Diagnosis of pulmonary injury and infection by exhaled breath analysis
Bos, L.D.J.

Citation for published version (APA):
Bos, L. D. J. (2014). Diagnosis of pulmonary injury and infection by exhaled breath analysis

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 10.

General discussion

Published in part in
Frontiers in Genetics 2014

Published in part in
Intensive Care medicine 2014

Lieuwe DJ Bos
Introduction

Ventilated patients in the intensive care unit (ICU) can present with, or develop pulmonary injury (acute respiratory distress syndrome; ARDS) and/or pulmonary infection (pneumonia) [1-3]. ARDS and pneumonia are both known to be associated with an inflammatory response [4, 5]. In animal models, a response can be observed in the systemic and pulmonary compartment. In samples from patients, local markers of the innate immune response seem to allow for better discrimination with a clinically relevant critically ill control group than systemic markers [6-8]. Thus when investigating biological markers of ARDS or pneumonia sampling from the pulmonary compartment should be preferred above blood sampling. However, it remains difficult to obtain a sample from the lungs with broncho-alveolar lavage, as this can introduce burden and risk and certainly cannot be performed very frequently. Therefore, the following central postulate was at the basis of this thesis: there is a need for a non-invasive tool that can assess pulmonary molecular markers in order to predict and diagnose ARDS and pneumonia in intubated and ventilated ICU-patients [9-15].

Besides several abundant molecules such as nitrogen, oxygen and carbon dioxide, breath also contains volatile organic compounds (VOCs) at very low concentrations (parts-per-billion). These molecules represent (fragments of) metabolic products and can be of systemic origin and transported to the lung via the circulation or can be produced locally in the lung [16]. Several mechanisms could lead to the formation of VOCs. Some are produced with ketone formation and other processes in central metabolism [17]. Oxidation of fatty acids may result in alkanes, alkenes and aldehydes [18]. In vitro experiments suggest that leukocytes can produce VOCs directly [19]. Nitrogen and sulfide containing organic molecules are normally eliminated via the liver and the kidney and their concentration may rise in the breath if the clearing function of these organs is limited [20, 21]. Micro-organisms also contribute to the exhaled VOCs. Gut bacteria produce volatile compounds with fermentation. More relevant for pneumonia, bacteria in the lung may also produce VOCs [22] either or not affected by interaction with the host. Finally, many compounds in the breath may be derived from the environment; they were previously inhaled, represent metabolites derived from food or drinks or reflect...
volatile drugs or drug metabolites [23, 24].

VOCs can be detected in various manners. In this thesis we focused on two distinct approaches: a) gas-chromatography and mass-spectrometry for the separation, identification and quantification of potential volatile biomarkers and b) electronic nose analysis for the recognition of complex VOC mixtures by capturing “molecular fingerprints”, which present patterns for ARDS or pneumonia and may be used for empirical classification. To allow for breath analysis in ventilated ICU-patients, a new, simplified breath collection system needed to be developed. The latter technology was used to identify volatile biomarkers in breath that may be useful for early recognition and diagnosis of ARDS and pneumonia. We also investigated the accuracy of a commercially available electronic nose for the empirical diagnosis of ARDS and ventilator-associated pneumonia (VAP). The results on ARDS and pneumonia were described in two separate parts of the thesis and will be discussed separately hereafter (part I on page 213, part II on page 225), followed by the general implications for breath analysis in ventilated patients (part III on page 231).

**Part I: Acute respiratory distress syndrome**

**Summary**

A new, simplified breath collection method was developed that allows for safe and quick collection of breath in mechanically ventilated patients (chapter 3). A side-stream flow allows for the trapping of VOCs onto sorbent material that can be used for GC-MS analysis. Alternatively, direct analysis with an electronic nose can be performed (chapter 4). The method provided reproducible results. Several VOCs were identified that emenated from the ventilator and tubing in high concentrations. These VOCs could *a priori* be excluded as potential biomarkers. In two animal experiments (chapter 5) in which we administrated lipopolysaccharide (LPS) intravenously or intra-tracheally we found that VOCs in the breath of rats change with the development of pulmonary inflammation. Several VOCs overlapped between the two animal models for lung injury and may therefore be considered candidate markers for an inflammatory response. Interestingly, hexanal, pentadecane and 6,10-
dimethyl-5,9-undecadien-2-one were found in lower concentrations after LPS instillation. This was unanticipated as these VOCs are typically associated with oxidative stress, potentially inducing an increase after LPS administration. In chapter 4, we described that breath analysis with a commercially available electronic nose can be used to discriminate patients with ARDS from patients without ARDS with moderate accuracy. The diagnostic accuracy increased with increasing severity of ARDS. However, we had to conclude that as yet the currently used eNose technology is not applicable in clinical practice, because the test accuracy was too low and the sensors showed marked drift over the inclusion period. In chapter 6 we used a similar study design as in the previous chapter but applied gas-chromatography and mass-spectrometry instead of electronic nose analysis as the index test. We found that the combined signal of three metabolites (octane, 3-methyl heptane and acetaldehyde) in the breath could be used to discriminate patients with ARDS from patients without ARDS with good accuracy. This accuracy was maintained in a temporal external validation cohort, which limits the chances that discrimination resulted from a type-II error. Furthermore, the exhaled breath signal improved diagnostic accuracy on top of a clinical prediction score (the lung injury prediction score; LIPS).

Taken together, these data demonstrate that specific VOCs in the exhaled breath can be used to discriminate patients with ARDS from patients without ARDS, suggesting that exhaled breath analysis may be used for the diagnosis of ARDS.

### Bayesian intermezzo

This section serves the purpose to explain the elaborate description of biological mechanisms that possibly lead to the formation of VOCs on the following pages. The research described in this thesis is performed with an “omics” perspective, meaning that no target markers were selected *a-priori*. That approach is frequently criticized with arguments such as “the fishing expedition”, referring to the chances on type I errors with this kind of research. Therefore, it is highly important for the community performing these studies, to be maximally self-critical. As described in every chapter, we have to
limit the chances of false-discovery (thus type I errors) by means of statistical procedures such as cross-validation and confirmation studies in independent cohorts (external validation).

In this intermezzo I want to carefully make the point that this may not be sufficient. All studies described in this thesis and most studies published in literature use variations of Fisher’s statistics and the inevitable p-values. Although most chapters focus on effect sizes and not statistical significance per se, the a priori (before experiment) probability of the hypothesis being true were never taken into account. However, with an a-priori 19 to 1 odds against the hypothesis (the long shot) even a p-value of 0.01 gives a post-test probability of only 30% on a “real” effect. An a-priori odds of 1:1 (coin toss) gives a post-test probability of 89% using the same p-value. Only a good bet (9:1 odds) gives a 99% post-test probability with p-value of 0.01 [25]. Why is this important? Because we make no assumptions on pre-test probability in “omics” research and typically apply the same cut-off to identify significantly associated markers, our conclusions about post-test probability ought to be heavily influenced by biological plausibility [26]. For example, if breath analysis reveals two compounds that are similar in discrimination for ARDS (same area under the receiver operating characteristics curve) and reject the NULL hypothesis with the same confidence (similar p-value and confidence interval) but one compound is an aromatic, fluorated compound whereas the other is a C8 alkane (octane) the post-test probabilities would differ dramatically. As aromatic fluorated compounds are not known to be produced by human metabolism the pre-test probability is close to 0 and even a very small p-value will increase that probability to not more than 10-30%. Alternatively, octane is known to be produced by peroxidation of fatty acids (see following section) and there is literature that involves that process in the pathogenesis of ARDS. Therefore, the pre-test probability was good for that compound and this biological marker should be favored for additional experiments. Importantly, the systems biologists within our community will certainly criticize these clauses. They state that if a priori probability would always be taken into account, radically new things would never be discovered.
That reaction is valid to a certain extend. However their defence is also partly based on two errors in the estimation of probability: availability (overestimating the frequency of vivid or easily recalled events) and probability transformations (small probabilities are overweighted and large probabilities are underweighted) [27]. In fact, this Bayesian-like approach could allow for radical discoveries. If something has a very low pre-test probability but is shown in several replications within different populations, the likelihood for a true positive finding steadily increases. In other words, to find the middle ground before continuing, this intermezzo describes a moderate view on things. Let us not discard hypotheses directly that have a very high a-priori probability if we fail to reject the NULL hypothesis, but let us also not claim things if we find a low p-value, without careful consideration of proper replication and biological plausibility.

**Biological mechanisms**

*Animal versus patient studies*

In the present thesis, the volatile organic compounds that are associated with ARDS were investigated in a translational setting. However, the VOCs that were found in the animal model were different from those found in the human studies. As mentioned in the summary, the decreased concentrations of hexanal, pentadecane and 6,10–dimethyl–5,9–undecadien–2–one in the animal study were peculiar as the opposite would be expected with increased oxidative stress. In the interpretation of these findings, the mild nature of the lung injury found in our animal model should be taken into account. Although we did observe marked inflammation in the animals that received LPS, markers of permeability were not different between the two groups. Additionally, we analyzed the changes in of breath VOCs after LPS administration in the animal studies but in the human study we took a single sample in established ARDS. Therefore, direct comparison of these findings is difficult. In hindsight, it could be argued that an animal model with severe lung injury could have provided more information in this stage. We could conclude from the animal model that LPS administration results in decreased exhaled concentrations of hexanal, pentadecane and
6,10-dimethyl-5,9-undecadien-2-one. However, in patients these VOCs may very well not be the best markers for discrimination between two patient populations of which one developed (diagnosis) or will develop (prediction) ARDS. In that sense, the animal study addressed the issue of monitoring whereas the human studies focused more on diagnosis.

**Lipid peroxidation**

Octane, a C8 unbranched hydrocarbon, discriminated best between patients with ARDS and without ARDS. Octane was higher in patients with ARDS. Lipid peroxidation of oleic acid (C18:1 n-9) is the most likely mechanism by which octane is produced in the body. As shown in figure 1, that mechanism is highly predictable based on the chemical structure of the fatty acid. Besides octane, several other hydrocarbons (heptane, 1-nonene and 1-decene) would result from peroxidation of oleic acid [28]. Other relatively abundant fatty acids give rise to similar VOCs, that could also be detectable in the breath of patients (e.g. hexane and 1-octane as a result of n-7 unsaturated fatty acid peroxidation and pentane from peroxidation of n-6 unsaturated fatty acids, such as linoleic acid). Two pathophysiological mechanisms can result in the observed increased concentration of octane in the breath of patients with ARDS. The first postulate is that ARDS increases oxidative stress, which increases the likelihood that fatty acids are peroxidized and thereby leading to all alkanes (and other peroxidation products) being increased in ARDS. The second, alternative, postulate is that the concentration of oleic acid is increased in ARDS and, irrespective of the peroxidative flux, thereby increasing the concentration of octane. A third possible explanation might be decreased degradation of alkanes. Humans cannot metabolize alkanes, but some micro-organisms can [29]. However, these enzymes are not specific for octane and typically cover a spectrum of C5-C10 [29]. It should be stated that these hypotheses are not completely mutually exclusive. We were able to test these hypotheses by a combination of post-hoc analysis and literature review. If the first or third scenario were true, other alkanes would also be different between ARDS and controls. Therefore, a post-hoc analysis was performed for hexane and pentane in the study that was described in chapter 6. Figure 2 shows that the concentrations of alkanes not produced by oleic acid peroxidation were not significantly increased in ARDS, thus rejecting the first hypothesis.
Figure 1: Peroxidation of fatty acids to alkanes and alkene

Fatty acid (C18 n-9) Oleic acid

\[ \text{CH}_2-(\text{CH}_2)_7-\text{CH}=\text{CH-(CH}_2)_7-\text{COOH} \]

10-Hydroperoxide 8-octadecenoic acid

\[ \text{CH}_2-(\text{CH}_2)_7-\text{CH}=\text{CH-CH-(CH}_2)_6-\text{COOH} + \cdot\text{OOH} \]

Alkoxy radical

\[ \text{CH}_2-(\text{CH}_2)_7-\text{CH}=\text{CH-CH-(CH}_2)_6-\text{COOH} + \cdot\text{O} + \cdot\text{OH} \]

Figure 2: Peroxidation products in ARDS

The upper row displays the concentration of alkanes and alkenes that are produced during the peroxidation of oleic acid. The bottom row for products of peroxidation of n-7 and n-6 unsaturated fatty acids. The concentration is given in pg/L.
The alternative scenario would result in increased concentrations of heptane, 1-nonene and 1-decene in ARDS. Indeed, the concentration of these three VOCs was increased in patients that suffered from ARDS, albeit not as markedly as for octane. These results are in line with the hypothesis that the high exhaled octane concentration that we observed in ARDS is most likely explained by an increased concentration of oleic acid.

*Oleic acid*

A review of the literature reveals that an increased plasma concentration of oleic acid (also called oleate and C18:1 n-9 in the literature) can be found in ARDS, even before its clinical manifestation [30, 31]. Actually, that was the motivation for testing lisofylline, a reducer of free fatty acids [31, 32], as a pharmaceutical treatment in patients with ARDS [33]. Unfortunately, the latter study showed no survival benefit for the intervention group [33]. As stated in the accompanying editorial [34], that might not be unexpected as the intervention was given 36 hours after the patient met the consensus criteria while no animal study showed benefit when administered more than 4 hours after the initial insult [35]. Lisofylline did also not decrease the plasma concentration of a range of fatty acids, including oleic acid [33]. These findings imply that if oleic acid is involved in the pathogenesis of ARDS it is probably in the early phases of disease.

In experimental studies, intravenous administration of oleic acid is a well-known model for lung injury that shows several pathological signs of human ARDS, such as hyaline membranes [36]. However, the dose that is injected is several orders higher than the natural occurring concentrations and mimics the situation of a fat embolism, a very rare cause of ARDS in humans. One paper reports the effects of oleic acid in plasma concentrations that can be observed in patients at risk for ARDS on type I alveolar cells [37]. Oleic acid inhibits Na-channels and Na/K/ATPase *in vitro*, which results in less fluid clearance from the lungs and could contribute to the exudative phase of ARDS. These results were confirmed in an *ex vivo* mouse lung [38]. Specific inhibition of Na/K/ATPase by ouabain also resulted in alveolar permeability, although less pronounced than with oleic acid injection. Thus, there is probably a strong role for Na/K/
ATPase in the mechanism by which oleic acid induced pulmonary edema, but additional mechanisms could contribute. Leukotriene B4 (LTB4) and prostaglandin E2 (PGE2) are suggested as a down-stream coupling with inflammation and thereby to injury [38]. Indeed, LTB4 and PGE2 were found to be elevated in trauma patients that were to develop ARDS [39]. However, the mechanism by which PGE2 is up-regulated are unclear as oleic acid does not increase production directly through arachidonic acid [40] but rather via production in the lung itself [41]. Taken together, there is strong evidence for a contribution of oleic acid to the development of ARDS, at least in animal models of lung injury and possibly in human ARDS. Together with the biochemical mechanism that links octane to oleic acid, this may imply that we are able to non-invasively measure a very early pathophysiological mechanism in the development of ARDS.

Several steps have to be taken before breath octane may be used as a clinical predictor for ARDS. The implications for clinical trials will be discussed in the next section. With the currently available data from the study that is described in chapter 6, we can already try to formulate basic assumptions for future studies. In the context of the concurrent MARS (Molecular Diagnosis and Risk Stratification of Sepsis) study, aimed to find biological markers of sepsis and organ failure in ICU-patients, we had cytokine levels in plasma and mini-BAL in ventilated ICU-patients (n=93 for plasma and n=13 for mini-BAL) to our disposal. Figure 3 shows the correlation between several peroxidation products of oleic acid (octane, heptane and 1-nonene) and products of other fatty acids (hexane and pentane) with plasma and pulmonary markers of inflammation. None of the breath alkanes show a correlation with systemic inflammation. Pulmonary inflammation is correlated with several breath alkanes. The products of oleic acid peroxidation show a marked inverse correlation with IL-6 and a positive correlation with IL-10, which suggests that oleic acid peroxidation is associated with anti-inflammatory processes.
**Figure 3:** Correlation between alkanes/alkenes and inflammatory markers

The other alkanes also show negative associations with other pro-inflammatory markers, primarily with IL-8. Although these pre-liminary results do not provide us with a mechanistic coupling between fatty acid peroxidation and inflammation, we can safely conclude that there is no association with systemic inflammatory response. Therefore further efforts should focus on inflammation in the lung. Additional measurements in these samples will give us a profile of fatty acids in the plasma and in the lung. If the correlation between absolute oleic acid concentration and breath octane concentration that we postulated on biochemical grounds can be confirmed, we can more confidently state that octane is indeed a surrogate for oleic acid and not for lipid peroxidation *per se*.

**Future clinical implications**

In chapter 6 we showed that exhaled breath analysis provides information on the presence or absence of ARDS and this is complementary to pre-test risk assessment by the lung injury prediction score (LIPS). This may
not be unexpected as the LIPS consists of clinical risk factors, whereas exhaled breath analysis is based on biochemical alterations that occur during ARDS. Hence, a patient can have a very high \textit{a priori} risk for disease (LIPS), but in absence of the biochemical presentation we may still conclude the patient probably does not have that specific condition (yet). In line with that, if a patient exhibits the biochemical profile with a negligible risk for disease such patient probably does not have the condition either. This may have implications for targeted therapies (figure 4). Until now, most pharmacological (targeted) interventions have been tested in a patient population that was selected based on clinical characteristics. It should be noted that every single pharmacological trial in patients with established ARDS was negative [42].

Our results further extent existing evidence that the clinical syndrome diagnosis of ARDS is not a good surrogate for specific biochemical alterations. However, as long as we cannot measure biological markers that sufficiently represent the biochemical pathway that the pharmacological agent intervenes with it remains to include the patient population that benefits most in clinical trials. Of note, an intervention that is highly effective in a subpopulation but slightly harmful in most patients will probably result in a negative (non-inferiority) clinical trial when an unselected cohort is included. Additionally, prevention of ARDS has probably more potential than treatment of established ARDS. If the assumption that oleic acid is a central mediator in the pathogenesis of ARDS in some patients will show to be correct and octane is a good marker of oleic acid peroxidation, this could be the first molecular marker that can be measured non-invasively, rapidly and at the bedside.

Pre-clinical research could help to further understand the mechanisms that couple oleic acid peroxidation and octane formation, via Koch’s postulates [43, 44]. Questions about dynamics and the potential of therapeutic intervention need to be addressed to allow for optimal design of clinical studies. First of all, octane formation after infusion of oleic acid has never been confirmed \textit{in vivo}. If such biochemical mechanism is confirmed, several \textit{in vitro} and \textit{in vivo} experiments should be conducted to evaluate the dynamics of that association.
Figure 4: Model for development of ARDS and timing of therapeutic intervention.

Relevant questions are whether oleic acid injection is always followed by alkane formation, what is the delay and can the association be uncoupled in certain conditions (excess super-oxide formation)? Most of these questions can be answered by straightforward experimental studies, with the injection of oleic acid and frequent measurement of exhaled octane concentrations. When we want to separate the contribution of endogenous and the infused oleic acid, the infused fatty acid could be labeled with C13, after which octane originating from that fatty acid can easily be separated by means of gas-chromatography and mass-spectrometry.

The influence of therapeutic interventions could also be measured in these models. Despite the negative results of clinical trials, lisofylline and beta agonists may be re-evaluated. Lisofylline lowers the concentration of free fatty acids and could thereby be used as a targeted therapy in...
patients with high exhaled octane concentrations (and thus, if the previous assumptions are confirmed, high oleic acid concentration). If given before alveolar flooding occurs, lowering of the serum oleic acid concentration may prevent inhibition of the Na/K/ATPase and thereby the lung may be able to compensate for increased exudate formation in the lung and stop alveolar flooding. Beta agonist can regulate Na/K/ATPase via MAPK/ERK pathways, without the influencing fatty acid concentrations [45] and could thereby increase alveolar fluid reabsorbance. Despite promising pre-clinical studies, randomized controlled trials in patients with established ARDS did not show benefit of beta agonists [46]. ARDS was also not prevented by beta agonists in a high-risk surgical population [47]. However, the latter trial did show less extra-vascular lung water in the intervention group. The regulation of Na/K/ATPase by beta agonist also makes this result biologically plausible. Before dismissing these two simple, biologically plausible preventive measurements, post-hoc analysis may provide additional evidence if a subgroup does benefit from the intervention. Subsequently, that can be prospectively evaluated in a new clinical trial.

In line with the previous comments on pharmacological prevention of ARDS, biological markers have most potential if used for the prediction of ARDS. Therefore, the diagnostic accuracy we found in chapter 6 should be confirmed in large, multi-center, longitudinal observational study. A cohort of patients at risk for ARDS should be included. Samples should be collected frequently so that the moment of biochemical presentation of ARDS is not missed. Two settings can be anticipated. First, mechanically ventilated patients undergoing major surgery or admitted to the ICU are at high risk to subsequently develop ARDS. These patients could be monitored using the sample methodology that is described in chapter 3. Another patient group that is at high risk are those admitted to the emergency department, with a lung injury prediction score above three. Sample collection is more cumbersome in this population as it currently requires forced exhalation in a sample collection bag; something that is unpleasant for dyspnoeic patients, time consuming and requires a special infrastructure. On page 217 of this discussion other methodologies for the direct analysis of breath are described.
To summarize, patients with clinical risk factors for ARDS could benefit most from preventive interventions, especially if the nature of the “second hit” could be detected with biological markers. Breath analysis of octane may capture oxidative stress and the process of oleic acid damaging the Na/K/ATPase leading to less alveolar fluid clearance and could hereby provide information for targeted treatment.

Part II: Ventilator associated pneumonia

Summary

We, eukaryotes, are biochemically just a differentiated fusion of two prokaryotes [48]. Bacteria are versatile biochemical factories [49]; they can obtain energy from almost any process perceivable [50]. We hypothesized in chapter 7 that the volatile metabolic products of bacteria may be used to identify the presence of any pathogen and of specific strains of pathogens. We systematically reviewed the literature on associations between volatile organic compounds and six of the most common pathogens on the intensive care unit. All six bacteria of interest produce isopentanol, formaldehyde, methyl−mercaptan and trimethyl−amine. The following volatile biomarkers were found for identification of specific strains: isovaleric acid and 2−methyl−butanal for *Staphylococcus aureus*, 1−undecene, 2,4−dimethyl−1−heptane, 2−butanone, 4−methyl−quinazoline, hydrogen−cyanide and methyl−thiocyanide for *Pseudomonas aeruginosa* and methanol, pentanol, ethyl−acetate and indole for *Escherichia coli*.

The major limitation of all included studies is that evaluation of the VOCs was done in vitro without providing the bacteria with a growth medium that partly represents in vivo conditions. Therefore, we evaluated the diagnostic accuracy of VOC analysis for the diagnosis of VAP without using culture media in chapter 8. In this study, tracheal aspirates were collected every three days from ventilated patients without signs of pneumonia. Some of these patients developed VAP and were included as cases. Tracheal aspirates were thawed and analysed by electronic nose and GC-MS. The electronic nose analysis showed that patients with VAP could be discriminated from patients without VAP and that the “VAP signal” already
increased before the moment of clinical diagnosis. However, patients with colonization but without signs of infection were classified as having no VAP in this study. This suggests that the in vivo markers of VAP are not related to bacterial presence or growth per se. That could mean the VOC signature that was detected by the electronic nose is mostly attributable to host response, or that bacteria produce other VOC profiles during invasive growth. In that study, we also attempted to identify the VOCs that were higher/lower during VAP. However, the GC-MS analyses were frequently unsuccessful due to absorption of water on the Tenax material and the extremely wide range of concentrations of VOCs. Some samples heavily overloaded the mass-spectrometer, whereas others showed only very low intensity peaks. In chapter 9 we further evaluated the association between VOCs, pneumonia and bacterial presence, albeit in vivo. In this study, breath analysis was performed following the methodology that was described in chapter 3. Patients were separated in those that had a definite pneumonia at admission and those without any signs of pneumonia. 1-Propanol was found to be significantly lower in the breath of patients with pneumonia than in those without pneumonia. Patients with positive and negative cultures of tracheal secretes were also compared. 1-Pentanol and heptanal were found to be significantly lower in the breath of patients with positive cultures. Finally, VOCs in breath were correlated with markers of inflammation measured in plasma. None of the VOCs passed the threshold for statistical significance for the correlation with IL-1b, IL-6, TNFa and GM-CSF. We observed significant correlated between several VOCs and IL-8, IL-10, IL-13 and interferon-gamma. The positive correlation with interferon gamma was most profound.

Taken together, these data demonstrate that specific VOCs are produced by different bacterial strains, but that these results are not easily translatable into a clinical test for the diagnosis of pneumonia.

**Biological mechanisms**

Our review described in chapter 7 showed that there was sufficient theoretical ground to hypothesize that bacteria contribute to the VOCs in exhaled breath. We did observe changes in the VOC (profile) during pneumonia (chapter 8 & 9), but these could not be contributed to bacterial presence alone. The simplistic view of: bacteria produce VOCs
there are more bacteria during pneumonia – hence pneumonia can be detected by VOC analysis does not capture the complexity of our results. Therefore, the theoretical framework for pneumonia detection by means of exhaled breath analysis has to be reconsidered.

In vitro vs. in vivo

Almost all studies on the association between VOCs and bacteria have been performed in vitro with reference strains [22]. Several important factors limit the translation of the results of these studies into the clinical. Firstly, not all studies used the exact same subtype of bacterial species. In one study, phage types of SA influenced headspace volatile organic compounds [51]. Genomic variation between subtypes could result in differences in efficacy of enzymes within a specific metabolic pathway. Certainly, this could hamper the clinical applicability of volatile biomarkers for strain identification. Secondly, the growth medium is the source of building blocks for the produced VOCs and may incude phenotypic alterations in the micro-organisms and it therefore is a confounding factor [51-54]. Without knowledge of the substrates that are available to the pathogen it is uncertain whether an association likely to be observed in vivo. In line with that, most studies do report on decreased concentrations of VOCs with bacterial growth, and this could be very relevant in clinical studies as observed in chapter 9. Thirdly, measurements were obtained at different moments in the growth of bacteria. Depletion of metabolites and growth phase (log or stationary) influence headspace metabolites [55-58]. Lastly, patient samples are less well defined than laboratory produced cultures of reference strains and are different in the following aspects: CFU’s, growth phase, host response, viscosity [59], confounding co–morbidities and medications (e.g. antibiotics [60]). Taken together, these limitations suggest that direct translation of in vitro results to in vivo testing will be difficult. This may explain some of the differences we observed.

Imperfect reference standard in clinical trials

Another important difference between most in vitro studies and samples collected from patients is that a known amount of a specific bacterial strain is added to a growth medium in the former, whereas an unknown
amount of an unknown bacterial strain is present in the latter. Therefore, the reference test is highly important in studies with clinical materials. We used bacterial culturing methods to first grow and subsequently identify the most common bacteria in the airways. However, it is now widely recognized that there is a wide taxonomy of bacteria present in the lung under normal conditions as well, which cannot be cultured and which may change during disease [61-63]. Therefore, we can hypothesize that our reference standard does not capture the process that we measure with VOC analysis; namely the metabolism of all bacteria in the lung. Several novel, PCR based technologies allow for untargeted analysis of the lung microbiome. This results in a number of copies per bacterial strain (frequently based on the sequence 16S) and can thus be used to evaluate the composition of the bacterial ecology in the lung. Bacterial metagenomic analysis not only replicated and sequences 16S RNA but also detects other bacterial DNA sequences [64]. Therefore, this method can be used to identify whether the bacteria in a sample have the genomic potential to take certain metabolic pathways. This approach was combined with breath analysis in patients with cystic fibrosis. Increased concentrations of 2,3-butanedione were observed in patients with cystic fibrosis and this was linked to an increased number of sequences encoding for proteins that are involved in the biosynthesis of 2,3-butanedione [64]. In combination with data from in vitro experiments and the abundance of different bacterial species (assessed by 16S replication) the authors concluded that Streptococcus species are most likely responsible for the production of 2,3-butanedione. In conclusion, future clinical studies on the association between bacteria and VOCs may benefit from novel sequencing based technologies that allow for a untargeted analysis of the pulmonary microbiome and can quantify the presence of certain sequences encoding for proteins involved in the biosynthesis of these VOCs.

Interplay between bacteria and the host

Bacteria are mostly investigated in mono-culture during in vitro or in vivo experiments. In the clinical situation however, bacteria grow within the ecology of the lung. Within that setting, volatile organic compounds may be used to interact with other bacteria or with the host. For example, several volatile metabolites are known to be involved in quorum sensing [65, 66]; a system that coordinates bacterial gene expression as a consequence of...
ecological factors such as population density and substrate availability. Furthermore, bacteria can influence the metabolism of the host for their own advantage [67], which could in theory also result in other volatile metabolic products.

*An updated theoretical framework*

The VOCs observed in the headspace of mono-cultures of bacteria give an insight into the volatile metabolites that these bacteria can biosynthesize. However, the VOCs that are produced are influenced by growth phase, antibiotic pressure, the availability of substrates and the biochemical pathways that are encoded in the genome of that specific pathogen. Furthermore, VOCs can be used to communicate between bacteria or may be influenced by the host response. Thus, the VOCs measured in the breath of patients with pneumonia could result from bacterial metabolism or host response and are likely influenced by the growth phase, the bacterial ecology and the genome of the pathogen (figure 5). This is the good news for discovering complex mechanisms in host-pathogen interaction, but the bad news for developing unambiguous test in clinical diagnosis and monitoring.

**Figure 5:** Updated theoretical framework (alveolar environment)
**Future clinical implications**

The theoretical framework described in the previous section has important consequences for future clinical trials. First of all, two explicit separate aims should be differentiated: studies that focus on diagnosis and clinical decision and pathophysiological studies that further explore the associations between VOCs and specific processes within the pulmonary microbiome.

*Diagnostic test for pneumonia*

Studies focused on diagnostic accuracy of a breath test should include only patients with clinical suspicion of the disease (intention to diagnose), where a clinical decision is to be taken (do or do not administer antibiotics) [68]. The reference test should be highly reliable in these patients. For example, if patients with suspected VAP are to be included quantitative cultures of broncho-alveolar lavage fluid could be used as reference standard. This is in stark contrast to studies (such as that described in chapter 9) that use semi-quantitative cultures in a non-selected population as reference standard. These type cultures are far more likely to become positive and the unselected nature of the population increases the probability of (false-)positive cultures, that are clinically irrelevant. For a detailed description of the steps that should be taken with discovery, validation [69] and clinical implication [70] see the section on page 222.

*Association between VOCs and microbiome*

Studies that aim for the discovery of VOCs that represent a specific process within the microbiome of the lung should focus on a complete description of the phenotypic characteristics of the patient that is sampled, with special emphasis on the microbiome that he/she carries around. This includes but is not limited to careful sampling, elaborate analysis of the micro-organisms that are present in the lung and further differentiation of the metabolic pathways that are available to those micro-organisms. The design of such a study could be case-control as the question of diagnostic accuracy is not addressed.
Part III: Exhaled breath analysis

Gas-chromatography and mass-spectrometry

Gas-chromatography and mass-spectrometry is widely recognized as the gold standard for the detection of volatile organic compounds and the discovery of biomarkers in breath [71]. GC-MS can be used to measure a broad range of compounds, for identification of unknown compounds and is semi-quantitative. However, several limitations of the technology should be noted. First, because the technology is mostly not available in the hospital, the sample needs to be stored, for example on an adsorption tube, which will always alter the constitution of the mixture. Second, guiding the sample over one chromatographic column will not separate all VOCs in retention time. As a consequence, low abundance VOCs can be missed and mass spectra are not derived from a single compound, which limits the identification of the compound. Furthermore, choices in absorption material, packing of the column and mass-spectrometer bias the measurement towards certain types of VOCs. For example, in our studies we typically used Tenax as a sorbent. VOCs with a carbon backbone of C4 or shorter do typically not retain very well on Tenax and several very volatile compounds without carbon do not retain at all (e.g. Ammonia). Therefore it may not be surprising that our results typically point to relatively long alkanes as biomarkers. The described limitations do not even include the problems with maintaining a constant retention time per compound, adjustment of the sensitivity of the system and interpretations of the data. Because we lack an alternative for volatile biomarker discovery, gas-chromatography and mass-spectrometry will probably remain the standard for future studies that focus on the identification of volatile metabolites that are associated with certain disease states. The use of more than one GC-MS platform could allow for validation procedures and an extended range of covered metabolites by using partly orthogonal set-ups. Such an approach has been used in clinical labs that specialize in environmental monitoring and also provide excellent, reproducible results in breath research [72-75]. For validation studies, other technologies may be more attractive than GC-MS. These are discussed in the paragraph on “alternative technologies”.
Electronic nose

As previously stated, electronic nose analysis is attractive from a clinical point of view as it can be available at the bedside and provide results instantly. It does not allow for the measurement of individual VOCs, but uses pattern recognition to capture composite VOC mixtures by cross-reactive sensors, thereby identifying clinically relevant groups of patients. Humans can discriminate approximately 1 trillion olfactory stimuli with only 400 cross-reactive receptors [76, 77], which is the highest resolution of any human sense. For breath analysis, sensitivity towards very low concentrations is equally important as discrimination between different gaseous mixtures. Therefore, a sensor array should meet two important criteria: the sensors should have orthogonal chemical selectivities and should be responsive to a wide range of concentrations (low parts per billion to high parts per million).

In chapter 8 we described the discrimination between tracheal aspirates from patients with and without VAP with an electronic nose. The diagnostic accuracy was good in that study. However, breath analysis in patients with and without ARDS provided only moderate discrimination, as shown in chapter 4. Accuracy did increase when we only considered patients with moderate/severe ARDS as cases. Based on the GC-MS study in patients with ARDS (chapter 4) and the literature study of VOCs that could be used for the diagnosis of pneumonia (chapter 7) we can postulate that pneumonia is a disease that is more likely to be discriminated by eNose analysis than ARDS. To explain that, we need to assume that sensor technology will be further developed to the border of what is technologically feasible. We will make three very specific assumptions to create an interpretable scenario, although it has to be stated that even this scenario is not technologically possible at the moment. First, sensors have a very large orthogonal chemical selectivity; each sensors reacts with a separate functional group but cannot deduce the length of the carbon backbone as this would require analysis of the mass. Second, sensors are unlimited in sensitivity. Third, we assume a limited number of VOCs: a 12x20 matrix, with carbon backbone length on the y-axis and functional chemical group on the x-axis, resulting in 188 VOCs (52 combinations between carbon backbone and functional group are chemically impossible). Figure 6 shows that the breathprint is expected to be more different in pneumonia than
in (early) ARDS especially if we keep the naturally occurring intra-person variation due to genetics and environmental exposure in mind [78].

**Figure 6:** Hypothetical eNose breathprint

The top figure shows the matrix of potential biological markers that were considered in this hypothetical example. The x-axis shows the different functional groups (= double binding, C carbon, N nitrogen, H hydrogen, S sulfide, "1" gives the location of the functional group). The y-axis shows the carbon backbone (Arom = aromatic compound, for this example only one type). The lower left figure shows the breathprint that was identified for ARDS based on the GC-MS results in chapter 6. The lower right figure shows the breathprint of that was identified for bacterial pneumonia based on the literature study described in chapter 7. Even with a perfect chemically orthogonal sensor array based on these two dimensions the discrimination of pneumonia will outperform that of ARDS.
Multiple solutions are thinkable. For ARDS we could focus on other analytical technologies that allow for identification of single compounds or possibly increased understanding of olfactation in animals may provide new insights that can further enhance detection of VOCs by means of different sensor arrays. For example, most electronic nose technology relies on van der Waals forces and physical absorption [79]. These are the weakest and least selective of all intermolecular interactions. Novel concepts, such as optoelectronic noses based on chemoresponsive colorants utilize the chemical reactivity of the VOCs and thus allow for more sensitive and selective detection [79].

**Alternative technologies**

Several other technologies, besides GC-MS and eNose analysis, may be attractive for clinical validation studies in the near future. All technologies described here aim to detect a single compound. If a technology does not aim at that, I could include it in the section on electronic nose analysis as it is not relevant what type of “sensor input” is provided for pattern recognition. It should be noted that different advantages and limitations than those previously discussed could apply to those systems.

In GC-MS, chromatography separates the VOCs based on boiling point and chemical interaction with the stationary phase. If mass-spectrometry is to be used alone, separation has to be performed by another method. Additionally, electron ionization is traditionally used to generate ions. However, as described in chapter 2, this generates tens of fragments per VOC. Thus without chemical separation, this method of ionization would lead to chaos. Several methods have been used to overcome the problems of separation and ionization.

Proton transfer-reaction mass-spectrometry (PTR-MS) uses protons for chemical ionization [80]. This increases the mass of the VOC with one and adds a charge, which can be used for mass-spectrometric detection. Thus the molecule is not fragmented. Because multiple VOCs have the same nominal mass, they cannot be separated with this method. One solution is the use of time-of-flight mass-spectrometry (TOF-MS), which can more accurately separate masses [81]. As the elements do not have nominal masses exactly (e.g. oxygen has nominal mass 16 but the actual
The accurate mass can be used to deduce the chemical formula of the molecule (but not the structural formula). One of the main limitations of PTR-MS is that not all molecules efficiently react with a proton, which limits the detection of some compounds. Furthermore, the PTR-MS machines that allow for accurate mass detection are very large and are not bedside devices.

Other ion-molecule reaction mass-spectrometry systems are also available and typically have other reactant molecules. For example, continuous breath analysis has been performed using krypton, mercury or xenon [82]. These systems have several advantages over PTR-MS with regard to the range of molecules that can be detected.

Selective ion-flow tube mass-spectrometry (SIFT-MS) also relies on chemical ionization but, in contrast to PTR-MS, uses multiple reactant molecules (H3O+, NO+ and O2+ ) [83, 84]. Furthermore, the addition of an ion-flow tube results in selective detection of molecules, because each molecule has an individual reaction rate coefficient and product ions. The use of Collision theory with reaction rate coefficients allows for absolute quantification of VOCs, which is a major advantage compared to any other analytical technique. SIFT-MS is a very versatile breath analyzer as many compounds can be detected specifically and quantitatively in real time. Furthermore, because the reactant molecules can be produced from water and air, a bottle of Helium gas is the only consumable that is required. Finally, the device is currently the size of a small table, is movable and can be used at the bedside. The cost per machine is currently the major drawback of this technology. With ion mobility spectroscopy (IMS) a sample of ionized VOCs is let into a drift chamber and driven towards a detector by an electric field [85, 86]. At the same time the ions are pushed towards the inlet by a drift gas (typically Helium). The time it takes an ion to travel through the drift tube, to the detector is a surrogate for the identity of the compound. Because the drift time is mostly non-specific for a VOC, detection of single molecular structures is not possible with traditional IMS. Several technological advances have increased the sensitivity and selectivity of the technique; a capillary column as separator [85, 87], a mass-spectrometer as detector [88] or changing field strengths for variable time periods [89] (High-field asymmetric-waveform ion-mobility spectrometry: FAIMS). Without a mass-spectrometer, this technique
can be miniturized and used as point of care machine. Therefore, IMS or similar technologies could be highly suitable for clinical practice if the problems with sensitivity and selectivity are overcome.

**Future applications**

The success of breath analysis as a clinical test for pulmonary injury or infection on the ICU is highly dependent on the availability, sensitivity, selectivity and ease of use analytical techniques. Several stages can be identified in the discovery, validation and valorization of any biological marker in general [68, 90]. In omic research, additional steps in the process of discovery are required as the markers that will be used in the diagnostic test are unknown at the beginning [91]. However, for breath research even these steps may be insufficient. Additional considerations should be given to the unknown biosynthesis of most volatile organic compounds and the difficulties with quantitative detection at the bedside. A 5 step program is proposed; from untargeted discovery to clinical test. Importantly, these steps are designed for analysis of specific VOCs only. For electronic nose approaches, without any considerations of the underlying changes in the VOC concentrations themselves, the traditional steps in assessment of diagnostic accuracy are sufficient [68, 90].

**Step 1: Discovery and model training**

The first step is identification of potential markers of the disease/syndrome. Most studies described in this thesis tried to take this step. I now recognize that several aspects are extremely important during discovery: (1) a (near) perfect reference standard for the clinical condition, (2) a clinically relevant control group, (3) semi-quantification of a wide range of VOCs in multiple platforms, (4) overlap in VOCs measured between the analytical platforms to minimize device errors and (5) a sufficient sample size. The clinical applicability of the analytical technique is of less importance in this step.

The combination of a good reference standard with a clinically relevant control group requires strong methodological efforts from clinicians. In the setting of the intensive care unit, comparing any ICU-patient to a healthy control does not contribute anything to discovery of clinically relevant biomarkers as any of thousands molecular pathways will be up- or down-
regulated during critical illness. This means that finding a significant difference does lead to the conclusion that this has anything to do with the disease of interest. The other way around, if a molecule is not found to be different between healthy individuals and patients with the disease this does not imply that it can never be used for clinical decision making within this specific population. This is in contrast to discovery studies where cases come from the general population (e.g. asthma or COPD). In that situation the likelihood of finding a true-positive or true-negative result when comparing cases to healthy controls are high compared to the above described ICU-situation. Thus even in this first phase, the included cohort should be a representation of a population that is clinically relevant (patients on the ICU) and the gold standard reference test should be applied to all included patients.

This step is also very challenging from an analytical chemistry perspective. The potential biomarkers are unknown and therefore an as wide range of molecules as possible should be measured. Furthermore, as described in this thesis, every analytical technique and platform has its selectivities and limitations. Therefore, preferably multiple methods and platform should be used on the same samples. In this thesis we always used one GC-MS platform under the assumption that this would be sufficient for discovery. However, I have to acknowledge that we could have missed potentially useful biomarkers by limiting our scope in that way. The usage of multiple GC-MS platforms would broaden the scope and limit the influence of measurement errors within a device. Addition of other analytical techniques such as SIFT-MS could also allow for the absolute quantification of some compounds and may be used to calibrate the other measurements.

In contrast to transcriptomics or genomics analysis that can reach low false-discovery rates with a small sample size because of relatively low heterogeneity, VOC analysis suffers from very high heterogeneity [92]. Furthermore, most VOCs in breath probably do not contain information on endo-genous processes as they come from the environment [92]. Thus, the sample size should be rather large compared to other “omics” approaches because of the high dimensionality of the predictor matrix, low signal-to-noise ratio and inter-person variation of exhaled VOCs.
Importantly, the research question in the first step is not: "can breath analysis accurately discriminate between patients with "X" and "Y"", but rather: "which VOCs in the breath are potential biomarkers to discriminate between "X" and "Y" and what algorithm can be used to discriminate between "X" and "Y" with these potential biomarkers?". The statistical considerations that should be taken with “omics” data in general and with breathomics specifically are recently described in excellent reviews and are outside the scope of this discussion [69, 91, 93]. Step 2a deals with assessment of the real diagnostic accuracy of a model, based on the biological markers that were discovered in step 1.

**Figure 7:** From untargeted discovery to clinical test

*Step 2a: Validation of the diagnostic accuracy.*

This is essentially an extension of the previous step and should be performed in a population that is selected with the same inclusion and exclusion criteria, using the same reference standard and the same
analytical techniques. However, in this step, biomarker discovery is no longer allowed. The only required analysis in this step is the blind application of the model that was developed in step 1. Therefore, this step could be published together with step 1 as one paper (see chapter 4 and chapter 6). In other words, the goal is to apply the classification algorithm on the discovered biomarkers to assess the real diagnostic accuracy within this population. It should be noted that reporting of diagnostic accuracy should be strictly following STARD guidelines [94].

**Step 2b: Biological translation**

A statistical association has been shown for several VOCs in step 1. However, as stated previously in this general discussion, the biological mechanism that leads to the formation of VOCs is important to estimate the post-test probability of a “real” discovery and to identify the population that is actually discriminated with this marker. A search in Pubmed, Pubchem, the Kyoto Encyclopedia of Genes and Genomes (KEGG) and other databases with information on biochemical and metabolic pathways can be used for identification of the biochemical pathways available to different cell types, which can be explored using a translational approach.

**Step 3: Technological translation**

VOCs were analyzed by sophisticated laboratory technologies that allow for identification of unknown compounds and can detect as many compounds as possible but are not available at the bedside. Furthermore, these technologies require specialized personal, storage and transport of breath samples and the analysis of the data is difficult. In this step, the technological translation from these laboratory techniques to a bedside test should be made. This is possible because the volatile biomarkers have been identified and analysis can be targeted rather than unbiased (as in step 1). Point of care (bedside) analysis removes the need to storage and transport of the sample and could thereby reduce the noise of the signal. Several technologies might be used as a rapid, bedside test. The most applicable technology mostly depends on the VOCs that were found in step 1. Several candidate techniques are: PTR-MS, SIFT-MS, IMR-MS, spectroscopy and semi-selective sensor technology.
Step 4: Clinical validation of the point of care device

In this step the diagnostic accuracy of the breath test is validated in a large, non-selective cohort. The technology adapted or developed in step 4 is used for detection of the VOCs. The primary outcome of this study is the discrimination that can be obtained by the breath test. This is the last opportunity to re-calibrate the breath test so that the post-test probabilities reflect the observed probabilities. Using this re-calibrated algorithm, test characteristics can be obtained at several cut-of values. If the measures of diagnostic accuracy are sufficient, they can be implemented in clinical decision support. In other words, what action is recommended if the test scores are above a certain level (in combination with several clinical parameters)?

Step 5: Implications for patient care

The implications for patient care are the most important evaluation of the added value of a diagnostic test. For example, a test that can discriminate between patients with and without positive BAL-cultures within a population that is clinically suspected of VAP could reduce the use of antibiotic therapy. Another clinical dilemma is the early recognition of lung injury and the start of tailored therapy as described previously in the general discussion of this thesis. It could be argued that diagnosis alone is a sufficient endpoint for a diagnostic test. However, providing a name for a condition is not the goal of medicine; we aim to improve outcomes. Therefore, the implications for patient care should not be assumed from the diagnostic accuracy of the test but should be empirically evaluated in a randomized clinical trial [95].

Summary

In this thesis we described the analysis of exhaled breath for the diagnosis of ARDS and pneumonia in ventilated intensive are unit patients. A novel, simplified sample methodology was developed that allows for collection of breath in large populations of ventilated patients (chapter 3). We relied on gas-chromatography and mass-spectrometry for the discovery of volatile biomarkers. In an animal model of acute pulmonary inflammation animals that did or did not receive lipopolysaccharide could be discriminated within hours (chapter 4). We also identified octane, 3-methyl-heptane and
acetaldehyde as markers of ARDS (chapter 6). We confirmed that ARDS could be discriminated from controls by means of electronic nose analysis of the breath, although the diagnostic accuracy was less than with GC-MS analysis (chapter 5). A review of the literature showed that multiple VOCs may be used to identify the presence of pathogens (chapter 6). Although we observed that pneumonia is associated with a changed VOC-profile of tracheal aspirates (chapter 8) and with a decreased concentration of 1-propanol in breath (chapter 9), the specific findings from chapter 6 could not be confirmed.

In this chapter, a biological interpretation of the results of the previous chapters was given. Furthermore, a roadmap for future studies on breath analysis for the diagnosis of ARDS and pneumonia in ventilated patients was sketched.

References


