Reliability of Urinary Charged Metabolite Concentrations in a Large-Scale Cohort Study Using Capillary Electrophoresis-Mass Spectrometry

Yoshiki Ishibashi
Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan
Sei Harada
Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan
Ayano Takeuchi
Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan
Miho Iida
Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan
Ayako Kurihara
Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan
Suzuka Kato
Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan
Kazuyo Kuwabara
Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan
Aya Hirata
Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan
Takuma Shibuki
Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan
Tomonori Okamura
Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan
Daisuke Sugiyama
Faculty of Nursing And Medical Care, Keio University, Fujisawa, Kanagawa, Japan
Asako Sato
Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata, Japan
Kaori Amano
Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata, Japan
Akiyoshi Hirayama
Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata, Japan
Masahiro Sugimoto
Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata, Japan
Tomoyoshi Soga

Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata, Japan

Masaru Tomita

Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata, Japan

Toru Takebayashi (ttakebayashi@keio.jp)

Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan

---

Research Article

Keywords: Reliability of Urinary, Electrophoresis-Mass, Spectrometry, metabolome

Posted Date: November 25th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-110925/v1

Version of Record: A version of this preprint was published at Scientific Reports on April 1st, 2021. See the published version at https://doi.org/10.1038/s41598-021-86600-9.

Abstract

Currently, large-scale cohort studies for metabolome analysis have been launched globally. However, only a few studies have evaluated the reliability of urinary metabolome analysis. This study aimed to establish the reliability of urinary metabolomic profiling in cohort studies.

In the Tsuruoka Metabolomics Cohort Study, 123 charged metabolites were identified and routinely quantified using capillary electrophoresis-mass spectrometry (CE-MS). We evaluated approximately 750 quality control (QC) samples and 6,720 participants' spot urine samples. We calculated inter- and intra-batch coefficients of variation in the QC and participant samples and technical intraclass correlation coefficients (ICC). A correlation of metabolite concentrations between spot and 24-hour urine samples obtained from 32 sub-cohort participants was also evaluated.

The coefficient of variation (CV) was less than 20% for 87 metabolites (70.7%) and 20–30% for 19 metabolites (15.4%) in the QC samples. There was less than 20% inter-batch CV for 106 metabolites (86.2%). The 96 metabolites (78.0%) was above 0.75 for the estimated ICC. Among individuals, the Pearson correlation coefficient of 24-hours and spot urine was more than 70% for 59 of the 99 metabolites. These results show that the profiling of charged metabolites using CE-MS in morning spot human urine is suitable for epidemiological studies.

Introduction

The use of large-scale metabolomics for prospective epidemiological studies is becoming more common, and various metabolomics epidemiological studies, such as Cooperative Health Research in the Region of Augsburg (KORA), Twins UK registry, and American cohorts such as the Framingham Heart Study (FHS) Offspring, are being conducted [1-6]. Such large-scale metabolomics studies may allow for the prediction of chronic diseases such as Alzheimer's disease [7], cardiovascular disease [8], and chronic kidney disease (CKD) [9]. Thus, we initiated the Tsuruoka Metabolomics Cohort Study (TMCS) in Japan and have enrolled 11,002 participants since April 2012 [10-14]. In this study, capillary electrophoresis–mass spectrometry (CE-MS) was used to assess charged metabolites and liquid chromatography-mass spectrometry (LC-MS) for lipid metabolites.

In metabolomics studies, the use of urine as a sample has several advantages such as its ready availability, ease of obtaining, and lower complexity than other body fluids such as blood [15,16]. Metabolomic studies generally need to report reliability in measurement, as measurement accuracy is more variable than in routine tests [17-19]. We have already reported the reliability and accuracy of plasma metabolites in cohort studies through another study conducted by us [13]. However, there have been few studies conducted on the reliability of urine metabolite measurement accuracy.

Thus, in this study, we aimed to examine the reliability of large-scale urinary metabolomic profiling, assessed with CE-MS platform, using approximately 750 quality control (QC) samples and 6,720 participants' spot urine samples over a 56-month baseline period in a population-based cohort study. We

also compared the concentration of the metabolites in spot urine samples with that obtained in 24-hour collected urine samples for reference, using a sub-cohort.

Materials And Methods

Study population and sample collection

TMCS is a Japanese cohort study, initiated in April 2012 (Tsuruoka City, Yamagata Prefecture, Japan), involving 11,002 participants aged 35 to 74 years [10-14]. The participants were attendees of the annual municipal or workplace health checkup programs held at four sites in the city at baseline (April 2012–March 2015). The study design is illustrated in Fig1.

TMCS was particularly designed to discover metabolomics biomarkers for common diseases and disorders related to environmental and genetic factors. All participants completed a comprehensive questionnaire on lifestyle, dietary habits, and medical history. Biological samples including serum, plasma, urine, and deoxyribonucleic acid (DNA) were collected, and medical examination data recorded during the health checkup programs were also collected at recruitment. Information on alcohol consumption and smoking habits, dietary pattern, stress, and level of physical activity was collected through a standardized self-management questionnaire, and these data were verified in person. The procedures for recruitments of the TMCS have also been described in detail in previous studies [10-14]. To avoid variation due to fasting state and circadian rhythm, urine samples were collected from each participant in the morning between 8:30 am and 10:30 am after an overnight fast.

QC samples were prepared by mixing approximately 10 randomly selected participant samples, and analyzed every 10 runs in each batch to evaluate analytical validation. Finally, the QC samples were measured 752 times in 69 batches for cation analysis, and 768 times in 71 batches for anion analysis. Baseline participant samples were analyzed to evaluate the variance of participants' metabolites in the population. A total of 6,720 samples from included participants were analyzed from the TMCS baseline cohort: first and second year (April 2012–March 2014). Follow-up participants were not included in these samples.

To compare the metabolite concentrations in the spot urine samples with that in the 24-hour collected urine samples as a reference, a sub-cohort consisting of 32 TMCS participants within the cohort was set in 2013. These participants also answered a questionnaire on lifestyle, dietary habits, and medical history. In the sub-cohort analysis, ascorbic acid was added to the samples during pretreatment.

Sample preprocessing, instruments and analytical conditions

The samples were initially vortexed for 30 s, followed by centrifugation at 2,300 μ*g* for 5 min at 4 °C. They were then diluted according to the creatinine concentration (Table S1) with Milli-Q water and Milli-Q water containing internal standards (2 mM each of methionine sulfone and camphor-10-sulfonic acid). We

found that a creatinine concentration of < 10 mg/dl does not cause ion saturation in the mass spectrometer. Hence, we set this as the upper limit for diluting urine.

Mass spectrometry-based metabolomic profiling was performed with fasting urine samples using capillary electrophoresistime-of-flight mass spectrometry (CE-TOFMS). CE-TOFMS analysis of cationic and anionic metabolites was performed as described previously [20,21,22]. Briefly, cationic metabolites were separated on a fused-silica capillary column (50 μm i.d. × 100 cm total length) filled with 1 M formic acid as the electrolyte, and a methanol/water (50%, v/v) containing 0.01 μM hexakis(2,2-difluoroethoxy)phosphazene (Hexakis) was delivered as a sheath liquid at a rate of 10 μL/min. The capillary temperature was maintained at 20°C. The sample solution was injected at 5 kPa for 5 s, and a positive voltage of 30 kV was applied. ESI-TOFMS was conducted in the positive ion mode, and the capillary, fragmentor, skimmer, and Oct RF voltages were set at 4,000, 75, 50, and 500 V, respectively. The nebulizer gas pressure was configured at 7 psig and the heated nitrogen gas (300 °C) was supplied at a rate of 10 l/min. Anionic metabolites were separated using a commercially available COSMO(+) capillary (50 μm i.d. × 105 cm, Nacalai Tesque, Kyoto, Japan) filled with 50 mM ammonium acetate (pH 8.5) as the electrolyte, and ammonium acetate (5 mM) in 50% (v/v) methanol/water containing 0.01 μM Hexakis was delivered as sheath liquid at a rate of 10 μL/min. The sample solution was injected at 5 kPa for 30 s, and a negative voltage of 30 kV was applied. ESI-TOFMS was conducted in the negative ion mode, and the capillary, fragmentor, skimmer, and Oct RF voltages were set at 3,500, 100, 50, and 500 V, respectively. Other conditions were identical for the cationic metabolite analysis. In both modes, the automatic recalibration function was used to correct the analytical variation of exact masses for each run as described previously [22].

Mass spectra were acquired at a rate of 1.5 cycles/s over a 50-1,000 *m/z* range.

Statistical analysis

Since missing values were created by being below the measurement limit, half of the lowest detected values were imputed for metabolites that were not detected [23]. As we performed previously [13], inter- and intra-batch variance for each metabolite concentration in the QC samples was calculated to evaluate the reproducibility of the data. To control the effects of the batch, a linear mixed model was formulated, as shown in equation (1).

$$Y_i = \mu + B_i + \varepsilon_i \qquad (1)$$

The observed metabolite concentration (Y), inter - and intra-batch variance for each metabolite (μ), random effects common to each batch (B), and residual error (ε) are defined in the formula. We calculated the coefficient of variation (CV) by dividing the standard deviation as estimated from this model by the mean. Pearson correlation coefficients between the inter- and intra-batch CV were then calculated. These analyses were also conducted with participant samples to assess inter- and intra-batch variance. The intraclass correlation coefficient (ICC) was calculated to assess the reliability of the

metabolite concentrations [24,25]. This value was calculated from the variance of the measurement errors and the total variance,

$$ICC = 1 - \sigma_E^2/\sigma_T^2 \quad (2)$$

where σ_E^2 is the variance of the measurement errors and σ_T^2 is the total variance, as shown in equation (2). Although we could not calculate the ICC for participant samples as there were no replicates, we calculated technical errors from a large number of replicates for QC samples considered to be representative of the population samples. We made an approximate calculation of ICC, substituting the CV of QC samples for error variance and CV of participant samples for the total variance, as shown in equation (3).

$$\text{Approximate ICC} = 1-(\text{CV QC})^2 / (\text{CV participant})^2 \quad (3)$$

When creatinine correction was performed, it is well known that substances tend to be lower in concentrated urine samples than in diluted urine samples. Second, some diseases and medications can cause fluctuations in urine creatinine levels. Therefore, a sensitivity analysis was conducted excluding samples with creatinine values >3.0 g/g/L or <0.3 g / g/L from participants [26].

For sub-cohort analysis, metabolites in 24-hour urine samples and spot urine samples were compared among individuals, and Pearson's correlation coefficient was calculated for each individual. To account for major factors that may affect urinary metabolite concentrations, a regression analysis was performed, and the slope was compared between the spot and 24-hour urine samples. The explanatory variables included age, sex, alcohol consumption, and smoking, all of which are known to affect metabolite concentrations [10,27,28].

Statistical analyses were performed using R version 3.5.2 (2018-12-20) (R Core Team 2018, 2018, The R Foundation for Statistical Computing, Vienna, Austria).

Ethical approval

This study was approved by the Medical Ethics Committee of the School of Medicine, Keio University, Tokyo, Japan (Approval No 20110264 and No 20130207 for the entire cohort study and the sub-cohort one, respectively). Informed consent was obtained in written form from all the participants included in the studies. All research was performed in accordance with the relevant guidelines and regulations.

Results

CV for QC samples

Among the 123 metabolites, CV was less than 20% for 87 metabolites (70.7%), 20%–30% for 19 metabolites (15.4%), and more than 30% for 17 (13.8%) metabolites (Fig 2A). The median CV was 11.7%

and 17.8% for the cation and anion compounds, respectively. Table S2 shows the CV values of 123 metabolite concentrations in the QC samples. Inter-batch CV was ≦ 20% for 106 compounds (86.2%). Intra-batch CV was ≦ 20% for 103 compounds (83.7%) (Fig 2B). There were similar values between inter- and intra-batch CVs (medians, 6.9% and 9.0% for cations; 13.1% and 11.7% for anions), and they were also strongly correlated (Pearson's r = 0.89, N=123, P <0.001) (Fig S1).

CV in participant samples

Figs 3A and 3B show the distribution of the total, inter-, and intra-batch CVs among participants. The medians of total, inter-, and intra-batch CVs were 81.3%, 15.6%, and 79.9% for cations, respectively, and 61.9%, 19.6%, and 55.5% for anions, respectively. Predictably, the CV values of the participant samples were larger than those of the QC samples. There was a larger intra-batch CV than inter-batch CV in participant samples, in contrast to QC samples. Table S2 shows a statistical summary of 123 metabolite concentrations quantified in participant samples. Fig 4 shows the distribution of the estimated ICC. The estimated ICC was > 0.75 for 96 metabolites (78.0%), 0.40.0.75 for 18 metabolites (14.6%), and < 0.40 for 9 metabolites (7.3%) . The estimated ICC values are shown in Table S2 (metabolites of ICC < 0 are excluded in Fig 4). The results of the sensitivity analysis by exclusion based on the creatinine levels were similar to that seen in the overall participants (Table S3).

Comparison of spot and 24-hour urine samples

Among individuals, the Pearson's correlation coefficient for 24-hour and spot morning urine samples was > 0.7 for 59 (59.6%) of the 99 metabolites (Fig 5). The medians of the Pearson's correlation coefficients for 24-hour and spot morning urine samples were 0.75 and 0.80 for cations and anions, respectively. The association between 24-hour urine and spot urine samples was strong for the majority of the metabolites. A statistical summary of metabolites measured in the spot and 24-hour samples is shown in Table S4. The total CV distribution among metabolites is shown in Fig S2. The medians of the total CV for the spot and 24-hour urine samples were 37.0% and 41.9% for cations and 40.1% and 41.9% for anions, respectively. The regression coefficients of the basic demographic characteristics (age, sex, alcohol consumption, and smoking) were comparable between 24-hour urine and spot urine samples (Table S5) for most of the metabolites.

Discussion

In this large-scale epidemiological study, the reproducibility of 123 charged compounds in urine samples as quantified by CE-MS was reported to be good to high: QC CV for 106 compounds (86% of all) was less than 30%, while the measurement period was 56 months. Inter-batch CV for QC samples was less than 20% for 106 compounds (86.2%). A high CV was observed among the participants, and was caused by intra-batch CV rather than inter-batch CV. Therefore, the impact of inter- and intra-individual variability is likely to be greater than the impact of each measurement batch. In general, inter-batch CV values are often higher than the intra-batch values. However, as shown in our previous study [13], intra-batch CV was higher than inter-batch CV in the participants. This is evidence of the diversity of metabolites among the

participants, and the same phenomenon has been observed in urine. Metabolomic cohort studies have been used to measure urinary metabolites, [29,30] but the reliability of metabolite assessments and the variance between each batch and all of the participants have yet to be reported in most studies. In previous metabolomics studies, QC sample methods have been widely used to evaluate reproducibility [31]. Plasma QC CV, which has been reported previously, is lower than that of urine [13]. However, features with a QC CV < 20% are often considered to have good reproducibility, as recommended by the US Food and Drug Administration [32]. A QC CV < 30% is also considered acceptable [31,33].

This is the first study to report inter- and intra-batch CVs in urine samples from a cohort study. Reducing inter-batch variability is an important issue in large-scale metabolomics studies [13,34,35]. The results showed that a metabolomics assay of cations and anions effectively controlled the batch-to-batch effects of many measured compounds in this study. Our estimated ICC was greater than 0.40 for most measured metabolites, except for nine (N-acetylneuraminate, 2-oxoglutarate, cysteine S-sulfate, o-hydroxyhippurate, 4-hydroxy-3-methoxymandelate, 3-hydroxypropionate, 3-phosphoglycerate, malonate, and itaconate). Metabolites with a low reproducibility but an ICC > 0.40 may be of value as biomarkers, provided careful evaluation of their measurement errors is done [13]. The metabolites with low ICCs didn't belong to the same class or have low concentrations. It may be possible to obtain good ICCs for these metabolites in the future by changing the analytical conditions.

The metabolite concentrations in spot urine samples following overnight fasting conditions and 24-hour urine samples were comparable in majority of the metabolome, even after considering the basic demographic variables. Since 24-hour urine samples are known to be stable and aid in the quantification of metabolites [36,37], early morning fasting spot urine metabolomic measurements can be used as a surrogate index for 24-hour urine measurements; thus, they are suitable for large-scale biomarker studies.

Although recent studies have used blood metabolites to detect many diseases [1-6], studies using urinary metabolites are also on the rise [15,16] due to the ease of availability and collection of samples as well as its low complexity. In a nuclear magnetic resonance (NMR)-based metabolomics cohort study, CVs of 43 urinary metabolites in QC samples were reported to be low [17]; however, CE-MS allows us to quantify a larger number of metabolites at lower concentrations compared to NMR. In a cohort, the accuracy of measurement of metabolites is important for personalized medicine and disease onset prediction because the ability to measure a larger number of metabolites will increase the chance to explore new biomarkers for various diseases. For example, metabolites such as symmetric dimethylarginine, asymmetric dimethylarginine, and ethanolamine, which were detected using CE-MS in this study but not in NMR, not only had low CVs but also showed similar relationships between spot urine and 24-hour urine samples across the various groups. This indicates that these spot urine metabolites may be useful in predicting CKD and rheumatic disease, which are conditions where biomarkers such as these have already been reported [38, 39].

This study had some limitations. The CVs of some metabolites (17 metabolites) were more than 30% in the QC samples. Thus, further improvement in the quality of measurement is required for these

metabolites. Second, this population can be considered representative of the Japanese population with a homogenous genetic background; however, the diversity of environmental factors must also be taken into account. Therefore, further research is needed in Japan, and in other countries, to enable generalization of the findings. Although international comparative studies have been conducted in plasma samples, further studies with urine samples are needed. Since the genetic characteristics of the population are also not clear, newer multi-omics technologies, including genome-wide association studies, will be able to better predict disease development in the future.

In conclusion, this study showed that the CE-MS platform provides reliable values for urine metabolites, as assessed in a large-scale cohort study. CE-MS provides high-quality metabolomics data to help us understand the relationship between metabolites and disease risk.

Declarations

Acknowledgments

We thank the residents of Tsuruoka City for their interest in our study and the members of the Tsuruoka Metabolomic Cohort Study team for their commitment to the project.

This study was supported in part by research funds from the Yamagata Prefectural Government (http://www.pref.yamagata.jp/) and the city of Tsuruoka (https://www.city.tsuruoka.lg.jp/), Grant-in-Aid for Scientific Research (B) (grant numbers JP24390168 and JP15H04778), Grant-in-Aid for Challenging Exploratory Research (grant number 25670303), and Grant-in-Aid for Scientific Research on Innovative Areas (grant number JP 16H06277) from the Japan Society for the Promotion of Science (http://www.jsps.go.jp/). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Y.I contributed conceptualization, formal analysis, investigation, methodology, project administration, software, validation, visualization, writing – original draft and writing – review & editing. S.H contributed data curation, formal analysis, funding acquisition, investigation, methodology, project administration, software and writing – review & editing. A.T contributed methodology and software. M.I. contributed data curation and investigation. A.K. contributed data curation and investigation. S.K. contributed data curation and investigation. K.K. contributed data curation and investigation. A.H. contributed data curation, investigation and writing – review & editing. T.S. contributed data curation and investigation. T.O. contributed investigation and writing – review & editing. D.S. contributed investigation and writing – review & editing. A.S. contributed investigation and data curation. K.A. contributed investigation and data curation. A.H. contributed data curation, investigation, methodology, data curation and writing – review & editing. M.S. contributed data curation, methodology, resources and software. T.S. contributed funding acquisition, methodology, resources, supervision. M.T. contributed funding acquisition, methodology,

resources, supervision. TT. contributed conceptualization, data curation, funding acquisition, investigation, methodology, project administration, supervision, writing – review & editing.

All authors reviewed the manuscript.

Additional information

Competing interests

The author(s) declare no competing interests.

Data Availability

The most relevant data are within the paper. Raw data cannot be made publicly available, as study participants did not consent to have their information freely accessible. Based on these consents, the Ethics Committee for Tsuruoka Metabolomics Cohort Study inhibits any public data sharing because data contain potentially identifying or sensitive disease information. This committee includes representatives of Tsuruoka citizens, the administration of Tsuruoka City, a lawyer, and expert advisers. Data accession requests may be sent to the administration of the Ethics Committee for Tsuruoka Metabolomics Cohort Study. The data will be shared after a review of the purpose and with permission from the ethics committee. Data requests can be made to Sei Harada, seiharada@keio.jp. The source code and analysis generated during the current study are not publicly available, but are available from the corresponding author upon reasonable request.

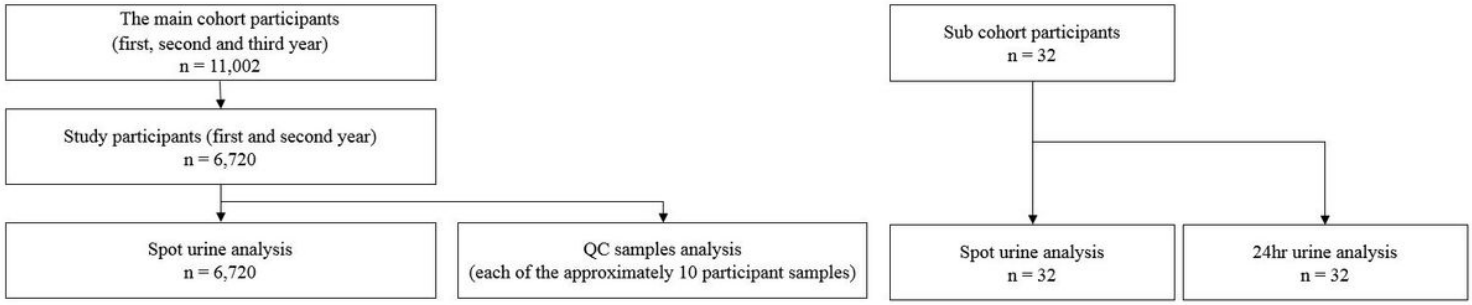
References

1. Suhre, K. et al. Human metabolic individuality in biomedical and pharmaceutical research. *Nature*. **477,** 54-60 (2011).
2. Sekula, P. et al. A metabolome-wide association study of kidney function and disease in the general population. *J. Am. Soc. Nephrol*. **27,** 1175-1188 (2016).
3. Long, T. et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nat. Genet*.**49,** 568-578 (2017).
4. Wang, T. et al. Metabolite profiles and the risk of developing diabetes. *Nat. Med*. **17,** 448-454 (2011).
5. Cheng, S. et al. metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation*. **125,** 2222-2231 (2012).
6. Iida, M., Harada, S. & Takebayashi, T. Application of metabolomics to epidemiological studies of atherosclerosis and cardiovascular disease. *J. Atheroscler. Thromb*. **26,** 747-757 (2019).
7. Wilkins, J. M. & Trushina, E. Application of metabolomics in Alzheimer's disease. *Front. Neurol*. **8,** 719 (2018).
8. Shah, S. H. et al. Baseline metabolomic profiles predict cardiovascular events in patients at risk for coronary artery disease. *Am. Heart J*. **163,** 844.e1-850.e1 (2012).

9. Kobayashi, T. et al. A metabolomics-based approach for predicting stages of chronic kidney disease. *Biochem. Biophys. Res. Commun*. **445,** 412-416 (2014).
10. Harada, S. et al. Metabolomic profiling reveals novel biomarkers of alcohol intake and alcohol-induced liver injury in community-dwelling men. *Environ. Health. Prev. Med*. **21,** 18-26 (2016).
11. Iida, M. et al. Profiling of plasma metabolites in postmenopausal women with metabolic syndrome. *Menopause*. **23,** 749-758 (2016).
12. Fukai, K. et al. Metabolic profiling of total physical activity and sedentary behavior in community-dwelling men. *PLoS One*. **11,** e0164877 (2016).
13. Harada, S. et al. Reliability of plasma polar metabolite concentrations in a large-scale cohort study using capillary electrophoresis-mass spectrometry. *PLoS One*. **13,** e0191230 (2018).
14. Sasaki, M. et al. Gender-specific association of early age-related macular degeneration with systemic and genetic factors in a Japanese population. *Sci. Rep*. **8,** 785 (2018).
15. Ryan, D., Robards, K., Prenzler, P. D. & Kendall, M. Recent and potential developments in the analysis of urine: a review. Anal. Chim. Acta. **684,** 8-20 (2011).
16. Zhang, A., Sun, H., Wu, X. & Wang, X. Urine metabolomics. *Clin. Chim. Acta*. **414,** 65-69 (2012).
17. Tynkkynen, T. et al. Proof of concept for quantitative urine NMR metabolomics pipeline for large-scale epidemiology and genetics. *Int. J. Epidemiol*. **48,** 978-993 (2019).
18. White, E. Measurement error in biomarkers: sources, assessment, and impact on studies. *IARC Sci. Publ*. **163,** 143-161 (2011).
19. Floegel, A. et al. Reliability of serum metabolite concentrations over a 4-month period using a targeted metabolomic approach. *PLoS One*. **6,** e21103 (2011).
20. Sugimoto, M., Wong, D. T., Hirayama, A., Soga, T. & Tomita, M. Capillary electrophoresis mass spectrometry-based saliva metabolomics identified oral, breast and pancreatic cancer-specific profiles. *Metabolomics*. **6,** 78-95 (2010).
21. Soga, T., Igarashi, K., Ito, C., Mizobuchi, K., Zimmermann, H. P. & Tomita, M. Metabolomic profiling of anionic metabolites by capillary electrophoresis mass spectrometry. *Anal. Chem*. **81,** 6165-6174 (2009).
22. Hirayama A, Sugimoto M, Suzuki A, Hatakeyama Y, Enomoto A, Harada S, Soga T, Tomita M, Takebayashi T. Effects of processing and storage conditions on charged metabolomic profiles in blood. *Electrophoresis*. 2015 Sep;36(18):2148-2155. doi: 10.1002/elps.201400600. Epub 2015 May 18. PMID: 25820922.
23. Hornung, R. W. & Reed, L. D. Estimation of average concentration in the presence of nondetectable values. Appl. Occup. Environ. Hyg. **5,** 46-51 (1990).
24. Townsend, M. K. et al. Reproducibility of metabolomic profiles among men and women in 2 large cohort studies. Clin. Chem. **59,** 1657-1667 (2013).
25. Sampson, J. N. et al. Metabolomics in epidemiology: sources of variability in metabolite measurements and implications. Cancer Epidemiol. Biomarkers Prev. **22,** 631-640 (2013).

26. ACGIH. Biological Exposure Indices (BEI®) Introduction. https://www.acgih.org/tlv-bei-guidelines/biological-exposure-indices-introduction (1998).
27. Rist, M. J. et al. Metabolite patterns predicting sex and age in participants of the Karlsruhe Metabolomics and Nutrition (KarMeN) study. PLoS One. **12,** e0183228 (2017).
28. Gu, F. et al. Cigarette smoking behaviour and blood metabolomics. Int. J. Epidemiol. **45,** 1421-1432 (2016).
29. Gil AM, Duarte D, Pinto J & Barros AS. Assessing Exposome Effects on Pregnancy through Urine Metabolomics of a Portuguese (Estarreja) Cohort. J Proteome Res. **17,** 1278–1289. (2018).
30. Playdon, M. C. et al. Comparing metabolite profiles of habitual diet in serum and urine. Am. J. Clin. Nutr. **104,** 776-789 (2016).
31. Dunn, W. B., Wilson, I. D., Nicholls, A. W. & Broadhurst, D. The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans. Bioanalysis. **4,** 2249-2264 (2012).
32. US FDA. Guidance for Industry, Bioanalytical Method Validation. https://www.fda.gov/downloads/Drugs/Guidance/ucm070107.pdf (2011) .
33. Brunius C, Shi L, Landberg R. Large-scale untargeted LC-MS metabolomics data correction using between-batch feature alignment and cluster-based within-batch signal intensity drift correction. Metabolomics. **12,** 173 (2016).
34. Kuligowski, J. et al. Detection of batch effects in liquid chromatography-mass spectrometry metabolomic data using guided principal component analysis. Talanta. **130,** 442-448 (2014).
35. Wehrens, R. et al. Improved batch correction in untargeted MS-based metabolomics. Metabolomics. **12,** 88 (2016).
36. Stella, C. et al. Susceptibility of human metabolic phenotypes to dietary modulation. J. Proteome Res. **5,** 2780-2788 (2006).
37. Rezzi, S. et al. Human metabolic phenotypes link directly to specific dietary preferences in healthy individuals. J. Proteome Res. **6,** 4469-4477 (2007).
38. Oliva-Damaso, E. et al. Asymmetric (ADMA) and symmetric (SDMA) dimethylarginines in chronic kidney disease: A clinical approach. Int. J. Mol. Sci. **20,** 3668 (2019).
39. Cuppen, B. V. et al. Exploring the inflammatory metabolomic profile to predict response to TNF-α inhibitors in rheumatoid arthritis. PLoS One. **11,** e0163087 (2016).

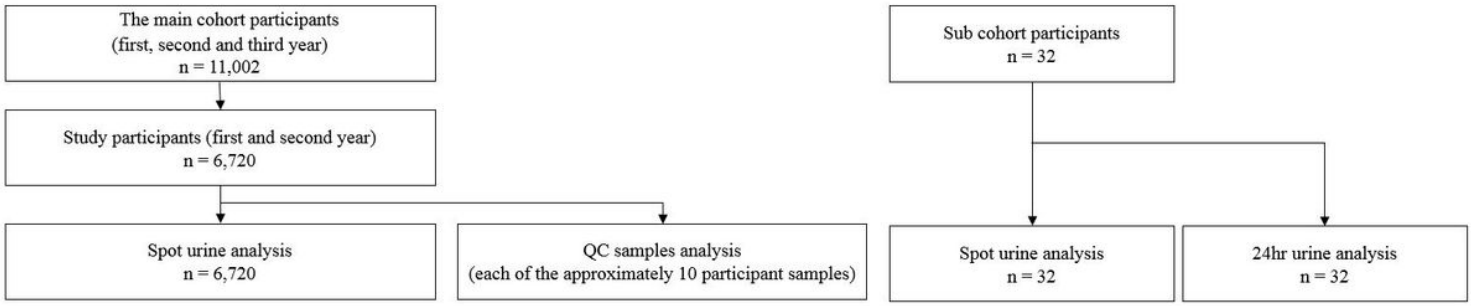
Figures



	Number of QC samples	Number of participants samples	Number of batches	Number of QC samples per batches (mean, (sd))	Number of participants samples per batches (mean, (sd))
Cation	752	6,720	69	10.90, (3.07)	97.39, (33.1)
Anion	768	6,720	71	10.82, (3.43)	94.65, (33.6)

Figure 1

Overview of Study design.



	Number of QC samples	Number of participants samples	Number of batches	Number of QC samples per batches (mean, (sd))	Number of participants samples per batches (mean, (sd))
Cation	752	6,720	69	10.90, (3.07)	97.39, (33.1)
Anion	768	6,720	71	10.82, (3.43)	94.65, (33.6)

Figure 1

Overview of Study design.

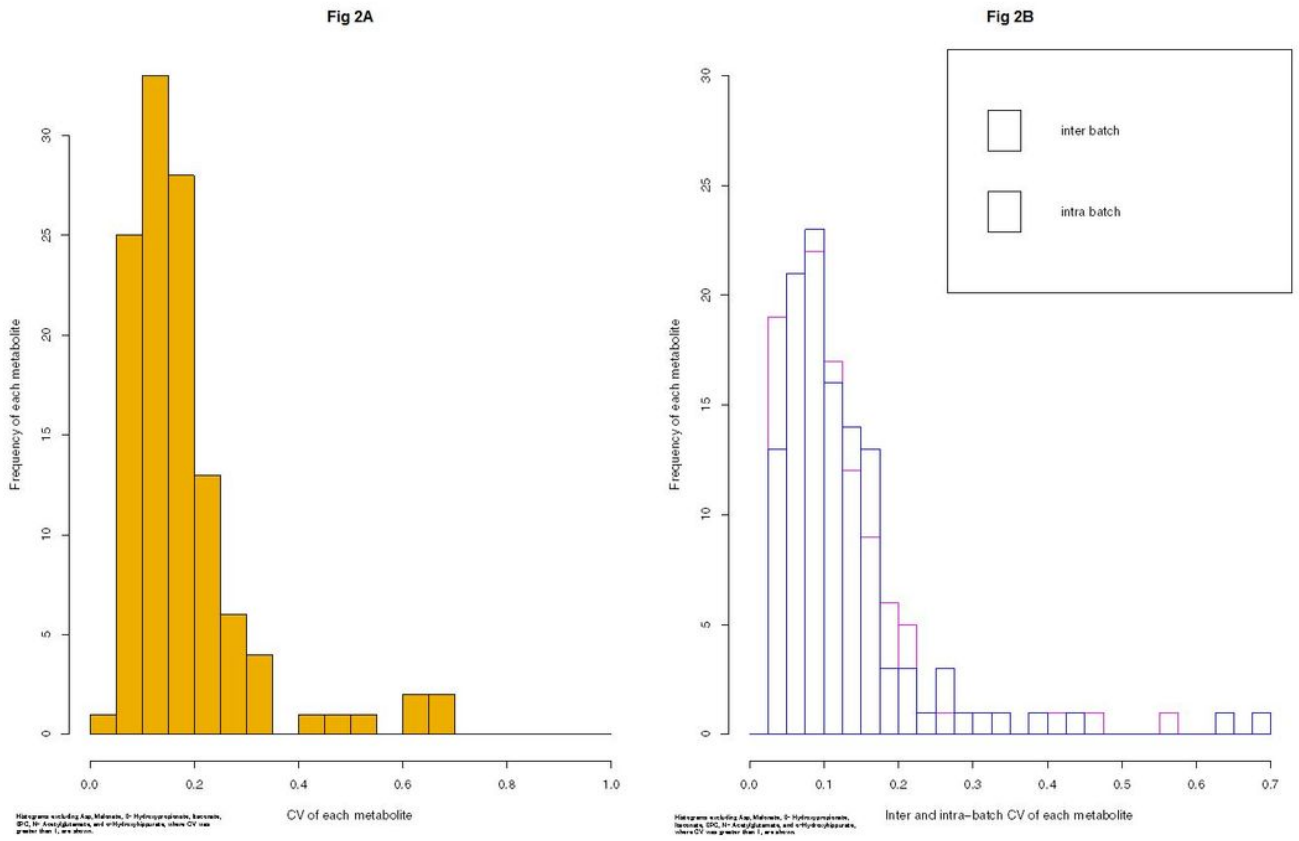


Figure 2

Histogram of CV for each metabolite in QC samples. A. Total CV of each metabolite. B. Inter and Intra-batch CV of each metabolite.

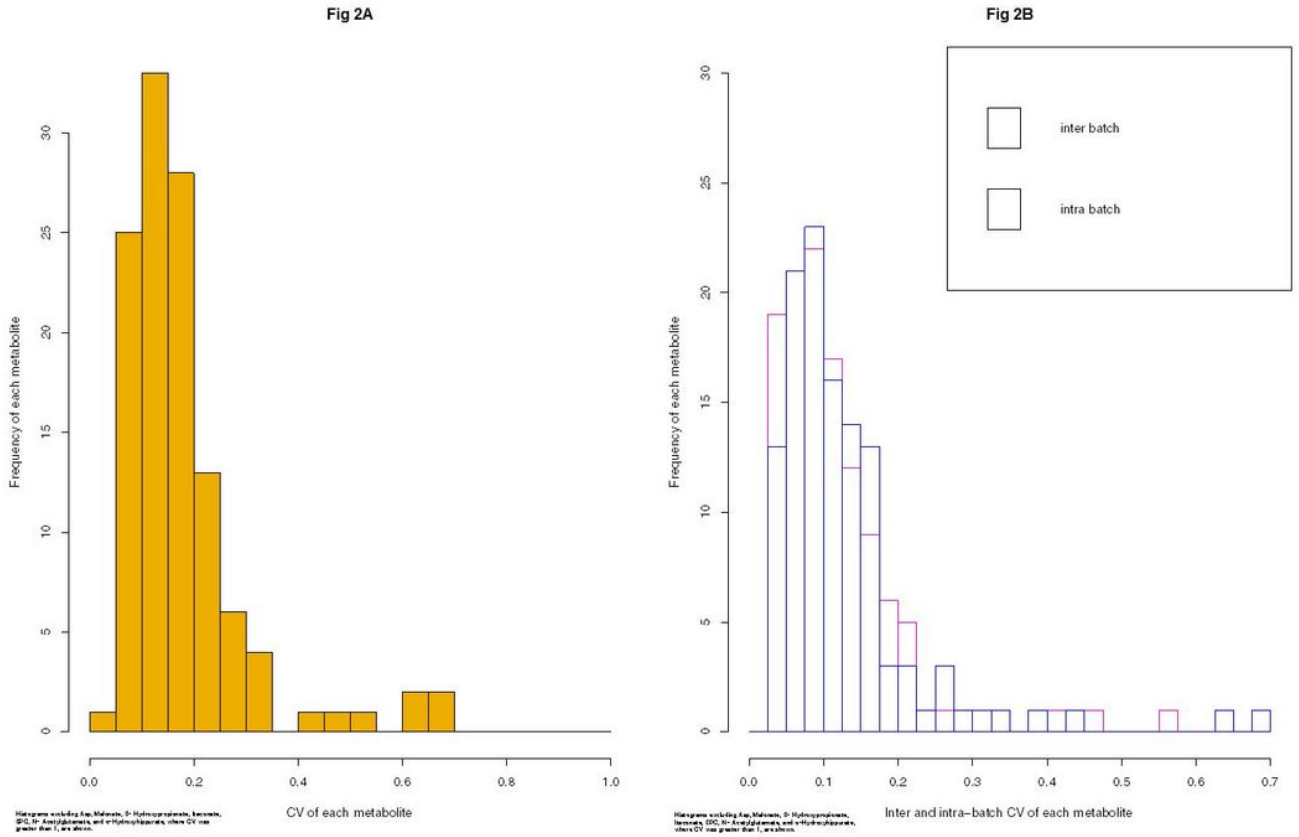


Figure 2

Histogram of CV for each metabolite in QC samples. A. Total CV of each metabolite. B. Inter and Intra-batch CV of each metabolite.

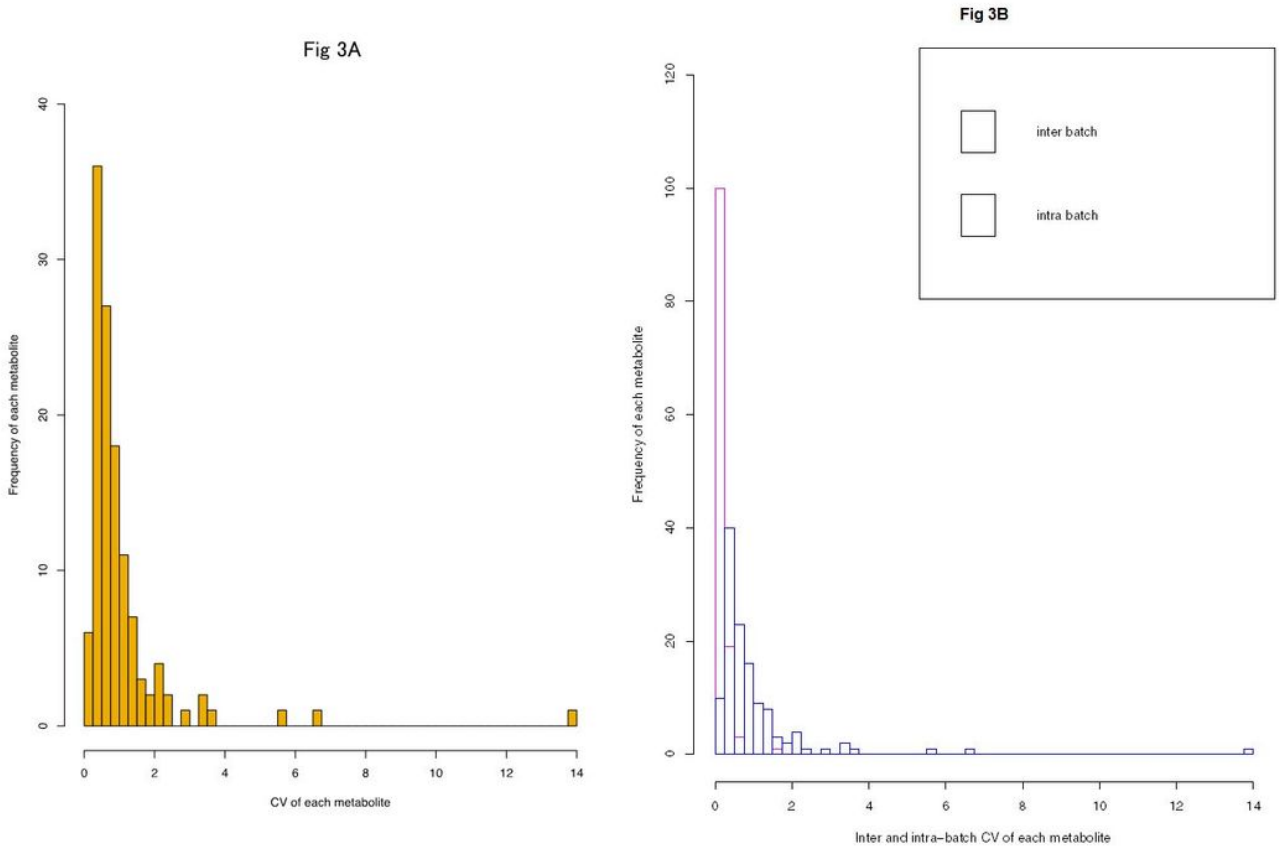


Figure 3

Histogram of CV for each metabolite in participants samples. A. Total CV of each metabolite. B. Inter and Intra-batch CV of each metabolite.

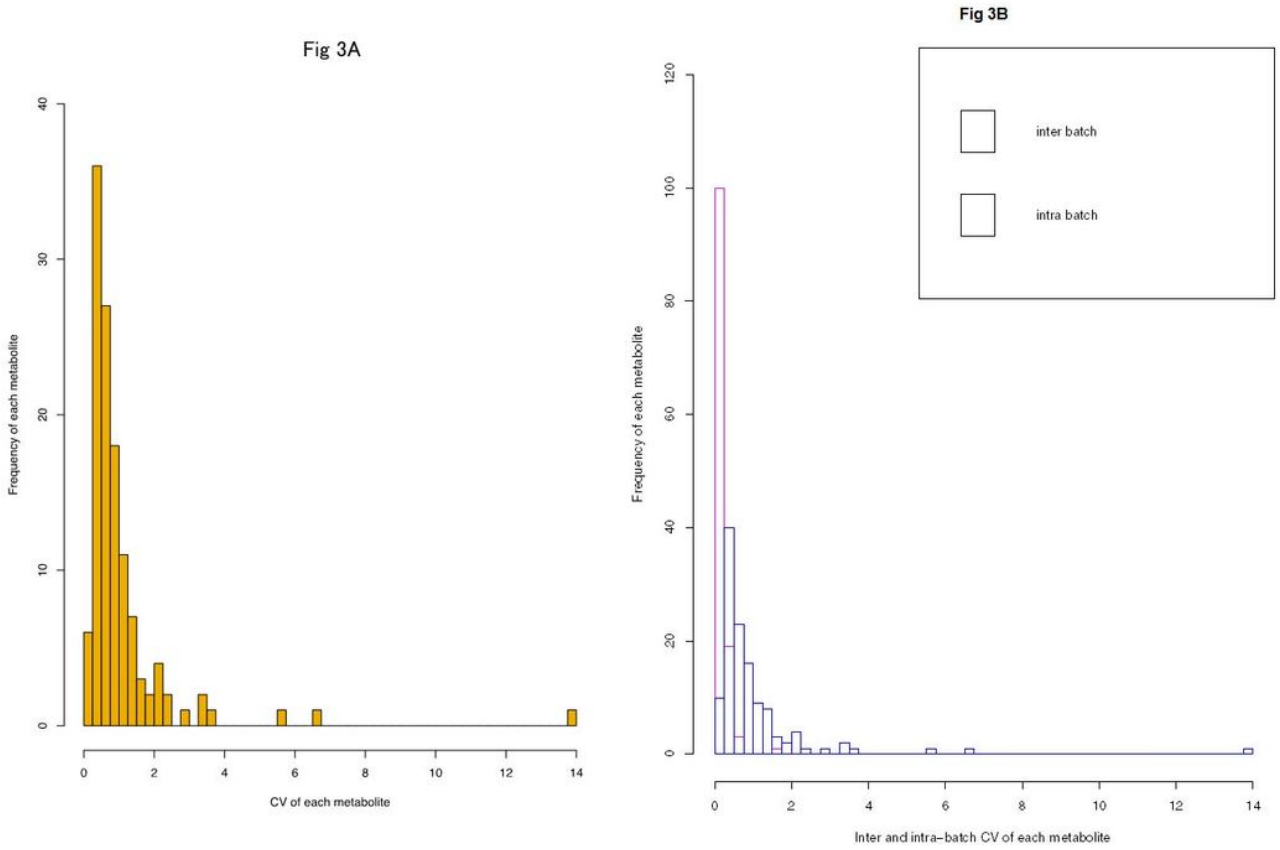


Figure 3

Histogram of CV for each metabolite in participants samples. A. Total CV of each metabolite. B. Inter and Intra-batch CV of each metabolite.

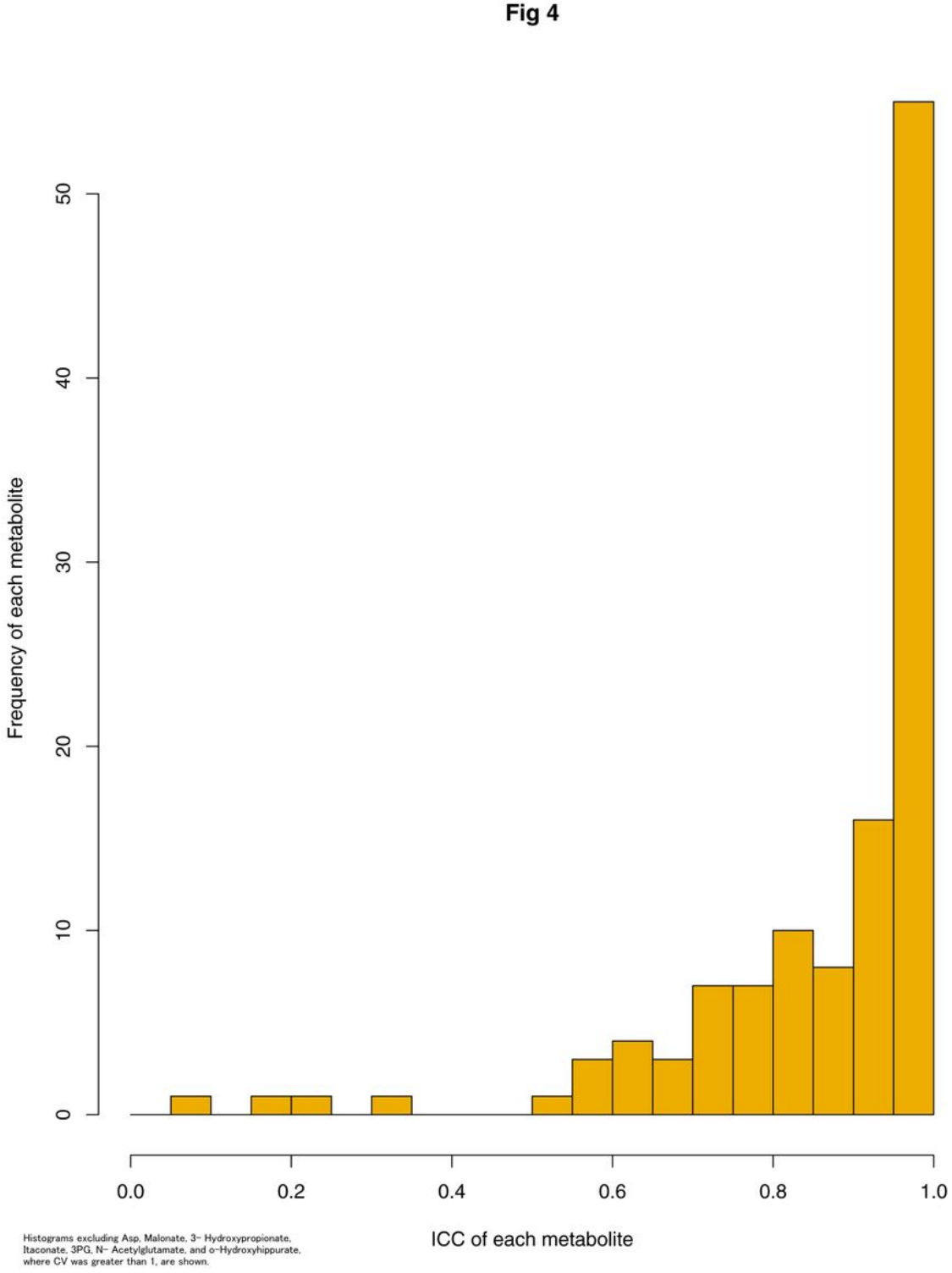


Figure 4

The distribution of the estimated ICC for each metabolite.

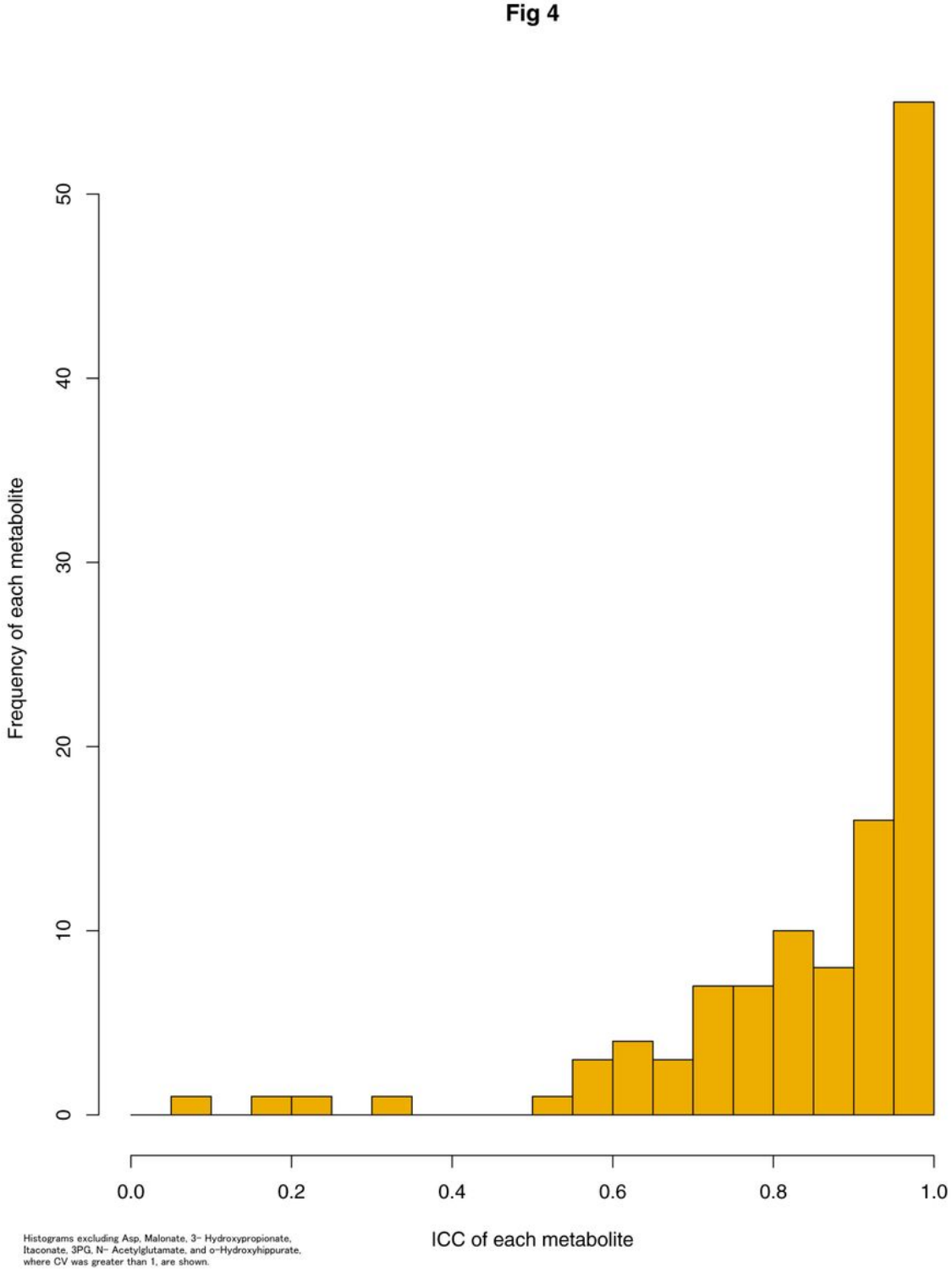


Figure 4

The distribution of the estimated ICC for each metabolite.

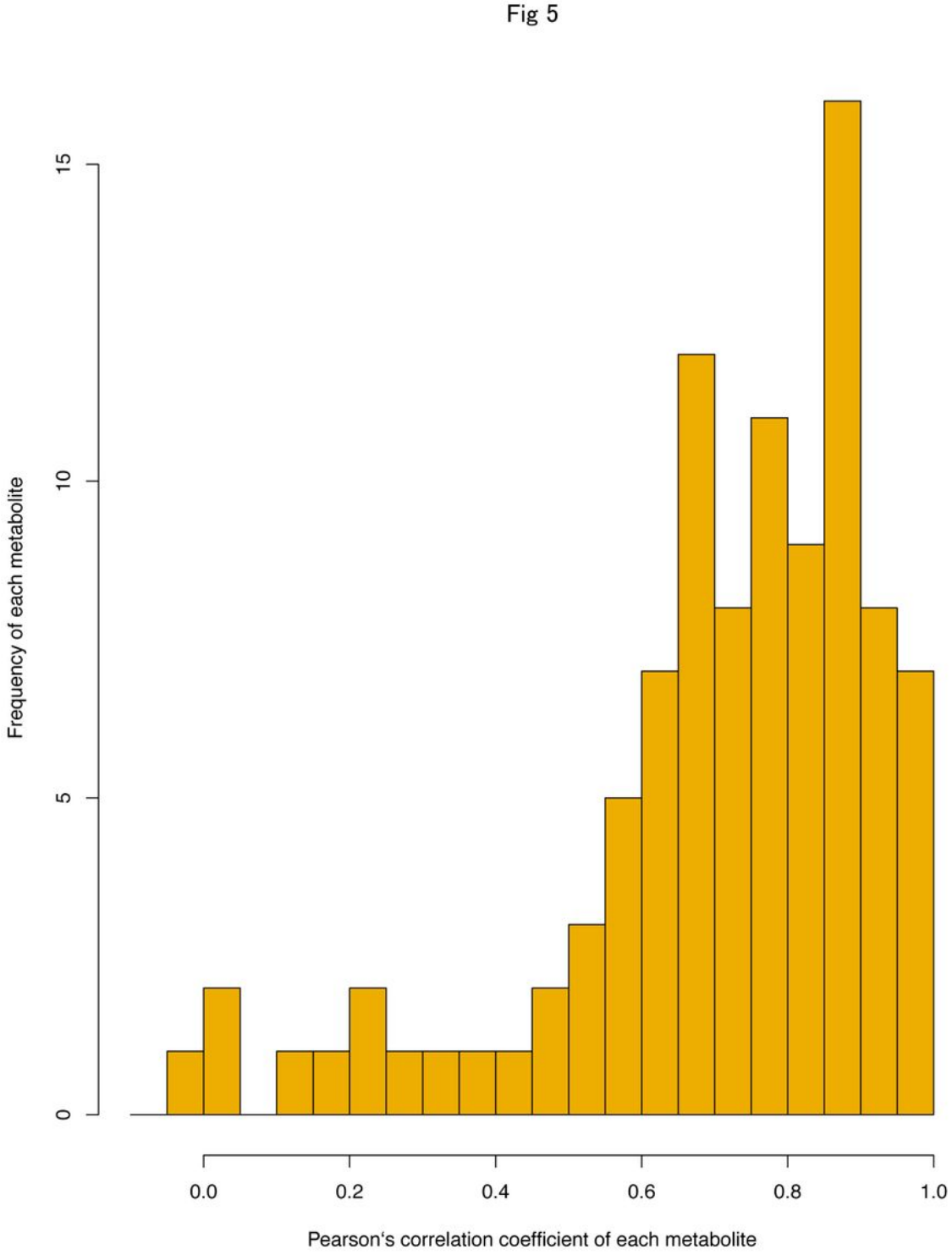


Figure 5

Pearson's correlation coefficient for 24-hour and spot morning urine samples.

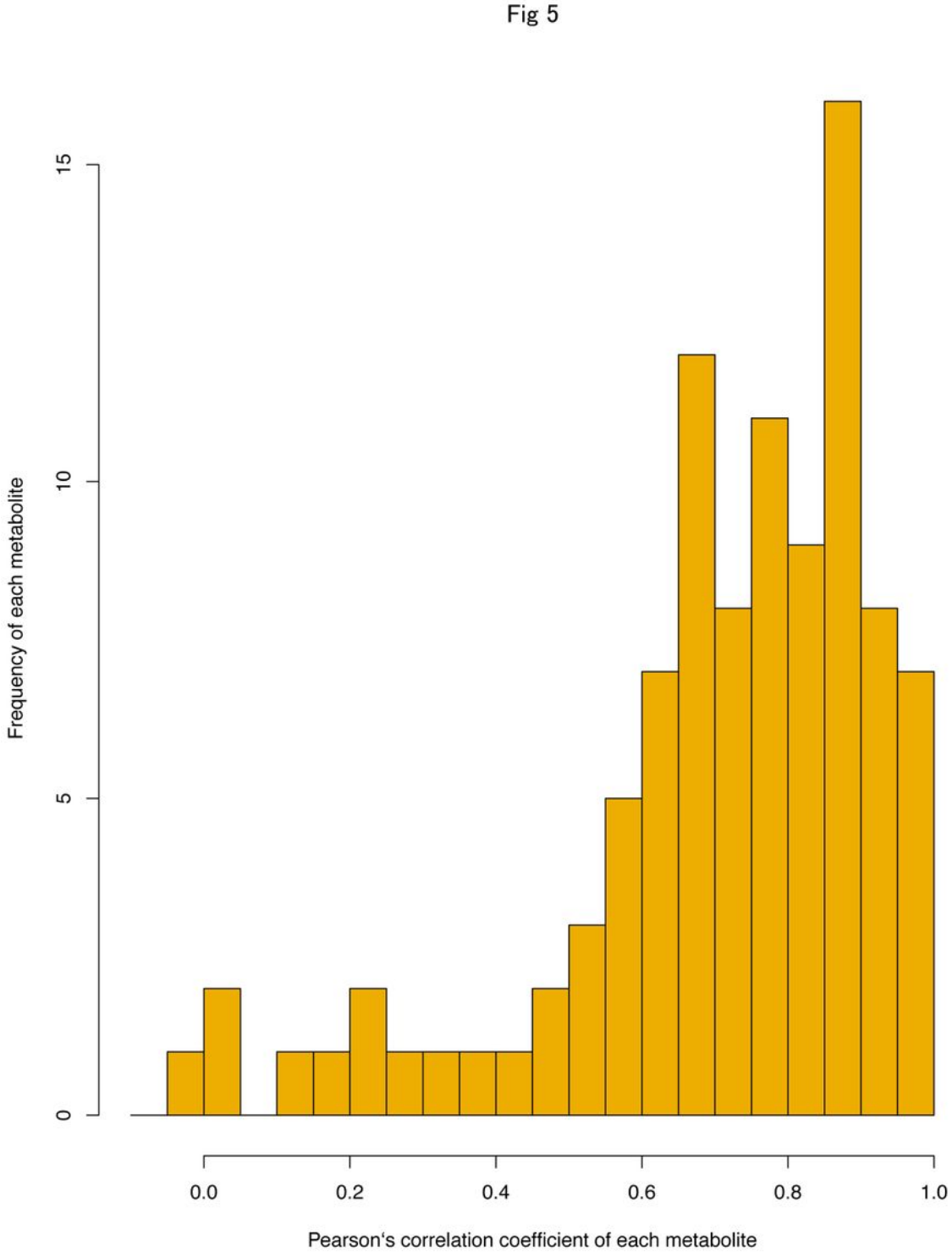


Figure 5

Pearson's correlation coefficient for 24-hour and spot morning urine samples.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS2.xlsx
- TableS2.xlsx
- Supplementarymaterials.pdf
- Supplementarymaterials.pdf
- TableS1.xlsx
- TableS1.xlsx
- TableS5.xlsx
- TableS5.xlsx
- TableS4.xlsx
- TableS4.xlsx
- TableS3.xlsx
- TableS3.xlsx
- SFig3.pdf
- SFig3.pdf
- SFig1.pdf
- SFig1.pdf
- SFig2.pdf
- SFig2.pdf