Chapter 8

Summary, general discussion and recommendations
Summary

In this thesis, we focus on the use of scent detection for diagnosis of Clostridium difficile infections (CDI).

Many different diagnostic tests for CDI are available, as described in chapter 1, all of which have their own advantages and disadvantages. Common limitations include turnaround time, specialised laboratory requirements, variable and limited sensitivity and inability to discriminate asymptomatic carriage from infection (limited specificity). On top of these limitations, several factors delay the diagnosis of C. difficile infections. These include patient’s and doctor’s delay, delays in sampling stool, and time required to process samples in the laboratory. Nonetheless, early and rapid identification of CDI is important for timely treatment and to prevent transmission by prompt isolation measures. Previous research including nurses and our own experience with colleagues claiming to be ‘C. diff sniffers’ incited us to make use of the unique CDI smell for early diagnosis.

Our sense of smell depends on the ability of specialised sensory cells of the nasal cavity to perceive volatile (organic) compounds (VOCs) that bind to receptors with specific affinity in the nasal mucosa. Diseases such as infections and malignancies can be associated with specific VOC profiles that, for example, originate from microorganisms or from changes in host metabolism, and thus have a different odour. In short, some diseases literally stink.

The growth of C. difficile is strongly associated with the production of specific VOCs, e.g. certain short chain fatty acids (like iso-caproic acid) and p-cresol (as an end-product of tyrosine metabolism). However, none of these compounds is unique for C. difficile, and the smell of C. difficile is more likely the result of a mix of characteristic VOCs rather than of one unique ingredient, just like the sweet smell of a rose is due to a blend of odorous components (2-phenylethanol, β-ionone, β-damascone, β-damascenone, citronellol, rose oxide, geraniol, and nerol).

The diversity of VOCs in C. difficile- infected stool is considerably smaller compared to that of stool from healthy controls. C. difficile- infected stool is less likely to contain butyric acid, but more likely to contain butanol; the metabolic profile of C. difficile infected stool consists largely of furans and acids rather than aldehydes and indoles as found in stool from healthy controls.

A key factor in the pathogenesis of CDI is the decline in faecal microbial diversity, mostly caused by antibiotic therapy. In particular the proportion of the bacterial phyla Bacteroidetes and Firmicutes that are so abundant in healthy controls is decreased, and there is an excess of Proteobacteria. For patients with recurrent CDI, these differences seem even more pronounced and the microbial community is not only less diverse but also less variable over time. It is likely that the reduction in microbial diversity results in a reduction of VOC diversity, and as a consequence the characteristic C. difficile VOCs are more prominent in a C. difficile- infected stool sample. This assumption was strengthened by the
observation that many microbiology technicians easily recognize the characteristic smell of a *C. difficile* culture, and many nurses easily recognize the characteristic smell of a patient with full blown CDI. This led us to investigate whether a biological or electronic olfactory system could be used as a diagnostic instrument.

In chapter 2 we give a narrative review of key studies in medical scent detection performed with human, animal and electronic noses. The ability of humans to detect disease by smelling has only rarely been formally assessed, with exception of recognising stool samples with *C. difficile* and rotavirus infection. Animal scent detection studies are more prevalent. Most suggest similar or even superior accuracy compared to standard diagnostic tests. Examples include the use of dogs for detecting lung cancer in breath samples, or rats for *Mycobacterium tuberculosis* detection in sputum. Studies using different types of electronic noses for pulmonary disease and cancer also often show high accuracy. Although the overall results of studies presented in this chapter are encouraging, the medical field remains largely sceptical. Several factors are to blame: results of different types of noses are not easily generalizable and confirmatory studies are seldom performed.

Chapter 3 is a proof of principle study that investigates the diagnostic accuracy of a trained detection dog to detect *Clostridium difficile* in stool samples and in hospitalised patients. A 2 year old beagle (Cliff) was trained by a professional detection dog trainer to identify the smell of *C. difficile* by a reward based training method. The dog was trained to sit or lie down when *C. difficile* was detected. After a 2 month initial period of practice on culture plates, we formally tested his accuracy with 50 *C. difficile*-positive and 50 *C. difficile*-negative faecal samples. Depending on the interpretation of an inconclusive response, the dog’s sensitivity and specificity were 100% (95% confidence interval (CI): 91-100%) and 94- 100% (95%CI: 83-100%), respectively. Subsequently, we tested his ability to detect CDI infection in hospitalised patients. We included 30 consecutive hospitalised CDI cases with diarrhoea, a positive toxin enzyme immunoassay and *C. difficile* culture, plus 270 control participants from the hospital ward of the index case. Depending on the interpretation of an inconclusive response, sensitivity varied between 86 and 93% (95% CI: 67-99%) and specificity varied between 97 and 98% (95% CI: 94-100%).

We concluded it is feasible to use a dog to detect *C. difficile* in stool samples and infected patients. For the purpose of detection, the dog did not need a stool sample or physical contact with patients. In addition, dogs are quick and efficient: an entire hospital ward can be screened for the presence of patients with *C. difficile* infection in less than 10 minutes. This promising result raised further questions: e.g. how will the dog do during an outbreak in comparison to the situation in this case-control design, when several patients in one room or adjacent rooms could be affected, or none at all? How do training and fatigue influence his performance when screening several wards during an outbreak rather than only one in the pilot study? And does the dog actually make a quicker point-of-care diagnosis when compared to standard laboratory tests?
Three years after the pilot study, an outbreak of CDI ribotype 027 occurred in the VU University medical centre (VUmc), a 750-bed tertiary care centre. Hence, we took the opportunity to assess the diagnostic accuracy of a trained sniffer dog in detecting CDI patients in an outbreak setting, as further described in Chapter 4. During a total of 9 hospital visits, screening 371 participants 651 times cumulatively, the dog had almost comparable sensitivity and specificity as during the pilot study.

What stood out here is that the dog did not consistently indicate all CDI cases as positive on every consecutive visit. A logical explanation could be that the dog response is related to the severity of symptoms, which usually wane during treatment. This means it is not necessarily ‘false negative’ if the dog identifies a case as CDI negative after several days of treatment; but this explanation could not account for all inconsistencies.

On the other hand, 2 of 11 (18%) CDI negative participants that were ‘falsely’ indicated by the dog as positive, did actually developed CDI during the three months of in-hospital follow-up, compared to only 12 of the 346 CDI negative participants (3.5%) that the dog correctly identified as negative (p=0.06). There were a few exceptions in which the dog signalled an infection before it was confirmed by the laboratory, but in general this study did not show quicker CDI diagnoses and treatment with the dog than with conventional laboratory tests.

Chapter 5 describes a case-control investigation of this outbreak of C. difficile ribotype 027, to identify risk factors for acquisition of the outbreak strain. Individual risk factors for CDI, like antibiotic use, older age, length of stay (LOS) etc., are well-known as described in chapter 1. In this study we focused on hospital-associated risk factors (like admittance to a specific ward, undergoing a specific intervention, number of in-hospital transfers etc.): which of these factors increased the risk of developing CDI and therefore contributed to the hospital-wide spread of the outbreak strain?

The outbreak of a single clone of C. difficile ribotype 027, as confirmed by whole genome sequencing, occurred between May 2013 and July 2014, involving 19 departments and 79 patients. For every CDI ribotype 027 case (n=79), four non-CDI control patients were included (n = 316) matched for age, ward and attending specialty. All consecutive patients diagnosed with CDI due to other ribotype strains (n = 70) between May 2013 and March 2014 constituted a second control group.

Antibiotic use, LOS >10 days, and admission to the intensive care unit (ICU) prior to the date of CDI diagnosis were significantly associated with ribotype 027 CDI in multivariable analysis compared to non-CDI controls; cases were less likely to have been admitted to a ward with a known patient with CDI. For those who had stayed at the ICU prior to CDI diagnosis (35 ribotype 027 cases; 51 non-CDI controls), the use of selective decontamination of the digestive tract (SDD; adjusted odds ratio (OR) [95% CI]: 7.90 [0.71–88.02]) and a longer ICU LOS (adjusted OR [95% CI]: 1.53 [0.75–3.10]) were associated with CDI risk. Only seven of the non-027 CDI patients had been admitted to the ICU during the outbreak period, which made it impossible to discern whether SDD and/ or
ICU LOS increased the risk of any ribotype CDI, or specifically of the outbreak strain. Although we did adjust for APACHE IV score and LOS in the ICU in the multivariable analysis, it is impossible to rule out residual confounding due to disease severity in the development of CDI. Nonetheless, the use of SDD is aimed at eradicating potentially pathogenic microorganisms from the digestive tract. Considering the crucial pathophysiological role the imbalance of gut microbiota plays in the development of CDI, the concept of SDD use as an important risk factor for CDI seems very biologically plausible.

In chapter 6 we focus on olfactory detection by means of an electronic nose. We studied whether the profile of VOCs emanating from stool samples can differentiate between *C. difficile* positive and -negative samples. To analyse the faecal VOC profile we used ion mobility spectrometry (Field Asymmetric Ion Mobility Spectrometry (FAIMS); Lonestar, Owlstone, Cambridge, UK). This technique separates molecules in a gaseous mixture, e.g. in the headspace of a stool sample, based on their size, mass and mobility in varying electric fields. Rather than identifying individual molecules in the headspace (or ‘air above’) a stool sample, it creates a composition profile of the chemical components (‘chemical fingerprint’). The faecal VOC profile is translated into a calculated ‘CDI-VOC score’ between 0 and 1, indicating the probability of CDI based on the VOC profile. We analysed 135 stool samples to create this CDI-VOC score, and subsequently validated it on 78 different samples, of which 26 were *C. difficile* positive. The discriminatory ability of this CDI-VOC score was good, with a C-statistic of 0.86 (0.75–0.97) on the validation sample set. For fresher samples, i.e. with less time between stool sampling and analysis, the C statistic was even higher. The benefit of this diagnostic instrument is that it is quick, relatively cheap, accurate and compact. Provided the mathematical analysis of the FAIMS reading, which was now performed afterwards with special software on a separate computer, is integrated into the FAIMS computer, it would make a very useful point-of-care test.

Chapter 7 describes whether the CDI-VOC score also correlates to the risk of CDI recurrence. In patients with recurrent CDI the intestinal microbiota appears more affected than in patients with non-recurrent CDI. Faecal VOC composition is importantly influenced by the gut microbiota. Therefore, for 46 CDI patients whose samples were included in chapter 6, we determined whether a high CDI-VOC score is predictive of high recurrence risk. After adjustment for known risk factors for CDI recurrence (age, ribotype 027, the use of antibiotics or proton pump inhibitor after CDI treatment, and whether the index episode was a first episode or a first recurrence) every 0.10 rise in the CDI-VOC score was significantly associated with increased risk of recurrence (adjusted hazard ratio 1.4; 95% CI 1.1–1.8; p<0.01). This means that at the time of initial CDI diagnosis, a high CDI-VOC score suggests high recurrence risk. Provided other studies confirm our findings, VOC profiling can signify a marked improvement in recurrence prediction, and enable tailored therapeutic strategies for CDI patients.
GENERAL DISCUSSION AND RECOMMENDATIONS.

In this thesis we describe the use of scent detection for diagnosis of *Clostridium difficile* infection (CDI).

**FAIMS and Cliff: what do they sniff?**

The distinctive odour of faeces is the result of the gut microbiota. Human faecal flora produces compounds such as ammonia, aliphatic amines, branched chain fatty acids, indole, phenol and volatile sulphur containing compounds, which are responsible for the odour of faeces and flatulence.\(^{(11)}\) Both the gut microbiota and the faecal chemical composition can be affected during disease, e.g. there are significant differences between CDI patients and non-CDI patients, as described in the introduction of this chapter. There have been attempts to pinpoint the significant smell of *C. difficile* down to one unique ingredient, but it seems more likely the smell is the result of a mix of characteristic volatile organic compounds (VOCs). For stool samples it is likely that the reduction in microbial diversity observed in CDI patients results in a reduction of VOC diversity, and as a consequence the characteristic *C. difficile* VOCs are more prominent.

We have not formally tested whether the dog responds to asymptomatic carriers of non-toxigenic or toxigenic *C. difficile* strains. However, it is unlikely that he does, given the dog’s high specificity (combined from chapter 3 and 4: 607 of 627 CDI negative control participants (97%) were correctly identified as CDI negative) and the carriage rate of toxigenic and non-toxigenic hospitalised patients (recently >20% on hospital admission in a North American hospital, of which 15% were toxigenic and 6% non-toxigenic strains\(^{(12)}\)). In asymptomatic carriers the quantity of *C. difficile* probably constitutes a smaller portion of the gut microbiota when compared to CDI patients. The characteristic mix of *C. difficile* associated VOCs is therefore likely overshadowed by VOCs of the common microbiota in asymptomatic carriers.

**The use of animal olfactory detection: pro’s and con’s.**

Animal noses are unbeatable in their smelling capacity, either by man or machine. Nonetheless, and despite the overwhelming enthusiasm in our hospital generated by Cliff’s visits and performance, the use of detection animals has major drawbacks. Their training takes time and expertise and they need regular practice to maintain their skills. Animals, like humans, have good and bad days (limited intrinsic reproducibility). Despite a high level of training, their behaviour is not fully predictable: we noticed during our hospital visits the dog would be distracted by other stimuli (for example, a real-size gorilla balloon attached to the side of a hospital bed). Additionally, all dogs are different (limited generalisability), which means all trained animals need an individual assessment of performance, and their calibration needs to be repeated over time to guarantee a consistent performance. This makes them difficult to “mass produce”.

For hygienic purposes there are obvious restrictions with regards to allowing animals into healthcare facilities. Dogs can be carriers of *C difficile* strains and other pathogens. Similar to hospital staff, the dog could be a source of
transmission. This risk can be minimised by using strict preventive measures such as avoiding physical contact with patients and their surroundings. Furthermore, the medical community often has difficulty to accept anything that is considered too ‘unconventional’ or that does not come with a Standard Operating Procedure. It does not help that for the majority of animal medical detection studies, there are no confirmatory studies and occasional spectacular results do not get any follow-up.

So what is needed to make better use of animal noses for medical purposes? First and foremost, more and better studies are warranted, particularly confirmatory studies. Experiments need a large number of samples, including new, different validation samples to ensure the dog recognises the specific condition rather than the individual patient. Positive and negative specimens should be sampled, stored and transported separately, but under identical circumstances to avoid picking up a confounding scent. Preferably, multiple animals should be included in order to assess the generalisability of their skills and their training. Appropriate blinding of both the dog and handler should get high priority in order to avoid unintentional bias. Subsequently, positive results should be scrutinised to see whether the use of medical detection animals is cost-effective and actually improves our current diagnostic arsenal.

With regards to our *C. difficile* sniffer dog Cliff, we demonstrated its impressive sensitivity and specificity for finding CDI cases, but using the detection dog did not systematically improve time to diagnosis in respective study designs. Whether canine *C. difficile* detection can lead to a quicker diagnosis depends on diagnostic accuracy, but also on surveillance frequency and CDI incidence. Currently the incidence of CDI is at an all-time low in the VUmc after the 2013 CDI outbreak (figure 1). On the contrary, North American healthcare facilities encounter an ever increasing CDI incidence. A confirmatory study in this geographical area may prove the efficacy and cost-effectiveness of canine CDI surveillance to be more favourable.

**The use of electronic noses: pro’s and con’s.**
For electronic noses (e-noses) similar drawbacks apply with regards to intrinsic reproducibility, generalisability, individual validation, flaws in study design, lack of confirmatory studies etc.
There are many different e-nose instruments and techniques, the results of which are not interchangeable. Traditional e-noses have an array of chemical sensors that have specific affinity for a group of chemicals, just like a canine or human nose, but far more limited with regards to number and specificity of sensors. When triggered by the molecules in a gaseous mixture, the sensor responses form a specific pattern or ‘odorous fingerprint’. The problem with these traditional e-noses includes their need for calibration, since affinity and selectivity of sensors can differ between instruments (limited generalisability), plus their sensors often drift over time (limited intrinsic reproducibility). Furthermore, the sensors and therefore the test results are heavily influenced by environmental factors like humidity and temperature.
Summary, general discussion and recommendations

For newer techniques, e.g. ion mobility spectrometry as used in this thesis, these technical limitations are mostly overcome. The amount of data generated allows for very complex and detailed headspace analyses, while the device still fulfils the requirements necessary for a point of care tool with regards to speed, size and non-invasiveness. Still, in the field of e-noses again we find that the diagnostic properties of observed discriminatory features are only sporadically validated with new, blinded samples, and promising small studies are often not followed by independent, confirmatory studies.

In this thesis we have attempted to address a number of these issues. The researcher training, testing and validating the algorithms was unaware of (or blinded to) the \textit{C. difficile} status of the validation samples. What happened during the research for this thesis stresses the importance of validation: after FAIMS analysis of about 250 samples we repeatedly found that the observed (and often very significant) differences in odorous profile between the negative and positive \textit{C. difficile} samples had no discriminating value when applied on the new validation samples. After about a year of analysing samples, we discovered the generated odorous fingerprint was determined by the residue of cleaning product in the glass jars, rather than by the stool sample itself. We were forced to discard all the gathered data and start again from scratch, and finally, with new samples and using trace-free jars for FAIMS analysis, we were able to build an algorithm that did have discriminatory value in the validation samples. It turns out is not very difficult to find differences in VOC profiles between sample sets when using a sensor that generates > 50,000 data points per sample: by chance alone some of the VOC characteristics are bound to be consistently different between the sets. The proof of the pudding in this case is in validating the observed differences on new samples.

This thesis demonstrates faecal VOC profiling can be used to discriminate between CDI positive and negative samples. The potential benefit lies in the speed of analysis, low costs, and it has the potential of point-of-care use. For this thesis the mathematical analysis of the VOC profile was performed afterwards on a separate computer with extra software. The mathematical analysis should
first be integrated into the FAIMS computer to make it ready for use in our daily practice, where we would aim to have the complete stool sample analysis and test result instantly available.

Interestingly, in diagnostic studies like the ones presented in this thesis, a new diagnostic test can only be as good as its reference standard. Faecal VOC profiling does not depend on conventional diagnostic principles, e.g. demonstrating the presence or absence of specific *C. difficile* markers in faeces (toxins, cell wall components, etc.), rather it analyses whether or not the VOC composition fits a *C. difficile* infection profile. This unique approach could prove an excellent test to diagnose true CDI at the bedside as a stand-alone test and might actually be able to discriminate between asymptomatic carriage and infection. Yet as for every new test, we judge its merits within the framework of existing techniques, which for *C. difficile* infections does not offer an infallible gold standard. As described in the introduction of this chapter, the presence of toxins in a faecal sample best predicts increased mortality and therefore best defines serious *C. difficile* infection. ‘Only’ culture or nucleic acid amplification tests (NAAT) positive samples, without the presence of toxins in stool, can represent either asymptomatic carriage or milder *C. difficile* infection. An unequivocal reference test that can discriminate between asymptomatic carriage and infection might also include clinical parameters (response to treatment, complications, endoscopic findings etc.), but these can be subject to interpretation, or again fail to define milder cases, etc. Lack of a perfect reference standard complicates the appraisal of a new technique, in this case faecal VOC profiling.

For now, we hope larger studies confirm the encouraging results in using faecal VOC profiling for diagnosing CDI. Additionally it would be very interesting to focus research on cases with intermediate test results in conventional diagnostics (e.g. can faecal VOC profiling elucidate the implications of a culture/ NAAT positive, toxin negative sample in a patient with diarrhoea?) or in faecal VOC profiling (e.g. could an intermediate CDI- VOC score signify asymptomatic carriage?). Finally, in our small study sample the VOC profile seems indicative of recurrence risk, which could aid clinicians in their choice of CDI treatment; a finding that also calls for reproduction with a larger number of patients to clarify its potential and limitations.

**Role of scent detection**

The use of detection animals and e-noses could considerably contribute to the field of medical diagnostics; however, at this moment in time virtually none are available for routine clinical use. The principle of scent detection offers many advantages: sniffing excreta is of a very non-invasive nature, and both this thesis and other publications underscore that the diagnostic accuracy of scent detection techniques is sometimes even superior to other diagnostic modalities. Specifically, we feel scent detection has important potential for diagnosing CDI. It may take unconventional research, that perhaps goes against the grain of mainstream scientific beliefs, similar to that presented in the studies in this thesis, to overcome the gap between concept and implementation into daily practice.
Loosely adapted from:

REFERENCE LIST
