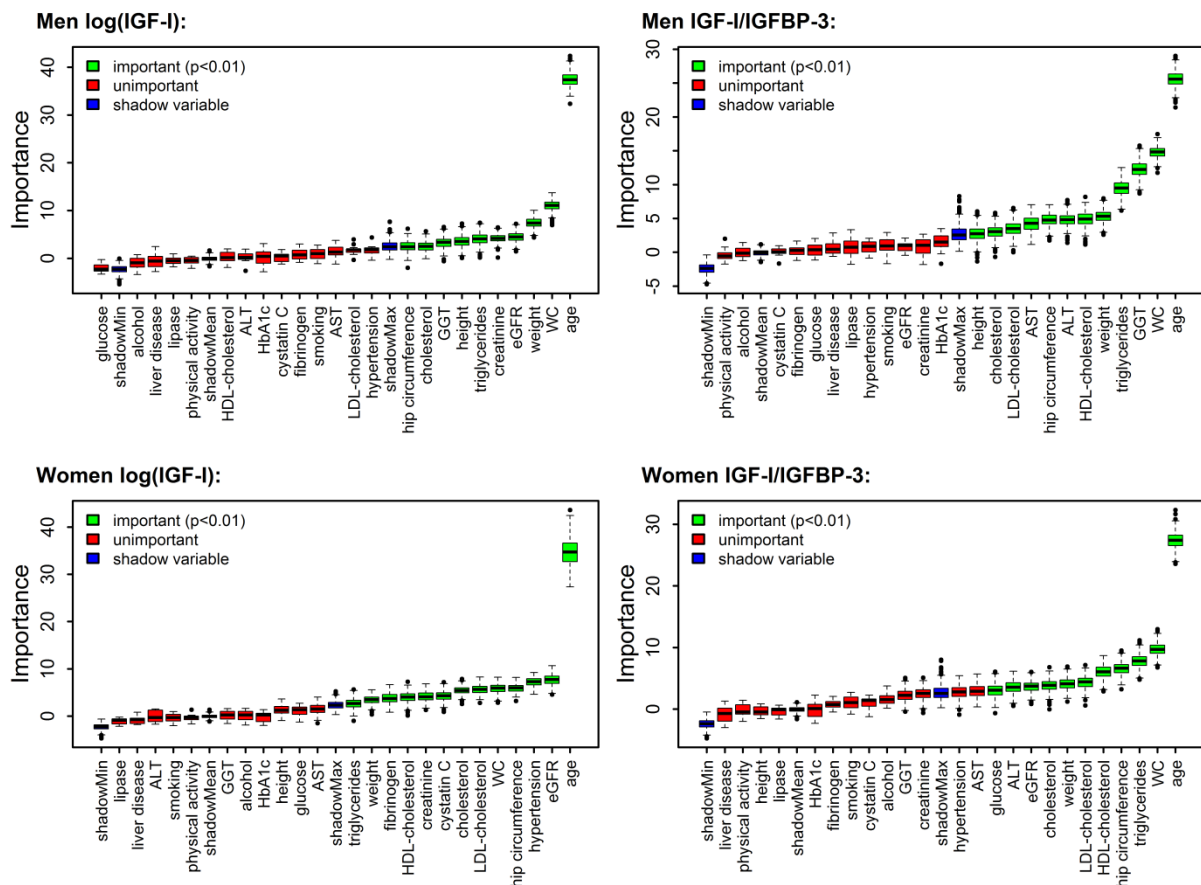
Supplemental Figure 6. Boxplots of importance measures from Boruta feature selection with a random forest as classifier for prediction of log(IGF-I) (left panel) and the IGF-I/IGFBP-3 ratio (right panel) in men (upper panel) and women (lower panel), respectively. Green boxes indicate significant (p<0.01) important variables to explain serum hormone concentrations. Blue boxes indicate variation in shadow variables used to assess random noise. WC = waist circumference; GGT = γ-glutamyl transpeptidase; AST = aspartate transaminase; HbA1c = glycated hemoglobin; ALT = alanine transaminase.

Supplemental tables

Supplemental Table 1. Metabolites with variable importance in projection (VIP) > 1 resulting from OPLS models in women.

Plasma			Urine		
	IGF-I/IGFBP3 ratio	IGF-I		IGF-I/IGFBP3 ratio	IGF-I
5alpha-pregnan-3beta,20alpha-diol disulfate	2.476	2.201	andro steroid monosulfate 2	3.888	3.263
pregnanediol-3-glucuronide	2.286	1.716	DHEA-S	3.336	2.787
DHEA-S	2.223	2.290	4-androsten-3beta,17beta-diol disulfate (2)	3.162	2.956
pregn steroid monosulfate	2.115	1.928	androsterone sulfate	3.090	2.826
pregnenolone sulfate	1.877	1.901	epiandrosterone sulfate	2.918	2.390
epiandrosterone sulfate	1.873	1.976	andro steroid monosulfate (1)	2.698	2.351
1-oleoylglycerol (1-monoolein)	1.773	1.705	etiocholanolone glucuronide	2.521	2.461
androsterone sulfate	1.760	1.921	21-hydroxypregnenolone disulfate	2.372	2.149
eicosapentaenoate (EPA; 20:5n3)	1.547	1.054	pregnen-diol disulfate	2.080	2.104
docosapentaenoate (n3 DPA; 22:5n3)	1.542	1.383	4-androsten-3beta,17beta-diol disulfate (1)	1.986	2.226
CMPF	1.517	1.268	histidine	1.913	1.664
palmitoleate (16:1n7)	1.514	1.254	X - 17736	1.865	1.388
bilirubin (Z,Z)	1.463		CMPF	1.821	1.585
10-heptadecenoate (17:1n7)	1.453	1.314	X - 16087	1.736	1.731
X - 11372	1.424	1.029	salicyluric glucuronide	1.656	1.357
eicosenoate (20:1n9 or 11)	1.362	1.146	X - 18838	1.499	1.603
andro steroid monosulfate 2	1.357	1.323	X - 12258	1.459	1.037
X - 11880	1.316	1.101	X - 12122	1.402	
hexanoylcarnitine	1.288	1.267	X - 12704	1.360	1.043
butyrylcarnitine	1.287		X - 01911	1.359	1.407
pregnen-diol disulfate	1.275	1.428	tryptophan	1.354	
X - 02269	1.273	1.124	X - 16397	1.344	
10-nonadecenoate (19:1n9)	1.258	1.219	X - 20620	1.313	1.156
X - 12096	1.241	1.093	N-acetyl-aspartyl-glutamate (NAAG)	1.295	1.593
docosahexaenoate (DHA; 22:6n3)	1.239		X - 13726	1.280	1.213
hydroxybutyrylcarnitine	1.239	1.174	X - 13844	1.274	1.326
X - 11299	1.210		X - 16774	1.267	1.002
1-arachidonoylglycerophosphoinositol	1.191		X - 11357	1.251	
X - 11469	1.189		tyrosine	1.230	
X - 11564	1.185	1.041	4-ethylphenylsulfate	1.228	
gamma-glutamylvaline	1.185		X - 12329	1.211	
21-hydroxypregnenolone disulfate	1.176	1.050	N-acetylaspartate (NAA)	1.206	1.443

Metabolite			Metabolite		
myristoleate (14:1n5)	1.161	1.113	N-acetyl-1-methylhistidine	1.200	
C-glycosyltryptophan	1.161	1.083	glycocholenate sulfate	1.195	
cyclo(leu-pro)	1.160		homovanillate sulfate	1.194	1.228
bradykinin, des-arg(9)	1.158		glycylproline	1.171	1.196
bradykinin	1.151		N6-acetyllysine	1.139	1.088
1-stearoylglycerophosphoethanolamine	1.106		X - 11593	1.136	
palmitate (16:0)	1.101		N-acetylthreonine	1.126	1.123
1-palmitoylglycerophosphate	1.082		uracil	1.121	1.143
4-acetamidobutanoate	1.079		X - 17353	1.120	1.350
gamma-glutamylisoleucine	1.072		X - 17185	1.109	
urea	1.071		X - 12722	1.093	
1-eicosapentaenoylglycerophosphoethanolamine	1.071		5-hydroxyhexanoate	1.090	
dihomo-linolenate (20:3n3 or n6)	1.059		4-vinylphenol sulfate	1.081	
glutamate	1.057		hydroquinone sulfate	1.073	1.085
N1-Methyl-2-pyridone-5-carboxamide	1.045		X - 12511	1.045	
oleate (18:1n9)	1.043		N-acetylhistidine	1.043	
3-(4-hydroxyphenyl)lactate	1.034		X - 17348	1.034	1.235
1-palmitoylglycerophosphoinositol	1.030		phenylcarnitine	1.033	1.286
cortisone	1.021	1.015	X - 17303	1.028	
X - 12095	1.018		X - 17320	1.026	
myristate (14:0)	1.015		X - 12636	1.024	
4-androsten-3beta,17beta-diol disulfate (2)	1.015	1.162	dihydroferulic acid	1.006	
15-methylpalmitate	1.009		X - 17453	1.000	
gamma-glutamylphenylalanine	1.007		sucrose		1.271
xanthine	1.004		creatine		1.270
X - 02249	1.000		3-methylhistidine		1.208
glycoursodeoxycholate		1.257	X - 11440		1.175
X - 11315		1.183	X - 12753		1.113
X - 17323		1.150	guanine		1.072
pyroglutamine		1.091	X - 17361		1.061
gamma-glutamyltyrosine		1.090	7,8-dihydroneopterin		1.058
palmitoylcarnitine		1.065	X - 13709		1.045
X - 17612		1.029	N-acetyl-beta-alanine		1.023
oleoylcarnitine		1.004	1,7-dimethylurate		1.016

3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF), dehydroisoandrosterone sulfate (DHEA-S), 15-methylpalmitate (isobar with 2-methylpalmitate).

Supplemental Table 2 and **3** are presented in an extra excel file (Supplemental tables.xlsx).

Supplemental methods

Metabolomics Measurements

Non-targeted metabolomics analysis for metabolic profiling was conducted at the Genome Analysis Center, Helmholtz Zentrum München. Two separate LC-MS/MS analytical methods mere used as previously published, i.e. in positive and in negative ionization modes, were used to detect a broad metabolite panel (Evans et al. 2009 19624122). In this study, samples were divided into two sets according to the biological matrices of the samples, i.e. plasma and urine. On the day of extraction, samples were thawed on ice. A 100µL of the sample were pipetted into a 2mL 96-well plate. In addition to study samples, a human pooled reference plasma sample (Seralab, West Sussex, United Kingdom) and another pooled reference matrix of each sample set (Seralab, West Sussex, United Kingdom) were extracted and placed in 1 and 6 wells, respectively, of the 96-well plate. These samples served as technical replicates throughout the data set to assess process variability. Beside those samples, 100µL of water was extracted as samples and placed in 6 wells of the 96-well plate to serve as process blanks. Protein was precipitated and the metabolites were extracted with 475µL methanol, containing four recovery standards to monitor the extraction efficiency. After centrifugation, the supernatant was split into 4 aliquots of 100µL each onto two 96-well microplates. The first 2 aliquots were used for LC-MS/MS analysis in positive and negative electrospray ionization mode. Two further aliquots were kept as a reserve. The extracts were dried on a TurboVap 96 (Zymark, Sotax, Lörrach, Germany). Prior to LC-MS/MS in positive ion mode, the samples were reconstituted with 0.1% formic acid (50µl for plasma, 100µl for urine). Whereas samples analyzed in negative ion mode were reconstituted with 6.5mM ammonium bicarbonate (50µl for plasma, 100µl for urine), pH 8.0. Reconstitution solvents for both ionization modes contained internal standards that allowed monitoring of instrument performance and also served as retention reference markers. To minimize human error, liquid handling was performed on a Hamilton Microlab STAR robot (Hamilton Bonaduz AG, Bonaduz, Switzerland). LC-MS/MS analysis was performed on a linear ion trap LTQ XL mass spectrometer (Thermo Fisher Scientific GmbH, Dreieich, Germany) coupled with a Waters Acquity UPLC system (Waters GmbH, Eschborn, Germany). Two separate columns (2.1 x 100 mm Waters BEH C18, 1.7 µm particle-size) were used either for acidic (solvent A: 0.1% formic acid in water, solvent B: 0.1% formic acid in methanol) and or for basic (A: 6.5mM ammonium bicarbonate, pH 8.0, B: 6.5mM ammonium bicarbonate in 95% methanol) mobile phase conditions, optimized for positive and negative electrospray ionization, respectively. After injection of the sample extracts, the columns were developed in a gradient of 99.5% A to 98% B over an 11 min run time at 350µLl/min flow rate. The eluent flow was directly run through the ESI source of the LTQ XL mass spectrometer. The mass spectrometer analysis alternated between MS and data-dependent MS/MS scans using dynamic exclusion and the scan range was from 80-1000 m/z. Metabolites were identified by Metabolon, Inc. from the LC-MS/MS data by automated multiparametric comparison with a proprietary library, containing retention times, m/z ratios, and related adduct/ fragment spectra

(Lawton et al. 2008 18384253). Identification criteria for the detected metabolites are described in Evans et al. (Evans et al. 2009 19624122). Quality control methods and normalization of metabolite levels are explained in detail in the supplement.

Metabolomics Measurements: Quality Control and Normalization of Metabolite Levels

To correct for daily variations of platform performance, the raw ion count of each metabolite was rescaled by the respective median value of the run day. Valid estimation of the median was ensured by keeping only metabolites with at least three measured values on more than the half of the run days. This procedure resulted in 475 and 558 metabolites for plasma and urine, respectively, available for the present analysis. 263 metabolites were measured in both bio fluids. We chose probabilistic quotient normalization (PQN) (Dieterle et al. 2006 16808434) to account for diurnal variation of urine samples, since this procedure was shown to be superior to the common creatinine scaling. For this purpose we calculated a mean-pseudo-spectrum depending on metabolites with measurements for all participants (131 urine metabolites). Subsequently, we calculated a dilution factor as the median quotient between the reference spectrum and each sample. Of note, urine creatinine and the estimated dilution factor were highly correlated (r=0.91, p<0.001) within the present study sample. Afterwards all metabolite levels were $\log_2$-transformed. Separately for plasma and urine samples we performed multivariate outlier detection using an algorithm proposed by Filzmoser *et al.* (Filzmoser et al. 2008) as implemented in the *pcout* function within the R package *mvoutlier*. The algorithm provides an outlier score for each sample based on a weighted combination of location and scatter estimations using principle component analysis and the Mahalanobis distance on a robustly scaled data matrix. The default parameters were used for the identification process, except the critical value for the location outliers was set to 4, as it corresponds to a 4 SD exclusion criteria. The minimum score was used as cut-off for outlier identification. As a result 13 and 8 samples from plasma and urine were excluded, respectively.

Gaussian graphical models (GGMs)

Briefly, GGMs are based on partial correlations, which represent the correlation between two metabolites correcting for all remaining metabolites. We additionally included age, sex and BMI as covariates to account for major confounding factors. Edges in the GGMs were declared significant if both partial and Pearson correlation were significant at $\alpha = 0.05$ after Bonferroni correction for all possible edges (correcting for $\binom{p}{2}$ tests, where p is the number of metabolites). Since GGM calculation requires a full data matrix, imputation of missing values was necessary. The influence of the imputation was minimized by exclusion of all metabolites with more than 20% missing values, resulting in 263 and 399 metabolites in plasma and urine, respectively. Assuming missing values mainly due to low concentrations of metabolites, the distribution of each metabolite on a run day could be estimated as a left-censored log-normal distribution prior normalization. Hence, we reconstructed

these distributions based on maximum likelihood estimation and sampled missing values from the censored part of the distribution. This procedure was only applied for metabolites with at least ten observations on the specific run day. Remaining missing values were imputed by multiple chained equations using the R-package '*mice*'. Both approaches are expected to rather lower (truncated sampling) or maintain (*mice*, using predictive mean matching) correlations between metabolites than falsely increasing them. Since each data set (plasma and urine) contained a ratio of observations to metabolites of about 2:1 we decided to use a shrinkage estimator based approach as implemented in the R-package '*GeneNet*' (Opgen-Rhein et al. 2007 17683609) to generate the GGMs following previous work (Do et al. 2015 25434815). The GGMs derived for plasma and urine were overlaid to visualize inter-fluid dependencies comprising 576 unique nodes and 681 edges. The network was visualized using the freely available software Cytoscape 3.2.1 (http://www.cytoscape.org/). Subsequently, results from linear regression analyses were mapped on the graph to visually inspect altered clusters of metabolites.

Variable Importance on PLS projection (VIP)

The VIP score summarizes the influence on the dependent variables ($_Y$, here IGF-I or IGF-I/IGFBP-3) of every predictor variable (X_k, here all plasma or urine metabolites and age) in a given PLS model. The VIP score for the variable X_k is defined as

$$VIP_{Ak} = \sqrt{\sum_{a=1}^{A} \left(w_{ak}^2 + * (SSY_{a-1} - SSY_a) * \frac{K}{(SSY_a - SSY_A)} \right)}$$

Where K is the number of predictors, A the number of total dimensions, $(w_{ak})^2$ the squared PLS weight of variable X_k for dimension a, ($SSY_{a-1} - SSY_a$) the explained sum of squares of the PLS dimension a and ($SSY_a - SSY_A$) the total explained sum of squares of the PLS model. The Sum of squares of all VIP's is equal to the number of terms in the model hence the average VIP is equal to 1. Terms with large VIP, larger than 1, are the most relevant for explaining Y.

Random Forest Regression

Predictive signatures for log(IGF-I) and the IGF-I/IGFBP-3 ratio were built on the following phenotypic characteristics using random forest regression: age, waist circumference, height, weight, hip circumference, glycated hemoglobin, fibrinogen, creatinine, glucose, cystatin C, estimated glomerular filtration rate, presence of hypertension or liver disease, smoking behavior, amount of alcohol consumption, physical activity, LDL-cholesterol, HDL-cholesterol, triglycerides, cholesterol, serum activities of gamma-glutamyl transpeptidase, alanine aminotransferase, aspartate aminotransferase and lipase. Feature selection and validation of the performance were assessed in a shuffled two-stage procedure implying the separation of 100 subjects each for independent validation. The inner loop consisted of a feature selection approach as the random forest (500 trees) was trained

with all variables and prediction was performed on the left data. Subsequently, the importance of each variable during this process was obtained based on the permutation of the out-of-bag data and weighted by the achieved root mean square error of prediction. The ten highest ranking variables were now used to build a new random forest which was validated on the unseen remaining subjects (outer loop). Once more weighted variable importance was obtained and finally the ten top ranking candidates were presented (Supplemental Figure 5). Both, the inner and outer loops were repeated ten times. The random forest was implemented in R using the package '*randomForest*'. To compile whether the selected variables were significant predictors of either log(IGF-I) or the IGF-I/IGFBP-3 ratio we employed Boruta feature selection (Kursa et al. 2010). Briefly, this procedure adds noisy variables to the feature matrix which are derived by randomly shuffling original variables. Importance of the variables is obtained using Z-scores derived from the loss of accuracy when omitting a variable. Finally, Z-scores of original features are compared with their noisy shadows to test for a significant contribution.

References

Dieterle, F., et al. (2006). "Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics." Anal Chem **78**(13): 4281-4290.

Do, K. T., et al. (2015). "Network-based approach for analyzing intra- and interfluid metabolite associations in human blood, urine, and saliva." J Proteome Res **14**(2): 1183-1194.

Evans, A. M., et al. (2009). "Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems." Anal Chem **81**(16): 6656-6667.

Filzmoser, P., et al. (2008). "Outlier identification in high dimensions." Computational Statistics and Data Analysis **52**(3): 1694-1711.

Kursa, M. B. and W. R. Rudnicki (2010). "Feature Selection with the Boruta Package." Journal of Statistical Software **36**(11).

Lawton, K. A., et al. (2008). "Analysis of the adult human plasma metabolome." Pharmacogenomics **9**(4): 383-397.

Opgen-Rhein, R. and K. Strimmer (2007). "From correlation to causation networks: a simple approximate learning algorithm and its application to high-dimensional plant gene expression data." BMC Syst Biol **1**: 37.