Biomedical Scent Detection Dogs: Would They Pass as a Health Technology?

Mirkka Koivusalo¹ and Catherine Reeve²*

Abstract

Biomedical scent detection dogs identify the scent profiles of diseases, such as cancer, diabetes or pathogenic micro-organisms. What the field of biomedical scent detection has been lacking, however, is the assessment of the method from the point of view of a health technology. All health technologies undergo a thorough evaluation of safety, clinical effectiveness and costs, as well as ethical, social, organizational and legal evaluations in some cases. Passing these regulatory controls is a pre-requisite before a technology is approved for use in decision-making about patient outcomes. Biomedical scent detection has a lot of attractive qualities, such as the sensitivity and specificity of the dogs’ noses, safety and relative cost-effectiveness. However, the method also has various challenges, in particular regarding its clinical effectiveness. The most pertinent issues to address before the dogs would pass as a health technology are standardization of the training techniques, both intra- and inter-dog reproducibility, and generalization of the detection task to the early stages of disease progression. We recommend setting realistic goals in terms of what the dogs can and cannot do and adopting a collaborative approach between clinicians and animal psychophysicists.

Highlights

• Because biomedical scent detection is a potential tool to identify diseases, in the eyes of regulatory bodies it is comparable to a health technology.

• This is the first report evaluating the canines’ sense of smell from the point of view of a health technology.

• It is important for researchers to be aware of the required regulatory controls before dogs can have an accredited role in the clinic.

• We discuss the current advantages and disadvantages of the method to help guide researchers towards attainable goals.
INTRODUCTION

Canines’ sense of smell has been utilized successfully for a long time for different purposes, such as detection of explosives, narcotics, and bed bugs, to name just a few. Biomedical scent detection (or medical scent detection) is an emerging method that uses the canine olfactory capacity for identifying the scent profiles of diseases, such as different types of cancers (Moser and McCulloch 2010; Jezierski et al. 2015), diabetes (Gadbois and Reeve 2014; Hardin et al. 2015) or the presence of pathogenic micro-organisms (Bomers et al. 2014; Bryce et al. 2017; Koivusalo et al. 2017; Koskinen et al. 2017). It is the combination of the acuity of the sense of smell with the ability to learn by operant conditioning that makes dogs potential biodetectors for different tasks (Pirrone and Albertini 2017). The ultimate goal in most cases is to involve dogs as a fast method for the diagnosis or screening of human patient disease. Research studies illustrate that dogs are capable of distinguishing the volatile organic compounds characteristic of diseases (Rudnicka et al. 2014). An important topic to take into consideration when training and testing dogs for this purpose is a comparison with conventional technologies used in screening and diagnosis. If a dog's indication is to be used in making decisions on a patient's health and treatment outcomes, canine scent detection would be comparable to a health technology in the eyes of health organizations and regulatory bodies. The purpose of this review is to evaluate canine scent detection as a medical device, to assess whether it would pass a health technology assessment (HTA) as it currently stands and what the most pertinent technical challenges are.

DEFINITION AND QUALITY ASSESSMENT OF HEALTH TECHNOLOGIES

The World Health Organization (WHO) defines a health technology as (World Health Organization 2018):

“A health technology is the application of organized knowledge and skills in the form of devices, medicines, vaccines, procedures and systems developed to solve a health problem and improve quality of lives.”

Amongst the different technologies, the dog’s nose could be categorized as a medical device: a highly sensitive, built-in olfactory detection system used to identify diseases. However, before a medical device or any health technology is granted permission for official clinical use in any country, an approval is required by national health organizations, such as Food and Drug Administration (FDA) in the US. The requirements for approval include passing several regulatory controls (US Food and Drug Administration, 2018) that set a high standard for all health technologies.

HEALTH TECHNOLOGY ASSESSMENT (HTA) IS AN EVALUATION TOOL

A practical tool for evaluating the quality of health technologies is a process known as Health Technology Assessment (HTA), which uses scientific research to inform decision-makers on the introduction and use of a health technology. Each country’s regulatory bodies implement HTA in their own way, but international organizations have set guidelines for the process (EUnetHTA Core Model® 2016; Health Technology Assessment International 2018; International Network of Agencies for Health Technology Assessment 2018).

The aspects to be evaluated depend on the technology in question and the assessing regulatory body. The following information, which is pertinent also to biomedical scent detection, is always asked:

• Description of the health problem and currently available technology for solving it.

• Technical description of the new technology in detail.

• Clinical effectiveness: Does the new technology work equally well or better than a standard method currently in use? The requirement before granting a permission for marketing is to demonstrate that the technology is at least “substantially equivalent” (FDA) to a current standard technology.
• Safety: The technology has to be safe not only for the patients but also for the people operating it. Potential risks and harms are assessed that will determine the risk classification of the technology.

• Economic evaluation: Value-for-money assessment is conducted to evaluate the costs and benefits of the technology in relation to health-related outcomes as well as the economic burden on the healthcare system and society.

• Ethical, social, organizational and legal aspects may also be evaluated.

HOW DOES BIOMEDICAL SCENT DETECTION RANK IN EVALUATION OF THE MOST FUNDAMENTAL CRITERIA?

CLINICAL EFFECTIVENESS

If we ask whether dogs can smell cancer or a pathogen, the answer would be a ‘yes’, as dogs can detect diseases with very high levels of sensitivity and specificity (Jezierski et al. 2015). But to assess the clinical effectiveness of the technology the question is: can dogs do it with the same reproducibility and accuracy as an analytical instrument already in use, and can patient treatment outcomes be planned on the basis of the results?

To be approved by health organizations, several aspects need to be addressed before the answer is ‘yes’ for canine scent detection (Fig. 1).

Here are some of the main challenges that scent detection researchers are trying to solve:

1) Standardization of the Training and Testing Methods

Currently there is no standardized way to train and test the dogs, which leads to variability in results (as discussed by Elliker et al. 2014; Jezierski et al. 2015; Gadbois & Reeve, 2016). Traditionally dogs have been trained and tested using an alternative forced choice (AFC) method, where a line-up or a scent wheel contains one target scent (S+, here, a “disease odour”) and several non-target or distractor scents (S–, a healthy control odour), and then running double-blind trials, where neither the dog nor the dog handler knows the sample positioning and the dog is required to locate and indicate the target, diseased, odour. This method is appropriate for a proof-of-principle study, but in a real screening situation the number of S+ and S– samples is unknown. If the dog has been trained to always find one ‘win’ in a line-up, what happens in a screening situation?

Studies have shown that when dogs were trained using the AFC task of one S+ and four S– samples, and then tested with randomized numbers of S+ and S– samples,
the sensitivity and specificity of their performance deteriorated significantly (Amundsen et al. 2014; Hackner et al. 2016). This illustrates the importance of training and testing the dogs with randomized samples as early in the process as possible.

The AFC method also has a disadvantage due to a memory load it may impose on the detection dog. Gadbois and Reeve (Gadbois and Reeve 2014; 2016) have shown that the accuracy of canines decreases significantly the further down the target is in the line-up because of a tax on the dogs’ memory processing. The goal of biomedical scent detection is to assess the sensory-perceptual abilities of dogs to distinguish scents extremely similar to each other, and not make the process a memory load task of a past event. Therefore, to avoid the tax on memory, the use of discrimination tasks of 2 or 3 choices is more favourable (Gadbois and Reeve 2014; 2016; Koivusalo et al. 2017). An even more elegant solution is a pure detection task, i.e. a yes/no, go/no-go system, where the dog is presented with one stimulus at a time, which requires a ‘yes’ or ‘no’ answer (Gadbois and Reeve 2014, 2016). In this approach the dog always gets a win and the task is a pure sensory task, which increases accuracy.

2) Reproducibility

An analytical laboratory instrument analyzes hundreds of samples daily with high accuracy and throughput, albeit a certain degree of variability (as discussed by Moser & McCulloch 2010). Dogs cannot be considered as instruments, and they will never obtain an accuracy of 100% due to their own inherent variability (Hackner and Pleil 2017). Furthermore, the published medical scent detection studies are different from real screening situations, where larger sets of samples would be processed daily, and the reproducibility will be harder to accomplish. This stems from several factors.

The work of an ‘analytical’ dog is highly repetitive, and it can be difficult to keep up their motivation, as dogs are subject to boredom, fatigue and external distractions. The job of these dogs is different from the scent detection work in the field (e.g. drugs, explosives, search and rescue), where the dogs get to use more of their natural ‘hunting’ behaviors and the task remains more of a game to them.

In contrast to laboratory devices, dogs may try and change their strategies in order to be rewarded at minimum effort leading to inaccuracies during double-blind trials.

The handler’s positive and negative sensations are easily transmitted to the dogs.

Also, in contrast to analytical devices that detect compounds very specifically, each individual dog can perceive complex scent patterns differently leading to variability between different dogs. Finding dogs with an ability to generalize the complex detection task is important.

3) Detection of Diseases in True Unknown Samples

Most published studies have been performed with samples from already diagnosed patients. However, it is crucial to improve the overall survival: earlier cancer detection may permit earlier intervention and detection of colonization by a pathogen prior to an actual infection may enable early treatment and containment as well. Therefore, the next step after standardization of training/testing methods is to obtain results with completely unknown samples.

It is likely that the scent profile changes between samples from patients with an advanced vs. an earlier stage of disease. Amundsen et al. (Amundsen et al. 2014) showed that the dogs’ performance is affected when they are tested with patients suspected to have different stages and forms of lung cancer. This finding highlights the importance of training the dogs not only with samples from patients with an advanced or malignant disease but with samples from benign stages of disorders.

To teach the dogs the subtle discrimination among different stages of a disease may not be a trivial task, as what truly is a positive sample and what is a negative one can turn out to be a grey area. This is where drawing realistic cut-offs for what dogs can and cannot do should be considered.
4) Sample presentation

One experimental factor to consider and standardize is the way the biological samples are presented and preserved (Hackner and Pleil 2017; Reeve et al. 2017). Also the protocol used (or not used) to clean and handle the scent containers between trials differs widely, but ultimately the goal is that the dogs discriminate the target scent amongst any background scent. What is important to acknowledge in biomedical scent detection is that the dog is constantly learning to discriminate subtle differences between samples, and without due diligence they may accidentally learn to indicate something superfluous or superstitious. To circumvent this, many researchers settle for a set cleaning protocol or using disposable scent containers only (Jezierski et al. 2015; Koivusalo et al. 2017; Koskinen et al. 2017).

SAFETY

This is one of most attractive qualities of biomedical scent detection: it is very non-invasive in nature, and therefore safe for the patients. For some patients a physical interaction with animals may be a concern, but if the detection work takes place in a laboratory, there is no need for contact between the patients and the dogs.

An important risk assessment to be made is the evaluation of the consequences of false positive (dog indicates a S– sample) and false negative (dog does not indicate S+) results. False negative results would impose a more serious safety concern, as it would involve missing patients at risk.

COST EVALUATION

Another appealing quality of biomedical scent detection is its potential to be more cost-effective than many expensive laboratory methods. For example, Arnaud (2016) determined that biopsy-based methods of non-small cell lung cancer biomarker testing can cost upwards of $2,500. The analysis of volatile organic compounds (VOCs, Amann et al. 2014) in exhaled breath samples is a much less expensive, and less invasive procedure, and has proven a promising avenue for cancer diagnosis (Shirasu and Touhara, 2011). The analytical techniques used to identify disease-specific VOCs, such as SPME and GC-MS, however, can be expensive and extensive training is required to perform the analysis and analyze the results (Shirasu and Touhara, 2011).

Alternatively, the identification of diseased breath samples by detection dogs requires minimal equipment and the analysis (interpretation of the dogs’ indication behaviour), is straightforward (Jezierski et al. 2015). McCulloch et al. (2006) examined canine detection of non-small cell lung cancer in donated breath samples and showed that the dogs were 99% sensitive and specific, providing evidence for a highly promising yet cheaper detection system. Also, dogs would be a rational solution in areas where high-tech instrument-based analysis is unavailable (Rooney et al. 2013; Jezierski et al. 2015)

The disadvantage is that the process of training a dog is a time-consuming and hence a money-consuming process because it can take up to a year to train a dog; fully trained medical alert dogs can cost anywhere between $8000 - $20,000 USD (beyontype1.org). Even then, it does not mean that every individual dog is suitable for the task, and even at later stages of testing, some dogs may not show the desired sensitivity and specificity (Jezierski et al. 2015).

CONCLUSION

To be approved as a clinically valid health technology by regulatory bodies, biomedical scent detection has not reached the required standard yet. Whether it will, depends on the crucial research that needs to be done. In the current age of high-tech medical devices competing in the market, setting a realistic target for the canines might speed up the process as it is a race of who has the most clinically effective and cost-effective method first. In order to make biomedical scent detection a reliable health technology, understanding of the canine olfactory learning processes is needed both through a clinical research and psychophysical approach.

ACKNOWLEDGEMENTS

We thank Dr. Hanni Uusi-Kerttula for critically reading the manuscript.
This research was supported by an NSERC PGS-D to Catherine Reeve.

REFERENCES


