This is a post-print of the original article. Please cite as: Khamis, M.M., Holt, T., Awad, H. et al. (2018) Comparative analysis of creatinine and osmolality as urine normalization strategies in targeted metabolomics for the differential diagnosis of asthma and COPD, Metabolomics (2018) 14: 115. https://doi.org/10.1007/s11306-018-1418-9.

# Comparative analysis of creatinine and osmolality as urine normalization strategies in targeted metabolomics for the differential diagnosis of asthma and COPD

Mona Khamis[1], Teagan Holt[1], Hanan Awad[2], Anas, El-Aneed[1], Darryl J. Adamko[3]

1. College of Pharmacy and NutritionUniversity of SaskatchewanSaskatoonCanada
2.Calgary Laboratory ServicesAlberta Health ServicesCalgaryCanada
3.Department of Pediatrics, College of MedicineUniversity of SaskatchewanSaskatoonCanada

## Abstract

**Introduction** Urine is an ideal matrix for metabolomics investigation due to its non-invasive nature of collection and its rich metabolite content. Despite the advancements in mass spectrometry and $^{1}$H-NMR platforms in urine metabolomics, the statistical analysis of the generated data is challenged with the need to adjust for the hydration status of the person. Normalization to creatinine or osmolality values are the most adopted strategies, however, each technique has its challenges that can hinder its wide application.
**Objective** Assessment of whether the statistical model established using a targeted urine metabolomics dataset for the differential diagnosis of asthma and chronic obstructive pulmonary disease (COPD) would be improved by normalization to osmolality instead of creatinine.
**Methods** A metabolomics dataset from 51 patient urine samples previously analyzed using two liquid chromatography-tandem mass spectrometry methods was used. The data was normalized to creatinine and osmolality. The statistical analysis was achieved using partial least square discriminant analysis and models of separation were generated and compared.
**Results** Creatinine and osmolality values in asthma and COPD patients were strongly correlated. Using the same metabolites, we found that normalization to osmolality did not significantly change the results. The metabolites of importance in separation remained the same for both methods. The statistical strength of the creatinine model was somewhat better than the osmolality normalized model ($R^2Q^2$=0.919, 0.705 vs $R^2Q^2$ =0.929, 0.671).
**Conclusion** Our findings suggest that targeted urine metabolomics data can be normalized to creatinine or osmolality with no significant impact on the diagnostic accuracy of the model.

Key words: urine normalization, metabolomics, PLS-DA, creatinine, osmolality, targeted

1. Introduction

Metabolomics aims at identifying and quantifying metabolites that are altered in a biological system in response to illness or drug therapy (Khamis et al., 2015). Among the various biological samples, urine is appealing for metabolomic investigation. It does not demand extensive sample preparation steps with its lower protein content, and it is collected non-invasively in large volumes (Nobakht M. Gh et al., 2014, Fernández-Peralbo and Luque de Castro, 2012). Urine has been widely used in studying cancer (Wu et al., 2009, Issaq et al., 2008), hepatobiliary diseases (Wang et al., 2012), psychiatric disorders (Zheng et al., 2010) and respiratory illnesses (Saude et al., 2011, Saude et al., 2009, Adamko et al., 2015, Adamko et al., 2012). Using $^{1}$H-NMR, Adamko et al. (Adamko et al., 2015) has found that the urinary metabolomic profile of asthma patients is different from that of chronic obstructive pulmonary disease (COPD) patients. This difference was attributed to a set of metabolites identified through partial least square discriminate analysis (PLS-DA). PLS-DA is a supervised multivariate statistical tool widely used in metabolomics (Szymańska et al., 2012). It investigates the relationship between multiple predictor variables (metabolomics data) and dependent categorical variables (classes of participants) (Szymańska et al., 2012, Wold et al., 2001). If a statistically significant discrimination model can be found between classes, PLS-DA then uses the model parameters to predict subjects of unknown origin, given their metabolomics data (Szymańska et al., 2012, Wold et al., 2001).

One important challenge in urine metabolomic analysis is the variation in urine volume, which inversely affects metabolite concentration (Ryan et al., 2011, Wu and Li, 2016). Accordingly, prior to statistical comparison between participant groups, it is necessary to normalize the data to a common denominator to account for the hydration state of the individual (Warrack et al., 2009, Ryan et al., 2011). Creatinine (CRNN) has been widely used for normalization since its glomerular filtration is relatively constant under normal conditions (Warrack et al., 2009, Slupsky et al., 2007, Ryan et al., 2011, Wu and Li, 2016). However, there have been concerns with its validity since its excretion can vary with age, gender, race and muscle mass difference (Slupsky et al., 2007, Ryan et al., 2011). Osmolality (OSM) values can be also used for normalizing urine data. It is affected by the total number of solute particles in urine and can be used to reflect urine concentration (Chen et al., 2015, Fernández-Peralbo and Luque de Castro, 2012, Wu and Li, 2016). Unlike CRNN, OSM is not affected by age, gender, stress or diet (Warrack et al., 2009, Chen et al., 2015, Wu and Li, 2016). The widespread use of OSM is hindered by the absence of the necessary equipment in typical laboratories (Chadha et al., 2001, Wu and Li, 2016). In addition, insoluble particles and sample heterogeneity can affect the accuracy of the measurements (Chen et al., 2015, Wu and Li, 2016). While CRNN and OSM are the most widely accepted normalization strategies (Chen et al., 2015), other strategies have been proposed, such as dansylation and MSTUS (Wu and Li, 2016, Wu and Li, 2012). Each technique has its own limitations, as such, there is no universally accepted procedure for urine volume normalization. Moreover, comparative studies between all methods of normalization are not available (Fernández-Peralbo and Luque de Castro, 2012, Vogl et al., 2016). Comparison of normalization techniques using OSM and CRNN was reported in untargeted urine metabolomic platforms (Warrack et al., 2009, Chen et al., 2015, Vogl et al., 2016, Chetwynd et al., 2016, Kennedy et al., 2016) or with the use of unsupervised multivariate statistical analysis such principle component analysis (PCA) (Warrack et al., 2009, Chetwynd et al., 2016). In the scarce studies using targeted analysis, only few metabolites were investigated (less than 5) (Reid et al., 2012, Barber and Wallis, 1986). In general, results within available literature are conflicting, and are confounded with the recruitment of patients with established compromised kidney functions. Khamis et al. and Hanan et al. (Khamis et al., 2017, Awad et al., 2016) have developed

liquid chromatography (LC)- tandem mass spectrometric (MS/MS) methods for the absolute quantification of 32 potential urine biomarkers previously identified by $^{1}$H-NMR (Adamko et al., 2015). Differential isotope labeling was used to provide accurate absolute quantitative data along with the use fully validated methods as per regulatory agencies guidelines (Khamis et al., 2017, Awad et al., 2016, 2011, 2013). PLS-DA was used for the differentiation of diseases based on post-acquisition CRNN normalized metabolite values quantified by LC-MS/MS (Khamis et al., 2017, Awad et al., 2016). In this work, we hypothesized that the method of post-acquisition urine normalization, CRNN or OSM, might have an impact on the metabolomic results. Accordingly, we used metabolomic data from asthma and COPD patient urine samples analyzed using established methodologies (Khamis et al., 2017, Awad et al., 2016) to compare the accuracy of CRNN- and OSM- post acquisition normalized models. In addition, since the ultimate goal is to identify a smaller subset of biomarker metabolites, we investigated the effect of urine normalization on the biomarkers statistically identified as significant for group separation. This work represents the first comparison between CRNN- and OSM- post acquisition normalization strategies using a larger set of targeted metabolites (32 in total) that are quantified using differential isotope labeling-LC-MS/MS platform.

## 2. Experimental

### 2.1. Patients characteristics and sample collection

Fifty-one mid-stream urine samples from consenting patients were collected. Ethics board approval has been obtained (Bio#1389) (Khamis et al., 2017, Adamko et al., 2015). Subjects had a previous physician based diagnosis of asthma (n=25; female=14; mean age ±SD of 54.1±10.92) or COPD (n=26, female=13; mean age ±SD of 60.8± 9.30).

### 2.2. Creatinine measurement

Urinary CRNN concentration was determined using QuantiChrom™ CRNN assay kit (QC, CA) (Chen et al., 2015, Kit). The colored complex formed with picric acid in alkaline medium by Jaffe's reaction was detected at 510 nm using BioTek Synergy™ HT Multi-Mode Microplate Reader (VT, USA).

### 2.3. Osmolality measurement

Urine samples were vortexed for 10 sec and aliquots of 200 µL were transferred into new sample tubes. The OSM was measured using a freezing point depression osmometer (Model 3250, Advanced Instruments, MA, USA).

### 2.4. Statistical analysis

The concentrations of 32 metabolites were previously measured using two LC-MS-MS methods (Khamis et al., 2017, Awad et al., 2016). Patients' data was normalized (post acquisition) by dividing metabolite concentration (g/L) by CRNN (g/L) or OSM (Osm/Kg) values. Data was then exported to SIMCA® software (SIMCA-P 11, Umetrics, Sweden) for PLS-DA. Two statistical models of separation were created (32 metabolites/urine sample). Each model is represented by an $R^2$ value (multiple correlation coefficient), which explains the data variation, and by a $Q^2$ value (cross validated $R^2$), which predicts the data variation (Wold et al., 2001). In the CRNN model, 10 asthma and 10 COPD patients from CRNN-normalized metabolite values were used as a training set, while the remaining CRNN-normalized patients' data was withheld from modeling in order to be run blindly for diagnostic accuracy. In the OSM statistical model, the OSM-normalized dataset of the same 10 asthma and 10 COPD patients was used as a training test. Likewise, the remaining OSM-normalized patient's samples were withheld as a test set for accuracy assessment. Metabolites with consistently greater differences in concentration between asthma and COPD groups are displayed as a Variables of Importance Plot (VIP). The difference in each metabolite

for each group of patients is shown as a Coefficient of Variation (COV) plot. PLS-DA generates prediction scores based on the separation of data between subject groups. GraphPad Prism software (version 6; GraphPad Software, San Diego), was used for the generation of prediction score plots.

3. Results and discussion

CRNN and OSM values were measured in urine samples of 25 asthma and 26 COPD patients. Values of CRNN and OSM were not statistically different between asthma and COPD subjects (CRNN $p$=0.77 or OSM $p$=0.11; Fig. 1). The correlation of OSM and CRNN in each patient was also compared. There was a strong correlation between OSM and CRNN values (Spearman rank correlation coefficients (Mukaka, 2012) of 0.9006 and 0.8976 ($p$<0.001) in asthma and COPD patients, respectively) (Fig.2). This agrees with the literature, in which the relationship between CRNN and OSM is proportional to kidney functions (Vogl et al., 2016, Reid et al., 2012).

PLS-DA was performed using 32 metabolites normalized to CRNN or OSM in the asthma and COPD subjects and the prediction score plots were generated (Fig.3). The CRNN based PLS-DA model could separate asthma from COPD subjects with an $R^2Y$ value of 0.919 and a $Q^2$ value of 0.705. The same subjects from the OSM-normalized dataset were then used for the generation of the OSM model, which had an $R^2Y$ value of 0.929 and a $Q^2$ value of 0.671. As seen from the generated data, the $R^2$ values, which indicate how well the models explain the dataset (Szymańska et al., 2012, Wold et al., 2001), are similar between the CRNN and OSM models. These findings showed insignificant impact of the normalization approach on the statistical quality of the PLS-DA generated models. Similarly, $Q^2$ values also suggest the minimal difference in the predictive power between the CRNN based- and OSM based- PLS-DA models.

In addition, the effect of normalization between disease groups for each of the 32 metabolites was assessed. The method of urine normalization had no significant impact on the coefficient of variation values, being positive (COPD) or negative (Asthma) (Fig. 4). With the exception of one metabolite (tryptophan), all metabolites consistently increased or decreased between diseases regardless of the method of normalization (Fig.4). The marginal exception of tryptophan (Fig.4) is attributed to its larger than average variation between subjects and wide error bars. In addition, tryptophan variation was not a concern for the intended purpose as this amino acid was not a key metabolite used in disease separation based on the VIP plot of each model (data not shown).

The metabolites used in the model are ranked and displayed as a VIP. Metabolites with a VIP score >1 are more important to the strength of separation, and removing these metabolites will weaken a model. We investigated if the method of normalization might have an impact on the metabolites chosen as important in each PLS-DA model. Twelve metabolites had a VIP score>1 in CRNN model, while 11 metabolites had a VIP score>1 in OSM model. All but one metabolite were similar between the two models (data not shown). Accordingly, it can be conferred that CRNN or OSM normalization has no significant effect on the metabolites identified as significant between patient groups.

The last aim was to compare the predictive ability of each normalization technique for blinded sample analysis. CRNN- or OSM- normalized data of 15 asthma and 16 COPD patients were shown to the respective PLS-DA models without a diagnosis. The model generates a prediction score for the disease category in which the samples best fit. For the CRNN normalized dataset, the model was able to correctly diagnose 14 out of 15 (93% accuracy) asthma patients and 15 out of 16 COPD patients (93% accuracy). Similarly, when the OSM-normalized PLS-DA model was used to predict the diagnosis of the same blinded patient set using their OSM-normalized urine metabolomics data, the model demonstrated an accuracy of 87% (n=13 out of 15) and 75% (n =12

out of 16). Both models were reasonably good given the small sample size and thus the larger weight imposed by each participant on the overall accuracy %. The marginally improved prediction accuracy of the CRNN model might be explained by the stronger $Q^2$ value.

Similar to Kennedy *et al.* (Kennedy et al., 2016), we have demonstrated a consistency in biomarkers regardless of CRNN or OSM normalization. Our results are different than Vogl *et al.*(Vogl et al., 2016) or Warrack *et al.* (Warrack et al., 2009). Vogl *et al.*(Vogl et al., 2016) reported that OSM normalization resulted in the higher percentage of significant features in an untargeted urine metabolomics platform. It is important to note that the patient cohort had established chronic kidney disease, which would greatly affect creatinine measurements (Vogl et al., 2016). Warrack *et al.* (Warrack et al., 2009) used rats treated with doses of a lung/liver toxicant before PCA analysis, however, their kidney functions were not assessed prior to sample collection.

4. Conclusion

Overall, CRNN and OSM post acquisition normalized values were similar between asthma and COPD patients. The statistical parameters of the PLS-DA models did not differ significantly between the normalization strategies. As such, we suggest that either normalization strategy could be used for targeted urine metabolomics using differential isotope labeling if kidney disease is not a significant co-morbidity in the subject cohort.

**Acknowledgements**

The authors would like to thank Dr. Paul Lee, College of Medicine, for providing help with osmolality measurements

**Compliance with ethical standards**

**Conflict of interest**

The authors declare that they have no conflict of interest and no conflict of financial interest.

**Ethical approval**

Collection of human urine samples from consenting participants was approved by the University of Saskatchewan Biomedical Research Ethics Board and the University of Alberta and St. Joseph's Healthcare Hamilton Health Research Ethics Boards, (Bio#1389).

**Authors' contribution**

Ms. Holt did the experimental work, compiled the data and generated graphs using excel. Ms. Khamis prepared the manuscript and statistically analyzed the data using SIMCA. Dr. El-Aneed supervised Ms. Holt and Ms. Khamis, facilitated the equipment for measurements, revised the manuscript and proposed reviewers for the submission. Dr. Adamko conceived the idea of the work, contributed to the writing and revision of the manuscript, supervised Ms. Holt and Ms. Khamis and statistically analyzed the data using graph prims.

## 1. References


European Medicines Aagency (EMA), Committee for Medicinal Products for Human Use (CHMP), Guidelines on bioanalytical method validation. 2015

Food and Drug Administration, FDA Guidance for Industry:Bioanalytical Method Validation, DRAFT GUIDANCE. US Department of Health and Human Services. 2013. FDA, Center for Drug Evaluation and Research, Rockville, MD, USA. *https://www.fda.gov/downloads/drugs/guidances/ucm368107.pdf.*

ADAMKO, D. J., NAIR, P., MAYERS, I., TSUYUKI, R. T., REGUSH, S. & ROWE, B. H. 2015. Metabolomic profiling of asthma and chronic obstructive pulmonary disease: A pilot study differentiating diseases. *Journal of Allergy and Clinical Immunology,* 136**,** 571-580. e3.

ADAMKO, D. J., SYKES, B. D. & ROWE, B. H. 2012. The Metabolomics of AsthmaMetabolomics and AsthmaNovel Diagnostic Potential. *CHEST Journal,* 141**,** 1295-1302.

AWAD, H., ALLEN, K., ADAMKO, D. J. & EL-ANEED, A. Detection and quantification of 17 organic acid metabolites excreted in the urine of respiratory illness patients using a novel LC-MS/MS method. The 21st International Mass Spectrometry Conference (IMSC), 2016 Toronto, ON, Canada.

BARBER, T. & WALLIS, G. 1986. Correction of urinary mercury concentration by specific gravity, osmolality, and creatinine. *Journal of Occupational and Environmental Medicine,* 28**,** 354-359.

CHADHA, V., GARG, U. & ALON, U. S. 2001. Measurement of urinary concentration: a critical appraisal of methodologies. *Pediatric Nephrology,* 16**,** 374-382.

CHEN, G.-Y., LIAO, H.-W., TSENG, Y. J., TSAI, I.-L. & KUO, C.-H. 2015. A matrix-induced ion suppression method to normalize concentration in urinary metabolomics studies using flow injection analysis electrospray ionization mass spectrometry. *Analytica chimica acta,* 864**,** 21-29.

CHETWYND, A. J., ABDUL-SADA, A., HOLT, S. G. & HILL, E. M. 2016. Use of a pre-analysis osmolality normalisation method to correct for variable urine concentrations and for improved metabolomic analyses. *Journal of Chromatography A,* 1431**,** 103-110.

FERNÁNDEZ-PERALBO, M. & LUQUE DE CASTRO, M. 2012. Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis. *TrAC Trends in Analytical Chemistry,* 41**,** 75-85.

ISSAQ, H. J., NATIV, O., WAYBRIGHT, T., LUKE, B., VEENSTRA, T. D., ISSAQ, E. J., KRAVSTOV, A. & MULLERAD, M. 2008. Detection of bladder cancer in human urine by metabolomic profiling using high performance liquid chromatography/mass spectrometry. *The Journal of urology,* 179**,** 2422-2426.

KENNEDY, A. D., MILLER, M. J., BEEBE, K., WULFF, J. E., EVANS, A. M., MILLER, L. A., SUTTON, V. R., SUN, Q. & ELSEA, S. H. 2016. Metabolomic profiling of human urine as a screen for multiple inborn errors of metabolism. *Genetic testing and molecular biomarkers,* 20**,** 485-495.

KHAMIS, M. M., ADAMKO, D. J. & EL-ANEED, A. 2017. Development of a validated LC-MS/MS method for the quantification of 19 endogenous asthma/COPD potential urinary biomarkers. *Analytica Chimica Acta,* 989**,** 45-58.

KHAMIS, M. M., ADAMKO, D. J. & EL-ANEED, A. 2015. Mass spectrometric based approaches in urine metabolomics and biomarker discovery. *Mass spectrometry reviews*.

KIT, Q. C. A. https://www.bioassaysys.com/Creatinine-Assay-Kit.html.

MUKAKA, M. M. 2012. A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal,* 24**,** 69-71.

NOBAKHT M. GH, B. F., ALIANNEJAD, R., REZAEI-TAVIRANI, M., TAHERI, S. & OSKOUIE, A. A. 2014. The metabolomics of airway diseases, including COPD, asthma and cystic fibrosis. *Biomarkers***,** 1-12.

REID, C. N., STEVENSON, M., ABOGUNRIN, F., RUDDOCK, M. W., EMMERT-STREIB, F., LAMONT, J. V. & WILLIAMSON, K. E. 2012. Standardization of diagnostic biomarker concentrations in urine: the hematuria caveat. *PLoS One,* 7**,** e53354.

RYAN, D., ROBARDS, K., PRENZLER, P. & KENDALL, M. 2011. Recent and potential developments in the analysis of urine: a review. *Analytica chimica acta,* 684**,** 17-29.

SAUDE, E. J., OBIEFUNA, I. P., SOMORJAI, R. L., AJAMIAN, F., SKAPPAK, C., AHMAD, T., DOLENKO, B. K., SYKES, B. D., MOQBEL, R. & ADAMKO, D. J. 2009. Metabolomic biomarkers in a model of asthma exacerbation: urine nuclear magnetic resonance. *American journal of respiratory and critical care medicine,* 179**,** 25-34.

SAUDE, E. J., SKAPPAK, C. D., REGUSH, S., COOK, K., BEN-ZVI, A., BECKER, A., MOQBEL, R., SYKES, B. D., ROWE, B. H. & ADAMKO, D. J. 2011. Metabolomic profiling of asthma: diagnostic utility of urine nuclear magnetic resonance spectroscopy. *Journal of Allergy and Clinical Immunology,* 127**,** 757-764. e6.

SLUPSKY, C. M., RANKIN, K. N., WAGNER, J., FU, H., CHANG, D., WELJIE, A. M., SAUDE, E. J., LIX, B., ADAMKO, D. J. & SHAH, S. 2007. Investigations of the effects of gender, diurnal variation, and age in human urinary metabolomic profiles. *Analytical chemistry,* 79**,** 6995-7004.

SZYMAŃSKA, E., SACCENTI, E., SMILDE, A. K. & WESTERHUIS, J. A. 2012. Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics,* 8**,** 3-16.

VOGL, F. C., MEHRL, S., HEIZINGER, L., SCHLECHT, I., ZACHARIAS, H. U., ELLMANN, L., NÜRNBERGER, N., GRONWALD, W., LEITZMANN, M. F. & ROSSERT, J. 2016. Evaluation of dilution and normalization strategies to correct for urinary output in HPLC-HRTOFMS metabolomics. *Analytical and bioanalytical chemistry,* 408**,** 8483-8493.

WANG, X., ZHANG, A., HAN, Y., WANG, P., SUN, H., SONG, G., DONG, T., YUAN, Y., YUAN, X. & ZHANG, M. 2012. Urine metabolomics analysis for biomarker discovery and detection of jaundice syndrome in patients with liver disease. *Molecular & Cellular Proteomics,* 11**,** 370-380.

WARRACK, B. M., HNATYSHYN, S., OTT, K.-H., REILY, M. D., SANDERS, M., ZHANG, H. & DREXLER, D. M. 2009. Normalization strategies for metabonomic analysis of urine samples. *Journal of Chromatography B,* 877**,** 547-552.

WOLD, S., SJÖSTRÖM, M. & ERIKSSON, L. 2001. PLS-regression: a basic tool of chemometrics. *Chemometrics and intelligent laboratory systems,* 58**,** 109-130.

WU, H., XUE, R., DONG, L., LIU, T., DENG, C., ZENG, H. & SHEN, X. 2009. Metabolomic profiling of human urine in hepatocellular carcinoma patients using gas chromatography/mass spectrometry. *Analytica chimica acta,* 648**,** 98-104.

WU, Y. & LI, L. 2012. Determination of Total Concentration of Chemically Labeled Metabolites as a Means of Metabolome Sample Normalization and Sample Loading Optimization in Mass Spectrometry-Based Metabolomics. *Analytical chemistry,* 84**,** 10723-10731.

WU, Y. & LI, L. 2016. Sample normalization methods in quantitative metabolomics. *Journal of Chromatography A,* 1430**,** 80-95.

ZHENG, S., YU, M., LU, X., HUO, T., GE, L., YANG, J., WU, C. & LI, F. 2010. Urinary metabonomic study on biochemical changes in chronic unpredictable mild stress model of depression. *Clinica Chimica Acta,* 411**,** 204-209.