Cochrane Library

Cochrane Database of Systematic Reviews

Electronic and animal noses for detecting SARS-CoV-2 infection (Protocol)

Leeflang MMG, Bell K, Deeks JJ, Dinnes J, Doust J, Korevaar DA, Lord SJ, Spijker R

Leeflang MMG, Bell K, Deeks JJ, Dinnes J, Doust J, Korevaar DA, Lord SJ, Spijker R.
Electronic and animal noses for detecting SARS-CoV-2 infection (Protocol).
Cochrane Database of Systematic Reviews 2021, Issue 6. Art. No.: CD015013.
DOI: 10.1002/14651858.CD015013.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Protocol]

Electronic and animal noses for detecting SARS-CoV-2 infection

Mariska MG Leeflang[1], Katy Bell[2], Jonathan J Deeks[3], Jacqueline Dinnes[4], Jenny Doust[5], Daniël A Korevaar[6], Sarah J Lord[7], René Spijker[8]

[1]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands. [2]Wiser Healthcare, Sydney School of Public Health, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia. [3]NIHR Birmingham Biomedical Research Centre, University Hospitals Birmingham NHS Foundation Trust and University of Birmingham, Birmingham, UK. [4]Test Evaluation Research Group, Institute of Applied Health Research, University of Birmingham, Birmingham, UK. [5]Centre for Research in Evidence Based Practice, Bond University, Gold Coast, Australia. [6]Department of Respiratory Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands. [7]NHMRC Clinical Trials Centre, The University of Sydney, Camperdown, Australia. [8]Cochrane Netherlands, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht , Netherlands

Contact address: Mariska MG Leeflang, m.m.leeflang@amsterdamumc.nl.

Editorial group: Cochrane Infectious Diseases Group.
Publication status and date: New, published in Issue 6, 2021.

Citation: Leeflang MMG, Bell K, Deeks JJ, Dinnes J, Doust J, Korevaar DA, Lord SJ, Spijker R. Electronic and animal noses for detecting SARS-CoV-2 infection (Protocol). *Cochrane Database of Systematic Reviews* 2021, Issue 6. Art. No.: CD015013. DOI: 10.1002/14651858.CD015013.

ABSTRACT

Objectives

This is a protocol for a Cochrane Review (diagnostic). The objectives are as follows:

1. To assess the diagnostic test accuracy of eNoses to screen for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection in public places, such as airports.
2. To assess the diagnostic test accuracy of sniffer animals, and more specifically dogs, to screen for SARS-CoV-2 infection in public places, such as airports.
3. To assess the diagnostic test accuracy of eNoses for SARS-CoV-2 infection or COVID-19 in symptomatic people presenting in the community, or in secondary care.
4. To assess the diagnostic test accuracy of sniffer animals, and more specifically dogs, for SARS-CoV-2 infection or COVID-19 in symptomatic people presenting in the community, or in secondary care.

Secondary objectives

If sufficient data are available, we will investigate the accuracy (either by stratified analysis, or by subgroup analysis) according to specific eNose technology or animal, and according to whether those who are tested are symptomatic or not. We will also investigate whether eNose brand, reference standard, and healthcare setting are associated with differences in diagnostic test accuracy.

BACKGROUND

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the resulting COVID-19 disease present important diagnostic evaluation challenges. Given the massive scale of the pandemic, one of these challenges is to determine an efficient testing algorithm for large groups of people with or without symptoms, for example for population screening, or to screen travellers entering a country via an airport. Similar challenges apply in healthcare settings, such as emergency departments and general practice, where point-of-care tests may be required that are non-invasive, easy to use, and that provide a test result within minutes. Ideally, there would be a test available that is relatively cheap, easy and quick to perform, provides immediate results, and is sufficiently accurate for diagnosing SARS-CoV-2 infection.

Breath analysis devices, such as electronic noses, smell sensors, and sniffer dogs or other trained animals, seem to fulfil these criteria. Some governments have invested in the development and purchase of electronic devices (eNoses), without an apparent, rigorous evaluation before implementation (RIVM 2021). Similarly, media have reported that sniffer dogs may be able to detect SARS-CoV-2 infection (e.g. BBC 2020). A thorough overview of the current diagnostic properties of both eNoses and sniffer dogs will inform decision makers about the potential for these tests.

Target condition being diagnosed

COVID-19 is the disease caused by infection with SARS-CoV-2. The key target conditions for this review is current SARS-CoV-2 infection. Different reference standards may be used for the diagnosis of this target condition, including molecular assays, such as reverse transcription polymerase chain reaction (RT-PCR), or internationally recognised clinical guidelines for diagnosis of SARS-CoV-2 infections. Although the severity of the disease is important for a person's outcome, the role of point-of-care tests – examples of which are eNoses and sniffer animals – is to detect SARS-CoV-2 infection of any severity.

COVID-19 public health interventions focus on reducing disease transmission, thus, it is important to identify and isolate infected people before and while they are infectious. This could be relevant, for example, for testing done in airports and other entry points to a country. However, there is no reference standard for being 'infectious'. Using RT-PCR status as a reference standard (as is done for the target condition of infection) will ensure that infectious people are not missed, but as RT-PCR continues to detect viral RNA days and weeks after the onset of infection, it may wrongly classify some people as infectious. Therefore, we focus here on the target condition of SARS-CoV-2 infection.

Index test(s)

Electronic devices

An electronic nose (eNose) mimics the olfactory system of mammals. It is a chemical analyzer, containing multiple sensors that react to a multitude of volatile organic compounds (VOCs) in air and vapour (Röck 2008; Wilson 2015). The sensors' responses are measured and quantified, and combined into a signal. The underlying hypothesis is that people with a certain disease or target condition, emit a different smell (i.e. a different mix of VOCs), than people without this disease. This smell can come from breath, and other samples, such as urine or faeces (Bajtarevic 2009; Di Natale 2003; Peng 2010; Van de Goor 2018). The sensor signals can be combined into a 'breathprint' or 'smellprint', with a unique print for each disease or person. As the combination of these signals into a specific print requires statistical modelling, an algorithm must be developed, evaluated, and calibrated for each particular setting. After this, validation of the algorithm in a separate population is crucial, as it is often unknown what the eNose picks up as a signal, and artefacts may compromise its usefulness. The disadvantage of eNoses is that they may be very sensitive to environmental air, or smoking or alcohol use by the participants. Another disadvantage may be that the eNoses are seen as a 'black box', and the underlying algorithm is often not transparent. Therefore, validation of the eNose device, including the underlying algorithm should be completed in the setting in which the device will be used. If necessary, the algorithm can be adjusted to that particular practice situation.

A systematic review, published in 2019, summarized the diagnostic accuracy of eNoses for airway obstructive diseases, infectious and inflammatory diseases, several types of cancer, cystic fibrosis, and a number of other diagnoses (Farraia 2019). Although the authors claimed that "More than a half of the selected studies showed good accuracy", most of the included studies estimated sensitivity in a different groups of people (severe cases) than the group of people used to estimate specificity (health controls), which may not be representative of clinical practice. The review cited nine primary studies, which used exhaled breath or fecal gas to detect infections, the results of which were promising; for example, a study in which an eNose could predict the diagnosis of sinusitis in at least 72% of the samples correctly (Thaler 2006).

Dogs and other animals

Dogs can be trained to discriminate between different smells, and are well known for their ability to detect illegal drugs, corpses, or living people in damaged buildings. More and more, dogs are also being trained to detect people with certain conditions or infectious diseases (Bomers 2014; Hackner 2016). Sniffer dogs and their handler form a team, and the role of the handler is crucial in the evaluation of a dog for disease detection. Dogs may also play a role in the detection of people with SARS-CoV-2 infection or COVID-19, for example in airports, where they are already standard equipment for the detection of illegal drugs. Other animals may also be used for sniffing disease, but are socially less acceptable, such as cane rats, which are used to detect land mines, and have been trained to detect tuberculosis (Kanaan 2021). More recently, a Dutch research group trained bees to detect SARS-CoV-2 infections in humans (Reuters 2021). However, these results have not yet been made public. Disadvantages of sniffer animals include the possibility of potential health issues, such as zoonotic diseases, or anxiety among those who are afraid of dogs.

Clinical pathway

The clinical pathway depends on the place where the eNoses or sniffer animals will be used, which in turn, depends on their anticipated diagnostic test accuracy. Their role in detecting SARS-CoV-2 infection will most likely be as either a screening or triage test. As an example of an intended test use for triage in persons with mild symptoms, in The Netherlands, eNoses have been purchased by the government with the intention to lower the burden in public health service test locations, where people with symptoms of SARS-CoV-2 infection currently all undergo RT-PCR testing. The

intended role of the eNose would be to detect people in whom SARS-CoV-2 can be ruled-out with certainty, and who can be sent home safely. Using this approach, researchers anticipate that 70% to 75% of the RT-PCR tests may be avoided (De Vries 2021). This could be beneficial, as RT-PCR is relatively invasive and costly, and it may take between less than a day to more than a week for the results to become available, especially in locations with high levels of infection. Reducing the testing burden could also improve turnaround time for the remaining 25% who still need a nose-throat swab sample taken and sent to the laboratory for RT-PCR. A similar pathway, as a triage test, may be installed in other settings, such as in hospitals, and for case-and-contact tracing.

As an example of an intended test use for screening of asymptomatic people, sniffer animals, such as dogs, may be used at airports, and other places where a formal testing line would take too much time. Dogs could walk independently along a line of waiting people, and signal when they suspect someone is infected. Those who the dog found to be test-positive, may be referred for confirmation of infection by RT-PCR, or they may be isolated directly.

In more formal healthcare settings, such as a general practice or a hospital, fewer people (in absolute numbers) will be tested, so the use of expensive eNoses or dogs may not be feasible or practical. In these settings, both eNoses and sniffer animals need to have higher accuracy than other point-of-care tests, such as the antigen rapid tests, which can be purchased in bulk, and stored without many precautions.

Depending on the setting in which the eNoses and sniffer animals will be used, the minimally acceptable diagnostic accuracy may vary, as well as the actual accuracy of these tests. We will address these differences in the review.

Alternative test(s)

Several tests are currently used to diagnose SARS-CoV-2 infection and COVID-19. Signs and symptoms may be used to select people should be tested, for example, by RT-PCR. Routine laboratory tests are often used to assess the severity of disease and identify alternative diagnoses, but are less useful in diagnosing SARS-CoV-2 infection. Chest imaging is mainly used in hospitals - for example in emergency departments and intensive care units. RT-PCR is the most commonly used confirmatory diagnostic test for SARS-Cov-2 infection. To screen large groups of people, other novel tests have been considered, such as thermal imaging. We will not address these tests in this review, unless we find comparative designs where eNoses are compared to other tests.

The tests that are the most logical comparators for eNoses and sniffer dogs, are other point-of-care tests, such as rapid antigen and antibody tests. These are increasingly being used, although their diagnostic test accuracy varies. If we find sufficient studies that directly compare the diagnostic accuracy of eNoses or sniffer animals to other point-of-care tests, we will analyse these separately.

Rationale

Several tests are currently used to diagnose SARS-CoV-2 infection and COVID-19. To screen large groups of people, eNoses and sniffer dogs are being considered. Before clinical application, it is necessary to understand the diagnostic test accuracy of the currently available electronic and animal nose-driven tests for the detection of COVID-19.

This protocol is one in a series of protocols and reviews that covers the full series of Cochrane DTA Reviews for the diagnosis of COVID-19 (Deeks 2020; Dinnes 2020; Islam 2021; Stegeman 2020; Struyf 2021). Therefore, the background and methods sections of this review use some text that overlaps with some of our other reviews.

OBJECTIVES

1. To assess the diagnostic test accuracy of eNoses to screen for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection in public places, such as airports.
2. To assess the diagnostic test accuracy of sniffer animals, and more specifically dogs, to screen for SARS-CoV-2 infection in public places, such as airports.
3. To assess the diagnostic test accuracy of eNoses for SARS-CoV-2 infection or COVID-19 in symptomatic people presenting in the community, or in secondary care.
4. To assess the diagnostic test accuracy of sniffer animals, and more specifically dogs, for SARS-CoV-2 infection or COVID-19 in symptomatic people presenting in the community, or in secondary care.

Secondary objectives

If sufficient data are available, we will investigate the accuracy (either by stratified analysis, or by subgroup analysis) according to specific eNose technology or animal, and according to whether those who are tested are symptomatic or not. We will also investigate whether eNose brand, reference standard, and healthcare setting are associated with differences in diagnostic test accuracy.

METHODS

Criteria for considering studies for this review

Types of studies

For all questions, we will keep the eligibility criteria broad. We will include any study that produces estimates of diagnostic accuracy, both those using single-gate (also referred to as cohort) and multi-gate (also referred to as case-control) designs. We will include both algorithm development studies with internal validation only, and external validation studies. We will include studies regardless of their methodological or reporting quality, but we will carefully consider the limitations of different study designs in the assessment of methodological quality, the analysis, and the interpretation of findings.

We will exclude algorithm development studies without any evaluation.

Participants

Studies recruiting people who present with suspicion of COVID-19 are eligible, as well as asymptomatic people. This includes any group of people, with or without symptoms (e.g. adults and children). We will not exclude studies if the study population is unclear.

Index tests

We will include studies of any eNose device, and any sniffer animal. Any sample type will be eligible, including breath, saliva, respiratory secretions, urine, or sweat (wipes or cloths rubbed over the skin). We will include both hand-held devices and bench top models of eNoses; we will investigate the difference in accuracy between the two, when possible.

Target conditions

To be eligible, studies will need to report on the identification of:

(a) Current severe acute respiratory syndrome coronavirus 2 (SARS-COV-2) infection (either symptomatic or asymptomatic)

(b) COVID-19 disease, including COVID-19 pneumonia

Reference standards

We anticipate that a variety of reference standards will be used in studies within this review, and across the suite of reviews. Most studies are likely to use reverse transcription polymerase chain reaction (RT-PCR) of respiratory samples, which is generally considered to be the best available test for diagnosing SARS-CoV-2 infection. Although this test is considered to have excellent specificity, sensitivity is suboptimal, missing a substantial proportion of cases. For this reason, many studies repeat RT-PCR testing at least one more time in people who test negative, thereby increasing sensitivity. Alternatively, RT-PCR may be used in combination with other tests. To enable investigation into the role of the different reference standards, we will include studies regardless of the reference standard used, as defined by the authors.

Search methods for identification of studies

Electronic searches

The Cochrane COVID-19 Study Register is a freely-available, continually-updated, annotated reference collection of human primary studies on COVID-19, and can be found at covid-19.cochrane.org/.

The register contains records from:

1. Cochrane Central Register of Controlled Trials (CENTRAL) via the Cochrane Register of Studies
2. PubMed
3. Embase.com, provided under license from Elsevier
4. ClinicalTrials.gov
5. WHO International Clinical Trials Registry Platform (ICTRP)
6. medRxiv

For more information, please visit community.cochrane.org/about-covid-19-study-register.

We will use the Cochrane COVID-19 Study Register for study retrieval and will develop a search strategy to retrieve studies specifically for this review from this study register. The search strategy will be designed to reach maximum retrieval sensitivity, using the following terms:

1. (enose):AB OR (e-nose):AB OR (artificial NEAR3 nose):AB OR (spironose):AB OR (breathomix):AB OR (aeonose):AB OR (breath NEAR3 analys*):ab OR (electronic NEAR3 nose):ab OR (enose):TI OR (e-nose):TI OR (artificial NEAR3 nose):TI OR (spironose):TI OR (breathomix):TI OR (aeonose):TI OR (breath NEAR3 analys*):TI OR (electronic NEAR3 nose):TI OR (breath NEAR3 analys*):TI OR (electronic NEAR3 nose):TI OR (dog):AB OR (dogs):AB OR (canine):AB OR (RAT):AB OR (rats):AB AND COVID19:INREGISTER
2. ((volatile-organic NEAR1 compound*)):AB AND COVID19:INREGISTER
3. (VOC):AB AND COVID19:INREGISTER
4. MESH DESCRIPTOR Electronic Nose EXPLODE ALL AND COVID19:INREGISTER
5. #1 OR #2 OR #3 OR #4

We will apply no language limits. We may revise strategies, as indicated, to account for changes to the COVID-19 Study Register's eligibility criteria, changes to database interfaces, and search performance assessments.

Searching other resources

We will contact companies to request further information about studies. We will scan citations from included studies for relevance, and use Scopus to identify citations that reference included studies.

Data collection and analysis

Selection of studies

We will independently select studies, in duplicate. We will resolve disagreements by involving a third experienced review author for initial titles and abstract screening, and through discussion between three review authors for eligibility assessments.

Data extraction and management

We will independently extract data, in duplicate. We will resolve disagreements by discussion between three review authors. We will write to study authors to check details and obtain necessary information

Assessment of methodological quality

We will independently assess methodological quality of included studies, in duplicate. We will resolve disagreements by discussion between three review authors. The QUADAS-2 operationalization is similar to the other DTA protocols on COVID-19 (Appendix 1).

Statistical analysis and data synthesis

We will present the two-by-two tables and the estimated sensitivity and specificity from each evaluation, in each study, using paired forest plots.

We will analyse eNoses and sniffer animals separately, wherever possible, based on the results per participant tested, and not per sample. In cases where results are only presented per sample, we will indicate that this was the case, and investigate the effect of these studies in a sensitivity analysis. If studies evaluated multiple tests or populations, we will include them multiple times. Where possible, we will present and meta-analyze the results for asymptomatic groups separately from symptomatic groups; if we find studies using different sniffer animals, we will analyse the different animals separately.

Where meta-analysis is possible, we will estimate average sensitivity and specificity using bivariate hierarchical models, where tests report binary results; when different studies report a different explicit threshold for the same test, we will use the HSROC model. We will present all estimates with 95% confidence intervals.

We will undertake meta-analysis in R (lme4, R) or using SAS software (NLMIXED, SAS 9.4), using existing validated macros as detailed in the *Cochrane Handbook for Diagnostic Test Accuracy Reviews* (Macaskill 2010).

When studies present only estimates of sensitivity or of specificity, we will fit univariate, random-effects, logistic regression models. We will clearly mark these analyses in the tables.

When studies only report positive or negative predictive values, we will fit univariate, random-effects, logistic regression models for those predictive values. We will clearly mark these analyses in the tables, and provide prevalence estimates. When too few studies are available for a bivariate meta-analysis, we will simplify models by first, assuming no correlation between sensitivity and specificity estimates, and secondly, by setting near-zero variance estimates of the random-effects to zero (Takwoingi 2017).

We will not make any formal comparisons between different tests, except when we find sufficient studies comparing eNoses or sniffer animals to a point-of-care test (and both are evaluated against the same reference standard, and in the same population). In that case, we will treat the different tests in the comparison as covariates in a bivariate meta-regression model.

Investigations of heterogeneity

If adequate data are available, we will investigate the sources of heterogeneity that are listed in the secondary objectives, using meta-regression models. If possible, we will investigate the eNose brand reference standard, and healthcare setting.

Sensitivity analyses

We will do a sensitivity analysis to investigate whether studies with low risk of bias provide different estimates of sensitivity and specificity than the estimates from all studies.

Assessment of reporting bias

We will contact researchers and manufacturers in the field, requesting they point us to ongoing and unpublished research.

ACKNOWLEDGEMENTS

The Cochrane Infectious Diseases Group (CIDG) Academic Editor is Dr Paul Hine, and Diagnostic Test Accuracy (DTA) Contact Editor is Dr Matthew Grainge.

We thank Dr Paul Brinkman for answering our technical questions about eNoses in this protocol.

We thank peer reviewers Dr Monica Staniek, Dr Susan Gould, and the DTA editorial team peer reviewers.

The CIDG editorial base is funded by UK aid from the UK government for the benefit of low- and middle-income countries (project number 300342-104). The views expressed do not necessarily reflect the UK government's official policies.

Jonathan Deeks is a UK National Institute for Health Research (NIHR) Senior Investigator Emeritus. Jonathan Deeks and Jacqueline Dinnes are supported by the NIHR Birmingham Biomedical Research Centre. This paper presents independent research supported by the NIHR Birmingham Biomedical Research Centre at the University Hospitals Birmingham NHS Foundation Trust and the University of Birmingham. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

Katy Bell is supported by a National Health and Medical Research Council (NHMRC), Australia Investigator grant (#1174523).

REFERENCES

Additional references

Bajtarevic 2009

Bajtarevic A, Ager C, Pienz M, Klieber M, Schwarz K, Ligor M, et al. Noninvasive detection of lung cancer by analysis of exhaled breath. *BMC Cancer* 2009;**9**(1):348.

BBC 2020

BBC Learning English. The sniffer dogs detecting coronavirus. www.youtube.com/watch?v=8dllVPYSArM 2 Oct 2020 (accessed 11 June 2021).

Bomers 2014

Bomers MK, van Agtmael MA, Luik H, Vandenbroucke-Grauls CM, Smulders YM. A detection dog to identify patients with Clostridium difficile infection during a hospital outbreak. *Journal of Infection* 2014;**69**(5):456-61. [DOI: 10]

Deeks 2020

Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Spijker R, Taylor-Phillips S, et al, Cochrane COVID-19 Diagnostic Test Accuracy Group. Antibody tests for identification of current and past infection with SARS-CoV-2. *Cochrane Database of Systematic Reviews* 2020, Issue 6. Art. No: CD013652. [DOI: 10.1002/14651858.CD013652] [PMID: 32584464]

De Vries 2021

De Vries R, Vigeveno RM, Mulder S, Farzan N, Vintges DR, Goeman JJ, et al. Ruling out SARS-CoV-2 infection using exhaled breath analysis by electronic nose in a public health setting. medRxiv 16 February 2021. [DOI: 10.1101/2021.02.14.21251712]

Di Natale 2003

Di Natale C, Macagnano A, Martinelli E, Paolesse R, D'Arcangelo G, Roscioni C, et al. Lung cancer identification by the analysis of breath by means of an array of non-selective gas sensors. *Biosensors and Bioelectronics* 2003;**18**(10):1209-18.

Dinnes 2020

Dinnes J, Deeks JJ, Adriano A, Berhane S, Davenport C, Dittrich S, et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database of Systematic Reviews* 2020, Issue 8. Art. No: CD013705. [DOI: 10.1002/14651858.CD013705]

Farraia 2019

Farraia MV, Cavaleiro Rufo J, Paciência I, Mendes F, Delgado L, Moreira A. The electronic nose technology in clinical diagnosis: a systematic review. *Porto Biomedical Journal* 2019;**4**(4):e42.

Hackner 2016

Hackner K, Errhalt P, Mueller MR, Speiser M, Marzluf BA, Schulheim A, et al. Canine scent detection for the diagnosis of lung cancer in a screening-like situation. *Journal of Breath Research* 2016 Sep 27;**10**(4):046003. [DOI: 10.1088/1752-7155/10/4/046003]

Islam 2021

Islam N, Ebrahimzadeh S, Salameh J-P, Kazi S, Fabiano N, Treanor L, et al. Thoracic imaging tests for the diagnosis of COVID-19. *Cochrane Database of Systematic Reviews* 2021, Issue 3. Art. No: CD013639. [DOI: 10.1002/14651858.CD013639.pub4]

Kanaan 2021

Kanaan R, Farkas N, Hegyi P, Soós A, Hegyi D, Németh K, et al. Rats sniff out pulmonary tuberculosis from sputum: a diagnostic accuracy meta-analysis. *Scientific Reports* 2021;**11**(1):1877. [DOI: 10.1038/s41598-021-81086-x]

Macaskill 2010

Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and Presenting Results. In: Deeks JJ, Bossuyt PM, Gatsonis C, editor(s). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2010. Available from http://srdta.cochrane.org/.

Peng 2010

Peng G, Hakim M, Broza YY, Billan S, Abdah-Bortnyak R, Kuten A, et al. Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors. *British Journal of Cancer* 2010;**103**(4):542-51.

R [Computer program]

R Foundation for Statistical Computing R: A language and environment for statistical computing. Version 3.4.2. Vienna, Austria: R Foundation for Statistical Computing, 2017. Available at www.R-project.org.

Reuters 2021

Reuters. Dutch researchers train bees to detect COVID-19. www.youtube.com/watch?v=1UgQK13VwoM 9 May 2021 (accessed 11 June 2021).

RIVM 2021

National Institute for Public Health and the Environment (RIVM). Testing. www.rivm.nl/en/coronavirus-covid-19/testing (accessed 11 June 2021).

Röck 2008

Röck F, Barsan N, Weimar U. Electronic nose: current status and future trends. *Chemical Reviews* 2008;**108**(2):705-25.

SAS 9.4 [Computer program]

SAS Institute SAS 9.4 for Windows. Version 9.4. Marlow, Buckinghamshire: SAS Institute, 2018.

Stegeman 2020

Stegeman I, Ochodo EA, Guleid F, Holtman GA, Yang B, Davenport C, et al, Cochrane COVID-19 Diagnostic Test Accuracy Group. Routine laboratory testing to determine if a patient has COVID-19. *Cochrane Database of Systematic Reviews* 2020, Issue 11. Art. No: CD013787. [DOI: 10.1002/14651858.CD013787]

Struyf 2021

Struyf T, Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MMG, et al. Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19. *Cochrane Database of Systematic Reviews* 2021, Issue 2. Art. No: CD013665. [DOI: 10.1002/14651858.CD013665.pub2]

Thaler 2006

Thaler ER, Hanson CW. Use of an electronic nose to diagnose bacterial sinusitis. *American Journal of Rhinology* 2006;**20**(2):170-2.

Van de Goor 2018

Van de Goor R, van Hooren M, Dingemans A-M, Kremer B, Kross K. Training and validating a portable electronic nose for lung cancer screening. *Journal of Thoracic Oncology* 2018;**13**(5):676-81.

Wilson 2015

Wilson AD. Advances in electronic-nose technologies for the detection of volatile biomarker metabolites in the human breath. *Metabolites* 2015;**5**(1):140-63. [DOI: 10.3390/metabo5010140] [PMID: 25738426]

APPENDICES

Appendix 1. QUADAS-2 operationalization

QUADAS-2		
Index test(s):	**eNose devices**	**Sniffer animals**
Participants (setting, intended use of index test, presentation, prior testing):	General practice, primary care, emergency care, hospital settings, and community test locations In people presenting with suspected COVID-19, or asymptomatic people No prior testing; sometimes selection based on signs and symptoms	Community test locations and airports In people presenting with suspected COVID-19, or asymptomatic people No prior testing; sometimes selection based on signs and symptoms
Reference standard and target condition:	The focus will be on the diagnosis of COVID-19 pneumonia or infection with SARS-CoV-2.	
Participants selection		
Was a consecutive or random sample of participants enrolled?	This will be similar for all index tests, target conditions, and populations. YES: if a study explicitly states that all participants within a certain time frame were included; that this was done consecutively; or that a random selection was done NO: if it is clear that a different selection procedure was used; e.g. selection based on clinician's preference, or based on institution UNCLEAR: if the selection procedure is not clear, or not reported at all	
Was a case-control design avoided?	This will be similar for all index tests, target conditions, and populations. YES: if a study explicitly states that all participants came from the same group of (suspected) people, with no differential selection by COVID-19 status NO: if it is clear that a different selection procedure was used for the participants, depending on their COVID-19 (pneumonia) status or SARS-CoV-2 infection status UNCLEAR: if the selection procedure is not clear or not reported at all	
Did the study avoid inappropriate exclusions?	Although the in- and exclusion criteria will be different for the different index tests, inappropriate exclusions will be similar for all index tests: e.g. only elderly people excluded, or children (as sampling may be more difficult). This needs to be addressed on a case-to-case basis.	

(Continued)

	YES: if all eligible participants were included, and if the numbers in the flow chart show not too many excluded participants (a maximum of 20% of eligible participants excluded without reasons) NO: if over 50% of eligible participants are excluded without providing a reason; if, in a retrospective study, participants without index test or reference standard results were excluded; if exclusion was based on severity assessment post-factum or comorbidities (cardiovascular disease, diabetes, immunosuppression) UNCLEAR: if the exclusion criteria are not reported	
Could the selection of participants have introduced bias?	HIGH: if one or more signalling questions were answered with NO, as any deviation from the selection process may lead to bias LOW: if all signalling questions were answered with YES UNCLEAR: all other instances	
Is there concern that the included participants do not match the review question?	HIGH: if accuracy of the index test was assessed using a case control design, or in an already highly selected group of participants LOW: any situation in which the index test is the first assessment or test to be done on the included participants UNCLEAR: if a description about the participants is lacking	
Index tests		
Index test(s):	eNose devices	Sniffer animals
Were the index test results interpreted without knowledge of the results of the reference standard?	YES: if blinding was explicitly stated, or the assessment was conducted before the COVID-19 status was known NO: if it was explicitly stated that the index test results were interpreted with knowledge of the results of the reference standard UNCLEAR: if blinding was unclearly reported.	YES: if blinding was explicitly stated for both the animal and the trainer, or the assessment was conducted before the COVID-19 status was known NO: if it was explicitly stated that the trainer was aware of the results of the reference standard; or if the participants or samples could have been traced in other ways UNCLEAR: if blinding was unclearly reported
If a threshold was used, was it pre-specified?	YES: if the device was stand-alone and used in a fixed way; i.e. if calibration was not necessary NO: if calibration was explicitly reported UNCLEAR: if threshold selection was unclearly reported	YES: if a definition for test positivity was given; i.e. the way through which the trainer sees that the dog has sniffed a 'case'. NO: if the definition for positivity was based on the results afterwards (this will usually not be the case) UNCLEAR: if the criteria for test positivity are unclear
Could the conduct or interpretation of the index test have introduced bias?	HIGH: if one or more signalling questions were answered with NO, as even in a laboratory situation, knowledge of the reference standard may lead to bias LOW: if all signalling questions were answered with YES UNCLEAR: all other instances	
Is there concern that the index test, its conduct, or interpretation differ from the review question?	HIGH: if the detection algorithm has not been appropriately validated for the setting in which the test is used	HIGH: if there are concerns that dogs or personnel have been trained in a way that would not apply in practice

(Continued)

	LOW: if the detection algorithm has been appropriately validated for the setting in which the test is used, or if the device is commercially and generally available UNCLEAR: if the current availability status of the device is unclear	LOW: if dogs or personnel have been trained in a way that applies to practice UNCLEAR: if no information was provided about the training of the dog
Reference standard		
Is the reference standard likely to correctly classify the target condition?	YES: for SARS-CoV-2 infection: RT-PCR, done by trained personnel, following guidelines for confirmed cases, and done with an assay targeting a minimum 2 targets NO: any other test UNCLEAR: if no reference standard was reported, or if it was just reported that RT-PCR was done	
Were the reference standard results interpreted without knowledge of the results of the index test?	YES: if it was explicitly stated that the reference standard results were interpreted without knowledge of the results of the index test, or if the reference test was conducted prior to the results of the index test being known NO: if it was explicitly stated that the reference standard results were interpreted with knowledge of the results of the index test, or if the index test was used to make the final diagnosis UNCLEAR: if blinding was unclearly reported	
Could the conduct or interpretation of the reference standard have introduced bias?	HIGH: if one or more signalling questions were answered with NO; If only the signalling question about blinding was answered with NO, and the only test used as a reference standard was the RT-PCR, and the target condition was infection, then this NO may be considered to have limited impact LOW: if all signalling questions were answered with YES UNCLEAR: all other instances	
Is there concern that the target condition as defined by the reference standard does not match the review question?	HIGH: if only RT-PCR was used for any of the target conditions, as highly probable COVID-cases with negative PCR would be missed by an Rt-PCR reference standard; if an alternative diagnosis is highly likely and not excluded (will happen in paediatric cases, where exclusion of other respiratory pathogens is also necessary); if tests used to follow up viral load in known positive tests LOW: if above situations not present UNCLEAR: if intention for testing is not reported in the study	
Flow and timing		
Was there an appropriate interval between index test(s) and reference standard?	YES: this will be similar for all index tests, populations, and target conditions: as the situation of a participant, including clinical presentation and disease progress, evolves rapidly, and new or ongoing exposure can result in case status change, an appropriate time interval will be within 24 hours. If the reference standard consists of multiple tests or test instances, then the first instance should be within 24 hours and the last within a maximum of two days. NO: if there is more than 24 hours between the index test and the reference standard, or if participants are otherwise reported to be assessed with the index versus reference standard test at moments of different severity UNCLEAR: If the time interval is not reported	
Did all participants receive a reference standard?	YES: if all participants received a reference standard (clearly no partial verification) NO: if only (part of) the index test positives or index test negatives received the complete reference standard	

(Continued)

	UNCLEAR: if it is not reported
Did all participants receive the same reference standard?	YES: if all participants received the same reference standard (clearly no differential verification) NO: if (part of) the index test positives or index test negatives received a different reference standard UNCLEAR: if it is not reported
Were all participants included in the analysis?	YES: if all included participants were included in the analyses as well NO: if after the inclusion/exclusion process, participants were removed from the analyses for different reasons: no reference standard done, no index test done, intermediate results of both index test or reference standard, indeterminate results of both index test or reference standard, samples unusable. UNCLEAR: if this is not clear from the reported numbers
Could the participant flow have introduced bias?	HIGH: if one or more signalling questions were answered with NO LOW: if all signalling questions were answered with YES UNCLEAR: all other instances

CONTRIBUTIONS OF AUTHORS

MMGL drafted the first version of the protocol.

RS designed the search strategy.

All other authors edited subsequent versions of the protocol, and approved the final version for publication.

DECLARATIONS OF INTEREST

MMGL has no known conflicts of interest.
KB has no known conflicts of interest.
JJD has no known conflicts of interest.
JaD has no known conflicts of interest.
JeD has no known conflicts of interest.
DAK has no known conflicts of interest.
SJL has no known conflicts of interest.
RS has no known conflicts of interest.

SOURCES OF SUPPORT

Internal sources

- Liverpool School of Tropical Medicine, UK
- University of Birmingham, UK

External sources

- Foreign, Commonwealth and Development Office (FCDO), UK

  Project number 300342-104
- NIHR Birmingham Biomedical Research Centre at the University Hospitals Birmingham NHS Foundation Trust and the University of Birmingham, UK
- National Institute for Health Research (NIHR), UK
- National Health and Medical Research Council (NHMRC), Australia

  Investigator grant #1174523 to Katy Bell

- The University of Queensland, Australia

  Provided support to Jenny Doust