Comparison of the performance of SARS-CoV-2 RNA qRT-PCR testing based on expectorated and drooled saliva samples

Olivia Kay

Public Health England

Matthias Erwin Futschik (✉ matthias.futschik@ukhsa.gov.uk)

UK Health Security Agency   https://orcid.org/0000-0002-6245-8071

Elena Turek

Deloitte

David Chapman

Deloitte

Simon Carr

Deloitte

Malur Sudhanva

King\'s College Hospital NHS Trust: King's College Hospital NHS Foundation Trust

Paul Klapper

University of Manchester Faculty of Medical and Human Sciences: The University of Manchester
Faculty of Biology Medicine and Health

Tony Cox

UK Biocentre Milton Keynes

Michael Hill

University of Oxford

Joanna Cole-Hamilton

UK Health Security Agency

Peter Marks

UK Health Security Agency

Sarah Tunkel

UK Health Security Agency

Tim Peto

University of Oxford

Lindsey Davies

UK Health Security Agency

Tom Fowler

UK Health Security Agency   https://orcid.org/0000-0002-7258-2279

Research Article

Abstract

Purpose

Saliva has been considered a suitable sample material for SARS-CoV-2 testing but uncertainty remained regarding the sensitivity and reliability of different saliva collection methods for community mass testing. This study aimed to investigate the potential utility of expectorated saliva (ES) and drooled saliva (DS) through a large cohort study.

Methods

ES and DS samples were collected in a two-stage non-randomized prospective cohort study. Their utility for SARS-CoV-2 RNA qRT-PCR testing was assessed by comparison with results for combined throat and nose (CTN) swabs. A total of 2,878 subjects were recruited, from which 2,747 were evaluable for statistical analyses.

Results

Using CTN swab-based results as reference,DS- and ES-based tests showed the same high level of concordance (98% vs 98%) or specificity (99% vs 99%). Sensitivity seemed to be higher for DS than for ES (93% vs 80%) but not significantly once viral concentration was taken into account. Multivariable analysis indicated however an inferior sensitivity of saliva-based testing for female compared to male subjects with low viral burden. Assuming no false positive qRT-PCR results, an unbiased comparison showed no significant difference in sensitivity between saliva- and swab-based testing.

Conclusion

SARS-CoV-2 RNA testing based on saliva showed high diagnostic accuracy and can be considered an alternative where swabbing may not be tolerated or operationally feasible. Drooled saliva yielded the same diagnostic performance compared to expectorated saliva and may present a preferred option with reduced aerosol risk and increased compliance. Observed sex-specific difference in detection performance however warrant further investigations.

Introduction

The emergence of the novel coronavirus SARS-CoV-2 and its associated disease COVID-19 in early 2020 led to a global surge in the development of diagnostic tests and analytical methodologies. The current "gold standard test" for SARS-CoV-2 is the detection of viral nucleic acid via nucleic acid amplification [1]. Previous experience from coronavirus non-COVID-19 related infections enabled the rapid development and implementation of quantitative reverse transcription PCR (qRT-PCR) based testing (hereafter simply referred to as 'qPCR') [2]. Nasopharyngeal (NP), mid-turbinate nasal (MT) and oropharyngeal (OP) swabs have been identified as optimal sample types for detecting upper respiratory tract pathogens by the European Centre for Disease Prevention and Control (ECDC) [3] and US Centre for Disease Control and

Prevention (CDC) [4]. At the time of this study, the standard method of sample collection for community testing for SARS-CoV-2 in England was a combined throat and nose (hereafter referred to as 'CTN') swab sample; this could be collected by a trained professional or self-collected by the individual being tested [5].

Whilst this method is effective, other methods to collect samples for viral testing have been put forward [6]. Saliva may be a particularly suitable sample type for SARS-CoV-2 RNA testing, considering there might be multiple pathways for SARS-CoV-2 to be present in the saliva including from the salivary glands, respiratory tract and blood [7, 8]. Saliva collection may also overcome some of the challenges of MT/NP/OP swab methods and offers advantages such as the ease of self-testing, reduced need for specialist equipment such as swabs in VTM and increased adherence to continuous surveillance testing [9]. Thus, the less invasive nature of saliva collection could increase individuals' test-seeking behaviour, thereby allowing additional cases of COVID-19 to be diagnosed and onward community transmission reduced through self-isolation and other public health interventions. On the other hand, saliva as specimen poses challenges for sampling handling and processing in laboratories due to its inherent properties such as viscosity and heterogeneity [10]. Furthermore, saliva contains PCR inhibitors which may affect the effectiveness for PCR-based diagnostics especially if saliva is directly used in the PCR reaction [11].

Systematic reviews comparing the use of professionally collected NP/OP swab and saliva have indicated a similar diagnostic performance between the two sampling methods [6, 12–16]. However, the direct comparison between studies remained challenging because of great heterogeneity in terms of saliva collection. Saliva collection methodologies have included spit/drool saliva collection, direct saliva sampling and saliva-based swabbing, expectorated saliva and gargling saline [8, 17]. Saliva samples may contain whole saliva, sputum and gingival crevicular fluid [18]. Spit or drool-based samples are formed from whole saliva, including nasopharyngeal and posterior oropharyngeal secretions and can involve the unstimulated or stimulated accumulation of saliva [3, 12]. Expectorated saliva, collected through the process of throat-clearing and coughing followed by production of saliva, can contain both upper and lower respiratory tract secretions [12]. It remains unclear whether these diverse sample types and collection methods impact on the detection of SARS-CoV-2 by qPCR. There exists also uncertainty whether factors that are difficult to control in a community testing setting such as eating, drinking, smoking or vaping prior to providing a saliva specimen can affect the detection of viral RNA.

This study aimed to assess the suitability of saliva as a specimen for qPCR analysis in a community mass testing setting. Additionally, we wanted to examine the performance of two common saliva collection methods, namely expectorated saliva (ES) and drooled saliva (DS) in comparison to self-collected CTN swab as standard care.

Materials And Methods

Study Design

To accomplish the two aims of the study (suitability of saliva sampling and comparison of saliva sampling methods) a non-randomized prospective cohort study with 2-stage design was carried out (Fig. 1) using three testing methods (i) expectorated saliva (ES), (ii) drooled saliva (DS) and (iii) self-collected combined throat and nose (CTN) swab (standard of care reference test). The saliva collection methods were chosen, as they were considered to provide enough specimen to support the pandemic testing response. Self-collected CTN swabbing was the standard diagnostic sample collection method for qPCR testing within England's COVID-19 community testing programme at that time of the study. Stage 1 was designed to assess whether one saliva collection method showed a better concordance with CTN. In stage 2, recruitment would be continued for the superior method to increase statistical power and to assess effects of co-variates on performance. In case of equal performance in stage 1, the default was to select the DS method, which was considered the simpler and potentially safer collection method.

During the first part of stage 1, each subject was required to provide one ES sample (which involved clearing of the throat and expectoration of material into a collection vessel) and one self-sampled CTN swab. Once the required number of subjects were recruited for ES samples (19 August 2020), recruitment continued and subjects were requested to provide a DS sample, involving a passive "drool" of saliva into the collection vessel, as well as a self-sampled CTN swab. When the minimum recruitment target for stage 1 was reached (700 samples for each saliva collection method, see Sample Size section), recruitment to stage 2 was initiated to obtain at least 900 samples for the selected method i.e. DS (see Sample Size section). The combined data across both stages is reported in the analysis for the DS cohort when compared with the ES cohort.

Study recruitment and procedures

For both stages, data were collected from adult subjects across 21 NHS Test and Trace symptomatic ambulatory community testing sites in the East Midlands, Yorkshire and North West regions of England (including Leicester, Blackburn, Oldham, Rochdale, Nelson, Darwen, Kirkleen, Bradford, Bolton, Middleton, and Leeds) between 30 July 2020 and 8 October 2020. To resemble faithfully the heterogeneity typically found in community testing settings, no strict inclusion criteria were set for this study Participants were requested to abstain from food or drink 30 minutes prior to testing, but this was not enforced and participants ignoring this request were still included. The time of the last food or drink intake, as well as records of smoking or vaping prior to sampling was provided by 833 participants.

For both ES and DS groups, samples were collected via a plastic funnel into a polystyrene tube (International Scientific Supplies Ltd) to the minimal required level (1mL), to which 2.5mL of 0.9% $^{v}/_{v}$ saline was added. Subjects were required to add the saline from the pouch provided, remove funnel and seal the vial. The procedure was monitored by a trained testing operator and the method of taking the samples were recorded in a case report form.

Samples were collated and sent to the Milton Keynes UK Lighthouse Laboratory for processing. For both saliva and swab samples, RNA was purified using Thermo Fisher Kingfisher Flex instrument with MagMAX Viral/Pathogen II Nucleic Acid Isolation kit (ThermoFisher Scientific, UK). For detection,

ThermoFisher Applied Biosystems TaqPath COVID-19 RT-PCR assays were performed. Thes assay amplifies regions of the ORF1ab, N gene and S gene of the SARS-CoV-2 genome and uses a bacteriophage MS2 as internal PCR and extraction control. A qPCR test was determined as positive if a quantification cycle (Cq) value of less than 40 was observed for at least one of the three target genes. Both samples of a subject were stored and transported together and run together on the same PCR plate. Laboratory staff were blinded as to subject identity and mode of sample collection.

Primary and secondary outcomes

The primary outcomes were concordance for stage 1 and sensitivity for stage 2 determined as the ratios (TP + TN)/(TP + TN + FP + FN) and TP/(TP + FN), respectively, based the number of True Positives (TP), True Negatives (TN), False Negatives (FN) and the False Positives (FP) with self-sampled CTN swab based testing as reference. Primary outcomes were derived for (i) ES- and (ii) DS-based testing and compared. Secondary outcomes included specificity (TN/TN + FP), positive predictive value (PPV = TP/(TP + FP)), and negative predictive value (NPV = TN/(TN + FN))

Sample Size

Stage I

To estimate the required sample size for stage 1, a power calculation for a two sample test for proportions was carried out using the *power.prop.test* function implemented in the statistical programming language R. Seven hundred subjects per group were required to detect a 10% difference of concordance rates at a significant level of 0.05 with a power of 95%.

Stage II

To obtain a precision of 10% for sensitivity, 80 positive cases based on CTN swabs were needed when using the exact Clopper-Pearson method to derive confidence intervals (CIs). Assuming a prevalence rate of 5%, which was observed in stage 1, the overall sample size was estimated to be 1,600, which required at least an additional 900 subjects to be recruited in stage 2.

Statistical Analysis Methods

Statistical analyses were conducted on subjects with pairwise qPCR records of saliva collection (DS or ES) and CTN swab. The observed concordance, sensitivity, specificity, PPV and NPV were determined for ES and DS with respect to CTN swabs, along with exact 95% CIs using the Clopper-Pearson method. Statistical significance of differences in concordance, sensitivity, specificity, PPV and NPV was assessed using the χ^2-test with Yates' continuity correction. For analysis, Cq values for the different target genes (ORF1ab, N and S gene) were averaged for each sample. To test whether age, smoking, drink, or food intake could have influenced Cq values for saliva samples, univariable linear regression was carried out for the different factors (age, time to last smoke, vape, meal or drink). Viral concentrations (in digital copies/mL or dCopies/mL) were derived from Cq values based on previous calibration by the laboratory using Qnostics linearity analytical panel [19]. Univariate and multivariate logistic regression models were

used to model the relationship between the probability of a TP and $\log_{10}$ viral concentrations (Log VC), saliva collection, age in years, sex, and smoking/vaping status. For statistical analysis and visualisation of results, R version 4.0.3. was used.

Results

Study population and demographics

In the initial stage, 874 and 722 subjects were recruited to the (i) ES and (ii) DS cohorts, respectively. As concordance with CTN swab results was not statistically different (ES: 98.1%, DS: 98.0%, p = 0.8), the DS method was continued and a further 1,282 subjects were recruited. In total, 2,878 subjects took part in the study with 874 providing ES and 2,004 providing DS samples, while all participants underwent CTN swabbing (see CONSORT diagram, Fig. 1). 131 subjects were excluded because of missing paired samples or void samples. Statistical analysis was performed on the remaining 2,747 subjects, of which 1,904 were evaluable for comparison between DS vs CTN swabs and 843 between ES vs CTN swabs (Table 1). For ES vs DS, respectively, median (range) age was 40 (13−92) vs 36 (15−83) years (p < 0.001); Males 52% vs 47% (p = 0.07); smokers: 30% vs 28% (p = 0.28). There were 1.03% vs 2.15% void samples for ES vs DS respectively (p = 0.06). The positivity rate based on CTN swabs increased strongly from 5.2% for stage 1 to 18.9% for stage 2 in line with an increased prevalence of infection in the wider population over the relevant time period. The overall positivity rate was 10.4%.

Table 1
Demographics (age, sex, smoking/vaping status) and saliva qPCR results

		Stage 1		Stage 1 & 2	Total
Category	Statistic/Value	ES	DS	DS	
All Subjects	N	875	722	2003	2878
All Evaluable Subjects	N	843	699	1904	2747
Age (Years)	Median	40	42	36	37
	Min-Max	(13−92)	(15−83)	(15−83)	(13−92)
Sex (n, %)	Male	439 (52%)	351 (50%)	915 (48%)	1354 (49%)
	Female	395 (47%)	340 (49%)	958 (50%)	1353 (49%)
	Missing	9 (1%)	8 (1%)	31 (2%)	40 (1%)
Current smoker/vaper (n, %)	Yes	254 (30%)	195 (28%)	529 (28%)	783 (29%)
	No	582 (69%)	495 (71%)	1344 (71%)	1926 (70%)
	Missing	7 (1%)	9 (1%)	31 (2%)	38 (1%)
qPCR results (n, %)	Positive	40 (5%)	41 (6%)	245 (13%)	286 (10%)
	Negative	803 (95%)	658 (94%)	1700 (87%)	2747 (90%)
	Total	843 (100%)	699 (100%)	1904 (100%)	2747 (100%)

Comparison of Cq values from saliva and swab samples

We assessed Cq values obtained from the different specimens and found that saliva samples tended to have lower quantities of viral genetic material compared to swab samples (Fig. 2, **Supplementary Fig. 1**). This is reflected in significantly larger averaged Cq values for ES (mean: 25.04) and DS (24.17) with respect to CTN swabs (23.01) and the increase in Cq values held for the individual qPCR target genes (Fig. 2B-D). Despite differences in the underlying distributions, the correlation of Cq values between saliva and CTN swab was highly significant for all qPCR target genes ($p < 0.001$, Fig. 3). The strongest correlation ($R = 0.56$) was observed for averaged Cq values indicating that averaging over target genes tends to reduce variability of saliva Cq values with respect to CTN swab Cq values.

In univariable linear regression analyses, neither Cq values for saliva nor for CTN swabs showed statistically significant dependencies on age or time to last smoke, vape, meal or drink (**Supplementary Figures S2 and S3**). In addition, no significant difference ($p = 0.82$ or 0.77) between median Cq value for male (24.13 or 21.85) and female (24.0 or 21.95) was found for saliva or swab samples, respectively. A small proportion (10.7%) of our cohort did not comply with the advised nil by mouth period of 30 min. We

therefore analysed the Cq values from saliva samples of the 293 individuals who had their last meal (53), drink (211) or smoke (86) within 30 minutes of testing. Notably, no statistically significant difference in Cq values was observed between the samples taken within the 30 minutes window and those taken following the advised nil by mouth period (**Supplementary Figure S4**). There was also no significant difference detected between CTN swab Cq results from individuals with previous and without previous swabbing experience (p = 0.22).

Performance of saliva qPCR testing with CTN swab as reference

After merging data from stage 1 and stage 2, we assessed the saliva qPCR testing outcome with the testing outcome of CTN swabs as reference (Table 2). No statistically significant difference between ES and DS was found for concordance (98.1% vs 97.9%), specificity (99.0% vs 98.6%) and NPV (99.0% vs 99.0%). However, we observed a significantly higher sensitivity for DS (93.1%) compared to ES (80.0%, p = 0.007) as well as a higher PPV for DS (90.9%) compared to ES (80.0%, p = 0.04).

Table 2

Outcomes of saliva qPCR. Concordance rates, sensitivity, specificity, and predictive values were derived with CTN swabs as reference.

Statistic	ES (N = 843)	DS (N = 1904)	Saliva total (N = 2747)	DS vs ES P-value
True Positive	32	229	261	-
True Negative	795	1635	2430	-
False Positive	8	23	31	-
False Negative	8	17	25	-
Concordance (95% CI)	98.1 (96.4, 98.9)	97.9 (97.2, 98.5)	98.0 (97.4, 98.5)	0.73
Sensitivity (95% CI)	80.0 (64.4, 91.0)	93.1 (89.2, 95.9)	91.3 (87.4, 94.3)	**0.007**
Specificity (95% CI)	99.0 (98.1, 99.6)	98.6 (97.9, 99.1)	98.7 (98.2, 99.1)	0.42
PPV (95% CI)	80.0 (64.4, 91.0)	90.9 (89.2, 95.9)	89.4 (87.4, 94.3)	**0.04**
NPV (95% CI)	99.0 (98.1, 99.6)	99.0 (98.4, 99.4)	99.0 (98.5, 99.3)	0.94

Since elderly individuals are among COVID-19 high risk groups, the performance of saliva-based qPCR for individuals of age 60 or more was examined (**Supplementary table S1**). Based on 319 participants in this age group, a sensitivity of 90.9% was derived, which was not significantly different compared to the sensitivity of 91.5% observed for the 2311 younger participants in our cohort. We also did not detect a significant difference in sensitivity between smokers (92.86%) and non-smokers (91.08%, p = 0.82) (**Supplementary table S2**).

For 56 cases, we obtained divergent results for saliva- and swab-based testing. Thirty-one saliva samples were apparent false positives i.e. they were qPCR positive while the paired CTN swab sample was negative. Twenty-five saliva samples were apparent false negatives i.e. they were qPCR negative when the paired CTN swab sample was positive. It appeared that most of the divergent cases were associated with high Cq values (**Supplementary figure S5**). Some 90% of positive saliva samples paired with a negative swab sample had Cq values greater than 27.5, which corresponds to a viral concentration of 1000 SARS-Cov-2 dCopies/mL. In contrast, only 16% of positive saliva samples had paired positive swab sample with Cq values greater than 27.5. Similarly, 84% of the positive swab samples paired with a negative saliva sample had Cq values greater than 27.5, while only 15% of positive swab samples paired with positive saliva samples had Cq values greater than 27.5. Thus, discordant results between saliva and swab were predominantly from individuals with low viral burden suggesting that testing performance depends strongly on the viral concentrations of the samples. This observation motivated us to analyse the relationships between testing outcomes and viral concentration or other covariates in more detail using uni- and multivariable regression models.

Logistic models

The probability of observing a true positive saliva result was modelled with $\log_{10}$ viral concentrations (based on the corresponding swab qPCR), type of saliva specimen, sex, smoking status and age as independent continuous or categorical variables. Both univariable and multivariable logistic regression models were fitted to the data (Table 3). Univariable models led to significant odds ratios (OR) for $\log_{10}$ viral concentrations (p < 0.001) and type of saliva specimen (p = 0.008) in line with the previous observations that false negatives occur mainly when low viral concentration was low and that DS led to a higher overall sensitivity. Additionally, univariable modelling indicated a diminishing probability of a TP with increased age although the effect is marginal with an OR close to one. As viral concentration had the most significant effect, it was important to adjust for the viral concentration when assessing the impact of covariates on the saliva test outcome. For the multivariable model, $\log_{10}$ viral concentrations remains the most significant predictor variable. Notably, the OR for type of saliva sample was not significant after adjustment for the Cq value of the reference samples (Fig. 4A). Thus, the higher sensitivity for DS-based saliva qPCR can be explained by the higher viral concentrations observed in DS samples. Indeed, the mean Cq value for true positive DS samples was 21.9, while it was 24.5 for true positive ES samples. While age was not a significant predictor for a true positive saliva sample in the multivariable model, the adjusted OR of 3.7 for sex was significant (p = 0.02). The difference in predicted probabilities of a true positive is most notable for viral concentration below 10,000 dCopies/mL (Fig. 4B).

Logistic regression models. Odds ratios (OR) or adjusted odds ratios (aOR), their 95% CI and p-value from univariable and multivariable models. For multivariable models, the reference categories were "ES" for saliva sample type, "Female" for sex, and "No" for smoking/vaping status.

Independent variable	Univariable			Multivariable		
	OR	95% CI	p-value	aOR	95% CI	p-value
Log_{10} VC	2.80	1.95, 4.01	< 0.001	2.95	1.96, 4.45	< 0.001
Saliva sample (DS)	3.53	1.40, 8.91	0.008	1.06	0.31, 3.62	> 0.9
Age (Years)	0.97	0.97, 1.00	0.023	0.98	0.98, 1.01	0.14
Gender (Male)	2.17	0.92, 5.15	0.078	3.69	1.23, 11.1	0.02
Smoking Status (Yes)	1.28	0.46, 3.56	0.60	1.01	0.28, 3.61	> 0.9

Unbiased estimation of sensitivity of SARS-Cov-2 detection for saliva and swab samples

To assess the performance of saliva testing in community settings, we used self-sampled CTN swabbing as reference test. Despite having observed that age, sex or behavioural covariates do not significantly affect Cq values for swabs, we cannot fully exclude individual errors in the swabbing procedure. In particular, insufficient swabbing might lead to false negative results. To provide an unbiased estimation of sensitivities for qPCR tests based on different specimens, we assumed that all positive tests are true positives. This is equivalent to the assumption that the specificity is 100% for both saliva and swab, or that there are no false positives produced by qPCR tests. Under this assumption, we obtained a higher sensitivity for detecting SARS-Cov-2 by DS with 93.7% (95% CI: 90.1–96.3%) than by CTN swab with 90.3% (86.4–93.3%) or by ES with 83.3% (69.8–92.6%). The differences in sensitivity for the three types of specimens however were statistically not significant suggesting a similar capacity of SARS-Cov-2 detection.

Discussion

Saliva has been suggested as an alternative to NP/MT/OP swabs to obtain specimens for SARS-Cov-2 qPCR testing and various studies have compared their diagnostic performance [6, 8, 13, 15–17]. Most of these studies however were based on a small number of participants and only a few large cohort studies have been conducted. While meta-analyses have attempted to pool evidence across individual studies and indicate a sensitivity of around 85% and a specificity of 99% for saliva with NP swabs as reference [12, 17], the large heterogeneity between individual studies makes a comparison of saliva sampling methods or the analysis of potential confounding factors challenging. Our large prospective cohort study addressed this limitation and enabled a detailed statistical analysis of the performance of saliva-based

qPCR testing for SARS-Cov-2 qPCR. Importantly, we inspected potential relationships of Cq values with other factors and adjusted detection performance for confounding variables such as viral concentrations. These are crucially important aspects that have been frequently neglected in previous analyses.

Based on CTN swab as reference, the overall sensitivity was higher for DS than for ES. However, after taking variation in viral concentration into account, no significant difference in the probability to detect SARS-Cov-2 was found between the two saliva sample types. This highlights the importance of our study, as adjustments for difference in underlying viral concentrations have generally not been carried out and may not even be possible in the meta-analysis comparing different saliva testing approaches. Given the evidence of our study, DS might be considered as a preferrable sampling strategy over ES as it may reduce aerosolisation risk without compromising test performance. Assuming that qPCR testing for SARS-Cov-2 does not produce false positives, an unbiased estimate of sensitivity further indicated equal diagnostic performance of DS samples and self-sampled CTN swabs in the community testing setting.

While we detected a significant correlation between Cq values for saliva and swab samples from the same individual, higher Cq values for saliva samples compared to swab samples were recorded. This finding agrees with previous studies [8]. However, caution in the interpretation of the difference in Cq values is warranted, since the recorded viral concentration are directly affected by the applied protocols such as the used volume of viral transport medium for swabs or the volume of added saline to the saliva samples. More importantly, Cq values for saliva and swabs do not appear to be impacted by other factors such as age, sex or behaviour prior to testing. This indicates the potential strength of saliva-based testing in the presence of potentially confounding factors. In particular, we did not detect significant differences in Cq values between individuals who avoided eating, drinking or smoking/vaping in the 30 minutes prior to providing a sample and those that did not. This observation could support the saliva sampling without a food or drink free period, which may open opportunities for its use, for example, in social care and healthcare settings where medications and drinking are more difficult to restrict. However, further study is needed to clarify whether there exist specific contaminants which individuals need to avoid consuming prior to saliva sampling.

The overall concordance rate of saliva- and swab-based test was high (98%). Most of the discordant cases were associated with low viral concentration of less than 1000 dCopies per mL. For viral concentrations over 10,000 dCopies per mL, which have separately been associated with up to 85% of the infection events [20], we rarely found a swab qPCR positive that was missed by saliva testing. We detected only three saliva qPCR negative samples out of 178 that had with a paired swab positive samples equivalent to a sensitivity of 98.3%. Remarkably, our multivariable regression model indicated a higher probability of obtaining true positive result for male compared female participants. This difference is especially noticeable for low viral concentration. For a viral concentration of 100 dCopies/mL, the predicted probability of a true positive DS sample is 86% (95% CI: 72–93%) for a male while only 62% % (41–80%) for a female individuum. In contrast, the difference is minimal for high viral concentration. For 100,000 dCopies/mL, male participants would have a probability of 99% (97–100%), while female would have a probability of 98% (93–99%) to receive a true positive result. The reason for this difference

remains to be determined. Sex difference in the immune response of Covid patients have been reported previously [21]. Studies have also indicated sex-related differences in the propensities of saliva e.g. pH and protein content of healthy subjects [22] but the implications for qPCR testing is unclear. Thus, further confirmation and analysis of sex-specific differences in saliva qPCR testing are warranted.

A limitation of this study was the non-randomised sampling strategy. While this does not impact on the derived performance measure of saliva-based testing with respect to CTN swab-based testing as reference, it limited the comparability of the performances for ES and DS. Only after adjustment for differences in viral concentrations, a similar sensitivity for ES and DS was obtained. Furthermore, our study was conducted in an ambulatory setting, which may not be directly comparable to other diagnostic testing cohorts, as it required mobility by participants and thus less mobile, more vulnerable people were less likely to participate. Notably, this study was not designed to assess the user acceptability of saliva as a collection material. There may be specific population groups where saliva can be an easier and more suitable collection material than MT/NP/OP swabbing. For example, some studies indicated that saliva-based testing were the preferred option for children, particularly in younger age ranges [9, 23]. In contrast, there can be challenges to producing the appropriate volumes of saliva required, particularly in certain sub-populations including those in long-term care facilities[10]. Future studies might investigate the feasibility of saliva-based tests as an alternative to CTN swabbing within specific populations.

Conclusion

Our large prospective cohort study demonstrated that saliva-based SARS-CoV-2 RNA qPCR testing showed similar diagnostic performance compared with self-sampled CTN swabbing in community testing settings. Moreover, it indicates that a deep cough is not necessary for effective performance. Thus, our study can provide direct evidence for operational and policy decision-making. In practice, operational aspects such as compatibility with laboratory workflows (include automation, where the viscosity of saliva can pose problems), cultural acceptability, and other factors need to be considered in the decisions to use saliva as a specimen for SARS-CoV-2 RNA testing.

Statements And Declarations

Funding

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Author contributions

Data Availability

The datasets generated during for the presented study are not publicly available as they include sensitive data but will be made available after de-identification of participants from the corresponding author on reasonable request.

Ethics approval

Within the context of the pandemic public health response and roll out of testing interventions, after review using the Health Research Authority (HRA) tool and after further discussions with HRA it was determined that this evaluation would not require HRA research ethics approval. All study participants received routine care through receipt of an individual diagnostic swab test and result.

Consent to participate

Informed consent was obtained from all individual participants included in the study.

References

1. Beeching NJ, Fletcher TE, Beadsworth MBJ (2020) Covid-19: testing times. BMJ 369:m1403. https://doi.org/10.1136/bmj.m1403)

2. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK et al (2020) Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045. Euro Surveill 25

3. European Centre for Disease Prevention and Control (ECDC). Considerations for the use of saliva as sample material for COVID-19 testing (2021). [https://www.ecdc.europa.eu/en/publications-data/considerations-use-saliva-sample-material-covid-19-testing. Accessed 21 December 2022

4. Centre for Disease Control and Prevention (CDC). Nucleic Acid Amplification Tests (NAATs) (2021) [https://www.cdc.gov/coronavirus/2019-ncov/lab/naats.html. Accessed 21 December 2022

5. UK Health Security Agency (UKHSA). Coronavirus (COVID-19) PCR home test kit instructions 2022 [https://www.gov.uk/government/publications/testing-for-coronavirus-at-home. Accessed 21 December 2022

6. Tsang NNY, So HC, Ng KY, Cowling BJ, Leung GM, Ip DKM (2021) Diagnostic performance of different sampling approaches for SARS-CoV-2 RT-PCR testing: a systematic review and meta-analysis. Lancet Infect Dis 21:1233–1245. https://doi.org/10.1016/S1473-3099(21)00146-8)

7. Sabino-Silva R, Jardim ACG, Siqueira WL (2020) Coronavirus COVID-19 impacts to dentistry and potential salivary diagnosis. Clin Oral Investig 24:1619–1621. https://doi.org/10.1007/s00784-020-03248-x)

8. Kapoor P, Chowdhry A, Kharbanda OP, Bablani Popli D, Gautam K, Saini V (2021) Exploring salivary diagnostics in COVID-19: a scoping review and research suggestions. BDJ Open 7:8. https://doi.org/10.1038/s41405-021-00064-7)

9. Forster J, Streng A, Rudolph P, Rucker V, Wallstabe J, Timme S et al (2022) Feasibility of SARS-CoV-2 Surveillance Testing Among Children and Childcare Workers at German Day Care Centers: A Nonrandomized Controlled Trial. JAMA Netw Open 5:e2142057. https://doi.org/10.1001/jamanetworkopen.2021.42057)

10. Matic N, Stefanovic A, Leung V, Lawson T, Ritchie G, Li L et al (2021) Practical challenges to the clinical implementation of saliva for SARS-CoV-2 detection. Eur J Clin Microbiol Infect Dis 40:447–450. https://doi.org/10.1007/s10096-020-04090-5)

11. Morais O, Alves MR, Ramos C, Ferreira F, Fernandes P (2022) The Matrix Effect in the RT-PCR Detection of SARS-CoV-2 Using Saliva without RNA Extraction. Diagnostics (Basel) 12. https://doi.org/10.3390/diagnostics12071547)

12. Khiabani K, Amirzade-Iranaq MH (2021) Are saliva and deep throat sputum as reliable as common respiratory specimens for SARS-CoV-2 detection? A systematic review and meta-analysis. Am J Infect Control 49:1165–1176. https://doi.org/10.1016/j.ajic.2021.03.008)

13. Lee RA, Herigon JC, Benedetti A, Pollock NR, Denkinger CM (2021) Performance of Saliva, Oropharyngeal Swabs, and Nasal Swabs for SARS-CoV-2 Molecular Detection: a Systematic Review and Meta-analysis. J Clin Microbiol 59. https://doi.org/10.1128/JCM.02881-20)

14. Ibrahimi N, Delaunay-Moisan A, Hill C, Le Teuff G, Rupprecht JF, Thuret JY et al (2021) Screening for SARS-CoV-2 by RT-PCR: Saliva or nasopharyngeal swab? Rapid review and meta-analysis. PLoS ONE 16:e0253007. https://doi.org/10.1371/journal.pone.0253007)

15. Bastos ML, Perlman-Arrow S, Menzies D, Campbell JR (2021) The Sensitivity and Costs of Testing for SARS-CoV-2 Infection With Saliva Versus Nasopharyngeal Swabs: A Systematic Review and Meta-analysis. Ann Intern Med 174:501–510. https://doi.org/10.7326/M20-6569)

16. Butler-Laporte G, Lawandi A, Schiller I, Yao M, Dendukuri N, McDonald EG et al (2021) Comparison of Saliva and Nasopharyngeal Swab Nucleic Acid Amplification Testing for Detection of SARS-CoV-2: A Systematic Review and Meta-analysis. JAMA Intern Med 181:353–360. https://doi.org/10.1001/jamainternmed.2020.8876)

17. Duncan DB, Mackett K, Ali MU, Yamamura D, Balion C (2022) Performance of saliva compared with nasopharyngeal swab for diagnosis of COVID-19 by NAAT in cross-sectional studies: Systematic review and meta-analysis. Clin Biochem. https://doi.org/10.1016/j.clinbiochem.2022.08.004)

18. Subbarao KC, Nattuthurai GS, Sundararajan SK, Sujith I, Joseph J, Syedshah YP (2019) Gingival Crevicular Fluid: An Overview. J Pharm Bioallied Sci 11:S135–S9. https://doi.org/10.4103/JPBS.JPBS_56_19)

19. UK Health Security Agency (UKHSA). Evaluation of lateral flow device performance within the National Testing Programme (2022) [https://www.gov.uk/government/publications/lateral-flow-device-performance-data. Accessed 21 December 2022

20. Lee LYW, Rozmanowski S, Pang M, Charlett A, Anderson C, Hughes GJ et al (2022) Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infectivity by Viral Load, S Gene Variants and Demographic Factors, and the Utility of Lateral Flow Devices to Prevent Transmission. Clin Infect Dis 74:407–415. https://doi.org/10.1093/cid/ciab421)

21. Takahashi T, Ellingson MK, Wong P, Israelow B, Lucas C, Klein J et al (2020) Sex differences in immune responses that underlie COVID-19 disease outcomes. Nature 588:315–320. https://doi.org/10.1038/s41586-020-2700-3)

22. Prodan A, Brand HS, Ligtenberg AJ, Imangaliyev S, Tsivtsivadze E, van der Weijden F et al (2015) Interindividual variation, correlations, and sex-related differences in the salivary biochemistry of young healthy adults. Eur J Oral Sci 123:149–157. https://doi.org/10.1111/eos.12182)

23. McLaughlin HP, Worrell MC, Malone S, Dawson P, Maricque B, Halpin JL et al (2022) Acceptance of Saliva-Based Specimen Collection for SARS-CoV-2 Testing Among K-12 Students, Teachers, and Staff. Public Health Rep 137:557–563. https://doi.org/10.1177/00333549221074395)
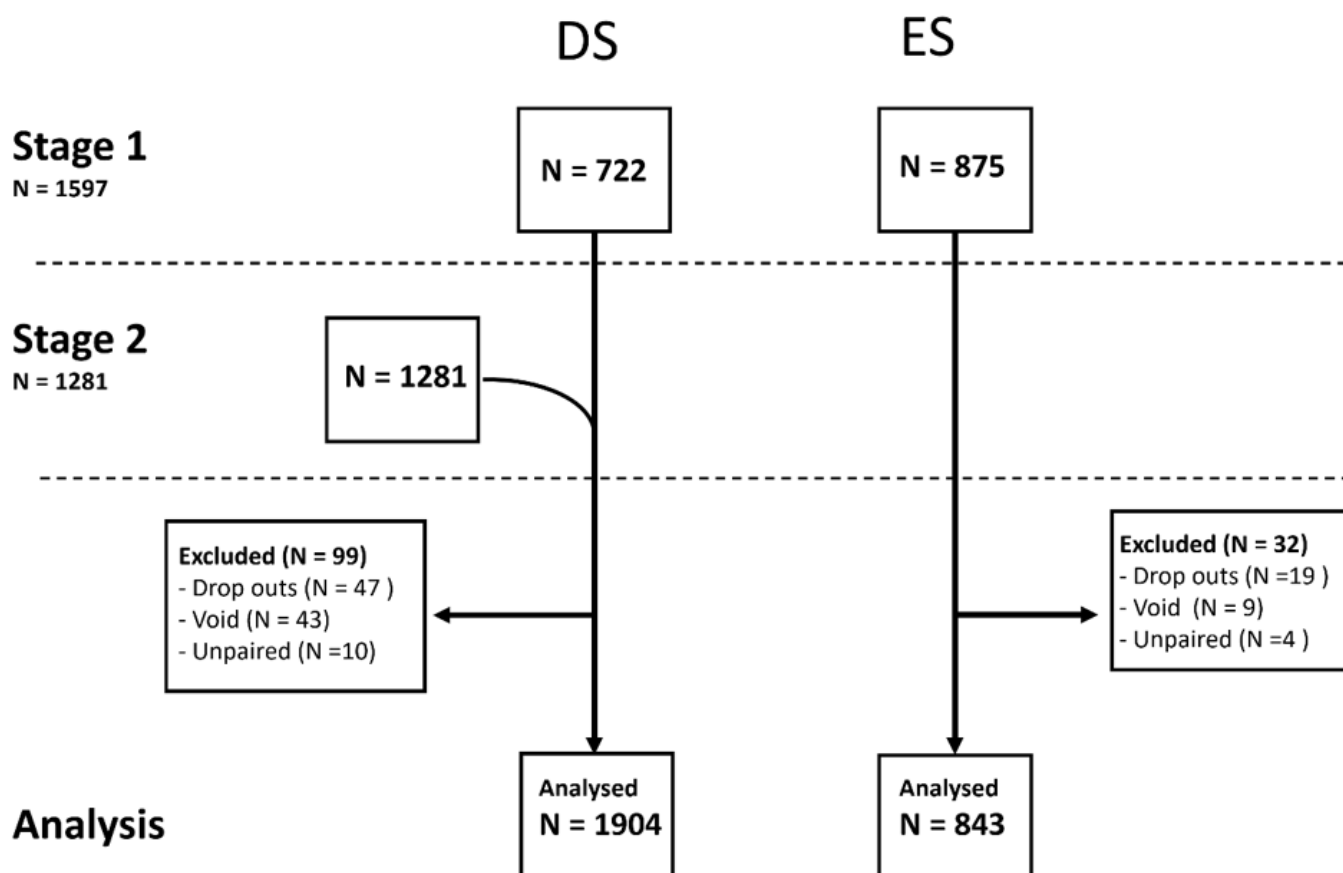
Figures

Figure 1

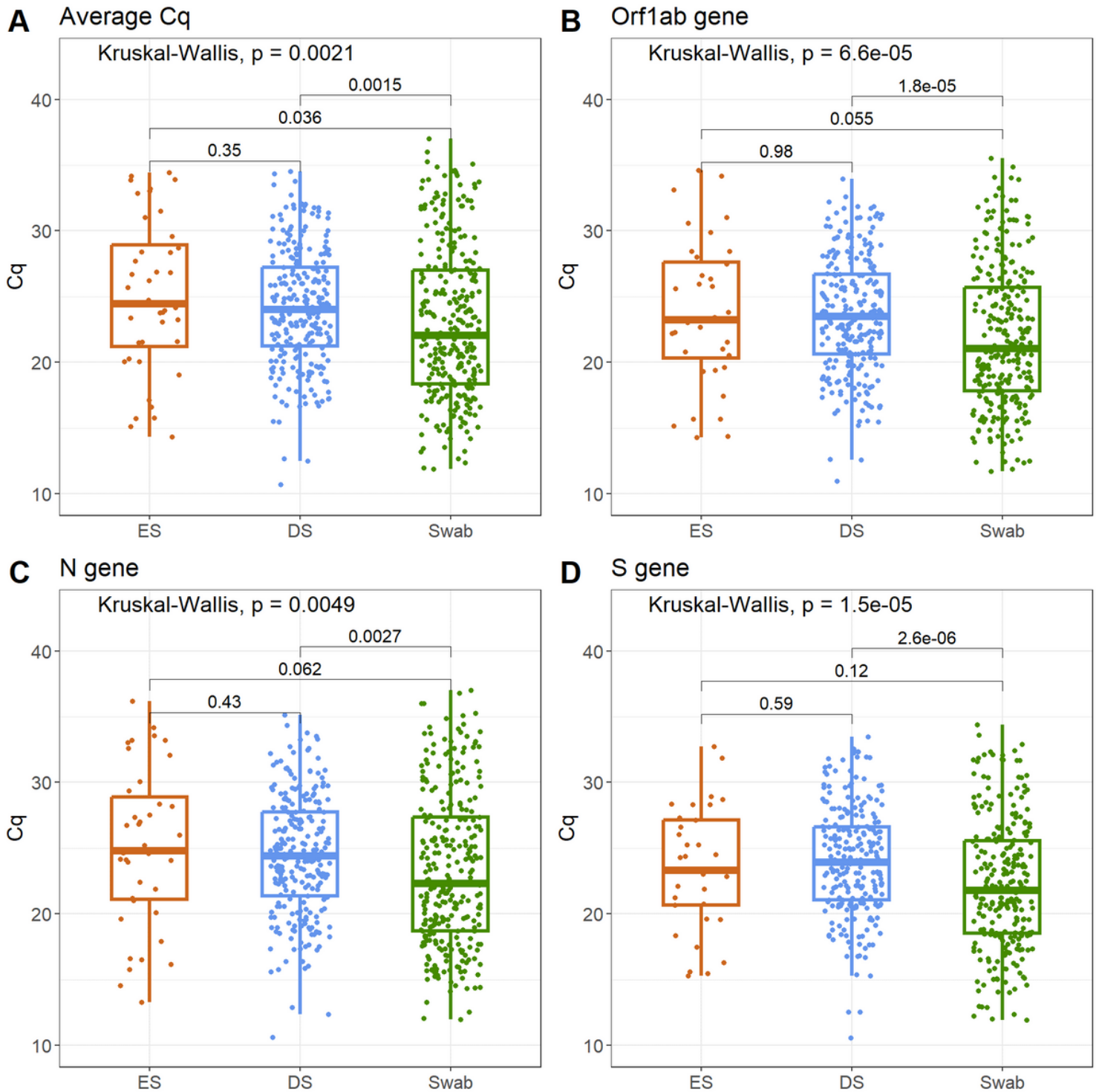CONSORT flow diagram. For DS excluded samples, 1 unpaired sample was also a drop out.

Figure 2

Distribution of Cq values of saliva and CTN swab samples. Display are box-and-whiskers plots for **A)** averaged Cq values, **B)** Cq values of Orf1ab gene, **C)** N gene and **D)** S gene. The lower and upper hinges correspond to the first and third quartiles. Significance (p-values) for pairwise difference between

sampling types are based on Wilcoxon test. Additionally, the result of the Kruskal-Wallis rank sum test comparing Cq values for the three sample types are shown.
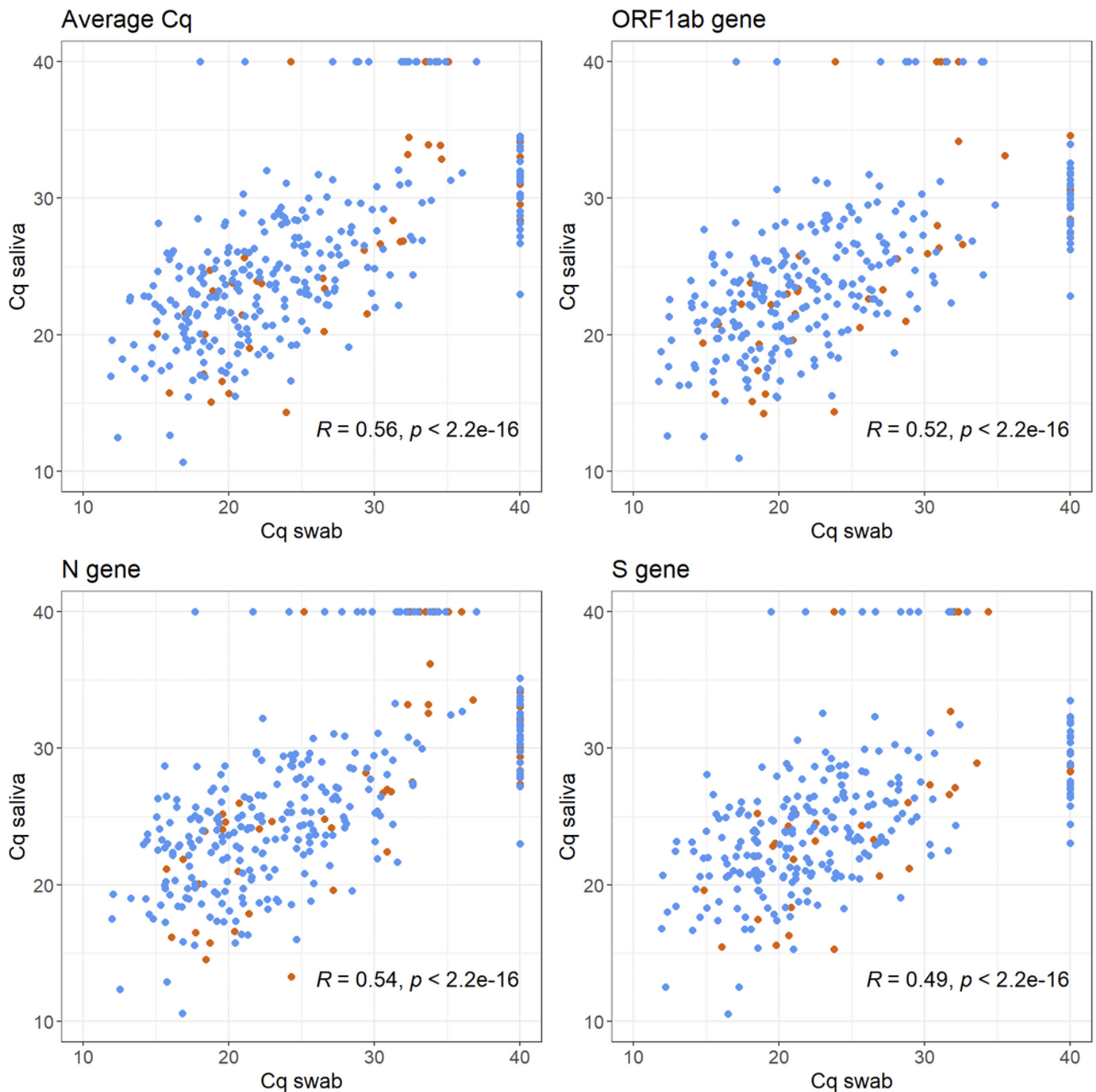

Figure 3

Correlation of Cq values between saliva and CTN swab samples. Scatterplots are displayed for (A) average Cq values, (B) Cq values for Orf1ab gene, (C) N gene and (D) S gene. For visualisation purpose, Cq values were set to 40 in cases where SARS-COVID2 was undetectable by qPCR in the sample. Thus,

dots on the Cq = 40 horizontal or vertical lines represent samples which were positive either for the saliva (vertical line) or swab (horizontal line) only. These values were not included for calculation of the Spearman correlation coefficient R and its significance.
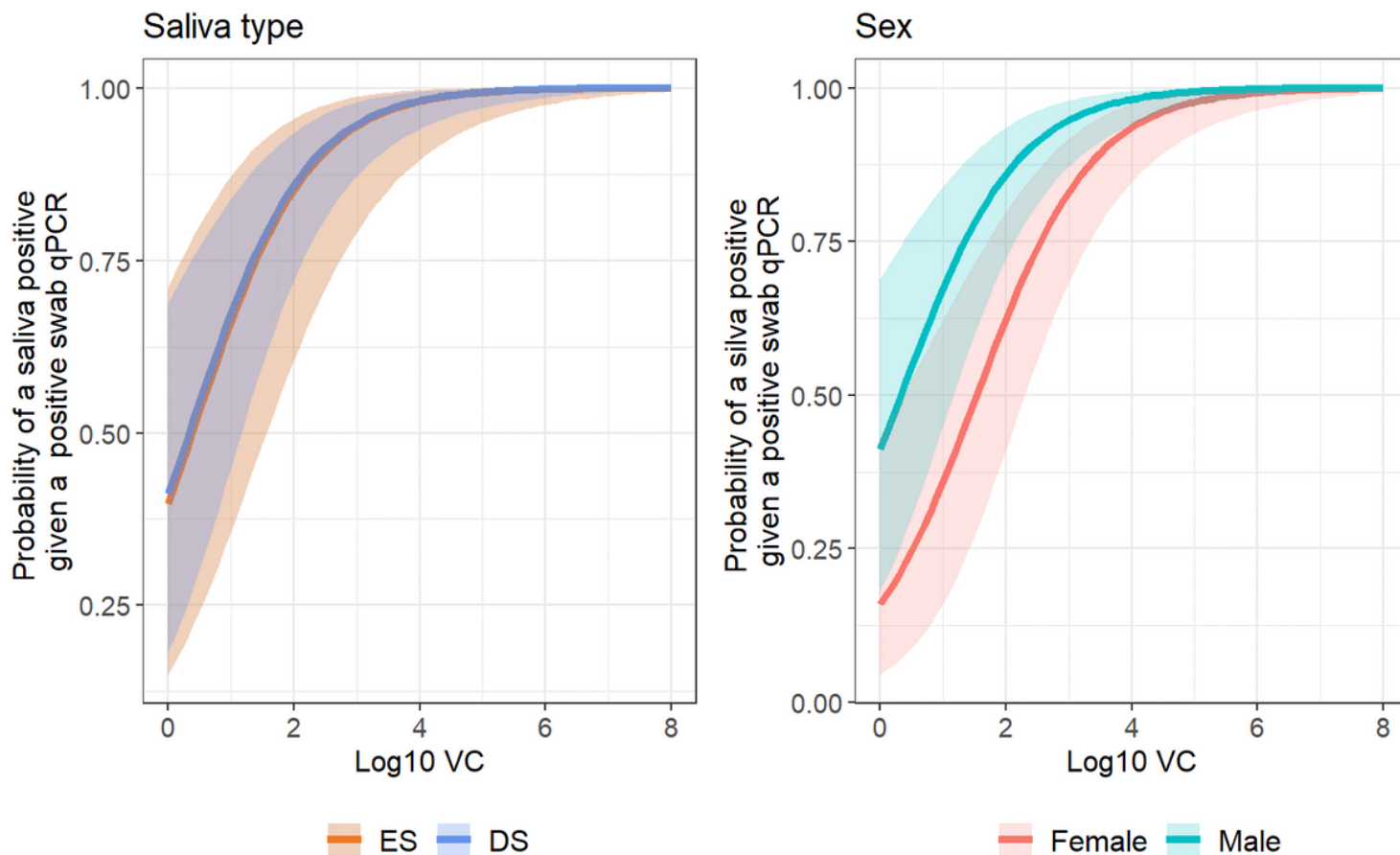

Figure 4

Predicted probabilities of TP by saliva specimen type (**A**) and by sex (**B**) based on multivariable logistic regression for a 40-year-old non-smoker. Swab $\log_{10}$ viral concentration (Log10 VC) was varied from 0 to 8 (i.e. viral concentration ranges from 1 to 10M dCopies/mL). For (**A**), the sex was set to "Male". For (**B**), the saliva sample type was set to "DS".

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SalivavsSwabSupplementaryMaterials.docx