


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

Anu Kantele <sup>1,2</sup>, Juuso Paajanen,<sup>3</sup> Soile Turunen,<sup>4,5</sup> Sari H Pakkanen,<sup>1,2</sup> Anu Patjas,<sup>1,2</sup> Laura Itkonen,<sup>6</sup> Elina Heiskanen,<sup>6</sup> Maija Lappalainen,<sup>7</sup> Loic Desquilbet,<sup>8</sup> Olli Vapalahti,<sup>7,9</sup> Anna Hielm-Björkman<sup>6</sup>

**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.<sup>1</sup> 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.<sup>2 3</sup> 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.<sup>4</sup> 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.<sup>5 6</sup>

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

During the current pandemic, scent detection dogs have been trained to identify samples from hospitalised patients with COVID-19<sup>13–17</sup> (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 un-sniffed 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.

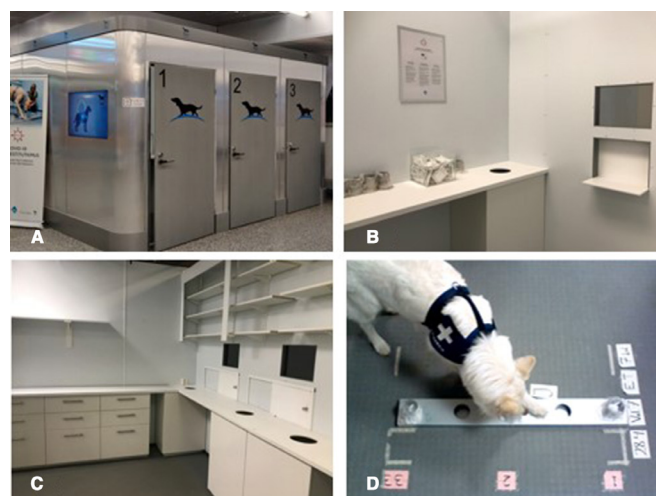
The scent dog approach appears particularly appealing for screening SARS-CoV-2-infected individuals in public places and among masses of travellers at airports and harbours. In the spring 2020, we started training dogs to see whether they could identify samples from SARS-CoV-2-infected individuals and, 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 Joustopakkauskset 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 Joustopakkauskset 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.<sup>18</sup> 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.<sup>19</sup> Based on epidemiological data,<sup>20</sup> 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

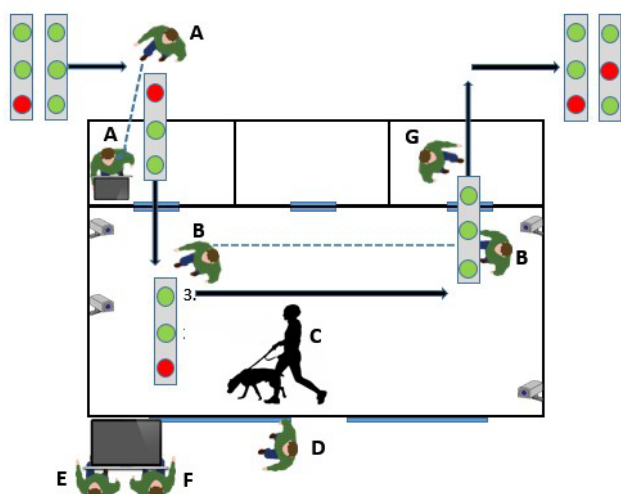
The initial training, aiming to provide the dogs with a clear scent picture of COVID-19, was carried out by skilled canine scent detection trainers using operant conditioning, with a clicker and treats used for positive reinforcement. In brief, first, the dogs were exposed to cans containing positive 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.





**Figure 2** Triple-blinded study. Assistant A gives the track through a hatch in the wall to assistant B, who places it on the floor and, after the dog and dog handler C have completed their work, gives it to assistant G. The dog handler C announces the result to data recorder D, who instructs whether to reward the dog. The external evaluator E and assistant F follow the setup from a video screen (four cameras inside the cubicle) and verify the triple-blinded study conduct. blinded: the dog, handler C, assistants B, E, G. circles: red, SARS-CoV-2 reverse transcriptase-polymerase chain reaction positive; green, negative sample.

and negative samples. Finally, confounding samples, including samples obtained from volunteers with other respiratory and viral diseases and samples from children, seniors or individuals with underlying diseases, such as asthma or allergies, cancer or diabetes, were introduced as controls. Once the dog and dog handler pair achieved a success rate of higher than 80% in detecting SARS-CoV-2 positive samples, the dog continued its training at Helsinki-Vantaa International Airport in the purpose-built working cubicle described earlier (figure 1A–D). The training was performed during a period when novel virus variants had not yet emerged in Finland.

The training was conducted using two different types of purpose-built metallic scent tracks each with either five or nine holes for the cans and/or triangular shaped metallic single can-holders. Cans used for positive samples were not mixed with cans used for negative samples. The cans and can-holders were washed in an industrial dishwasher at a temperature of around 85°C between every exercise.

## Part II: Triple-blinded validation study at Helsinki-Vantaa International Airport

The validation study was conducted according to the Helsinki University triple-blind validation protocol, as described in detail in the online supplemental notes; for the design and execution see [figure 2](#). In total, six investigators and one external controller were present at all validation sessions. Prior to the first validation day, the validation team and the dogs were familiarised with the study conditions and the protocol in a rehearsal session

identical to a validation session, here also introducing 60 novel samples. This was followed by seven validation sessions (VAL1–7). In these sessions, four parallel samples from the 420 individuals were randomised (samples: n=1680; dogs: n=4) into tracks of three samples, with 20 tracks in each of the seven sessions. Thus, each dog was exposed to 140 scent tracks. To allow comparisons between dogs, all four dogs received an identical set of samples. Thus, each parallel sample was used only once and for only one dog. The samples were assigned to the sessions (VAL1–7) in chronological order, the ones collected first in the VAL1 session and last in the VAL7 session. The order of the sessions was different for each dog. The dates of the sample collection and the order of the validation sessions per dog are shown in online supplemental table 2).

The dogs were rewarded for each positive result immediately after the correct indication. If a dog immediately selected the positive sample and skipped sniffing the other samples, the result was still recorded as successful.

The validation stage of the study was recorded using four cameras set up at different angles. A retired police sergeant from the Finnish K-9 police dog school was present during all the validation sessions as an external controller, confirming that the validations followed the predetermined protocol.

### Part III: Real-life cohort

The operational activity at Helsinki-Vantaa International Airport took place between 23 September 2020 and 30 April 2021. In total, 10 119 travellers (83.2%) and airport employees (16.8%) took part in the scent detection dog test, resulting in 48 (0.47%) samples indicated as positive. Part of these were recruited to the validation or the real-life study.

For collection of the skin swabs, see description above. The can with the sample was passed through a hatch to an assistant in the dogs' working space (figure 1C). The dog handler then placed the can in the five-hole scent track, together with a variable number of control samples (figure 1D). The dog handler interpreted the dog's indications as positive or negative for SARS-CoV-2 and a written test result was given to the participant.

## Statistical analysis

The power calculation suggested a minimum of 108 RT-PCR positive and 108 negative samples to achieve sensitivity (Se) and specificity (Sp) of 90%. This sample size was expected to have an 80% probability of obtaining an estimated Se and Sp of which the lower bound of the 95% CI would be greater than the minimal value of 80% (calculated using <https://www.stat.ubc.ca/~rollin/stats/ssize/b1.html>).<sup>21</sup>

Se and Sp were calculated according to Trevethan.<sup>22</sup> To cover incidences where the dogs directly marked a positive sample and skipped sniffing one or both of the other samples in the same track, Se and Sp were calculated using two separate methods. First, we calculated the

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 scenarios<sup>23</sup>: 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.<sup>24</sup> 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.

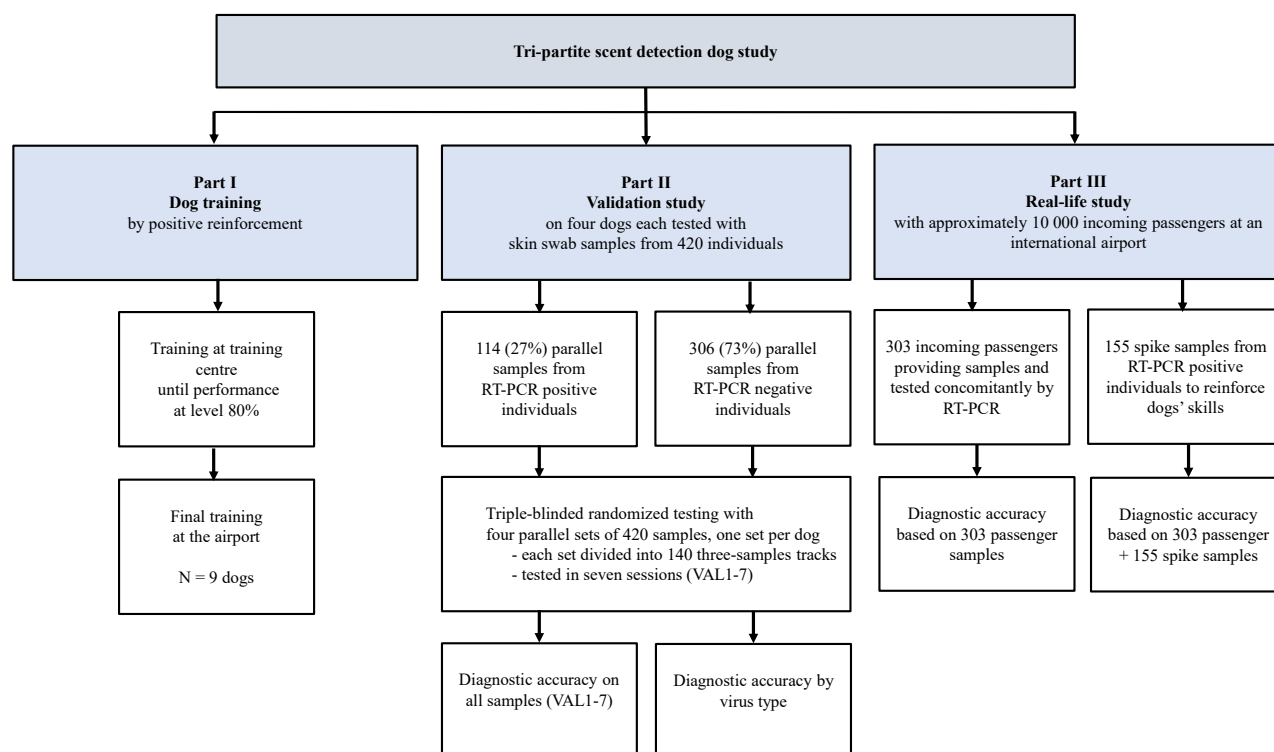
## RESULTS

### 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 <sub>sniff</sub> * (95% CI)	Sp <sub>sniff</sub> † (95% CI)	PPV <sub>sniff</sub> ‡ (95% CI)	NPV <sub>sniff</sub> § (95% CI)	Se <sub>all</sub> ¶ (95% CI)	Sp <sub>all</sub> (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<sub>sniff</sub>, sensitivity on sniffed samples only.

†Sp<sub>sniff</sub>, specificity on sniffed samples only.

‡PPV<sub>sniff</sub>, positive predictive value on sniffed samples. Based on the prevalence rate of COVID-19 positive samples in our data (27%).

§NPV<sub>sniff</sub>, negative predictive value on sniffed samples. Based on the prevalence rate of COVID-19 positive samples in our data (27%).

¶Se<sub>all</sub>, sensitivity on all randomised samples; 10, Sp<sub>all</sub>, specificity on all randomised samples.

RT-PCR, 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)		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,<sup>13–17</sup> 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<sup>13–17</sup> had some limitations: small samples sizes,<sup>13 16 17</sup> repeated use of the same samples,<sup>15</sup> use of inactivated samples,<sup>13 14 16</sup> use of empty cans or clean gauze swabs as controls<sup>15 16</sup> and conducting validation tests only in laboratory settings.<sup>13–17</sup> Perhaps even more importantly, in those studies almost all samples were collected from hospitals,<sup>13 15–17</sup> 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.<sup>25</sup> Apart from these published proof-of-concept studies, some non-peer reviewed preprints provide further data<sup>26</sup>: 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,<sup>20</sup> 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.<sup>26</sup> 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,<sup>14 15</sup> 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.<sup>27</sup>

### 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.<sup>28</sup> 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

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

<sup>2</sup>Human Microbiome Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland

<sup>3</sup>Department of Pulmonary Medicine, Heart and Lung Center, Helsinki University Hospital and University of Helsinki, Helsinki, Finland

<sup>4</sup>School of Pharmacy, Faculty of Health Sciences, University of Eastern Finland, Kuopio, Finland

<sup>5</sup>Nose Academy Ltd, Kuopio, Finland

<sup>6</sup>Department of Equine and Small Animal Medicine, University of Helsinki, Helsinki, Finland

<sup>7</sup>HUS Diagnostic Center, HUSLAB, Clinical Microbiology, Helsinki University Hospital and University of Helsinki, Helsinki, Finland

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

<sup>9</sup>Departments 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, Suvi 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 Sci* 2006;97:241–54.
- Angle C, Waggoner LP, Ferrando A, *et al*. Canine detection of the volatolome: 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.
- Jendry 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.
- Jendry 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.lshtm.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.