

There's something in the air : Volatile organic compounds in exhaled breath in pulmonary diseases

Citation for published version (APA):

van Berkel, J. J. B. N. (2010). *There's something in the air : Volatile organic compounds in exhaled breath in pulmonary diseases*. [Doctoral Thesis, Maastricht University]. Universitaire Pers Maastricht. <https://doi.org/10.26481/dis.20101103jb>

Document status and date:

Published: 01/01/2010

DOI:

[10.26481/dis.20101103jb](https://doi.org/10.26481/dis.20101103jb)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

CHAPTER 8

Summary

Introduction

Technical advances in analytical analysis during the last few decades have been responsible for the recent developmental improvements in diagnostics and partial understanding of metabolic and biological pathways. This could lead to the discovery of new biomarkers able to characterize and identify disease. This thesis aims to describe our efforts regarding design and validation of a diagnostic tool based on the analysis of exhaled air. Exhaled air contains a complex mixture of volatile organic compounds (VOCs), some of which could potentially be applied as biomarkers for lung diseases. To date only single markers in breath are used in clinical diagnostics lacking the degree of specificity and sensitivity in order for the methodology to be taken in clinical practice.

The rapid, accurate and non-invasive diagnosis of respiratory disease represents a challenge to clinicians while the development of new treatments can be confounded by insufficient knowledge of lung disease phenotypes. We have developed a sampling methodology for collecting concentrated samples of exhaled air from patients suffering from COPD and cystic fibrosis against which we employed two-stage thermal desorption gas chromatography-time-of-flight mass spectrometry (GC-MS) analysis. Data analysis tools have been developed enabling pipeline analysis of the generated GC-MS sample outputs. Informative VOCs were extracted from the compiled database and implemented into a classifier of which performance was evaluated.

Summary

Chapter 2 describes this developed methodology regarding the analysis of exhaled air. Both the instrumental setup and software are discussed, as well as the statistical analysis. Thermal desorption and gas chromatography coupled to time-of-flight mass spectrometry were used to analyze exhaled air samples. The VOC profiles obtained from each individual were combined into one final database based on similarity of mass spectra and retention indexes, which offers the possibility for a reliable selection of compounds of interest. The developed methodology was validated on a set of smokers and non-smokers. Support vector machine analysis identified 4 VOCs as biomarkers of recent exposure to cigarette smoke. Two of the selected VOCs were already mentioned in literature as possibly smoking behavior related compounds confirming the validity of the setup. After validation of the procedure a large study was setup to identify VOCs in bacterial headspace discriminating between different micro organisms. **Chapter 3** discusses the setup and results from the micro organism headspace analysis study. Different sets of VOCs were found enabling easy and fast identification of the different cultures. Several selected VOCs were previously mentioned in literature regarding bacterial headspace. We were able to identify a large number of compounds demonstrating a highly significant difference in availability in bacterial cultures compared to medium and in cul-

tures compared to one another. We have also determined highly significant compounds differing between the four *Escherichia coli* strains and between the two *Staphylococcus aureus* isolates: methicillin-resistant *Staphylococcus aureus* and methicillin-sensitive *Staphylococcus aureus*. SVM models were able to classify the microorganisms with very high degrees of sensitivity and specificity based on -on average- 6 VOCs from headspace. **Chapter 4** describes the developed methodology and search for biomarkers regarding chronic obstructive pulmonary disease (COPD). Fifty COPD patients and twenty-nine controls were sampled and VOCs were analyzed by gas chromatography-mass spectrometry to identify relevant VOCs. A classifier based on six VOCs correctly classified 92% of the subjects. Additionally several validation data sets were built and performance of the classifier was evaluated. **Chapter 5** investigates whether analysis of VOCs from exhaled air could discriminate between subject suffering from cystic fibrosis and controls, and between cystic fibrosis patients with and without *Pseudomonas aeruginosa* (*P. aeruginosa*) colonization. Statistical analysis identified 10 VOCs that combined into a SVM model correctly classified 92% of samples. This new technique was not only able to discriminate between cystic fibrosis patients and controls, but also between cystic fibrosis patients with and without *P. aeruginosa* colonization. **Chapter 6** emphasizes on application of the developed methodological exhaled air analysis technique on exacerbations in patients with cystic fibrosis. Prevention of exacerbations in cystic fibrosis is important since these pulmonary exacerbations are an important cause for the hospitalization of patients, respiratory symptoms, and decreases in lung function. In the described longitudinal study we studied whether changes in VOC profiles can predict pulmonary exacerbations of cystic fibrosis. Cystic fibrosis patients were sampled at 2-month intervals. Four extra samples were collected in case of an exacerbation. It appeared that 18 VOCs were differentially present in exhaled breath from patients not having an exacerbation versus patients suffering an exacerbation event, 15 VOCs were differentially present in exhaled breath from the comparison of baseline measurements to the first exacerbation measurement at the start of the exacerbation event. The study demonstrated that VOCs in exhaled breath are able to indicate cystic fibrosis exacerbations weeks before these adverse events are clinically manifest. Before these biomarkers can be of any clinical relevance however, the different classifiers need to be validated on external validation sets. **Chapter 7** presents the general discussion pointing out a methodological overview of the presented breath analysis technique and its shortcomings. It discusses the general biomarker properties and projects these onto the exhaled air biomarkers. Additionally several issues regarding the analysis and the biomarkers from exhaled air are described. Finally prospectives - regarding COPD heterogeneity and CF exacerbations - and perspectives are presented in short.

CHAPTER 9

Samenvatting

Introductie

De door technische en wetenschappelijke vooruitgang gedreven optimalisatie van analytische technieken heeft de laatste decennia geleid tot grote verbeteringen in de medische diagnostiek en de beschrijving van metabole pathways. Deze thesis beschrijft de ontwikkeling en validatie van een diagnostische tool gebaseerd op de analyse van componenten aanwezig in de uitademingslucht. Uitademingslucht is een complex mengsel van onder andere stikstof, zuurstof, koolstofdioxide en een groot aantal verschillende vluchtige organische verbindingen (VOCs). De huidige diagnostiek op basis van uitademingslucht is voornamelijk gebaseerd op aan- of afwezigheid van individuele componenten al dan niet aanwezig in de uitademingslucht. De lage sensitiviteit en specificiteit echter van deze huidige methoden vormt een drempel voor grootschalige toepassing van uitademingslucht analyses in de medische diagnostiek. Deze thesis geeft een beschrijving van de ontwikkeling en validatie van een meer nauwkeurige diagnostische tool op basis van uitademingslucht analyses. Er wordt zowel aandacht besteedt aan de manier waarop monsters verzameld worden en de chemische analyse alsmede de data-analyse en selectie van informatieve vluchtige organische verbindingen.

Samenvatting

Hoofdstuk 2 beschrijft de ontwikkelde methodologie van analyse van uitademingslucht. Zowel de instrumentele setup als de software en toegepaste statistiek worden uitgebreid behandeld. De analyse technieken zoals hier toegepast zijn gebaseerd op thermische desorptie en gas chromatografische scheiding van de aanwezige componenten. Uiteindelijk identificatie van de componenten werd uitgevoerd met massa spectrometrie (MS) in de vorm van time-of-flight MS. Na diverse data voorbewerkings stappen, als het verminderen van de ruis en uitvoeren van piekdetectie, zijn de overeenkomstige componenten aanwezig in de verschillende monsters vervolgens op basis van overeenkomst in retentietijd en massaspectrum aan elkaar gekoppeld. Dit resulteert in een database waarin alle in de uitademingslucht aanwezige componenten zijn opgenomen gekoppeld aan de mate van aanwezigheid in ieder monster. De voorgestelde methodologie is vervolgens getoetst op een populatie rokers/niet-rokers waarin informatieve componenten zijn geselecteerd. Vervolgens zijn classificatie algoritmen gebruikt om informatieve componenten te selecteren in de gegenereerde dataset. Het resulterende classificatie model gebaseerd op slecht 4 componenten was in staat om alle monsters correct te classificeren. De geïmplementeerde componenten waren reeds in de literatuur gerelateerd aan rookgedrag. Nadat de methodologie op deze manier is gevalideerd is een studie gestart om de onderscheidende VOCs van bacterie culturen te bepalen. **Hoofdstuk 3** beschrijft deze studie, waarbij de bovenstaande lucht van verschillende bacterie culturen in groei zijn gemeten. Het grote aantal verschillende VOCs aanwezig in die culturen zijn volgens de in hoofdstuk 2 uitgebreid beschreven protocol in een database onderge-

bracht en met de diverse statistische tools geanalyseerd. Zowel supervised als un-supervised learning algoritmen zijn toegepast om de data te verkennen. De verschillende culturen hadden een groot aantal in hoeveelheid significant verschillende componenten in de headspace aanwezig, wat differentiatie met SVM modellen tussen de culturen mogelijk maakte op basis van slechts een beperkt aantal componenten. Zelfs de 4 verschillende *Escherichia coli* stammen waren op basis van componenten aanwezig in de headspace te onderscheiden evenals de twee *Staphylococcus aureus* isolaten: methicilline-resistente *Staphylococcus aureus* en de methicilline-sensitieve *Staphylococcus aureus*. **Hoofdstuk 4** beschrijft toepassing van de ontwikkelde uitademingslucht-analyse methodologie op de inflammatoire luchtwegaandoening COPD. Uitademingslucht van 50 COPD patiënten en 29 niet-COPD personen werd verzameld en geanalyseerd. Een SVM classificatie model gebaseerd op slechts 6 componenten uit de uitademingslucht classificeerde 92 % van de monsters correct. Diverse validatie sets zijn vervolgens samengesteld, bestaande uit steroid-naive COPD patiënten om de invloed van medicatie te bestuderen en een externe validatie set bestaande uit monsters verkregen op andere locaties dan het Academisch Ziekenhuis Maastricht om de invloed van omgevingslucht nader te bestuderen. Classificatie van monsters in deze validatiesets resulteerde eveneens in een correcte classificatie van 92 %. De toepassing van de ontwikkelde methodologie in een klinische setting werd in **Hoofdstuk 5** voortgezet. Hier wordt gekeken naar de mogelijkheid patiënten met cystic fibrosis te onderscheiden van controle personen op basis van VOCs in de uitademingslucht. Statistische analyse resulteerde in een classificatie model gebaseerd op 10 VOCs in staat om 92% van de monsters correct te classificeren. Het bleek eveneens mogelijk om patiënten met en zonder *Pseudomonas aeruginosa* infectie te onderscheiden. In **Hoofdstuk 6** is gekeken naar de mogelijkheid om exacerbaties in patiënten met cystic fibrosis te voorspellen en te monitoren op basis van componenten aanwezig in de uitademingslucht. Het voorspellen en door tijdig handelen voorkomen of verminderen van exacerbaties is van groot belang omdat exacerbaties een belangrijke oorzaak vormen voor afname van longfunctie en kan leiden tot hospitalisatie van patiënten. Patiënten zijn met intervallen van 2 maanden bemonsterd en tijdens een exacerbatie werden 4 extra monsters afgenomen. Statistische analyse wees uit dat 18 componenten in significant verschillende hoeveelheden aanwezig waren in uitademingslucht van patiënten die 3 tot 8 weken na de meting een exacerbatie doormaakten uitgezet tegen patiënten die geen exacerbatie doormaakten. Vijftien componenten waren in significant verschillende hoeveelheden aanwezig in uitademingslucht van patiënten vóór vergeleken met tijdens een exacerbatie. Daarnaast zijn met SVM analyse de meest informatieve componenten geselecteerd welke gecombineerd in een classificatie model 93% van de exacerbatie monsters correct konden classificeren. De studie toont aan dat markers in uitademingslucht het mogelijk maken om exacerbaties bij patiënten met cystic fibrosis weken voor de daadwerkelijk exacerbatie te voorspellen. Er dient echter rekening mee gehouden te worden dat de totale patiëntenpopulatie beperkt was en validatie van de resultaten noodzakelijk zal zijn. Bijvoorbeeld door de modellen en componenten te toetsen in een

grotere validatie studie. Ten slotte geeft **Hoofdstuk 7** een beschrijving van de algemene discussie waarbij de nadruk ligt op de gekozen randvoorwaarden van de beschreven methodologie gebaseerd op analyse van uitademingslucht en de huidige tekortkomingen van deze methodologie.