Detection of Prostate Cancer by an Electronic Nose: A Proof of Principle Study

Antti Roine,*† Erik Veskimäe,† Antti Tuokko, Pekka Kumpulainen, Juha Koskimäki, Tuomo A. Keinanen,‡ Merja R. Häkkinen,† Jouko Vepsäläinen, Timo Paavonen, Jukka Lekkala, Terho Lehtimäki, Teuvo L. Tammela§ and Niku K. J. Oksala

From the School of Medicine, University of Tampere (AR, AT), Department of Surgery, School of Medicine, University of Tampere and Department of Urology, Tampere University Hospital (EV, JK, TLT), Department of Automation Science and Engineering, Tampere University of Technology (PK, JL), Department of Pathology, School of Medicine, University of Tampere and Finlab Laboratories, Tampere University Hospital (TP), Department of Clinical Chemistry, Finlab Laboratories and University of Tampere, School of Medicine (TL), Department of Surgery, School of Medicine, University of Tampere and Department of Vascular Surgery, Tampere University Hospital (NKJO), Tampere, Finland; and Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden (TAK), and School of Pharmacy, Biocenter Kuopio, University of Eastern Finland, Kuopio, Finland (MRH, JV)

**Abbreviations and Acronyms**

BPH = benign prostatic hyperplasia  
DRE = digital rectal examination  
eNose = electronic nose  
G/LC-MS = gas/liquid chromatography-mass spectrometry  
LOOCV = leave-one-out cross-validation  
NMR = nuclear magnetic resonance spectroscopy  
PCa = prostate cancer  
PSA = prostate specific antigen  
TURP = transurethral resection of prostate

**Purpose:** We evaluate the ability of an electronic nose to discriminate prostate cancer from benign prostatic hyperplasia using urine headspace, potentially offering a clinically applicable noninvasive and rapid diagnostic method.

**Materials and Methods:** The ChemPro® 100-eNose was used to discriminate prostate cancer from benign prostatic hyperplasia using urine sample headspace. Its performance was tested with 50 patients with confirmed prostate cancer and 24 samples from 15 patients with benign prostatic hyperplasia (15 patients provided urine preoperatively and 9 patients provided samples 3 months postoperatively) scheduled to undergo robotic assisted laparoscopic radical prostatectomy or transurethral resection of prostate, respectively. The patients provided urine sample preoperatively and those with benign prostatic hyperplasia also provided samples 3 months postoperatively to be used as a pooled control sample population. A discrimination classifier was identified for eNose and subsequently, sensitivity and specificity values were determined. Leave-one-out cross-validation was performed.

**Results:** Using leave-one-out cross-validation the eNose reached a sensitivity of 78%, a specificity of 67% and AUC 0.77.

**Conclusions:** The electronic nose is capable of rapidly and noninvasively discriminating prostate cancer and benign prostatic hyperplasia using urine headspace in patients undergoing surgery.

**Key Words:** prostatic hyperplasia, electronic nose, prostatic neoplasms

Prostate cancer is the second most common cancer in males and one of the leading causes of cancer mortality. The heterogeneity of PCas makes it difficult to diagnose and to predict tumor progression. Current cornerstones of the diagnosis, ie DRE and plasma PSA, have limited sensitivity and specificity, although free PSA adds to the performance. Histological examination of transrectal ultrasound guided biopsies leads to definitive diagnosis, but is associated with considerable costs, discomfort and risk of infectious complications. In addition, a significant proportion of diagnosed cancers is low grade and will not cause symptoms or disease.
specific mortality. Therefore, aggressive treatment can lead to decreased quality of life without affecting the longevity of the patient. Thus, there is a need for novel diagnostic tools.

The first report on the olfactory detection of cancer was a case study published in The Lancet. Since then, experimental studies of the use of trained dogs in the detection of cancer have confirmed the preliminary findings. Previous results of PCA detection were discouraging until Cornu et al showed that trained dogs are able to detect PCA using urine with high sensitivity and specificity. This finding sparked considerable attention.

The concern with canine olfaction is the heterogeneity of the performance of the dogs between and within studies. The eNose is a device that consists of a cluster of nonspecific sensors. When the device is exposed to the sample, it produces a profile or a smell print. eNoses are best suited for qualitative analysis of complex gaseous mixtures of molecules, and are routinely used in food and agricultural quality control and in military applications. eNoses have been studied in various medical applications, including the early detection of cancer, especially from exhaled air.

Exhaled air is a problematic sample material since it requires good cooperation and technique from the patient as well as immediate analysis, while urine is simple to attain and store and, therefore, more feasible in clinical practice. Urine has also been used in several metabolomics studies using GC-MS or NMR. Preliminary data suggest that the detection of urological malignancies using urine headspace is possible.

Recently we showed that the discrimination of PCa cells and cells from BPH is possible. These results encouraged us to launch a prospective clinical study recruiting patients undergoing robotic assisted laparoscopic radical prostatectomy and symptomatic patients with BPH undergoing TURP. In this study we test the hypothesis that the eNose system is capable of discriminating patients with PCA from those with BPH using urine headspace. We also evaluate potential sources of diagnostic error such as prostate volume and tumor size.

**MATERIALS AND METHODS**

**Patients and Samples**

The study population was prospectively formed of consecutive patients referred to a urologist for operative treatment of PCa or BPH at Tampere University Hospital (Tampere, Finland) or Hatanpää City Hospital (Tampere, Finland) between March 2011 and November 2012. Inclusion criteria were operative treatment of PCa or BPH. Exclusion criteria were patient refusal, material insufficient for histopathology, known malignancy other than PCa, persistent urinary infection and urinary catheter in place. Patients were divided into cases and controls, as in patients with biopsy proven PCa scheduled to undergo robotic assisted laparoscopic radical prostatectomy (50) and PCa-free patients (15). The PCa-free group comprised patients with BPH scheduled to undergo TURP, having later confirmed benign histology. Of the patients with BPH 3 had previously had negative prostate biopsies.

The control group consisted of 15 patients who gave a urine sample preoperatively, and of these patients 9 provided another sample 3 months postoperatively, resulting in a total of 24 samples. As the measurement of postoperative prostate volume was not included in the study protocol, we reduced prostate volume 40% from the postoperative results. Subjects provided a standard morning urine sample before the operation and 3 months after the operation. No standardization of diet, hydration status or bladder time was conducted. The samples were stored at −70°C until eNose analysis.

Written informed consent was acquired from all subjects. The study was approved by the ethical committee of Tampere University Hospital (code: R10066). The baseline clinical characteristics of the patients are provided in the table.

---

**Baseline clinical and pathological characteristics of study subjects**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Confirmed PCa Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>62</td>
<td>66</td>
</tr>
<tr>
<td>Median (range)</td>
<td>63.5 (49—73)</td>
<td>67 (53—72)</td>
</tr>
<tr>
<td><strong>Total PSA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Median (range)</td>
<td>6.3 (2—18.2)</td>
<td>1.75 (0.2—9)</td>
</tr>
<tr>
<td><strong>Prostate vol:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>26.7</td>
<td>48.6</td>
</tr>
<tr>
<td>Median (range)</td>
<td>33 (15—75)</td>
<td>40.3 (18.2—100)</td>
</tr>
<tr>
<td><strong>PSA density (PSA/prostate vol):</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.22</td>
<td>0.06</td>
</tr>
<tr>
<td>Median (range)</td>
<td>0.18 (0.04—0.89)</td>
<td>0.05 (0.02—0.2)</td>
</tr>
<tr>
<td><strong>No. postop Gleason score:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>Not applicable</td>
</tr>
<tr>
<td>7</td>
<td>34</td>
<td>Not applicable</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Not applicable</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>Not applicable</td>
</tr>
<tr>
<td><strong>No. pT status (%):</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pT1</td>
<td>27 (50)</td>
<td>Not applicable</td>
</tr>
<tr>
<td>pT3</td>
<td>23 (50)</td>
<td>Not applicable</td>
</tr>
<tr>
<td><strong>No. presence of Gleason 4 or greater (%):</strong></td>
<td></td>
<td>Not applicable</td>
</tr>
</tbody>
</table>

---
computer with software (Environics Inc.) for monitoring and logging the data.

The measurement chamber was similar to that described in our previous work. It consists of a polystyrene cell culture plate and a cover in which 3 holes are drilled. Two holes are for replacement air and the third hole is inserted with a modified 16G intravenous cannula. The injection port of the cannula is removed and a Teflon pipe is connected to it to provide sample air for eNose. The cover is secured using parafilm. The covers were recycled during the study and washed with sterile water between measurements. Culture plates were used only once. Sample material (5 ml) was defrosted and pipetted to the plate, resulting in a thin layer of fluid, maximizing the area for evaporation. The sample plate was heated and maintained at 37°C by a water poultice and kept in a laminar flow cabinet for the duration of the measurement. Each measurement lasted approximately 15 minutes and a recovery period of 10 minutes was used to avoid carryover. These intervals were previously determined to be sufficient. The maximum absolute values during the measurement interval were extracted for further analysis. Of 18 channels 2 had no response in the measurements and were excluded, effectively resulting in 16 channels for the analysis. The samples were analyzed in random order and the persons performing the measurements were blinded to the status of the patient.

**Statistical Analysis**

To emphasize the smell print profile produced by the eNose each sample was scaled by dividing each element (16 channels) with the L2-norm of the vector. Thus, the L2-norm of each scaled sample equaled one. Linear discriminant analysis was used to assign each sample into predefined classes of PCa and BPH. LOOCV was used to avoid overfitting the classifiers.

Pointwise confidence intervals for the ROC curves were computed as bias corrected and accelerated percentiles from 500 bootstrap samples. The optimal thresholds were identified as the upper leftmost contact points between a straight line that has a slope $S$ and the ROC curve. Assuming equal costs for false-positive and false-negative classes, the slope $S$ reduces to the ratio $N/P$, the number of negative samples divided by the number of positive samples.

Finally, the correlations between the eNose data and prostate volume in the BPH group as well as tumor volume in the PCa group were examined by a multilinear regression model, predicting prostate size using the eNose data. A stepwise regression was performed using $p$ values of 0.05 and 0.1 for including and removing variables from the model correspondingly. The same procedure was applied to examine the correlation between the eNose and tumor size.

**RESULTS**

The ability of eNose to differentiate the patients with operable PCa from those with symptomatic BPH requiring surgery was evaluated. The classifier comprised all the variables in the 16-dimensional data set without performing dimension reduction. A sensitivity of 82%, a specificity of 88% and AUC 0.92 were achieved using linear discriminant analysis. This represents the potential maximum performance of the methodology, but it is prone to show overly positive results in small sample cohorts due to overfitting. LOOCV was performed to minimize the effect of overfitting and to model an actual diagnostic implementation more accurately. In LOOCV analysis, a sensitivity and specificity of 78% and 67%, respectively, and AUC 0.77 were achieved. ROC curve is shown in figure 1. The stepwise regression model for predicting prostate size using eNose measurements resulted in the final model including only one eNose channel, number 9. Its correlation coefficient with prostate size was 0.34 ($p$ value 0.02 probability of zero correlation). The $R^2$ was 0.11 and, thus, the model captured 11% of the total variance. No significant model for predicting tumor size was found. The average profile of urine and characteristics of average smell prints of BPH and PCA are presented in figure 2.

**DISCUSSION**

For the first time to our knowledge we have shown that the eNose is capable of differentiating patients with PCa from subjects with symptomatic BPH by rapid analysis of urine headspace. Our results are in line with the preliminary findings of Bernabei et al. An important distinction is that while demonstrating the capability of the eNose to differentiate various urological malignancies (ie bladder and prostate cancer) from each other, they did not attempt to distinguish BPH from PCa. Cornu et al...
discriminated between BPH and PCa with excellent results by using trained dogs. The performance of the dog described in the study was superior to that of the eNose, potentially indicating that the dog’s olfactory apparatus is superior to the eNose or that during the training phase the dog learned to ignore distracting factors such as variable kidney function, diet or medication, which may alter the smell of urine. Cornu et al collected urine samples after attentive DRE, which might partially explain the higher sensitivity and specificity. However, DRE makes the testing more invasive.

NMR and G/LC-MS offer an approach analogous to eNose when used in a holistic way. Most studies focus on plasma or tissue as sample material and, instead of a holistic approach, aim to find new biomarkers for cancer. Sarcosine was found using this methodology but proved disappointing in further studies. In a recently published study Zhang et al attempted to distinguish patients with PCa from healthy controls using holistic LC-MS. Their biomarker panel reached a high AUC of 0.896. It should be noted that PSA reached an AUC of 0.94, a result that could be expected when comparing patients with metastasized PCa to healthy individuals. It should be noted that NMR and G/LC-MS are considerably bulkier and more expensive than eNose. NMR requires little sample preparation but has limited sensitivity, whereas mass spectrometry offers significantly higher sensitivity but requires costly and time-consuming sample preparation. With these limitations in mind, eNose, NMR and G/LC-MS are complementary, rather than competing methods.

PSA is known to correlate positively with prostate volume, which is a potential source of diagnostic error when comparing PCa with BPH. According to our current analysis prostate volume did not affect the eNose results, potentially indicating the high specificity of our sensor array to cancer. We also studied whether eNose signal correlates with tumor size and no such correlation was found.

To simulate real-life performance we conducted LOOCV in which the sample subject to discrimination is excluded from the data used to identify the classifier. Due to the limited sample number, splitting the sample into training and test populations was not feasible. LOOCV reduced the performance of eNose to a level similar to that reported by Thompson et al for PSA (AUC 0.678).

It should be noted that in our study the control patients with symptomatic BPH were confirmed cancer-free by histological examination of the material resected in TURP. This is noteworthy since the majority of cancers develop in the peripheral zone of the prostate. Thus, some patients in the control group may actually have subclinical cancer and this may negatively affect the discriminatory capability of the eNose. The false-positive rate of 21% is in line with the reported rate of 23% of biopsy negative PCa.

Another factor to be considered is that our BPH sample consisted entirely of patients sufficiently symptomatic to require surgical treatment. We hypothesize that discrimination efficacy could be enhanced in a population consisting of patients with BPH with mild symptoms and smaller prostates. An additional confounding factor is that patients with BPH may have more lower urinary tract symptoms such as increased voiding frequency and incomplete voiding, which may affect the time urine is exposed to the prostate.

While our study confirms our hypothesis of eNose technology use in the detection of PCa, further studies are needed to evaluate the effect of comorbid conditions, hydration, diet and medications. As Cornu et al pointed out another prospect is to pinpoint the specific molecules responsible for the different smell prints. In our previous work we showed that malignant prostatic cells modify their culture medium. One possibility is that the distinct smell of PCa is indeed caused directly by molecules secreted to the urinary tract by cancerous cells. One future prospect is to conduct prostate massage and to study whether it enhances diagnostic performance by stimulating prostatic excretion. Due to the small sample size and that most of the cancers were Gleason score 7 it was not possible to investigate reliably whether a positive eNose finding was associated with histological PCa aggressiveness. This would be especially valuable in PCa screening to differentiate the clinically significant cancers from those which would not need further investigations and treatment. A simple,
rapid urine test with sensitivity and specificity matching that of PSA as compared to the previous published literature would be worthy of further evaluation in PCa screening.

CONCLUSIONS

eNose is able to discriminate PCa from BPH using urine headspace. The performance matches that of PSA in the literature, and results are achieved rapidly and in a completely noninvasive manner. Future studies to enhance current technology and to identify the molecules behind the distinct odors are warranted.

ACKNOWLEDGMENTS

Mikko Malkamäki, Taru Poikonen, Minna Järvensivu, Kristiina Sippola, Sami Koskela and Mikko Moisander performed eNose measurements. Annamari Aitolahdi assisted with sample collection at Finlab laboratories.

REFERENCES


EDITORIAL COMMENT

Should you follow your eNose? Roine et al present some exciting new technology whereby an electronic nose might be able to sniff urine and discern prostate cancer from benign prostatic hyperplasia.
While the authors found the results encouraging, they acknowledged that the sample size was small and had other deficiencies. As in most studies of this nature, many open questions remain.

One concern is that in the BPH group there were 15 patients but 24 samples because 9 of the 15 patients provided 3-month urine samples in addition to the preoperative samples. In these 9 patients the eNose might have picked up a treatment effect (from TURP) rather than a difference between cancer and BPH. Moreover these 9 samples are not independent. They come from a subset of the other 15 patients, and as such, this clustering should be accounted for statistically rather than analyzed as truly independent samples.

Perhaps a larger issue is interpreting this study in the context of a real-world medical decision making scenario. Certainly it would be of great interest to use eNose to detect cancer from benign tissue in men before biopsy who are suspected of having prostate cancer. However, such a population differs from that used by Roine et al in that this population would contain men without prostate cancer or symptoms of BPH, and these men were not included by Roine et al. In men with symptoms of BPH who are destined for TURP, the question of which of those men will ultimately be found to have cancer is perhaps not that interesting clinically. In summary, the control group used in the study by Roine et al might not be the most clinically relevant control group, and one must wonder how well the eNose will perform when the control group is different. It is difficult to see the clinical usefulness in discerning men with biopsy proven prostate cancer headed for prostatectomy from those with BPH, scheduled to undergo TURP, who will not have cancer detected. These would seem to be 2 quite different groups that do not need to be distinguished from one another.

Of course, changing the question to detecting cancer in men suspected of having cancer, but without symptoms of BPH and destined for TURP, demands consideration of the numerous existing nomograms and risk calculators for predicting positive biopsy. These tools may discriminate as well or better than eNose. A head-to-head comparison is needed here, as comparing areas under the ROC curve across techniques and data sets simultaneously is not reliable. Thus, comparing eNose in one study to PSA in another study is problematic. However, it would be interesting to see how PSA alone would do in differentiating the 2 groups in the study by Roine et al, as it could have a performance comparable to that of the eNose (and this comparison would avoid the issues of comparison across data sets).

The authors noted that eNose channel number 9 was correlated with prostate size with a correlation coefficient of 0.34. The p value was 0.02, which they interpreted as the probability of a zero correlation. It sure would be nice if the p value were that easy to interpret. Instead, it is actually the probability of having obtained results as extreme as these if, in fact, the null hypothesis is true (that \( r = 0 \)), a much more abstract concept. However, this was really a secondary aim of this study.

In summary, as the authors indicate, more studies with eNose are needed. It would be particularly useful to compare eNose with current state-of-the-art algorithms for distinguishing clinically important groups from one another.

Of course, changing the question to detecting cancer in men suspected of having cancer, but without symptoms of BPH and destined for TURP, demands consideration of the numerous existing nomograms and risk calculators for predicting positive biopsy. These tools may discriminate as well or better than eNose. A head-to-head comparison is needed here, as comparing areas under the ROC curve across techniques and data sets simultaneously is not reliable. Thus, comparing eNose in one study to PSA in another study is problematic. However, it would be interesting to see how PSA alone would do in differentiating the 2 groups in the study by Roine et al, as it could have a performance comparable to that of the eNose (and this comparison would avoid the issues of comparison across data sets).

The authors noted that eNose channel number 9 was correlated with prostate size with a correlation coefficient of 0.34. The p value was 0.02, which they interpreted as the probability of a zero correlation. It sure would be nice if the p value were that easy to interpret. Instead, it is actually the probability of having obtained results as extreme as these if, in fact, the null hypothesis is true (that \( r = 0 \)), a much more abstract concept. However, this was really a secondary aim of this study.

In summary, as the authors indicate, more studies with eNose are needed. It would be particularly useful to compare eNose with current state-of-the-art algorithms for distinguishing clinically important groups from one another.

REFERENCE