This is the accepted manuscript of the article "Urine headspace analysis with field asymmetric ion mobility spectrometry for detection of chronic kidney disease", E. Jokiniitty, L. Hokkinen, P. Kumpulainen, Y. Leskinen, T. Lehtimäki, N. Oksala, A. Roine, which has been published in *Biomarkers in Medicine*, 2020, 14(8), 629-638. https://doi.org/10.2217/bmm-2020-0085

Urine headspace analysis with FAIMS for detection of chronic kidney disease

Short running title: FAIMS estimation of renal function

Author names:
Elina Jokiniitty[1*], Lauri Hokkinen[1], Pekka Kumpulainen[4], Yrjö Leskinen[2], Terho Lehtimäki[1,3], Niku Oksala[1,4], Antti Roine[1]

Author affiliations:
1. Department of Surgery, Faculty of Medicine and Health Technology, Tampere University, Finland
2. Department of Internal Medicine, Tampere University Hospital
3. Department of Clinical Chemistry, Fimlab Laboratories, and Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University
4. Centre for Vascular Surgery and Interventional Radiology, Tampere University Hospital, Finland

***Author for correspondence:** Elina Jokiniitty (Elina.jokiniitty@tuni.fi)

Author Contributions:
Elina Jokiniitty: study planning, data analysis, manuscript drafting, manuscript review
Lauri Hokkinen: study planning, data analysis, manuscript drafting, manuscript review
Pekka Kumpulainen: data analysis, manuscript dragting, manuscript review
Yrjö Leskinen: collection of the material, manuscript drafting, manuscript review
Terho Lehtimäki: study planning, manuscript drafting, manuscript review
Niku Oksala: study planning, funding, data analysis, manuscript drafting, manuscript review
Antti Roine: study planning, funding, data analysis, manuscript drafting, manuscript review

Acknowledgements: The authors want to thank medical laboratory technologist Ulla Kala for assistance in the laboratory analysis phase of the study.

Financial disclosure: This study has been financially supported by the Academy of Finland, grants 322098, 286284 (for T.L.); Competitive State Research Financing of the Expert Responsibility area of Tampere University Hospitals, grant X51001 (for T.L.) and grants 9s045, 151B03, 9T044, 9U042, 150618, 9V044, 9X040, 9AA057; Finnish Foundation for Cardiovascular Research; Finnish Cultural Foundation; Tampere Tuberculosis Foundation; Emil Aaltonen Foundation; Yrjö Johansson Foundation; Signe and Ane Gyllenberg Foundation; Diabetes Research Foundation of Finnish Diabetes Association; and Tampere University Supporting Foundation and Finnish Foundation for Technology promotion.

Niku Oksala and Antti Roine are shareholders of Olfactomics LtD., which commercializes technology for the detection of diseases by ion mobility spectrometry.

Information pertaining to writing assistance: N/A

Ethical disclosure: The study was approved by the Tampere University Hospital Ethical Committee (code 97081) and Tampere City Ethical Committee for the control subjects. All the study subjects gave an informed consent before their examination and biological sample collection.

Data sharing statement: N/A

Introduction:

Electronic noses (eNoses) are an emerging class of experimental diagnostic tools. They are based on the detection of volatile organic compounds (VOCs). Urine is used as sample medium in several publications but neither the effect of CKD on the analysis nor the potential to detect chronic kidney disease (CKD) has been explored.

Methods:

We utilized an eNose based on field asymmetric ion mobility spectrometry (FAIMS) technology to classify urine samples from CKD patients and controls.

Results:

We were able to differentiate extremes of kidney function with an accuracy of 81.4 %.

Conclusion:

In this preliminary study, applying eNose technology we were able to distinguish the patients with impaired kidney function from those with normal kidney function.

Chronic kidney disease is a growing global problem. An interesting novelty in the disease diagnostics are devices called electronic noses, which detect chemical compounds produced by the body metabolism and thus form a specific smell print of which diseases or other processes of the body can be detected. We analyzed urine of kidney patients with an electronic nose to discriminate those with a poor kidney function from the ones with a relatively normal kidney function. This study indicates that chronic kidney disease can in fact be smelled from urine samples.

Keywords: VOCs, chronic kidney disease, electronic nose, FAIMS, urine, diagnostics

Introduction

Chronic kidney disease (CKD) is a global health problem [1] and a significant risk factor for cardiovascular diseases, the leading cause of death in developed countries [2]. CKD has effects on all the other organs of the body and affects the treatment of comorbid diseases [2-4]. Early recognition of CKD allows more inexpensive and effective treatment, therefore the detection in primary health care should be made easy and inexpensive, possibly with non-invasive methods [5,6]. Estimation of glomelural giltration (GFR) is routinely used in the initial assessment of CKD [8].

While measuring GFR with inulin clearance is the gold standard, the complexity of the method makes it unfeasible in clinical practice. Therefore, GFR is nowadays estimated using serum creatinine –based formulas, commonly CKD-EPI equation with serum creatinine [4]. The equation is more accurate and less biased compared to the formerly used MDRD equation, especially at higher GFR levels and at the time being, also closest to the inulin measured GFR [7]. However, creatinine-based estimations can be confounded by muscle mass, diet and medication such as trimethoprim. In certain situations, the use of another biomarker, cystatin C can be utilized [8]. Other substances, such as beta2-microglobulin have also been under research for substitution for creatinine [9]. All current methods require an invasive blood sample.

The concept of olfactory detection of diseases has been demonstrated by a multitude of dog studies [10]. Olfactory sensation is caused by the adhesion of volatile organic compounds (VOC) to the olfactory epithelium. VOCs are a large group of compounds defined by their volatility at room pressure. In human body they are metabolites that are emitted through sweat, saliva, breath, milk, feces and urine. In addition to the physiological changes in metabolism, age, gender and diet, the human composition of VOCs is affected by diseases. [11]. An infection, cancer or a metabolic disease such as diabetes changes the smell of the body or the composition of VOCs emitted [12], as can be seen in patients with advanced renal failure having a urine-like smell in their breath, caused by organic amines [13]. Certain VOCs are typical for renal disease and uremia [14]. Pagonas et al. [15] were able to detect substances appearance of which correlated with deteriorating renal function in the exhaled breath of patients with a chronic renal dysfunction or end-stage renal disease (ERSD). These VOCs form the “uremic fingerprint”. Mochalski et al. identified a total of 60 VOCs in blood and breath of uremic dialysis patients, six of which were changed during dialysis.

VOCs related to uremia or CKD have not been analyzed from urine. The traditional method for identifying VOCs (in exhaled breath) has been gas chromatography – mass spectrometry (GC-MS). It has been used in several studies concerning VOCs in disease diagnostics, but the problem with this technique lies in the fact that the pathological metabolic pathways of the disease are often unknown. Therefore, the marker compounds are not always easily connected to a specific disease [16]. Furthermore, GC-MS procedure is relatively expensive requiring a vacuum and trained personnel to use it.

In attempt to mimic biological olfaction, devices called electronic noses (eNose) have been developed. Although the performance of dogs appears superior to the artificial devices [17], their performance depends on time-consuming training. In addition, dogs get tired, hungry, lose interest and can be distracted by external factors [18]. Machines perform concordantly and predictably, and their maintenance usually requires less resources than living creatures.

A typical eNose consists of a sampling system, an array of chemical gas sensors, an analog-digital convertor and a computer, which has an algorithm for classification of the samples [19]. The function of eNoses is not based on recognition of certain compounds or measuring their concentrations. They rely on the chemical fingerprint of the smell of the sample, which is constructed of the varying concentrations and composition of the VOCs in a sample. They recognize patterns and compare them to an algorithm in the previously constructed database. [16,20-22].

Several studies applying eNoses in disease diagnostics have been published with promising results [17]. The use of urine as a biological source of VOCs has been under research [23], yet to our knowledge, the only attempt to detect CKD from urine is the study of Di Natale et al [24]. They used an eNose based on quartz microbalance sensors to analyze urine samples from children with kidney disease and hematuria matched with healthy controls same age. Recently, in a proof-of-concept study [25], we applied an eNose with field asymmetric ion mobility spectrometry (FAIMS) technology to diagnose ovarian cancer from urine samples. Kidney function has not been taken into consideration in most eNose studies, yet we hypothesize that it is a considerable confounding factor.

In the current study we used an electronic nose with FAIMS -technology to analyze urine samples and stratify them according to the CKD-EPI classification (GFR class). The aim of the study was to find out if impaired kidney function can be detected from a urine sample based on detection of VOCs.

Materials and methods

Patient population and samples

In this study we used urine samples provided by patients participating in studies of CKD in relation to atherosclerosis and its complications in 1997 – 2000 in Tampere University Hospital and Tampere Municipal Hospital [26,27]. The cohort of the current study consisted of pre-dialysis patients with chronic renal failure aged under 70 years, and a matching (age, sex and BMI) control group from orthopedic outpatient clinic patients with no previously known chronic illnesses, representing normal creatinine levels and no albuminuria. Both groups were highly representative of their patient population in the area. All the participants had given a previous informed written consent for the use of their urine samples for biochemical research. The study was approved by the ethical committee of the University of Tampere.

At the time of the original study, the patients provided a 24-hour urine collection sample, and a 12-hour over-night fasting venous blood sample was also taken the same morning the urine collection was delivered, from which the serum creatinine levels were measured in a laboratory setting. Urine protein was measured from the 24-hour sample, and albuminuria was tested with a dipstick using the standard method of Tampere University Hospital laboratory testing. An amount of 20 ml of the urine was then stored in -70 °C in plastic test tubes without preservatives for future use. These samples were not thawed before the current study, which took place in 2017.

In the current study, we focused on the pre-dialysis patients and controls of the original study cohort. The patient population consisted of 108 patients, 13 of which were excluded from the study due to missing data, leaving a study population of 95 patients. The study population was stratified into five classes according to GFR (later: GFR classes 1-5). Patient demographics, including the

etiology of kidney disease, are shown in table 1. Due to the pilot nature of the study, the sample size was maximized based on feasibility, not statistical power based on pre-specified differences.

The stored urine samples were defrosted overnight, prior to the FAIMS analysis, in +4 °C and then mixed using vortex. An amount of 5 ml of each liquid urine sample was pipetted into a glass vial (Fisherbrand, Fisher Scientific, Finland) sized 24 or 28 ml for analysis. All glass vials were handled with clean nitrile gloves to prevent volatile contamination from skin. The samples were kept in room temperature for an average of 169 minutes (variation 104 – 219 minutes) before analysis.

GFR classes

A corresponding GFR –value for each patient was calculated using the CKD-EPI –formula [28] with the provided information of serum creatinine value, age and sex. After this the patients were stratified in GFR classes 1-5 (later: GFRc 1-5) according to their renal function (Table 1). The stratification was based on the clinical classification of CKD [3] as follows: GFR class 1 (GFRc 1) represents patients with eGFR > 90 ml/min per 1,73m2 (with no known kidney disease or albuminuria), GFRc 2 patients with eGFR 60 – 89, GFRc 3 patients with eGFR 30 – 59, GFRc 4 patients with eGFR 15 – 29 and GFRc 5 patients with eGFR < 15. GFR class 3 was originally to be divided in subgroups 3a and 3b, but after the exclusions, there was only one patient left in the group 3a, so 3a and 3b were merged into GFRc 3. Additionally, the presence of protein in the 24h urine sample and the results for a positive albumin dipstick test for each GFR class are also presented in the Table 1.

FAIMS analysis

We utilized an Owlstone Lonestar FAIMS chemical analyzer employing ATLAS headspace sampling system (Owlstone Ltd, Cambridge, UK) to analyze the urine samples. FAIMS is an ion mobility method based on accessing the physicochemical properties of VOCs. The liquid sample (urine) is heated to achieve a gaseous headspace which is subsequently ionized. The formed ions are referred to the sensor by carrier gas flow. In the sensor the sample flow is exposed to an alternating electrical dispersion field and compensation voltage. Due to the alternating current and collision of ions with the molecules of the carrier gas, the sample ions are separated according to their ion mobility and mass. The response of the ions colliding the detector as a function of the dispersion field provides a chemical fingerprint (Figure 1), which contains information about all the compounds present in a sample. The analysis takes place in room-temperature and atmospheric pressure. [29].

We conducted three consecutive scans from each sample and scanned approximately 20 samples per hour, 20-60 samples per day, on four separate days. The analyzing time for each sample was approximately 2 minutes. Between the urine samples the FAIMS system was cleaned with sterile water (5 ml). The cleaning was performed in a similar way than the analyzing of urine samples. We scanned a minimum of three scans with the water and changed for fresh water after ten urine samples.

The settings of the FAIMS scanning were provided by the manufacturer. The more detailed scan procedure is described in another paper [25], briefly: The gas (air) flow rate over the sample was 500 ml/min and this was mixed with clean air flow of 2000 ml/min, which makes a total flow of 2500 ml/min for the sensor. Dispersion field from 0 to 90% was scanned in 51 steps and compensation

voltage from −6 to +6 V was scanned in 512 steps. The scans were saved on Lonestar hard drive and a transportable USB drive for statistical analysis.

Statistical analysis

During measurement, for each matrix, we used variations consisting of 512 different compensation voltages and 51 different alternating current values and observed the amount of current (ions) reaching the detector plate. This was done for both positive and negative ions separately. Therefore, the original smell prints consisted of 52,224 measurement points per matrix of a sample. From each urine sample, the third measured matrix data was used for further analysis. We chose the third matrix for the mathematical analyzis in order to maximize the time the sample has to vaporize and generate a stronger signal. After measurements, the original resolution of 512 x 51 measuring points was condensed by averaging to 64 x 25, to significantly speed up the analysis time. Both, the negative and the positive responses were used and the L2 norm of each combined response was scaled to one. The analysis was conducted with Matlab R2020a (Mathworks). We used linear discriminant analysis (LDA) to create classification parameters. To avoid over-fitting, the results were cross validated using leave-one-out cross validation (LOOCV) [30].

The ability of the FAIMS electronic nose to detect and classify chronic renal disease was tested by comparing the groups of patients with a different CKD-EPI classification. The patients in the groups GFRc 4-5, and the patients in the GFRc 5 were compared to GFRc 1 (Table 2). Other combinations of testing were also performed, including comparingGFRc 4-5 to GFRc 1-2, and comparing GFRc 4-5 to GFRc 1-3. Furthermore, we compared GFRc 2-5 to GFRc 1. Stratification of the whole cohort in five classes was also attempted.

Same classifications were calculated for patients by utilizing 0.45g or more of protein per day in 24-hour urine as the threshold for positive finding, and positive finding for albumin in dipstick urine test.

We tested the groups for the main variables possibly affecting both the urine VOC status and CKD, including diabetes (DM), arterial hypertension (HA), smoking and gender. A Pearson's chi square - test was performed for the statistical dependency of the variables. In the original data, hypertension was diagnosed when a study subject had received medical treatment for hypertension or had a systolic blood pressure ≥ 160 mmHg or diastolic blood pressure ≥ 95 mmHg at the time of the examination (the urine sample collection). Respectively, diabetes was diagnosed when a previous or current 12 –hour fasting blood glucose level was 6.7 mmol/l or greater.

Results

The analysis was performed for 95 patients. The patient demographics and GFR estimates are presented in Table 1. Concerning the main variables, the GFR classes differed statistically significantly from each other in the incidence of diabetes ($p = 0.018$) and hypertension ($p = 0.000$), while no difference was found in history of smoking ($p = 0.360$) or gender ($p = 0.173$).

The ability to correctly classify patients was highest in comparison of the extremities: when GFRc 5 was compared to GFRc 1 (outcome of the test being GFRc 5) or even more accurately when GFRc 4-5

was compared to GFRc 1 (outcome being GFRc 4-5), which included more cohort. This resulted in a sensitivity of 97.7 %, a specificity of 33.3 % and an accuracy of 81.4 %. When five-class classification to five GFR classes was attempted, the classification accuracy of 30.5 % only slightly exceeded that of chance alone (28.4 %). The sensitivity, specificity and accuracy of each testing set is demonstrated in Table 2. Due to significant association of diabetes with CKD, the analyses were repeated after exclusion of patients with diabetes. After exclusion of patients with diabetes the results remained essentially the same. For example, with diabetic patients excluded, and distinguishing between GFRc 4-5 and GFRc 1-2, a correct classification rate of 69.6 % was achieved with a sensitivity of 76.7 % and a specificity of 61.5 %.

The sensitivity of 24-hour urine protein was 95.0 %, with specificity of 86.7 % and accuracy of 91.4 % for GFRc 5 vs GFRc 1. The results for dipstick urine albumin were 65.0 %, 100.0 % and 80.0 % accordingly. The sensitivity, specificity and accuracy of each testing set is demonstrated in Table 2.

Through all samples, Pearson correlation coefficients between signal intensity and kidney function were also calculated for the spectra. With this information, the linear dependence of each point in the spectra with kidney function could then be measured. In the positive channel, therefore, the higher the linear correlation of a given point, the more likely it is to gain higher values of signal intensity as GFR of the sample impairs. In the negative channel, as signal intensity is measured to negative direction, the method is interpreted in the reverse manner. In both channels, the extreme measured correlations were approximately ± 0.4, which show moderate linear correlation between the particular regions of spectra and kidney function. (Figure 2)

Discussion

Applying an electronic nose with FAIMS –technology we were able to distinguish patients with an impaired kidney function from controls. The stratification of all patients in five groups depending on the eGFR value could not be done reliably. The result is not surprising considering the nature of the disease: renal function impairment is a continuum characterized by a long asymptomatic phase and therefore a clear-cut classification is always artificial. Nonetheless, this study shows that chronic renal disease has a significant impact on urine VOC composition. This is the first study conducted from CKD point of view using urine samples and an electronic nose technology and underlines the importance of addressing renal function when analyzing urine gaseous headspace.

To our knowledge, the only study with urine VOCs related to CKD was focusing on pediatric hematuria [24]. The results of that study cannot be applied to adults, who rarely present with hematuria as the key symptom of CKD, since vast majority of adult CKD is caused by diabetes and hypertension. The study population size was limited, and no cross validation was used. Because of these limitations, the reference to this study should be considered anecdotal.

A limited number of studies focusing on the VOCs of chronic renal disease patients have been carried out using different methods and sample media compared to ours but resulting in similar outcomes. Voss et al. [31] were able to differentiate all healthy subjects from those with a renal failure using an electronic nose with conducting metal oxide -sensors attached to a patient's leg. Alike us, they were able to differentiate between the extremities, and had a correct classification of 95.2 % of

patients with end-stage renal dysfunction and less severe chronic renal failure. Lin et al. [32] applied an electronic nose constructed of six piezoelectric quartz crystal arrays to mimic the olfactory receptor protein to detect uremia and CKD in breath samples. They were able to differentiate these with a high accuracy of 86.78 %. In addition, the differentiation of the healthy and sick patients was done 100 % correctly. Also, in an animal model, Haick et al. [33] were able to identify rats with bilateral nephrectomy from the healthy individuals by 27 VOCs appearing in the breath of these rats. None of the mentioned studies, however, used urine as the biological source of VOCs.

Interestingly, the areas of the spectrum that correlate with kidney function in the FAIMS spectrum are very distinct, as demonstrated in Figure 2, although the correlation of the features is modest. The structure of the FAIMS spectrum consists of reactant ion peak from water and residual ion peak from molecules that the sensor is unable to discriminate. In analysis of air the molecules of interest are typically between these two peaks [34]. Since our analysis was explorative in nature, no standards were used to attempt explicit identification of molecules. The features highlighted by Pearson correlation in figure 2 seem to correspond the peaks of water (right image, strong yellow on the left) and ammonium (right image, blue on the extreme left). The positive correlation of CKD and water content is supported by impaired ability to concentrate urine [35] and decreasing urine concentrations of ammonium are connected to the development of metabolic acidosis and poorer outcomes as kidney failure progresses [36]. We therefore speculate that the ability of FAIMS to discriminate GFR classes is mostly explained by these two molecules.

Both breath and urine are alluring sources of VOCs, since their collection is non-invasive. However, breath analysis seems to be prone to confounding factors and is more challenging to obtain in a standardized manner. Age, gender and BMI of the individual as well as smoking habits have been shown to influence breath VOC profiles [37]. Furthermore, distinguishing exogenous contaminants from endogenous products can be challenging [13] and for example, hospital room air affects uremic breath patterns notably [14]. The patients in the Pagonas et al. study [15] had to refrain from eating, drinking or brushing their teeth for two hours prior to the breath analysis.

However, urine sample collection is less prone to environmental contaminants and does not require standardized settings. Urine is not in a gaseous form until it is heated up inside the headspace sampler and thus will not be as susceptible for contamination. Furthermore, giving a urine sample seems to be preferable to a blood sample and patients feel it is easier to give at home [38]. The compliance of the patient for a biological sample might be better when using urine instead of blood.

Our study involved a relatively large patient cohort and good demographic information of the patients which was collected with a structured interview and tests by the time the patients were recruited. The recruited CKD population contains a relatively wide range of etiologies. The urine analysis with FAIMS was done in a blinded manner and in a random order, which significantly reduces the risk of bias. The CKD status of the patients and healthy controls remained unknown until the analysis was completed. We used cross validation (LOOCV) for our mathematical analysis to minimize bias from over-fitting.

Even though urine is a stable medium for VOC analysis, the analytic window of the time for analysis after the initial collection might be limited. The results of Esfahani et al. [39] study indicate that the VOC concentration in urine samples decreases over time. The storage time of our samples was long, which may decrease their VOC output and thus the difference is likely to be under-estimated rather

than over-estimated. However, all the analyzed urine samples had been stored for nearly the same time, so that the variation between them was minor. Our study could be affected by the imperfect gold standard bias, aka copper standard bias, to some degree [41]. The "correct" classification of our index test (the Lonestar analysis) was based on a calculated GFR value with the use of measured serum creatinine, which itself has potential biases. Thus, our gold standard was imperfect. However, in that case, both sensitivity and specificity of our index test would be lowered, since it is independent of the gold standard test. We also compared the performance of FAIMS to the performance of dipstick albumin test and 24-hour urine albumin collection. Obviously, the performance of 24-hour urine albumin collection is superior to both dipstick and FAIMS analysis. The accuracy of FAIMS is roughly par with dipstick albumin but in contrast to low sensitivity of dipstick albumin, FAIMS is more sensitive but has poorer specificity.

The patients in different GFR classes differed from each other with the incidence of hypertension and diabetes, which is logical, since they are diseases that cause kidney dysfunction. Since diabetes results in significant metabolic alterations reflected in urine other than just presence of glucose and proteins alone [42,43], we performed additional analyses to exclude the possibility that the classification of CKD is rather due to diabetes-related factors and confirmed that the results are not related with diabetes. Additional limitation of our study is the lack of data on hematuria. Also, our study is limited to adult population.

Screening of impaired kidney function may allow for prevention of the development of severe renal insufficiency requiring dialysis, especially since chronic kidney disease is asymptomatic for a long time. However, screening is not routinely done. The U.S. Preventive Services Task Force has announced in 2012 that the screening for lowered eGFR or albuminuria of asymptomatic adults without a diagnosed CKD (ruling out the persons with diabetes or hypertension) is not effective considering the costs and harm caused by the screening. [44]. With the present clinical methods, screening for CKD is cost-effective only in high-risk populations (persons with diabetes) [45] and most guidelines therefore recommend screening only in these groups [8].

Conclusions

In conclusion, despite the limitations of our study, we were the first to show that FAIMS electronic nose technology is capable of differentiating the extremes of kidney function from urine samples and that the kidney function is reflected in the urinary VOC profile.This is likely explained by changes in ammonium and water content caused by CKD. Our study provides valuable information for other FAIMS studies, since CKD and impaired kidney function can be confounding factors when trying to diagnose other diseases. Since FAIMS analysis is not limited to single analyte, single non-invasive urine test could potentially be used to detect several different stages of illnesses or conditions at the same time.

Future Perspective:

Electronic noses have been in research use for over two decades and are getting closer to clinical use. They will likely compete with mass spectrometry in applications where affordable qualitative analysis is more convenient than expensive quantitative analysis.

The challenge in eNose research has been the wide variation in methodology. We demonstrated the significant effect of kidney function to urinary headspace and predict that in the future, normalization of sensor response to kidney function will be expected from all eNose studies that use urine as sample medium.

Topic of future studies: To establish, with a larger cohort, if different kidney diseases have specific VOC profiles and if the true differentiation of CKD could be done by utilizing an electronic nose.

Summary Points:

- Impairment of kidney function can be detected from the urine headspace utilizing FAIMS technology.
- We analyzed the urine of 95 patients with varying kidney function using Owlstone Lonestar FAIMS sensor.
- The discrimination of kidney function from urine headspace has not been previously attempted.
- Spectrums were analyzed in order to discriminate different kindey function groups and to see features that correlate with kidney function.
- Discrimination rate improved as the difference between the kidney function of the groups increased.
- Changes in ammonia and water content are speculated to be the reason for different smellprints.
- Kidney function has not been considered in most eNose studies and may be a significant confounder.

References:

Papers of special note have been highlighted: • of interest1. Hill NR, Fatoba ST, Oke JL *et al*. Global prevalence of chronic kidney disease - A systematic review and meta-analysis. *PLoS ONE [Electronic Resource]* 11(7):e0158765 (2016).

2. Gansevoort RT, Correa-Rotter R, Hemmelgarn BR *et al*. Chronic kidney disease and cardiovascular risk: Epidemiology, mechanisms, and prevention. *Lancet* 382 North American Edition(9889), 339-352 (2013).

3. Levey AS, Coresh J. Chronic kidney disease. *Lancet* 379(9811), 165-180 (2012).

4. Levin A, Stevens PE, Bilous RW *et al.* Kidney disease: Improving global outcomes (KDIGO) CKD work group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney International Supplements* 3(1), 1-150 (2013).

5. Eckardt K, Coresh J, Devuyst O *et al.* Evolving importance of kidney disease: From subspecialty to global health burden. *Lancet* 382(9887), 158-169 (2013).

6. James MT, Hemmelgarn BR, Tonelli M. Early recognition and prevention of chronic kidney disease. *Lancet* 375(9722), 1296-1309 (2010).

7. Al-Wakeel J. Accuracy and precision of the CKD-EPI and MDRD predictive equations compared with glomerular filtration rate measured by inulin clearance in a saudi population. *Ann Saudi Med* 36(2), 128-134 (2016).

8. Levey AS, Becker C, Inker LA. Glomerular filtration rate and albuminuria for detection and staging of acute and chronic kidney disease in adults: A systematic review. *JAMA* 313(8), 837-846 (2015).

9. Inker LA, Tighiouart H, Coresh J *et al*. GFR estimation using ß-trace protein and ß2-microglobulin in CKD. *Am J Kidney Dis*. 67(1), 40-48 (2016).

10. Edwards TL, Browne CM, Schoon A, Cox, C. & Poling, A. Animal olfactory detection of human diseases: Guidelines and systematic review. *Journal of Veterinary Behavior: Clinical Applications and Research* 20, 59-73 (2017).

11. de Lacy Costello, Amann A, Al-Kateb H *et al.* A review of the volatiles from the healthy human body. *J Breath Res*. 8(1), 014001-014001 (2014).

12. Shirasu M, Touhara K. The scent of disease: Volatile organic compounds of the human body related to disease and disorder. *J Biochem*. 150(3), 257-266 (2011).

13. Buszewski B, Kesy M, Ligor T, Amann A. Human exhaled air analytics: Biomarkers of diseases. *Biomedical Chromatography* 21(6), 553-566 (2007).

14. Mochalski P, King J, Haas M, Unterkofler K, Amann A, Mayer G. Blood and breath profiles of volatile organic compounds in patients with end-stage renal disease. *BMC Nephrol*. 15, 43-43 (2014).

15. Pagonas N, Vautz W, Seifert L *et al.* Volatile organic compounds in uremia. *PLoS One* 7(9), e46258-e46258 (2012).

16. Di Francesco F, Fuoco R, Trivella MG, Ceccarini A. Breath analysis: Trends in techniques and clinical applications. *Microchemical Journal* 79(1), 405-410 (2005).

17. Bijland LR, Bomers MK, Smulders YM. Smelling the diagnosis: A review on the use of scent in diagnosing disease. *Neth J Med*. 71(6), 300-307 (2013).

• A review article, that could be seen as an introduction to the use of smell in diagnosing diseases throughout time.

18. Hackner K, Pleil J. Canine olfaction as an alternative to analytical instruments for disease diagnosis: Understanding 'dog personality' to achieve reproducible results. *J Breath Res*. 11(1), 012001-012001 (2017).

19. Bhandare PB, Pendbhaje NS, Narang AP. Electronic nose: A review. *Research & Reviews: Journal of Engineering and Technology* 2(4) (2013).

• A report of the basics mechanisms of electronic noses.

20. Wilson AD, Baietto M. Advances in electronic-nose technologies developed for biomedical applications. *Sensors (Basel)* 11(1), 1105-1176 (2011).

21. Röck F, Barsam N, Weimar U. Electronic nose: Current status and future trends. *Chemical Reviews* 108(2), 705-725 (2008).

22. Casalinuovo IA, Di Pierro D, Coletta M, Di Francesco P. Application of electronic noses for disease diagnosis and food spoilage detection. *Sensors* 6(11), 1428-1439 (2006).

23. Capelli L, Taverna G, Bellini A *et al.* Application and uses of electronic noses for clinical diagnosis on urine samples: A review. *Sensors (Basel)* 16(10) (2016).

• A review of other studies with the application of electronic noses and urine

24. Di Natale C, Mantini A, Macagnano A, Antuzzi D, Paolesse R, D'Amico A. Electronic nose analysis of urine samples containing blood. *Physiol Meas*. 20(4), 377-384 (1999).

25. Niemi RJ, Roine AN, Eräviita E, Kumpulainen PS, Mäenpää J,U., Oksala N. FAIMS analysis of urine gaseous headspace is capable of differentiating ovarian cancer. *Gynecol Oncol*. 151(3), 519-524 (2018).

• A study that was conducted with the same FAIMS system and similar methods to this one.

26. Leskinen Y, Lehtimäki T, Loimaala A *et al.* Homocysteine and carotid atherosclerosis in chronic renal failure--the confounding effect of renal function. *Atherosclerosis* 175(2), 315-323 (2004).

27. Leskinen Y, Salenius JP, Lehtimäki T, Huhtala H, Saha H. The prevalence of peripheral arterial disease and medial arterial calcification in patients with chronic renal failure: Requirements for diagnostics. *Am J Kidney Dis*. 40(3), 472-479 (2002).

28. Levey AS, Stevens LA, Schmid CH *et al.* A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 150(9), 604-612 (2009).

29. Covington JA, van dS, Edge ASL, Boyle B, Savage RS, Arasaradnam RP. The application of FAIMS gas analysis in medical diagnostics. *Analyst* 140(20), 6775-6781 (2015).

• An overview of FAIMS in medical diagnostics.

30. Bishop CM. *Pattern recognition and machine learning.* New York: Springer Science+Business Media, LCC (2006).

31. Voss A, Baier V, Reisch R *et al.* Smelling renal dysfunction via electronic nose. *Ann Biomed Eng*. 33(5), 656-660 (2005).

• Another study to indicate that renal dysfunction could indeed be smelled with an electronic nose. Promising results, although with a different kind of eNose and sample media.

32. Lin Y, Guo H, Chang Y, Kao M, Wang H, Hong R. Application of the electronic nose for uremia diagnosis. *Sensors Actuators B: Chem*. 76(1–3), 177-180 (2001).

33. Haick H, Hakim M, Patrascu M *et al.* Sniffing chronic renal failure in rat model by an array of random networks of single-walled carbon nanotubes. *ACS Nano* 3(5), 1258-1266 (2009).

34. Cumeras R, Figueras E, Davis CE, Baumbach JI, Gràcia I. Review on ion mobility spectrometry. part 2: Hyphenated methods and effects of experimental parameters. *Analyst* 140(5), 1391-1410 (2015).

35. Combs S, Berl T. Dysnatremias in patients with kidney disease. *Am J Kidney Dis*. 63(2), 294-303 (2014).

36. Raphael KL, Carroll DJ, Murray J, Greene T, Beddhu S. Urine ammonium predicts clinical outcomes in hypertensive kidney disease. *J Am Soc Nephrol*. 28(8), 2483-2490 (2017).

37. Blanchet L, Smolinska A, Baranska A *et al.* Factors that influence the volatile organic compound content in human breath. *J Breath Res*. 11(1), 016013-016013 (2017).

38. Dhima M, Salinas TJ, Wermers RA, Weaver AL, Koka S. Preference changes of adult outpatients for giving saliva, urine and blood for clinical testing after actual sample collection. *J Prosthodont Res*. 57(1), 51-56 (2013).

39. Esfahani S, Sagar NM, Kyrou I *et al.* Variation in gas and volatile compound emissions from human urine as it ages, measured by an electronic nose. *Biosensors (Basel)* 6(1) (2016).

40. Turner C. Techniques and issues in breath and clinical sample headspace analysis for disease diagnosis. *Bioanalysis* 8(7), 677-690 (2016).

41. Kohn MA, Carpenter CR, Newman TB. Understanding the direction of bias in studies of diagnostic test accuracy. *Acad Emerg Med*. 20(11), 1194-1206 (2013).

42. Verbeke F, Siwy J, Van Biesen W *et al.* The urinary proteomics classifier chronic kidney disease 273 predicts cardiovascular outcome in patients with chronic kidney disease. *Nephrol.Dial.Transplant.* [Epub ahead of print] (2019).

43. Colombo M, McGurnaghan SJ, Blackbourn LAK *et al.* Comparison of serum and urinary biomarker panels with albumin/creatinine ratio in the prediction of renal function decline in type 1 diabetes. *Diabetologia* (2020).

44. Screening for chronic kidney disease: U.S. preventive services task force recommendation statement. *Ann Intern Med*. 157(8), 50 (2012).

45. Manns B, Hemmelgarn B, Tonelli M *et al.* Population based screening for chronic kidney disease: Cost effectiveness study. *BMJ* 341, 5869 (2010).

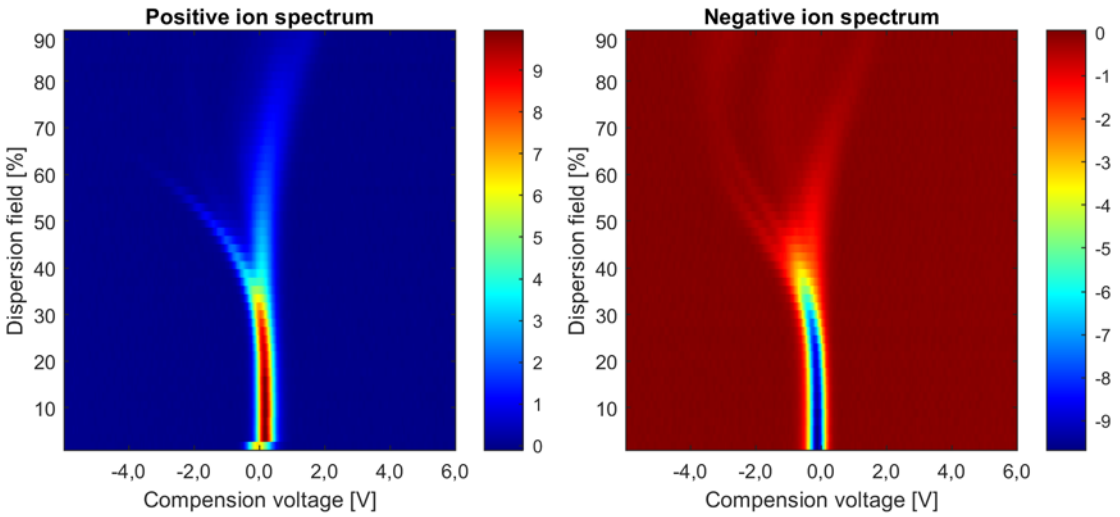

Figure 1: The FAIMS response of a typical urine sample of GFRc 1. *FAIMS generates separate response for positive ions (left) and negative ions (right).*

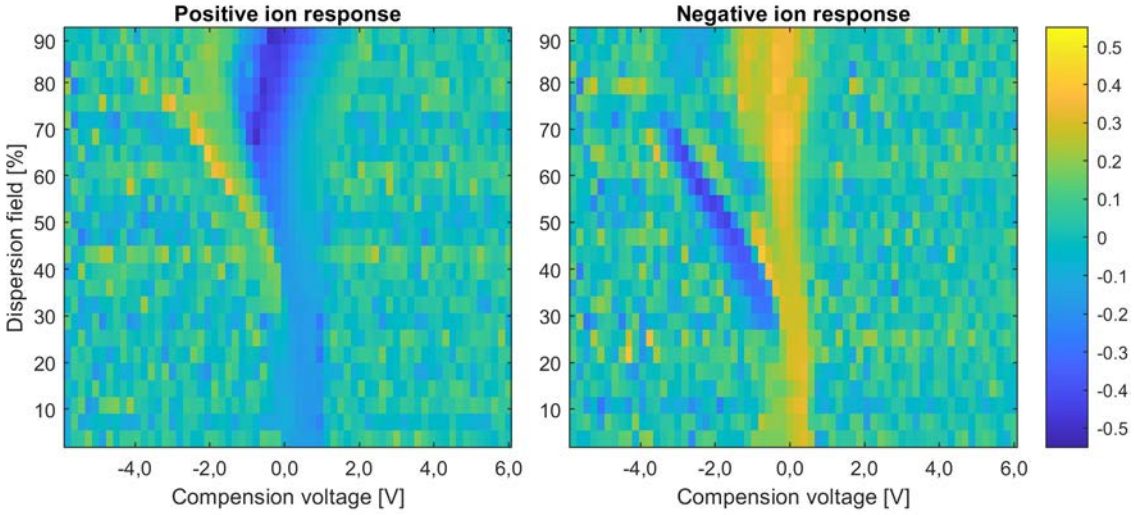

Figure 2: The correlation between the FAIMS response and the GFR class. *The positive ion response is on the left and negative ion response is on the right. Yellow areas are more prominent in kidney failure and blue areas are less prominent in kidney failure.*

Table 1: The patient demographics in accordance with the GFR class

	GFRc 1	GFRc 2	GFRc 3	GFRc 4	GFRc 5	Total
Age, mean (range)	49.3 (29-69)	56.7 (32-69)	50.2 (33-66)	53.6 (29-68)	57.8 (34-69)	54.4 (29-69)
Men	11 (73 %)	16 (59 %)	8 (89 %)	16 (67 %)	9 (45 %)	60 (63 %)
DM	0 (0 %)	1 (4 %)	1 (11 %)	7 (29 %)	7 (35 %)	16 (17 %)
HA	0 (0 %)	0 (0 %)	8 (89 %)	23 (96 %)	19 (95 %)	50 (53 %)
Smoking	7 (47 %)	4 (15 %)	3 (33 %)	5 (21 %)	3 (15 %)	22 (23 %)
Positive urine dipstick test for albumin	-	1 (4 %)	6 (67 %)	14 (58 %)	13 (65 %)	34 (36 %)
≥ 0,45 g protein in 24h urine sample	-	2 (7 %)	7 (78 %)	15 (63 %)	17 (85 %)	41 (43 %)
Etiology of kidney disease:						
Diabetic nephropathy	-	-	1 (11 %)	6 (25 %)	6 (30 %)	13 (14 %)
Chronic glomerulo-nephritis	-	1 (4 %)	3 (33 %)	3 (13 %)	2 (10 %)	9 (10 %)
Polycystic kidney disease	-	-	-	1 (4 %)	4 (20 %)	5 (5 %)
Hypertensive nephropathy	-	-	2 (22 %)	2 (8 %)	1 (5 %)	5 (5 %)
Chronic pyelonephritis	-	-	-	2 (8 %)	2 (10 %)	4 (4 %)
Amyloidosis	-	-	-	1 (4 %)	-	1 (1 %)
Other / unknown	-	-	3 (33 %)	9 (38 %)	5 (25 %)	17 (18 %)
Total	15	27	9	24	20	95

GFRc = glomerular filtration rate class, DM = diabetes mellitus, HA = hypertension

Table 2: Discrimination rates in comparison of different CKD groups

	GFRc 5 vs 1	GFRc 2-5 vs 1	GFRc 4-5 vs 1	GFRc 4-5 vs 1-2	GFRc 4-5 vs 1-3
FAIMS					
Accuracy	74.3%	75.8%	81.4%	69.8%	64.2%
Sensitivity	85.0%	94.4%	97.7%	68.2%	59.1%
Specificity	60.0%	20.8%	33.3%	71.4%	68.6%
24-hour urine protein					
Accuracy	91.4%	62.8%	84.9%	88.6%	82.9%
Sensitivity	95.0%	58.2%	84.2%	84.2%	86.7%
Specificity	86.7%	86.7%	86.7%	92.7%	80.0%
Dipstick urine albumin					
Accuracy	80.0%	62.8%	71.2%	78.8%	74.5%
Sensitivity	65.0%	58.2%	61.4%	61.4%	61.4%
Specificity	100.0%	86.7%	100.0%	97.6%	86.0%

FAIMS = Field Asymmetric Waveform Ion Mobility Spectrometry