


Scent dogs in detection of COVID-19: triple-blinded randomised trial and operational real-life screening in airport setting

Anu Kantele ,^{1,2} Juuso Paajanen,³ Soile Turunen,^{4,5} Sari H Pakkanen,^{1,2} Anu Patjas,^{1,2} Laura Itkonen,⁶ Elina Heiskanen,⁶ Maija Lappalainen,⁷ Loic Desquilbet,⁸ Olli Vapalahti,^{7,9} Anna Hielm-Björkman⁶

To cite: Kantele A, Paajanen J, Turunen S, *et al*. Scent dogs in detection of COVID-19: triple-blinded randomised trial and operational real-life screening in airport setting. *BMJ Global Health* 2022;**7**:e008024. doi:10.1136/bmjgh-2021-008024

Handling editor Seye Abimbola
AK and JP contributed equally.

Received 16 November 2021
Accepted 14 March 2022

ABSTRACT

Objective To estimate scent dogs' diagnostic accuracy in identification of people infected with SARS-CoV-2 in comparison with reverse transcriptase polymerase chain reaction (RT-PCR). We conducted a randomised triple-blinded validation trial, and a real-life study at the Helsinki-Vantaa International Airport, Finland.

Methods Four dogs were trained to detect COVID-19 using skin swabs from individuals tested for SARS-CoV-2 by RT-PCR. Our controlled triple-blinded validation study comprised four identical sets of 420 parallel samples (from 114 individuals tested positive and 306 negative by RT-PCR), randomly presented to each dog over seven trial sessions. In a real-life setting the dogs screened skin swabs from 303 incoming passengers all concomitantly examined by nasal swab SARS-CoV-2 RT-PCR. Our main outcomes were variables of diagnostic accuracy (sensitivity, specificity, positive predictive value, negative predictive value) for scent dog identification in comparison with RT-PCR.

Results Our validation experiments had an overall accuracy of 92% (95% CI 90% to 93%), a sensitivity of 92% (95% CI 89% to 94%) and a specificity of 91% (95% CI 89% to 93%) compared with RT-PCR. For our dogs, trained using the wild-type virus, performance was less accurate for the alpha variant (89% for confirmed wild-type vs 36% for alpha variant, OR 14.0, 95% CI 4.5 to 43.4). In the real-life setting, scent detection and RT-PCR matched 98.7% of the negative swabs. Scent airport prevalence (0.47%) did not allow sensitivity testing; our only SARS-CoV-2 positive swab was not identified (alpha variant). However, ad hoc analysis including predefined positive spike samples showed a total accuracy of 98% (95% CI 97% to 99%).

Conclusions This large randomised controlled triple-blinded validation study with a precalculated sample size conducted at an international airport showed that trained scent dogs screen airport passenger samples with high accuracy. One of our findings highlights the importance of continuous retraining as new variants emerge. Using scent dogs may present a valuable approach for high-throughput, rapid screening of large numbers of people.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Previous data suggest that scent dogs can discriminate between samples from individuals infected with SARS-CoV-2 and controls.

WHAT THIS STUDY ADDS

⇒ Scent dogs showed high diagnostic accuracy in a randomised, controlled, triple-blinded validation test with sample size based on power calculations.
⇒ Scent dogs trained with wild-type SARS-CoV-2 virus also mastered identification of other variants, although less accurately, revealing their robust discriminatory power and indicating a need for continual training to deal with emerging new variants of concern.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE AND/OR POLICY

⇒ Scent dog detection can serve as a prescreening method to save time and resources or even as the sole testing method when other approaches are not yet available—for example, at the early stages of a pandemic.
⇒ Scent dogs trained to screen SARS-CoV-2 carriers at a public international airport, and other similar mass gatherings, can provide a valuable tool to contain the pandemic.

INTRODUCTION

Containment of the COVID-19 pandemic necessitates rapid large-scale identification of infected individuals. Most patients with SARS-CoV-2 disease are either asymptomatic or have only mild symptoms, but can be contagious.¹ The test-and-isolate strategy has largely relied on the modern reverse transcriptase-polymerase chain reaction (RT-PCR) technique. Its practicality is hampered by inadequate availability, restricted testing capacity, high costs, long turnaround time, and prolonged positivity after infection.^{2 3} Rapid



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to
Dr Anu Kantele;
anu.kantele@hus.fi

screening methods, such as antigen tests are already in use.⁴ A fascinating screening strategy consists of detection by trained scent dogs, an approach not confined to laboratories, enabling large sample numbers with results in real time.^{5 6}

Dogs have an extremely sensitive olfactory system: their limit of detection reaches as low as one part per trillion concentrations,⁷ exceeding the instruments currently available.⁸ Dogs are presumed to detect distinct volatile organic compounds (VOCs)⁹ released by their hosts' metabolic processes in various conditions.⁵ Indeed, dogs have been reported to identify distinct VOCs elicited by various bacterial, viral and parasitic infections.^{10–12}

During the current pandemic, scent detection dogs have been trained to identify samples from hospitalised patients with COVID-19^{13–17} (online supplemental table 1). The preliminary data suggest that dogs can be trained within weeks to detect samples from SARS-CoV-2-infected individuals with an accuracy comparable to standard RT-PCR. However, stronger evidence is needed with power calculated sample sizes, better defined control groups, and above all, randomised double/triple-blinded research designs including previously unsniffed samples from the actual target population, outpatients. While proof-of-concept studies have been encouraging, scent dogs need to be taken from laboratory settings to real-life conditions.

The scent dog approach appears particularly appealing for screening SARS-CoV-2-infected individuals in public places and among masses of travellers at airports and harbours. In the spring 2020, we started training dogs to see whether they could identify samples from SARS-CoV-2-infected individuals and, in the autumn, started operational scent work at the Helsinki-Vantaa International Airport. Here, we present the results of a three-faceted study comprising (1) the dogs' training, (2) a prospective, randomised triple-blind validation study using four dogs and (3) a real-life prospective study using the same dogs in daily screening of incoming passengers at the airport, comprising a simultaneous scent detection dog test and nasopharyngeal SARS-CoV-2 RT-PCR.

METHODS

Study design

We explored whether scent dogs can be trained to identify humans with SARS-CoV-2 infection. The study was conducted at the scent detection dog training centre Wise Nose, Vantaa, Finland; University of Helsinki, Finland (Veterinary Faculty and Departments of Equine and Small Animal Medicine and Veterinary Biosciences and the DogRisk/Helsinki One Health and Medical Faculty, Department of Virology) and Meilahti Vaccine Research Centre, MeVac and Helsinki University Hospital Laboratory (HUSLAB), Helsinki University Hospital (HUH), Finland.

At the Helsinki-Vantaa International Airport, the study was conducted in a specifically designed cubicle



Figure 1 The purpose-built cubicle at the Helsinki-Vantaa International Airport. (A) The cubicle from the outside with the doors into the three sampling rooms. (B) Sampling room with a hatch for handing in the sample for the scent detection dog test. (C) A room for scent detection dog testing, showing two of the three hatches to the right. (D) White Shepherd, E.T., inside the test room, indicating the sample in the middle (No 2) as positive. During the validation, only three of the five scent track holes had cans with samples.

(figure 1A) built at the arrivals terminal. The cubicle setting was used in the dogs' final training, in the validation and as an area where the dogs screened incoming travellers in the real-life study. The cubicle had three sampling rooms for passengers' skin swab sampling (figure 1B), a working area for the dogs and sliding hatches on walls for samples and tracks (figure 1C).

Patient and public involvement

The participants were recruited from among the following groups: (1) inpatients in HUH, (2) outpatients and healthy individuals who were contacted by telephone or who contacted the study team in response to advertisements posted at PCR testing stations around Helsinki and (3) incoming flight passengers and personnel at Helsinki-Vantaa International Airport. For all inpatients and outpatients, a recent RT-PCR result was available at the time of recruitment. At the airport, recruited passengers were tested concomitantly by RT-PCR and a scent detection dog. Those with a pretravel negative RT-PCR test less than 72 hours old were not retested. For inclusion of airport employees, a RT-PCR test result within 72 hours was required.

All the volunteers gave written informed consent. They completed questionnaires on demographics, symptoms and PCR test results. If their results were not available at the time of form filling, these were obtained later. In addition, electronic medical records of hospitalised patients and personal interviews were used, when needed.

Individuals with incomplete questionnaires and/or non-availability of RT-PCR results were excluded. There were no restrictions of age, sex, nationality or concurrent diseases. In addition, samples from individuals

with asymptomatic or typical COVID-19 symptoms were included in all three arms of the study.

Skin swab sampling and specimen handling

All volunteers collected the skin swabs themselves using a sterile package containing five gauze eight ply swabs (Mölnlycke Healthcare AB, Göteborg, Sweden). They were instructed to separate the layers of the gauze and use a single layer to swab the skin of their neck, throat area, forehead, and wrists. They swabbed 5–20 gauze samples and, to avoid evaporation and cross-contamination during storage, they placed them in the smallest of three plastic zip lock bags (volumes 0.5 L, 1 L and 2 L) each of which they placed inside the next, larger one (Minigrip; Suominen Joustopakkaus Oy, Ikaalinen, Finland).

The outpatient samples to be used in the training, validation and as spike samples in the real-life cohort were collected from the volunteers at their homes shortly after their RT-PCR tests. The courier left the sampling kit at the front door, and after sampling, returned the samples to the dog training facility. All samples, positive and negative samples separated, were stored in plastic boxes (Smart Store; Orthex Sweden AB, Tingsryd, Sweden) in a dark place at 21°C–23°C until used. Samples with an unknown infection status were stored separately until the status was confirmed. The storage time for the validation samples was 0–5 months.

At the airport, in the sampling room (figure 1B), the gauze swab was placed in a 1 L plastic freezer bag (Pirkka; HP Rani Plast Ab, Teerijärvi, Finland) and then placed inside a metallic stainless steel can (85 mm high and 70 mm in diameter). Four extra swabbed gauze samples were placed in a 0.5 L plastic zip lock bag (Minigrip; Suominen Joustopakkaus Oy, Ikaalinen, Finland), enclosed in a 1 L zip lock bag and stored as previously described.

For each validation session, the scent tracks with samples were prepared in a separate location at the airport on the validation day. Sixty cans were prepared for each dog as follows: The zip lock bag containing the sample was opened, and a single gauze was transferred with sterile metallic tweezers into a can lined with a 1 L plastic freezer bag. To avoid sample odour contamination, the positive and negative samples were prepared on separate tables. The cans were loaded onto the scent tracks according to

a computer-generated randomisation list, and a trolley with the tracks transported to the cubicle.

The study team used adequate personal protective equipment, including a mask and powder-free nitrile gloves, when handling the study specimens.

Confirmation of a COVID-19 diagnosis

A COVID-19 diagnosis was based on a positive RT-PCR of a nasopharyngeal swab. Discrepancies in the validation results (at least two dogs giving a response different from that of the RT-PCR test) and in the real-life study were resolved by serum SARS-CoV-2 antibody test analysed by nucleocapsid protein and Spike IgG enzyme immunoassays, as described previously.¹⁸ When serum samples could not be obtained, viral load, symptoms and information about SARS-CoV-2 exposure were used to estimate the infection status.

The variant status of SARS-CoV-2 was determined by the SYNLAB for Diagnostics Centre of the Hospital District of Helsinki using data of S gene target failure (SGTF) for the alpha variant, HUSLAB using TaqPath COVID-19 PCR (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and TipMolBiol (Berlin, Germany) using N501Y mutation RT-PCR, which detects alpha or beta variants. The initial RT-PCR samples were subjected to genomic sequencing and bioinformatics analysis as previously described.¹⁹ Based on epidemiological data,²⁰ all samples obtained before 6 January 2020 were considered ‘wild-type,’ with reference to the D614G Wuhan-like strain.

Animals

All the dogs included in this study had previous experience of scent work (table 1).

Part I: Dog training

The initial training, aiming to provide the dogs with a clear scent picture of COVID-19, was carried out by skilled canine scent detection trainers using operant conditioning, with a clicker and treats used for positive reinforcement. In brief, first, the dogs were exposed to cans containing positive samples and taught to indicate a can with a positive sample. Second, they were introduced to a negative sample in parallel with a positive sample to allow scent discrimination. Third, the number of negative and positive samples was increased in the scent track to reinforce discrimination between positive

Table 1 The dogs’ characteristics, working history and indication behaviours

| Dog name | Breed | Age | Sex | Alert response for positive sample | Working history |
|----------|--------------------|-----|-----|------------------------------------|-----------------|
| Silja | Labrador retriever | 6 | F | Pawing | Narcotics |
| Rele | Labrador retriever | 5 | M | Sits | Dangerous goods |
| Kosti | Labrador retriever | 8 | M | Sits | Dangerous goods |
| E.T. | White Shepherd | 4 | F | Nose freeze + one paw | Canine cancer |

E.T. was initially trained to discriminate SARS-CoV-2 infection using urine samples and later, skin swab samples, while the other three were directly trained by skin swabs from patients with COVID-19 in the acute stage of the disease.

Se and Sp only for those samples the dogs truly sniffed (Se_{sniff} and Sp_{sniff}). Second, we calculated the Se and Sp for all the samples (Se_{all} and Sp_{all}). In this approach, we assumed that the dogs considered all un-sniffed samples negative, as they left them untouched. Positive predictive values (PPVs) and negative predictive values (NPVs) were calculated based on our data and on hypothetical prevalence scenarios²³: 40% reflecting a high prevalence setting such as a pandemic time hospital, and 1% reflecting a low prevalence setting such as an airport.

We also investigated whether some epidemiological/clinical variables were possibly associated with failure to identify positive samples. To do so, we restricted the calculations to include only positive samples. Positive samples were defined as a true positive (TP) if all four dogs correctly marked them, and as a false negative (FN) if even a single dog did not mark the sample. The candidate variables potentially associated with the outcome (FN) were: age, gender, concurrent chronic diseases (asthma, allergy, cancer, diabetes, migraine), presence of typical COVID-19 symptoms, duration of symptoms (time in days between symptom onset and sample collections), time between RT-PCR test and sample collection and type of virus (wild type vs alpha variant). Univariate logistic regression models were performed and ORs with their corresponding 95% confidence intervals (95% CI) were provided. For quantitative variables (such as age, duration of symptoms and time between RT-PCR test and sample collection), the linearity of the association was investigated using restricted cubic spline (RCS) functions with three knots located at the 5th, 50th, and 95th centiles of

the quantitative variable.²⁴ When the association with the outcome (on the log scale) was not considered as linear, the quantitative variable was included in the model with the RCS functions, and ORs were provided for arbitrary values. Because of the low number of FN samples, multivariate models could not be performed. However, in order to rule out the possible confounding effects of the candidate variables, bivariate logistic regression models were performed, exploring one by one each candidate variable as a potential confounder together with the variable of interest. ORs were considered significant (type-I error set at 0.05) when their corresponding 95% CI did not include the number 1. To obtain Se and Sp values for wild-type only samples, we used for calculations VAL1–2 (where no alpha variants had yet emerged). SAS University Edition (SAS Institute Inc., Cary, North Carolina, USA) was used for the statistical data analysis.

RESULTS

Participant and sample characteristics

For research design and number of samples in the three studies, see flowchart in figure 3. The particulars of the volunteers who provided samples for the validation and real-time studies are presented in table 2.

Part I: Dog training

The initial training relied on a positive reinforcement approach and predefined positive and negative samples. Having completed that phase at a training centre (qualifying by sensitivity and specificity exceeding 80%) and,

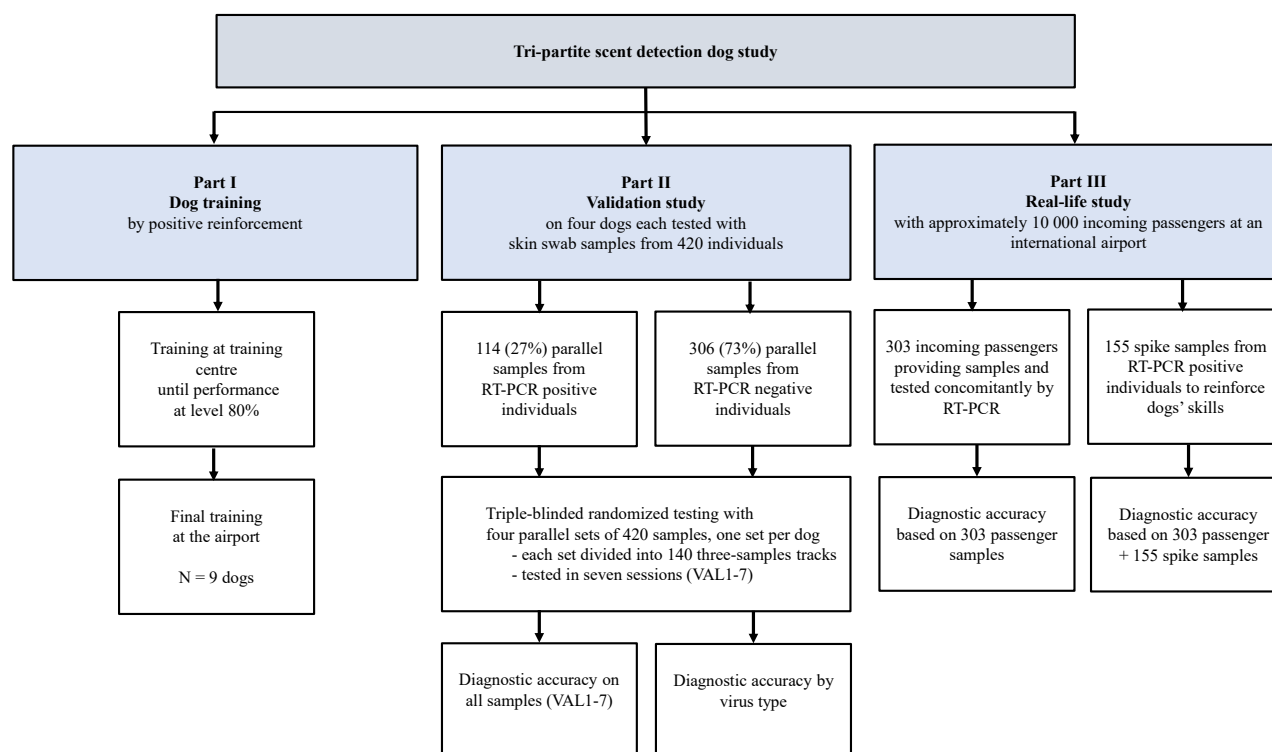


Figure 3 Flow chart of the study conduct.

Table 2 Data of volunteers providing skin swab samples with concomitant reverse transcriptase-polymerase chain reaction (RT-PCR) verification

| Characteristics | Skin swab samples used in the validation study | | | Skin swab samples sniffed by the validated dogs during operational work | | |
|---|--|-------------------------|-------------------------|---|-------------------------|-----------------------|
| | Overall (n=420) | RT-PCR negative (n=306) | RT-PCR positive (n=114) | Overall (n=303) | RT-PCR negative (n=300) | RT-PCR positive (n=3) |
| Age, median (IQR) | 38 (21) | 40 (23) | 34 (18) | 42 (22) | 42 (23) | 48 (NA) |
| Child, 0–12 years, n (%) | 15 (3.6) | 2 (0.7) | 13 (11.4) | 2 (0.7) | 2 (0.7) | 0 (0) |
| Sex, female, n (%) | 226 (53.8) | 168 (54.9) | 58 (50.9) | 192 (63.4) | 191 (63.7) | 1 (33.3) |
| Male, n (%) | 192 (45.7) | 137 (44.8) | 55 (48.2) | 111 (36.6) | 109 (36.3) | 2 (66.7) |
| Sample obtained, n (%) | | | | | | |
| Healthy screened | 301 (71.7) | 301 (98.3) | 0 (0) | 303 (100) | 300 (100) | 3 (100) |
| Hospitalised (non-COVID respiratory disease) | 2 (0.5) | 2 (0.7) | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| Outpatient | 117 (27.9) | 3 (1.0) | 114 (100) | 0 (0) | 0 (0) | 0 (0) |
| Days between PCR test and sampling, days, median (IQR) | 0 (2) | 0 (1) | 2 (1) | 0 (0) | 0 (0) | 0 (0) |
| Symptoms, n (%) | 304 (72.4) | 297 (97.1) | 7 (6.1) | 293 (96.7) | 291 (97.0) | 2 (66.7) |
| Asymptomatic | | | | | | |
| Respiratory infection | 116 (27.6) | 9 (2.9) | 107 (93.9) | 10 (3.3) | 9 (3.0) | 1 (33.3) |
| Days between start of symptoms and sampling, days, median (IQR) | 4 (3) | 3 (17) | 4 (3) | NA | NA | NA |
| SARS-CoV-2 variant*, n (%) | | | | | | |
| Wild-type† | 62 (14.8) | NA | 62 (54.4) | 2 (0.7) | NA | 2 (66.7) |
| Variant | 28 (6.7) | NA | 28 (24.6) | 1 (0.4) | NA | 1 (33.3) |
| Alpha | 25 (6.0) | NA | 25 (21.9) | 1 (0.4) | NA | 1 (33.3) |
| Beta | 1 (0.2) | NA | 1 (0.9) | 0 | NA | 0 |
| Alpha or beta | 2 (0.5) | NA | 2 (1.8) | 0 | NA | 0 |
| Unknown | 24 (5.7) | NA | 24 (21.1) | 0 | NA | 0 |
| Chronic disease, n (%) | | | | | | |
| Asthma, allergy | 28 (6.7) | 16 (5.2) | 12 (10.5) | 19 (6.3) | 18 (6.0) | 1 (33.3) |
| Cancer | 7 (1.7) | 2 (0.7) | 5 (4.4) | 1 (0.3) | 1 (0.3) | 0 (0) |
| Hypertension | 22 (5.2) | 14 (4.6) | 8 (7.0) | 25 (8.3) | 25 (8.3) | 0 (0) |
| Diabetes | 11 (2.6) | 6 (2.0) | 5 (4.4) | 8 (2.6) | 8 (2.7) | 0 (0) |
| Migraine | 3 (0.7) | 2 (0.7) | 1 (0.9) | 3 (1.0) | 3 (1.0) | 0 (0) |

*SARS-CoV-2 variant status was determined using S-Gene Target Failure (for alpha), N501Y Mutation PCR (for alpha or beta) and/or gene sequencing (for beta and some alphas) combined with epidemiological information (first alpha variant cases were detected 18 December in Finland).
†Wild-type refers to Wuhan-like lineages.

before starting operational work, the dogs were further coached in the purpose-built cubicle at the Helsinki-Vantaa International Airport (figure 1).

Part II: Validation

For the validation study, we selected the four dogs working at the airport during the study period. The conduct followed the detailed Helsinki University Scent detection validation protocol (online supplemental notes). Each dog was presented with an identical set of 420 parallel novel randomised samples, including 114 positives (27%) and 306 negatives (73%; table 2), in 140 fixed three-sample tracks, provided over seven validation trial sessions (VAL1–7). Of the 140 tracks, 26 (18.6%)

were randomised as not containing positive samples. The session and track order were not disclosed to personnel and varied by dog.

Overall, the diagnostic accuracy of all samples sniffed was 92% (95% CI 90% to 93%). The combined sensitivity Se_{sniff} and specificity Sp_{sniff} for all four dogs was 92% (95% CI 89% to 94%) and 91% (89%–93%), respectively (for un-sniffed samples and positive and negative predictive values, see table 3). Only minor variation was seen between the dogs: the best performance reached 93% (95% CI 85% to 96%) for Se and 95% (91% to 97%) for Sp , and the lowest 88% (80% to 94%) and 90% (85% to 93%), respectively. To obtain Se and Sp values for

Table 3 The diagnostic performance of the scent dogs in the triple-blind validation test

| Dog | Dogs' indication | Sniffed RT-PCR positive samples | Sniffed RT-PCR negative samples | Sniffed samples/all presented samples | Sniffed samples (% from all samples) | Se _{sniff} * (95% CI) | Sp _{sniff} † (95% CI) | PPV _{sniff} ‡ (95% CI) | NPV _{sniff} § (95% CI) | Se _{all} ¶ (95% CI) | Sp _{all} (95% CI) |
|----------|------------------|---------------------------------|---------------------------------|---------------------------------------|--------------------------------------|--------------------------------|--------------------------------|---------------------------------|---------------------------------|------------------------------|-------------------------------|
| Silla | Positive | 998 | 11 | 107/114 | 324 (77) | 93% (85% to 96%) | 95% (91% to 97%) | 90% (82% to 95%) | 96% (92% to 98%) | 87% (79% to 92%) | 96% (93% to 98%) |
| | Negative | | 206 | 217/306 | | | | | | | |
| Rele | Positive | 1047 | 18 | 111/114 | 326 (78) | 94% (87% to 97%) | 92% (87% to 95%) | 85% (77% to 91%) | 97% (93% to 98%) | 91% (84% to 95%) | 94% (91% to 96%) |
| | Negative | | 197 | 215/306 | | | | | | | |
| Kosti | Positive | 998 | 23 | 107/114 | 333 (79) | 93% (85% to 96%) | 90% (85% to 93%) | 81% (73% to 87%) | 96% (92% to 98%) | 87% (79% to 92%) | 92% (89% to 95%) |
| | Negative | | 203 | 226/306 | | | | | | | |
| E.T. | Positive | 90 | 23 | 102/114 | 326 (78) | 88% (80% to 94%) | 90% (85% to 93%) | 80% (71% to 86%) | 94% (90% to 97%) | 79% (70% to 86%) | 92% (89% to 95%) |
| | Negative | 12 | 201 | 224/306 | | | | | | | |
| All dogs | Positive | 392 | 75 | 427/456 | 1309 (78) | 92% (89% to 94%) | 91% (89% to 93%) | 84% (81% to 87%) | 96% (94% to 97%) | 86% (82% to 89%) | 94% (92% to 95%) |
| | Negative | 35 | 807 | 882/1224 | | | | | | | |

*Se_{sniff}, sensitivity on sniffed samples only.

†Sp_{sniff}, specificity on sniffed samples only.

‡PPV_{sniff}, positive predictive value on sniffed samples. Based on the prevalence rate of COVID-19 positive samples in our data (27%).

§NPV_{sniff}, negative predictive value on sniffed samples. Based on the prevalence rate of COVID-19 positive samples in our data (27%).

¶Se_{all}, sensitivity on all randomised samples; 10. Sp_{all}, specificity on all randomised samples.

||Se_{all}, reverse transcriptase-polymerase chain reaction.

detecting wild-type samples only, we used the data from VAL1–2 where no alpha variants had emerged, separately: the overall accuracy was 97% (95% CI 95% to 98%), the Se 99% (96% to 100%) and the Sp 96% (93% to 98%). The figures for each dog's individual validation session are provided in online supplemental table 2).

Discrepancies (at least two dogs' results differing from the RT-PCR) were observed for 19/420 samples (table 4), 79% of them in VAL6–7, with samples gathered in late February–March 2021. Eight of the 19 samples were RT-PCR positive. Our re-evaluation of sample status (based on RT-PCR viral load, symptoms, time since symptom onset and antibody data) confirmed all these eight as SARS-CoV-2 positive, of which six were alpha variants, one not known, and one was wild-type. Of the 11 RT-PCR negative samples detected as positive by the dogs, six were confirmed as SARS-CoV-2 negative, four as uncertain and one as a possible positive.

Based on the prevalence rate of COVID-19 positive samples in our data (27%), the overall PPV and NPV were 83.9% (95% CI 80.8% to 86.7%) and 95.8% (94.4% to 96.9%), respectively (table 3). In a population with 40% prevalence, the PPV and NPV were calculated as 87.8% (95% CI 85.3% to 90.0%) and 94.4% (92.4% to 95.8%), respectively. In a population with a prevalence of 1% the PPV and NPV were 9.8% (8.1% to 12.0%) and 99.91% (99.88% to 99.93%), respectively.

Of the 114 positive COVID-19 samples, 30 were FN and 84 were TP. Failure to identify a COVID-19 positive sample was associated with the SARS-CoV-2 variant status (alpha vs wild-type; OR=14.0; 95% CI, 4.5 to 43.4; table 5): the dogs indicated correctly 89% of the confirmed wild-type samples but only 36% of the alpha variant samples. Based on the OR values with the 95% CI, gender, concurrent chronic disease, time between start of symptoms and sampling, time since PCR test and increasing age of patients were found not to be associated with failure to identify a COVID-19 positive samples (table 5). None of the ORs presented in table 5 were modified in bivariate analyses (data not shown).

Part III: Real-life cohort

The dog identification and the RT-PCR result matched for 296/303 (97.7%) of the real-life samples of incoming passengers. The dogs correctly identified the samples as negative for 296/300 (98.7%) RT-PCR negative swabs. Table 6 provides details of the seven discrepant results. The dogs indicated three RT-PCR positive cases as negative. After re-evaluation with clinical and serological data, one was judged as SARS-CoV-2 negative, one as SARS-CoV-2 positive and one as a likely postinfectious positive RT-PCR result. Similarly, the dogs indicated four RT-PCR negative cases as positive. These were all judged as SARS-CoV-2 negative.

To maintain the dogs' screening skills in this low prevalence (0.47%) setting, a total of 155 novel RT-PCR positive 'spike' samples were provided to the dogs during working days (online supplemental table 3). They correctly indicated 98.7% of them as positive. Had the spike samples been calculated as part of the real-life study, the dogs' performance

Table 4 Validation participants with a discrepancy between SARS-CoV-2 RT-PCR and the response from two dogs or more

| Participants | Dog response | RT-PCR (viral load) | Symptoms | Time between symptom onset and skin swab/between RT-PCR test and skin swab, days | Antibody test (time between RT-PCR test and antibody test), days | Comment | SARS-CoV-2 infection status (variant*, if available) |
|--------------|--------------|---------------------|---|--|--|--|--|
| VL1 | Negative | Positive (medium) | Fever, rhinitis, diarrhoea | 9/5 | Positive (33) | Close relative PCR positive | Positive (alpha) |
| VL2 | Negative | Positive (medium) | Headache, cough, muscle pain | 4/3 | NA† | Relative and close contact PCR positive | Positive (alpha) |
| VL3 | Negative | Positive (low) | Asymptomatic | NA/5 | NA | 6-year-old child. Two relatives PCR positive | Positive (wild-type)‡ |
| VL4 | Negative | Positive (medium) | Fever, cough, loss of taste and smell, headache | 9/5 | Positive (48) | Relative PCR positive | Positive (alpha) |
| VL5 | Negative | Positive (high) | Sore throat, headache fever | 0/0 | NA | 11-year-old-child. Four relatives PCR positive | Positive (NA) |
| VL6 | Negative | Positive (medium) | Cough, muscle pain, loss of taste and smell, headache | 3/2 | NA | Two relatives PCR positive | Positive (alpha) |
| VL7 | Negative | Positive (medium) | Aches, tiredness, headache | 3/2 | Positive (108) | Close contact with PCR positive | Positive (alpha) |
| VL8 | Negative | Positive (medium) | Tiredness, sore throat, muscle pain, headache, loss of taste and smell, diarrhoea | 3/2 | NA | Close contact with PCR positive | Positive (alpha) |
| VL9 | Positive | Negative | Asymptomatic | NA/0 | Positive (99) | Three PCR negative within 2 weeks | Possible positive |
| VL10 | Positive | Negative | Asymptomatic | NA/3 | NA | NA | Uncertain |
| VL11 | Positive | Negative | Asymptomatic | NA/3 | Negative (85) | Two PCR negative within 3 days | Negative |
| VL12 | Positive | Negative | Asymptomatic | NA/0 | Negative (72) | Three PCR negative within 5 days | Negative |
| VL13 | Positive | Negative | Asymptomatic | NA/0 | NA | PCR negative month after dog test. | Uncertain |
| VL14 | Positive | Negative | Asymptomatic | NA/0 | NA | NA | Uncertain |
| VL15 | Positive | Negative | Asymptomatic | NA/0 | NA | Two PCR tests negative within 3 days | Uncertain |
| VL16 | Positive | Negative | Asymptomatic | NA/0 | Negative (82) | Three PCR negative tests within 3 days | Negative |
| VL17 | Positive | Negative | Asymptomatic | NA/0 | Negative (95) | Two PCR negative tests within 3 days | Negative |
| VL18 | Positive | Negative | Asymptomatic | NA/1 | Negative (83) | Two PCR positive | Negative |
| VL19 | Positive | Negative | Headache, cough | NA/2 | Negative (68) | Two PCR positive | Negative |

*Variant, SARS-CoV-2 variant status was determined using S-Gene Target Failure, N501Y Mutation PCR and/or gene sequencing combined with epidemiological information (first alpha variant cases were detected 18 December in Finland).
 †NA, not available.
 ‡Wild-type, refers to Wuhan like non-VoC lineages.
 RT-PCR, reverse transcriptase-polymerase chain reaction.

Table 5 Univariate analysis of associations between variables and failure to identify COVID-19 positive samples

| Variables suspected to be associated with the dog's performance | OR* | 95% CI† |
|---|------|-------------|
| Alpha variant (vs wild-type) | 14.0 | 4.5 to 43.4 |
| Presence of COVID-19 symptoms | 2.2 | 0.3 to 19.3 |
| Female (vs male) | 0.8 | 0.3 to 1.8 |
| Concurrent chronic disease | 1.0 | 0.4 to 2.9 |
| Days between start of symptoms and sampling (days) | 1.0 | 0.8 to 1.2 |
| Days between PCR test and sampling (days) | 1.1 | 0.8 to 1.5 |
| Age‡ | | |
| 30 | 1.0 | 0.7 to 1.3 |
| 40 (reference) | 1 | |
| 50 | 0.7 | 0.4 to 1.3 |
| 60 | 0.5 | 0.1 to 1.7 |

*OR, Odds Ratio.
 †CI, confidence interval, significant only when not including the number one.
 ‡Age was included by using restricted cubic spline functions (see text for details).

would have reached a sensitivity of 97% (95% CI 92% to 99%) and a specificity of 99% (96% to 100%).

In the operational real-life setup, we also used five non-validated dogs. Their results closely accorded with those of the validated dogs (data not shown).

DISCUSSION

Our study demonstrates that, in comparison with RT-PCR, scent detection dogs can be trained to identify SARS-CoV-2-infected individuals from skin swab samples with high diagnostic accuracy. In our real-life setting with a very low prevalence, the performance in identifying negative samples was very good (98.7%). Unfortunately, because of a low number of confirmed positive cases, accuracy with respect to positive samples could not be reliably assessed. However, ad hoc analysis also calculating the positive spike swabs showed a real-life performance of 98.5% for detecting positive samples. Below we discuss separately each of the three parts of the study.

Part I: Training of dogs

To keep the training short, we used dogs with previous scent work experience. Unlike studies conducted only in laboratory settings,^{13–17} we included two phases: initial training at a training centre, and—once the dogs were qualified—training in a challenging environment (Helsinki-Vantaa International Airport). Only one of nine dogs did not show high motivation for working in the test cubicle.

Part II: Validation test

Our validation experiment showed a high diagnostic accuracy with 92% sensitivity and 91% specificity. Several previous studies suggest that scent dogs can distinguish between samples from SARS-CoV-2-infected and uninfected individuals (reviewed in online supplemental table 1). However, although they demonstrate

Table 6 Real-life cohort participants with a discrepancy between SARS-CoV-2 RT-PCR and dog response

| Participants | Dog response | RT-PCR (viral load) | Symptoms | Time between symptom onset and skin swab/ between RT-PCR test and skin swab, days | Antibody test result (time between RT-PCR test and antibody test, days) | Comment | SARS-CoV-2 infection status (variant*, if available) |
|--------------|------------------------------|---------------------|---------------------------------------|---|---|---|--|
| RL1 | Negative | Positive (medium) | Asymptomatic | NA†/0 | Negative (81) | Two additional negative PCR tests | Negative (wild-type)‡ |
| RL2 | Negative | Positive (high) | Muscle aches, headache, fever | 1/0 | Positive (57) | | Positive (alpha) |
| RL3 | Negative | Positive (medium) | Asymptomatic | –10/0 | Positive (56) | Fever, cough, dyspnoea, headache 10 days before tests | Postinfectious prolonged PCR positivity (wild-type) |
| RL4 | Positive (two dogs) | Negative | Asymptomatic | NA/0 | Negative (178) | Negative scent dog test and PCR-test within 1 month | Negative |
| RL5 | Positive/Negative (two dogs) | Negative | Asymptomatic | NA/0 | Negative (97) | | Negative |
| RL6 | Positive (two dogs) | Negative | Chest pain, cough, tachycardia, fever | 1/0 | Negative (113) | Second PCR test negative 4 days after initial test | Negative |
| RL7 | Positive | Negative | Sore throat | –5/0 | Negative (55) | | Negative |

*Variant, SARS-CoV-2 variant status was determined using S-Gene Target Failure, N501Y Mutation PCR and/or gene sequencing combined with epidemiological information (first alpha variant cases were detected 18 December in Finland).
 †NA, not available.
 ‡Wild-type, refers to Wuhan like non_VoC lineages.
 RT-PCR, reverse transcriptase-polymerase chain reaction.

the dogs' diagnostic accuracy, these previous proof-of-concept studies^{13–17} had some limitations: small sample sizes,^{13 16 17} repeated use of the same samples,¹⁵ use of inactivated samples,^{13 14 16} use of empty cans or clean gauze swabs as controls^{15 16} and conducting validation tests only in laboratory settings.^{13–17} Perhaps even more importantly, in those studies almost all samples were collected from hospitals,^{13 15–17} failing to cover the actual target population, outpatients. Alternatively, positive samples were from hospitals and controls from outside hospitals—potentially misleading the dogs to identify hospital-associated odours as positive cues. Indeed, scent dog guidelines advise to watch out for systematic differences between positive and control samples.²⁵ Apart from these published proof-of-concept studies, some non-peer reviewed preprints provide further data²⁶: Guest *et al* collected 400 odour samples from patients with asymptomatic or mild COVID-19 and demonstrated in a randomised double-blind trial under laboratory conditions with six dogs a sensitivity range of 82%–94% and a specificity range of 76%–92%.

Based on the overall Se and Sp (92% and 91%, respectively), we calculated PPV and NPV according to two infection probability scenarios reflecting the prevalence of 40% and 1%. For a population with a prevalence of 40%, we estimated a PPV of 87.8% and a NPV of 94.4%. This means that the information provided by the dog (marking or not marking) increases the chances of detection to around 90%. For a population with a prevalence of 1%, by contrast, we estimated a PPV of 9.8% and an NPV of 99.9%. In both scenarios, high NPV supports the use of dogs for screening to exclude individuals not needing RT-PCR. We therefore suggest that dogs could be used both in sites of high SARS-CoV-2 prevalence, such as hospitals (to prescreen patients and personnel), as well as in low prevalence sites such as airports or ports (to prescreen passengers). Such prescreening could save considerably both time and PCR testing resources.

Our study design overcomes some of the major limitations of the previous studies: our sample size was based on a power calculation, our validation experiment was conducted outside laboratory conditions and our samples were collected in a random fashion as four parallel swabs, each used only once. We collected positive and negative swabs from both asymptomatic and symptomatic outpatients, children and seniors, and those with non-communicable diseases, and included samples collected during early and late phases of the disease. Unlike the previous validation studies, we randomly included tracks with no positive samples. This mimics better the real-life situation in low-prevalence settings.

In univariate analysis, the only variable strongly associated with failure to identify COVID-19 positive samples was the alpha variant (table 5). Indeed, according with the epidemiological situation in our country,²⁰ the virus variants started to emerge only at the end of our validation sample collection period, with 59% of the positive samples in VAL6–7 confirmed to represent virus variants

(mostly alpha) and only 3% the wild-type virus (the virus type of the others remained unknown). Importantly, the dogs had only been trained to detect samples of patients infected by the wild-type virus. The emergence of the new variants presumably explains the less successful performance by the dogs towards the end of the study period. In the bivariate analyses, after adjustment for all other variables possibly associated with the dogs' performance, the association between variant type and detection failure remained as strong as in the univariate analysis. Naturally, we cannot rule out confounding effects of other variables than the ones we investigated. Interestingly, Guest *et al* also had a small amount of alpha variant samples in their dataset. Their dogs correctly identified 38/48 (79%) of the alpha variant samples, the rate remaining lower than for the wild-type virus.²⁶ The difference was not significant, yet as their study was not designed to investigate the variants, it might have lacked statistical power. In our investigation the difference was highly significant according to the OR and its 95% CI, as the dogs correctly indicated 55/62 (89%) of the confirmed wild-type samples, but only 9/25 (36%) of the alpha variant samples. Thus, while the dogs indicated the alpha variant samples, their performance was lower than with the wild-type virus. Indeed, this observation is remarkable as it proves the scent dogs' robust discriminatory power. The obvious implication is that training samples should cover all epidemiologically relevant variants. Our preliminary observations suggest that dogs primed with one virus type can in a few hours be retrained to detect its variants (data not shown).

Another aspect to discuss is the low number of asymptomatic sample donors, which could have hampered the evaluation of the scent dogs' performance with samples from such individuals. In fact, the performance related to asymptomatic subjects is of particular importance, since in a real-world screening most individuals are asymptomatic. However, as we collected four samples from each of the sample donors, we ended up with 28 tests with samples from asymptomatic individuals. Only one was incorrectly identified as negative and two were left un-sniffed. Thus 25/28 (89.3%) were correctly identified as positive. In our analysis lack of symptoms was not associated with poorer performance.

Finally, since dogs may become tired or unfocused when working long hours, we ran the validation tests randomly in varying order over seven working days for each dog. Like in previous studies,^{14 15} all dogs did not perform equally. The differences were surprisingly small, however, particularly considering that the dog with the lowest results, E.T, was diagnosed with parotitis during her validation study, yet also her less successful days prior to diagnosis were included in her data. The low inter-dog variability observed in our study originates most probably from the consistent high-quality training performed both in the training centre and at the airport.

Part III: real-life air passenger study

While the validation experiment was successful, the real-life study at the airport met with some adversity. Although the dogs identified 98.7% of the negative samples as negative, they indicated four RT-PCR negatives as positives and did not identify three RT-PCR positive cases. Re-evaluations (the time span between symptom onset and sampling for RT-PCR, clinical symptoms, viral loads as cycle threshold counts and SARS-CoV-2 antibodies) of the RT-PCR positive cases suggested, however, that only one of the three represented the targeted group of early and potentially infectious cases. One of the RT-PCR positive individuals failed to seroconvert, suggesting a false-positive RT-PCR result. The samples of another three RT-PCR positive individuals had been collected as late as 10 days after symptom onset, presumably indicating a postinfectious positive RT-PCR result. Interestingly, the virus in the single verified case was identified as an alpha variant, possibly reflecting the dogs' lower sensitivity to detect it.

A major difference between the real-life and validation studies was seen in the rate of positive samples, which was over 50-fold lower in the real-life study than in validation (0.47% vs 27% of all samples)—that is, the dogs would have got only one positive sample to sniff each week at the most. Anticipating a low prevalence, the dogs' skills were kept up by providing them with a total of 155 novel (not sniffed by any dog), positive 'spike' samples over their shifts. Had these spike samples been included in the real-life study, the prevalence would have been 34%, not differing from that in the validation study, thus confirming the methods' potential for screening SARS-CoV-2 carriers. Similarly to any other diagnostic or screening tests, positive controls are needed to validate their accuracy. With the dogs, these spike samples serve as controls and also act as rewards, reinforcing the detection. In a low-prevalence setting, the use of spike samples needs to be preplanned before implementing scent dogs in the operational work. Of note, collection of spike samples from patients may no longer be needed in the future, as preliminary data suggest that spike material can be produced in the laboratory.²⁷

Limitations of the study

Some limitations deserve to be discussed. First, scent dogs previously trained to detect other substances such as drugs may also mark them, and the dog handler may record the marking falsely as positive for COVID-19. In this study, samples with false indications were not studied for narcotics and dangerous goods—that is, odours with which the three dogs were previously familiar.

Second, the age of the samples varied. The samples used in training and validation, as well as the 'spike' samples were older than in the real-life study, for they had to be verified before use: in the real-life operational setting, the samples were freshly collected and immediately presented to the dogs. We acknowledge that storage might have affected the VOCs.²⁸ Further studies have

been started to determine the precise nature of COVID-19-specific VOCs.

Third, the validation test also had some limitations. The low number of positive samples available led to a lack of tracks with multiple positive samples. This should not have had any greater effect on the results, as the dogs had practised both with blank tracks and tracks with multiple positive samples.

Finally, since variants did not emerge in Finland at the time of training, only wild-type samples were used. Many of the discrepant results were associated with the new variant. In the future, operational work skills should be kept up by simultaneous training with samples of emerging virus variants. Fortunately, once the dogs have received the basic training, retraining to cover new variants is expected to be easy as discussed above.

CONCLUSION

Employing a triple-blinded validation study setup, we provided evidence that trained scent dogs can master detection of samples from individuals infected with SARS-CoV-2 with good diagnostic accuracy. Interestingly, trained using samples only from individuals who had contracted the wild-type virus, the dogs' performance declined with samples of the variant era. We also provided some evidence that dogs can be trained to work at an international airport where large-scale rapid screening of crowds in a short period of time is required. In the real-life setting, we verified the results from our validation study for negative samples, but the dogs' ability to detect positive samples could not be confirmed owing to low prevalence of positive individuals. Ad hoc analysis also taking into account the positive spike samples, however, yielded convincing accuracy among the real-life cohort.

Author affiliations

¹Meilahti Vaccine Research Center, MeVac, Department of Infectious Diseases, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

²Human Microbiome Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland

³Department of Pulmonary Medicine, Heart and Lung Center, Helsinki University Hospital and University of Helsinki, Helsinki, Finland

⁴School of Pharmacy, Faculty of Health Sciences, University of Eastern Finland, Kuopio, Finland

⁵Nose Academy Ltd, Kuopio, Finland

⁶Department of Equine and Small Animal Medicine, University of Helsinki, Helsinki, Finland

⁷HUS Diagnostic Center, HUSLAB, Clinical Microbiology, Helsinki University Hospital and University of Helsinki, Helsinki, Finland

⁸Department of Biostatistics and Clinical Epidemiology, Ecole nationale vétérinaire d'Alfort, Maisons-Alfort, France

⁹Departments of Virology and Veterinary Biosciences, University of Helsinki, Helsinki, Finland

Acknowledgements The authors thank all scent dogs, dog handlers, trainers and study personnel who contributed to the study, particularly Ville Saksa, Eliisa Kekäläinen, Jaakko Kulonpalo, Sointu Mero, Simo Miettinen, Taru Miller, Suvu Kuivanen from Helsinki University Hospital and UH; the dog-handlers and assistants at the airport, Susanna Paavilainen, Anette Kare, Terhi Käppi, Marika Nöjd, Heli Niuro, Oona Pyykkö, Egil Björkman, Tanja Tall and Johanna Laitsaari; those collecting and handling the samples at Wise Nose Tuomas Stenius, Annika Tiainen, Kristina Luostarinen, Jon Mäkinen, Maiju Heikkinen, la Akkanen; as well as dog

trainer Tommy Wiren. We also thank the external controller Lauri Kanerva for his continuous participation in the validation trials. Finally, our most sincere thanks go to the scent dogs that we worked with in this project, and to all the volunteers who provided and delivered the samples used in the training, the validation test and the clinical application.

Contributors Study plan: AK, EH, LD, AH-B; ethics approval: AK, SHP, AP; sample acquisition: AK, JP, ST, SHP, AP, AH-B; laboratory analyses: OV, ML; statistical analysis: LD, AH-B; data curation: JP, ST, LI, EH, AH-B; funding acquisition: AK, OV, AH-B; supervision and project administration: AK, AH-B; writing of original draft: AK, JP, ST, AH-B; critical comments: OV; guarantor: AK, AH-B; review and approval of the final manuscript: AK, JP, ST, SHP, AP, LI, EH, LD, OV, AH-B.

Funding This study was supported by the veterinary clinic chain Evidensia, The Swedish Cultural Foundation in Finland (grant 138527), The Finnish Kennel Club, and The Finnish Cultural Foundation (grant 70287). The RT-PCR and antibody analyses were supported by the Finnish Cultural Foundation, the Finnish Governmental Subsidy for Health Science Research (TYH 2018322), the Finnish Medical Association, Jane and Aatos Erkko Foundation and the Academy of Finland (grants 1336490, 336 439 and 335527). The open access funded by Helsinki University Library. The funders of the study had no role in study design, data collection, data analysis or writing of the report.

Competing interests All authors have completed the ICMJE uniform disclosure form at <http://www.icmje.org/disclosure-of-interest/> and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval The study was part of a larger COVID-19 research protocol approved by the Helsinki University Hospital ethics committee (HUS/1238/2020). A research permit was obtained from the local authorities at HUH (HUS/157/2020) and Vantaa City (VD/8473/13.00.00/2020). All the participants provided written informed consent.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. All data relevant to the study are included in the article or uploaded as supplementary information. All data relevant to the study are included in the article or uploaded as supplementary information. Additional data are available upon reasonable request. Films are available through a link provided in the manuscript.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Anu Kantele <http://orcid.org/0000-0002-0004-1000>

REFERENCES

- Sayampanathan AA, Heng CS, Pin PH, *et al*. Infectivity of asymptomatic versus symptomatic COVID-19. *Lancet* 2021;397:93–4.
- Surkova E, Nikolayevskyy V, Drobniowski F. False-positive COVID-19 results: hidden problems and costs. *Lancet Respir Med* 2020;8:1167–8.
- Woloshin S, Patel N, Kesselheim AS. False negative tests for SARS-CoV-2 Infection - challenges and implications. *N Engl J Med* 2020;383:e38.
- Tang Y-W, Schmitz JE, Persing DH, *et al*. Laboratory diagnosis of COVID-19: current issues and challenges. *J Clin Microbiol* 2020;58:1–9.
- Angle TC, Passler T, Waggoner PL, *et al*. Real-time detection of a virus using detection dogs. *Front Vet Sci* 2015;2:1–6.
- Jones RT, Guest C, Lindsay SW, *et al*. Could bio-detection dogs be used to limit the spread of COVID-19 by travellers? *J Travel Med* 2020;27:1–3.
- Walker DB, Walker JC, Cavnar PJ, *et al*. Naturalistic quantification of canine olfactory sensitivity. *Appl Anim Behav* 2006;97:241–54.
- Angle C, Waggoner LP, Ferrando A, *et al*. Canine detection of the volatilome: a review of implications for pathogen and disease detection. *Front Vet Sci* 2016;3:1–7.
- Abd El Qader A, Lieberman D, Shemer Avni Y, *et al*. Volatile organic compounds generated by cultures of bacteria and viruses associated with respiratory infections. *Biomed Chromatogr* 2015;29:1783–90.
- Maurer M, McCulloch M, Willey AM, *et al*. Detection of bacteriuria by canine olfaction. *Open Forum Infect Dis* 2016;3:1–6.
- Bomers MK, van Agtmael MA, Luik H, *et al*. Using a dog's superior olfactory sensitivity to identify *Clostridium difficile* in stools and patients: proof of principle study. *BMJ* 2012;345:e7396–8.
- Guest C, Pinder M, Doggett M, *et al*. Trained dogs identify people with malaria parasites by their odour. *Lancet Infect Dis* 2019;19:578–80.
- Jendryn P, Schulz C, Twele F, *et al*. Scent dog identification of samples from COVID-19 patients - a pilot study. *BMC Infect Dis* 2020;20:1–7.
- Jendryn P, Twele F, Meller S, *et al*. Scent dog identification of SARS-CoV-2 infections in different body fluids. *BMC Infect Dis* 2021;21:1–14.
- Grandjean D, Sarkis R, Lecoq-Julien C, *et al*. Can the detection dog alert on COVID-19 positive persons by sniffing axillary sweat samples? A proof-of-concept study. *PLoS One* 2020;15:e0243122–19.
- Essler JL, Kane SA, Nolan P, *et al*. Discrimination of SARS-CoV-2 infected patient samples by detection dogs: a proof of concept study. *PLoS One* 2021;16:e0250158–14.
- Eskandari E, Ahmadi Marzaleh M, Roudgari H, *et al*. Sniffer dogs as a screening/diagnostic tool for COVID-19: a proof of concept study. *BMC Infect Dis* 2021;21:1–8.
- Rusanen J, Kareinen L, Levanov L, *et al*. A 10-minute “mix and read” antibody assay for SARS-CoV-2. *Viruses* 2021;13:143.
- Truong Nguyen PT, Plyusnin I, Sironen T, *et al*. HAVoC, a bioinformatic pipeline for reference-based consensus assembly and lineage assignment for SARS-CoV-2 sequences. *BMC Bioinformatics* 2021;22:373.
- Kant R, Nguyen PT, Blomqvist S, *et al*. Incidence trends for SARS-CoV-2 alpha and beta variants, Finland, spring 2021. *Emerg Infect Dis* 2021;27:3137–41.
- Rosner B. *Fundamentals of biostatistics*. 8th edn, 2016: 249–59.
- Trevethan R. Sensitivity, specificity, and predictive values: foundations, Pliabilities, and pitfalls in research and practice. *Front Public Health* 2017;5:p. 307.
- Altman DG, Bland JM. Diagnostic tests 2: predictive values. *BMJ* 1994;309:102.
- Desquilbet L, Mariotti F. Dose-response analyses using restricted cubic spline functions in public health research. *Stat Med* 2010;29:n/a–57.
- Edwards TL, Browne CM, Schoon A, *et al*. Animal olfactory detection of human diseases: guidelines and systematic review. *Journal of Veterinary Behavior* 2017;20:59–73.
- Guest C, Dewhirst SY, Allen DJ. Using trained dogs and organic semi-conducting sensors to identify asymptomatic and mild SARS-CoV-2 infections. Available: <https://www.lshmt.ac.uk/research/centres-projects-groups/using-dogs-to-detect-covid-19#results>
- Mendel J, Frank K, Edlin L, *et al*. Preliminary accuracy of COVID-19 odor detection by canines and HS-SPME-GC-MS using exhaled breath samples. *Forensic Sci Int Synerg* 2021;3:100155.
- Forbes SL, Rust L, Trebilcock K, *et al*. Effect of age and storage conditions on the volatile organic compound profile of blood. *Forensic Sci Med Pathol* 2014;10:570–82.

1 **Supplementary material**

2 **Table of Contents**

3 1. Supplementary Notes

4 A. Helsinki University scent dog validation protocol

5 B. Protocol violations

6 2. Supplementary Tables

7 3. Supplementary Appendix 1: Report of an external controller

8 4. Supplementary references

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31 **1. Supplementary notes.**

32 **A. Helsinki University scent dog validation protocol**

33 1. Positive/negative disease status of volunteers providing each of the samples is confirmed in advance
34 using an approved reference method.

35 2. Only novel samples i.e. new to the dog (no previous scenting), are used.

36 3. A power calculation is performed to obtain the lowest acceptable numbers of positive and negative
37 samples for determining the sensitivity and specificity while striving for a lower CI value of 80 %.

38 4. The ratio between positive and negative samples is predetermined but can be different in different
39 validation setups. All tracks are randomized using a computer-generated list, programmed to comprise
40 different amounts of positive samples per track as well as a predetermined proportion of “blank” tracks,
41 i.e. not including any positive samples.

42 5. The validation experiment is triple-blinded: the dog, the handler and the assistant presenting the
43 samples to the dog and handler inside the same testing space, are all blinded to the status of each track
44 (zero or more positive samples) and each sample (positive or negative).

45 6. Dogs are allowed to go directly to the positive sample and mark it. This is counted as a successful
46 result if there is a maximum of one single positive sample per track, thereby not necessitating sniffing
47 of the other samples (if agreed, after removing the positive sample, the dog may be allowed to continue
48 sniffing the remaining ones, but this part of the experiment is to be considered single-blinded (thereafter
49 only the dog is blinded).

50 7. To avoid regressing and to reinforce the target scent, the dogs are rewarded after each positive result.
51 Each sample will thus become unblinded onsite immediately after the validator has confirmed the result.

52 8. During validation each dog will sniff the predetermined number of samples. The working days
53 required for validation may vary in number, since the maximal amount of samples one dog can sniff in
54 one day depends on the dog and the setup.

55 9. All validations are video recorded using at least 3 cameras at different angles of the validation setup.
56 This allows any external evaluator to check the process and the results afterwards if needed.

57 10. An external pre-determined controller with societal credibility and expertise in scent dog work, has
58 to be present at all validation experiments, confirming that the process follows the pre-determined
59 protocol in detail. His/her report will be included in the study dossier.

60 11. The environment, such as sound and stress level, temperature and humidity should be optimal for
61 the dog.

62

63

64 B. Complete list of protocol violations

65

66 1. Four video cameras recorded the validations all the time. Cameras 1, 2 and 3 were linked
67 together and camera 4 was separate from the others, placed to show the runs and to allow the
68 external evaluator to see the interior of the cubicle from all angles, and to ensure that no one
69 acts in violation of the triple-blind protocol. There were occasions when the research team forgot
70 to start one or several of the cameras: 1) twice camera 4 was started late (E.T's, validations 6
71 and 7) yet the three other cameras were running from the start; 2) Rele's whole validation 6 was
72 recorded only on camera 4; 3) Silja's validation 3 had both camera systems off.

73 Evaluation: The researchers and the external evaluator considered these omissions of minor
74 significance to the implementation of the study.

75 2. The humidity in the working cubicle was suboptimal for the dogs: The temperature was quite
76 stable and not too warm (21-22°C) but the humidity was relatively low (21-25 %). Evaluation:
77 An undetermined, presumably moderate impact on performance, since humidity is known to
78 affect the volatility of the odour molecules and dogs adjust their behaviour to detect these
79 molecules when the humidity drops making it more demanding for scent work (1).

80 3. The Helsinki University scent dog validation protocol recommends a calm environment free of
81 external noise and stress factors.

82 Evaluation: An undetermined, presumably moderate impact on performance. As the validation
83 was conducted at the airport, we allowed normal airport stress factors (people rushing, loud
84 noises from suitcases and high heels, dogs coming to greet their owners, children crying,
85 loudspeakers giving corona advice every 3 minutes etc.) since we wanted to validate the dogs
86 in their operational work environment. Unfortunately, on top of all these factors, the Helsinki-
87 Vantaa airport had noisy construction work going on during the whole validation period, the
88 worst being concrete drilling on the next level, just above the cubicle. The noise level was
89 measured using a mobile phone app, showing values of up to 80 decibel during a validation.
90 This was unfortunate but neither the construction work nor the validation could be postponed.
91 The impact of this noise remains unclear.

92

93 Summary: none of the protocol violations were considered to have a major impact on conducting the
94 study or the dogs' performance.

95

96

97

98 **2. Supplementary Tables**

99

100

Supplementary Table 1. Review of the previously published COVID-19 scent dog validation studies

| Reference | Dogs (n) | Training period (weeks) | Sample type | RT-PCR ¹ positive samples (n) | RT-PCR negative samples (n) | Validation test | Success rate sensitivity / specificity |
|-----------|----------|-------------------------|--|---|---|---|--|
| 14 | 10 | 1 | Saliva, urine, sweat | Hospitalized and non-hospitalized (46) | Healthy non-hospitalized and other respiratory infections (47) | Blinding: dog, handler, Randomization:1+/6- Automated device (total 7 samples) | Overall success rate 92% Saliva: 93% / 82% / 96% Urine: 95% / 95% / 98% Sweat:93% / 91% / 94% |
| 13 | 8 | 1 | Saliva, trachea-bronchial secretion | Hospitalized (7) | Non-hospitalized (no recent signs of infection) (7) | Blinding: dog, handler, observer Randomization:1+/6- Automated device (total 7 samples) | 94% / 83% / 96% |
| 15 | 6 | 1-3 | Sweat | Hospitalized (95) | Hospitalized non-covid or healthy HCW ² (82) | Blinding: dog, handler Randomization:1+/2-3- /mocks Line of cones (total 3 or 4 samples) | 76-100% / NA ³ / NA |
| 17 | 6 | 7 | Pharyngeal secretion, face masks/clothes | ICU ⁴ (1. test 26, 2. test 50) | Hospitalized non-covid and healthy (1. test 54, 2. test 70) | Blinding: single-blind Randomization:1-3+/7-10- Training wheel (total of 10 samples) | 1. test: NA / 65% / 89% 2. test: NA / 86% / 93% |
| 16 | 9 | 5 | Urine, saliva | Hospitalized children (11) and adults (5) | Emergency department children (14) and asymptomatic, healthy, non-hospitalized adults (4) | Blinding: dog Randomization: 1+ / 3-4- / 7-8 distractors Training wheel (total of 12 samples) | Test 1: 94%/71%/ 99% Test 2: 94%/71%/98% Test 3:100%/75%/98% Test 4:62%/62%/98% Test 5:96%/68%/99% Test 6:11%/18%/41% Test 7:67%/22%/100% Test 8:100%/11%/94% |

101 1. RT-PCR, reverse transcription polymerase chain reaction, 2. HCW, healthcare worker, 3. NA, not available, 4. ICU, intensive care unit. +
 102 = PCR positive, - = PCR negative

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

Supplementary Table 2. Dogs performance by validation test

| VAL 1–7 and dates of sample collection | Order of validation days | Dog indication | PCR pos ¹ | PCR neg ² | Se _{sniff} ³ | Sp _{sniff} ⁴ | PPV _{sniff} ⁵ | NPV _{sniff} ⁶ |
|--|--------------------------|--------------------|----------------------|----------------------|----------------------------------|----------------------------------|-----------------------------------|-----------------------------------|
| SILJA | | | | | | | | |
| VAL-1 Dec 2020 | 1 | Pos Neg | 16 0 | 0 35 | 100 % | 100 % | 100 % | 100 % |
| VAL-2 Dec 2020–Jan 2021 | 2 | Pos Neg | 18 0 | 0 26 | 100 % | 100 % | 100 % | 100 % |
| VAL-3 Jan 2021 | 3 | Pos Neg | 15 1 | 2 25 | 94 % | 93 % | 88 % | 96 % |
| VAL-4 Jan–Feb 2021 | 4 | Pos Neg | 14 1 | 1 33 | 93 % | 97 % | 93 % | 97 % |
| VAL-5 Feb 2021 | 5 | Pos Neg | 15 0 | 0 26 | 100 % | 100 % | 100 % | 100 % |
| VAL1–5[†] | | Pos Neg | 78 2 | 3 145 | 98 % (90–100) | 98 % (94–99) | 96 % (89–99) | 99 % (95–100) |
| VAL-6 Feb–March 2021 | 6 | Pos Neg | 13 1 | 2 31 | 93 % | 94 % | 87 % | 97 % |
| VAL-7 March 2021 | 7 | Pos Neg | 8 5 | 6 30 | 62 % | 83 % | 57 % | 86 % |
| VAL6–7[†] | | Pos Neg | 21 6 | 8 61 | 78 % (57–90) | 88 % (79–95) | 72 % (52–87) | 91 % (81–96) |
| RELE | | | | | | | | |
| VAL-1 Dec 2020 | 5 | Pos Neg | 16 0 | 2 28 | 100 % | 93 % | 89 % | 100 % |
| VAL-2 Dec 2020–Jan 2021 | 1 | Pos Neg | 18 0 | 0 29 | 100 % | 100 % | 100 % | 100 % |
| VAL-3 Jan 2021 | 4 | Pos Neg | 17 0 | 0 29 | 100 % | 100 % | 100 % | 100 % |
| VAL-4 Jan–Feb 2021 | 2 | Pos Neg | 16 0 | 3 33 | 100 % | 92 % | 84 % | 100 % |
| VAL-5 Feb 2021 | 3 | Pos Neg | 15 0 | 1 31 | 100 % | 97 % | 94 % | 100 % |
| VAL1–5[†] | | Pos Neg | 82 0 | 6 150 | 100 % (94–100) | 96 % (91–98) | 93 % (85–97) | 100 % (97–100) |
| VAL-6 Feb–March 2021 | 6 | Pos Neg | 11 3 | 6 24 | 79 % | 80 % | 65 % | 89 % |
| VAL-7 March 2021 | 7 | Pos Neg | 11 4 | 6 23 | 73 % | 79 % | 65 % | 85 % |
| VAL6–7[†] | | Pos Neg | 22 7 | 12 47 | 76 % (56–89) | 80 % (67–89) | 65 % (46–80) | 87 % (74–94) |
| KOSTI | | | | | | | | |
| VAL-1 Dec 2020 | 4 | Pos Neg | 15 1 | 2 27 | 94 % | 93 % | 88 % | 96 % |
| VAL-2 Dec 2020–Jan 2021 | 7 | Pos Neg | 16 0 | 2 22 | 100 % | 92 % | 89 % | 100 % |

| | | | | | | | | |
|----------------------------|---|--------------------|-----------------|-------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| VAL-3 Jan 2021 | 3 | Pos Neg | 16 1 | 2 30 | 94 % | 94 % | 89 % | 97 % |
| VAL-4 Jan–Feb 2021 | 1 | Pos Neg | 13 2 | 3 37 | 87 % | 93 % | 81 % | 95 % |
| VAL-5 Feb 2021 | 2 | Pos Neg | 15 0 | 2 30 | 100 % | 94 % | 88 % | 100 % |
| VAL1–5† | | Pos Neg | 75 4 | 11 146 | 95 % (87–98) | 93 % (88–96) | 87 % (78–93) | 97 % (93–99) |
| VAL-6 Feb–March 2021 | 5 | Pos Neg | 15 0 | 2 33 | 100 % | 94 % | 88 % | 100 % |
| VAL-7 March 2021 | 6 | Pos Neg | 9 4 | 10 24 | 69 % | 71 % | 47 % | 86 % |
| VAL6–7† | | Pos Neg | 24 4 | 12 57 | 86 % (66–95) | 83 % (71–90) | 67 % (49–81) | 93 % (83–98) |
| E.T. | | | | | | | | |
| VAL-1 Dec 2020 | 3 | Pos Neg | 16 0 | 1 26 | 100 % | 96 % | 94 % | 100 % |
| VAL-2 Dec 2020–Jan 2021 | 5 | Pos Neg | 16 0 | 2 27 | 100 % | 93 % | 89 % | 100 % |
| VAL-3 Jan 2021 | 2 | Pos Neg | 13 3 | 3 27 | 81 % | 90 % | 81 % | 90 % |
| VAL-4 Jan–Feb 2021 | 6 | Pos Neg | 14 2 | 1 37 | 88 % | 97 % | 93 % | 95 % |
| VAL-5 Feb 2021 | 1 | Pos Neg | 11 2 | 4 31 | 85 % | 89 % | 73 % | 94 % |
| VAL1–5† | | Pos Neg | 70 7 | 11 148 | 91 % (81–96) | 93 % (88–96) | 86 % (77–93) | 95 % (91–98) |
| VAL-6 Feb–March 2021 | 4 | Pos Neg | 10 4 | 6 27 | 71 % | 82 % | 63 % | 87 % |
| VAL-7 March 2021 | 7 | Pos Neg | 10 1 | 6 26 | 91 % | 81 % | 63 % | 96 % |
| VAL6–7† | | Pos Neg | 20 5 | 12 53 | 80 % (59–92) | 82 % (70–90) | 63 % (44–78) | 91 % (80–97) |

125 1.Sniffed PCR positive, 2. Sniffed PCR negative, 3. Se_{sniff} , sensitivity on sniffed samples, 4. Sp_{sniff} , specificity on sniffed samples, 5. PPV_{sniff} ,
126 positive predictive value on sniffed samples, 6. NPV_{sniff} , negative predictive value on sniffed samples. †Shows summed sample numbers and
127 mean values with 95% Confidence Interval (bolded). The number of SARS-CoV-2 variants per validations: VAL-1 0 alpha/16 wild-type; VAL-2 0
128 alpha/15 wild-type; VAL-3 3 alpha/9 wild-type; VAL-4 1 alpha/13 wild-type; VAL-5 5 alpha/8 wild-type; VAL-6 7 alpha/1 wild-type; VAL-7 9 alpha/0
129 wild-type.

130
131
132
133
134
135
136
137
138

Supplementary Table 3. Data of volunteers providing spike skin swab samples with concomitant RT-PCR¹ verification used in real-life setting to maintain discrimination skills

| Characteristics | Spike samples (n=155) |
|--|--------------------------|
| Age, median (IQR) ² | 33 (27) |
| Child, 0–12 years, n (%) | 18 (11.6) |
| Sex, female, n (%) | 83 (53.5) |
| male, n (%) | 72 (46.5) |
| Sample obtained, n (%) | |
| Healthy screened | 0 (0) |
| Hospitalised | 0 (0) |
| Outpatient | 155 (100) |
| Sample delay from PCR-test, days, median (IQR) | 4 (4) |
| Symptoms, n (%) | |
| Asymptomatic | 12 (7.7) |
| Respiratory infection | 143 (92.3) |
| Chronic disease, n (%) | |
| Asthma, allergy | 22 (14.2) |
| Cancer | 6 (3.9) |
| Hypertension | 11 (7.1) |
| Diabetes | 5 (3.2) |
| Migraine | 4 (2.6) |

139 1. RT-PCR, reverse transcriptase polymerase chain reaction, 2. IQR, interquartile range

140
141
142
143
144
145
146
147
148
149
150

151 **3. Supplementary appendix. Report of the external controller**

152

Lauri Kanerva
Hautakankaantie 10-12, Orivesi, Finland

Orivesi 25.5.2021

I, Lauri Kanerva, retired K-9 police sergeant from the Finnish police K-9 unit and police dog training school, have served as an external controller during the validation of the COVID-19 scent detection dogs at the Helsinki-Vantaa airport during March-April 2021. The validation was performed as a co-operation between the DogRisk research group at the University of Helsinki, led by DVM, PhD Anna Hielm-Björkman, and the Helsinki University Hospital and Professor Anu Kantele.

My duties were to assure that the validation went according to the protocol, to see that the blinded dog handlers and assistant did not have a chance to see the track maps, and to see that there was no extra communication between the blinded dog handler and the blinded sample assistant and any of those, that had the "track facit" in their possession.

The setup was well organized and 6 of the 7 persons involved were the same during all 13 sessions. The seventh person was either of three veterinary students that were involved in the project and that had to share the evaluation work, as they had compulsory teaching. The seven persons that were present were:

Inside the validation cubicle: Dog, dog handler and sample assistant, all blinded (triple-blinded)
Outside the cubicle: One person giving the sample tracks to the sample assistant and one person taking the samples from the sample assistant. Further we were 3 persons watching the validation from a screen just outside the cubicle: One that informed the dog handler if the result was right or wrong so that the dog-handler could reward the dog, one that evaluated which of the samples the dog sniffed and which they marked, and the third was myself, as an external and independent controller.

The validation started with a "dress rehearsal" when all persons were present, and the setup was introduced to all dogs, handlers and validation crew. A whole validation with new samples was rehearsed. After that the 7 validation days per dog could be done in 12 days. It took longer that first planned since one of the dogs, E. T., got a lacrimal gland infection and had to be on antibiotics for 2 weeks. This infection also disturbed his working and the handler wanted to quit the validation 6 as she could see that the dog had no sense of smell and was merely guessing. So E. T continued his trials 15 days later. As one of the handlers also got ill, all validations stood still for 2 weeks.

From the point of dogs there were also some other disturbance: The building noise at the airport was at times so loud that even we could not hear each other. Also, the air was very dry (below 25% humidity) for dogs to be able to work at their best.

All validations were taken on video from 4 different angles (except two runs, when the researchers forgot to put on the video (7.4.2021 for dog Rele and 26.3.2021 for dog Silja). All videos can be obtained from the researchers on request.

But all in all, everything went as planned, following the protocol, and I can conclude that the validation can be considered valid.

If you have questions, please contact me on paukku.seri@gmail.com

153

Lauri Kanerva, Retired K-9 police sergeant

154

155

156

157 **4. Supplementary references**

158 1. Jinn, J., Connor, E. G. & Jacobs, L. F. How Ambient Environment Influences Olfactory Orientation in
159 Search and Rescue Dogs. *Chem. Senses* **45**, 625–634 (2020).

160

161

Trained sniffer dogs accurately detect airport passengers infected with SARS-CoV-2

Likely to be especially valuable in the early stages of any pandemic, suggest researchers

Trained sniffer dogs can accurately detect airport passengers infected with SARS-CoV-2, the virus responsible for COVID-19, finds research published in the open access journal **BMJ Global Health**.

This method of detection is likely to be especially valuable, not only in the early stages of a pandemic when other resources might not yet be available, but also to help contain an ongoing pandemic, suggest the researchers.

Dogs have a very keen sense of smell, and can pick up a scent at levels as low as one part per trillion, far exceeding any available mechanical techniques.

It is thought that they are able to detect distinct volatile organic compounds released during various metabolic processes in the body, including those generated by bacterial, viral, and parasitic infections.

Preliminary data suggest that dogs can be trained within weeks to detect samples from patients with COVID-19 infection, with a degree of accuracy comparable to that of a standard PCR nose and throat swab test.

While promising, these lab data results needed to be replicated in real-life conditions. The researchers therefore trained 4 dogs to sniff out SARS-CoV-2 in Spring 2020. Each of the dogs had previously been trained to sniff out illicit drugs or dangerous goods or cancer.

To test the dogs' detection skills, 420 volunteers provided four skin swab samples each. The 4 dogs each sniffed the skin samples from 114 of the volunteers who had tested positive for SARS-CoV-2 on a PCR swab test and from 306 who had tested negative. The samples were randomly presented to each dog over 7 trial sessions.

Overall, the diagnostic accuracy of all samples sniffed was 92%: combined sensitivity—accuracy of detecting those with the infection—was 92% and combined specificity—accuracy of detecting those without the infection—was 91%.

Only minor variation was seen among the dogs: the best performance reached 93% for sensitivity and 95% for specificity; the worst reached 88% for sensitivity and 90% for specificity.

Some 28 of the positive samples came from people who had had no symptoms. Only one was incorrectly identified as negative and two weren't sniffed, meaning that 25 of the 28 (just over 89%) were correctly identified as positive: the lack of symptoms didn't seem to affect the dogs' performance.

The 4 dogs were then put to work sniffing out 303 incoming passengers at Helsinki-Vantaa International Airport, Finland, between September 2020 and April 2021. Each passenger also took a PCR swab test.

The PCR and sniffer results matched in 296 out of 303 (98%) of the real-life samples. The dogs correctly identified the samples as negative in 296 out of 300 (99%) PCR negative swab tests and identified three PCR positive cases as negative.

After re-evaluation with clinical and serological data, one was judged to be SARS-CoV-2 negative, one SARS-CoV-2 positive, and one a likely post-infectious positive PCR test result.

Similarly, the dogs indicated 4 PCR negative cases as positive. These were all judged to be SARS-CoV-2 negative.

Because the prevalence of SARS-CoV-2 among the airport passengers was relatively low (less than 0.5%), 155 samples from people who had tested positive on a PCR swab test were also presented to the dogs.

The dogs correctly identified just under 99% of them as positive. Had these 'spike' samples been included in the real-life study, the dogs' performance would have reached a sensitivity of 97% and a specificity of 99%.

Based on these results, the researchers then calculated the proportion of true positive results (PPV) and the proportion of true negative results (NPV) in two hypothetical scenarios reflecting a population prevalence of SARS-CoV-2 of 40% and 1%.

For the prevalence of 40%, they estimated a PPV of 88% and an NPV of 94.5%. This means that the information provided by the dog increases the chances of detection to around 90%.

For a population prevalence of 1%, on the other hand, they estimated a PPV of just under 10% and an NPV of just under 100%.

In both scenarios, the high NPV backs the use of sniffer dogs for screening, with the aim of excluding people who don't need a PCR swab test, say the researchers.

And they suggest that: "dogs could be used both in sites of high SARS-CoV-2 prevalence, such as hospitals (to prescreen patients and personnel), as well as in low prevalence sites, such as airports or ports (to prescreen passengers)." This could save both considerable time and resource, they say.

The researchers acknowledge that dogs trained to sniff out other substances may mistakenly identify these substances as SARS-CoV-2 positive. The required storage period of the training and spiked samples may also have affected the viability of the volatile organic compounds, they say.

A key finding was that the dogs were less successful at correctly identifying the alpha variant as they had been trained to detect the wild type. But this just goes to show how good dogs are at distinguishing between different scents, say the researchers.

"This observation is remarkable as it proves the scent dogs' robust discriminatory power. The obvious implication is that training samples should cover all epidemiologically relevant variants. Our preliminary observations suggest that dogs primed with one virus type can in a few hours be retrained to detect its variants."