AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Clinical Microbiology®

VIROLOGY

Evaluation of Specimen Types and Saliva Stabilization Solutions for SARS-CoV-2 Testing

Sara B. Griesemer,[a] Greta Van Slyke,[a] Dylan Ehrbar,[a] Klemen Strle,[a] Tugba Yildirim,[a] Dominick A. Centurioni,[a] Anne C. Walsh,[b] Andrew K. Chang,[c] Michael J. Waxman,[c] Kirsten St. George[a,d]

[a]Division of Infectious Diseases, Wadsworth Center, New York State Department of Health, Albany, New York, USA
[b]Directors Office, Wadsworth Center, New York State Department of Health, Albany, New York, USA
[c]Department of Emergency Medicine, Albany Medical Center, Albany, New York, USA
[d]Department of Biomedical Sciences, School of Public Health, University at Albany, SUNY, Albany, New York, USA

Sara B. Griesemer and Greta Van Slyke shared equally in the work for the manuscript. Author order was determined based on laboratory membership.

ABSTRACT Identifying SARS-CoV-2 infections through aggressive diagnostic testing remains critical to tracking and curbing the spread of the COVID-19 pandemic. Collection of nasopharyngeal swabs (NPS), the preferred sample type for SARS-CoV-2 detection, has become difficult due to the dramatic increase in testing and consequent supply strain. Therefore, alternative specimen types have been investigated that provide similar detection sensitivity with reduced health care exposure and the potential for self-collection. In this study, the detection sensitivity of SARS-CoV-2 in nasal swabs (NS) and saliva was compared to that of NPS using matched specimens from two outpatient cohorts in New York State (total $n = 463$). The first cohort showed only a 5.4% positivity, but the second cohort ($n = 227$) had a positivity rate of 41%, with sensitivity in NPS, NS, and saliva of 97.9%, 87.1%, and 87.1%, respectively. Whether the reduced sensitivity of NS or saliva is acceptable must be assessed in the settings where they are used. However, we sought to improve on it by validating a method to mix the two sample types, as the combination of nasal swab and saliva resulted in 94.6% SARS-CoV-2 detection sensitivity. Spiking experiments showed that combining them did not adversely affect the detection sensitivity in either. Virus stability in saliva was also investigated, with and without the addition of commercially available stabilizing solutions. The virus was stable in saliva at both 4°C and room temperature for up to 7 days. The addition of stabilizing solutions did not enhance stability and, in some situations, reduced detectable virus levels.

KEYWORDS COVID-19, nasal swab, nasopharyngeal swab, saliva, sample type

Identifying acute COVID-19 infection through diagnostic testing and screening remains critical in our efforts to care for affected patients and curb the pandemic. Accordingly, laboratory services have worked to expand testing capacity, improve turnaround times, track the course of the outbreak, and provide data for patient management and the implementation of appropriate mitigation efforts and health care services (1).

The original standard for diagnosing COVID-19 was molecular testing for SARS-CoV-2 RNA on a nasopharyngeal swab (NPS) collected in virus transport medium (VTM). Unfortunately, performing NPS necessitates the use of a specifically designed swab and requires health care workers to wear extensive personal protective equipment (PPE). Within weeks of the pandemic onset in the United States, multiple aspects of the specimen collection and testing process became extremely difficult. The desired testing capacity quickly exceeded anything attempted in the country before, and vendors

Citation Griesemer SB, Van Slyke G, Ehrbar D, Strle K, Yildirim T, Centurioni DA, Walsh AC, Chang AK, Waxman MJ, St. George K. 2021. Evaluation of specimen types and saliva stabilization solutions for SARS-CoV-2 testing. J Clin Microbiol 59:e01418-20. https://doi.org/10.1128/JCM.01418-20.

Editor Alexander J. McAdam, Boston Children's Hospital

Address correspondence to Kirsten St. George, Kirsten.St.George@health.ny.gov.

Received 5 June 2020
Returned for modification 26 June 2020
Accepted 24 February 2021
Accepted manuscript posted online 5 March 2021
Published 20 April 2021

were unable to meet the supply demands for all required sampling components, including the specific swabs used for NPS specimen collection, VTM, PPE, reagents, and instruments needed for testing the samples at the laboratories (1–6). In response, a number of changes were implemented: guidelines were rapidly changed to require clinicians to limit the use of (or to reuse) PPE (4, 7–9), alternatives to VTM were tested and implemented (10–12), and numerous additional commercial diagnostic tests and instruments were authorized and distributed by the U.S. Food and Drug Administration (https://www.fda.gov/emergency-preparedness-and-response/mcm-legal-regulatory-and-policy-framework/emergency-use-authorization#covid19euas). Additionally, other sample types were assessed as alternatives to NPS (3, 13–19).

Testing of oropharyngeal swabs (OPS), nasal swabs (NS), midturbinate swabs (MTS), and saliva provided the opportunity to be less restrictive about the type of swab used, and the last three options also offer the additional advantage of self-collection. This has been demonstrated to be successful in previous studies on influenza detection, where self-collected specimens were found to have detection rates similar to those collected by health care personnel (20, 21). Major advantages of self-collection include limiting the exposure of health care workers to potentially infected patients, subsequently reducing the use of PPE, and limiting the exposure of patients to hospitals, where infection is likely more prevalent. Self-collection also allows for individuals to conveniently test themselves at home, which could lead to increased testing rates with a limited burden on health care systems.

Therefore, numerous recent studies have focused on the comparative sensitivity of upper respiratory swabs other than NPS, which has long been considered the preferred option for the optimal detection of most respiratory viruses (3, 15, 19, 22). Experience with saliva as a diagnostic sample type for respiratory viruses has been limited, although in recent months, several studies have suggested saliva is of acceptable or equal sensitivity to NPS or other specimen types for the diagnosis of SARS-CoV-2 (14, 17, 18, 23–25). Some studies have also proposed the use of saliva stabilizing solutions, and it has been unclear whether these are necessary or advantageous.

In this study, nasal swabs and saliva were compared to NPS for the detection of SARS-CoV-2 in outpatients from two distinct New York populations, one of which had high prevalence while the other had low. The stability of SARS-CoV-2 in saliva and the utility of three stabilizing solutions were investigated. Finally, methods for combining swabs and saliva in the laboratory were assessed to maximize detection of the virus without running multiple tests when more than one specimen type is collected.

MATERIALS AND METHODS

Specimen collection sites and patients. This study was granted nonresearch determination under the emergency response criterion by the New York State Department of Health Institutional Review Board.

Two sites were accessed for specimens collected between 20 and 26 March 2020. The first testing site was a large tent erected on the Albany Medical Center (AMC) Emergency Department parking lot in Albany, NY. Open from 7 a.m. to 11 p.m. for 2.5 days, this walk-in site enrolled 236 subjects. Patients without an appointment could be tested for COVID-19 and were not required to be symptomatic. No patients were excluded. The second testing site was a drive-through Regional Operations Center (ROC) in New Rochelle (NR), NY, where an identified superspreader event had taken place early in the pandemic and the COVID-19 positivity rate was known to be high. Patients were required to have an appointment and qualifying symptoms or exposures; 227 subjects were enrolled from this site.

Specimen types, collection, and transport. Three specimens were collected from each subject: nasopharyngeal swab (NPS), nasal swab (NS), and saliva. Collection of NS was by bilateral swabbing on flocked swabs with insertion to approximately one inch and gentle rotation for several seconds. For NPS swabs, appropriate flocked NPS swabs were inserted in the NP cavity and gently rotated for 10 to 15 s. Both NS and NPS were placed in separate tubes containing 1 ml molecular transport medium (MTM; Longhorn Vaccines and Diagnostics, LLC, San Antonio, TX). Saliva samples were collected in sterile 50-ml conical tubes. Patients were instructed to refrain from eating, drinking, chewing gum or tobacco, or smoking for 30 min prior to collection and to work up saliva in their mouth, not cough, and to drool or gently spit the saliva into the container. Specimens from AMC were brought to the Wadsworth Center Laboratory of Viral Diseases (LVD) approximately every 2 to 4 h for accessioning and testing. Those from NR were placed in coolers and transported to the LVD within 24 h for accessioning and testing.

TABLE 1 Experimental parameters for stability studies

Matrix	Virus concn spiked (gene copies, gc/ml)	Temp	Time point (day)
Saliva only	High, 1×10^6; low, 2×10^5	4°C, room temp, −80°C	0, 1, 3, 5, 7
VTM only	High, 1×10^6; low, 2×10^5	4°C, room temp, −80°C	0, 1, 3, 5, 7
Nasal Swab only	High, 1×10^6; low, 2×10^5	4°C, room temp, −80°C	0, 1, 3
Saliva + Nasal Swab only	High, 1×10^6; low, 2×10^5	4°C, room temp, −80°C	0
Saliva + VTM	High, 1×10^6; low, 2×10^5	4°C, room temp, −80°C	0, 1, 3, 5, 7
Saliva + 23andMe stabilization buffer	High, 1×10^6; low, 2×10^5	4°C, room temp, −80°C	0, 1, 3, 5, 7
Saliva + AncestryDNA stabilization buffer	High, 1×10^6; low, 2×10^5	4°C, room temp, −80°C	0, 1, 3, 5, 7
Saliva + AncestryDNA stabilization buffer, combined with nasal swab	High, 1×10^6; low, 2×10^5	4°C, room temp, −80°C	0, 1, 3

Specimens were held at 4°C from the time of collection to the time of processing into lysis buffer for molecular testing. All testing was performed within 24 to 72 h of the time of specimen collection.

SARS-CoV-2 testing. NPS and NS specimens collected in MTM did not require further processing before extraction. Saliva volumes varied from 1 to 10 ml. Many samples from the NR cohort were excessively mucoid and could not be aspirated through a 1,000-μl pipette tip for transfer into lysis buffer. To streamline the process, all samples from the NR cohort were digested prior to extraction with 0.5 to 1 ml Snap n' Digest (Scientific Device Laboratory, Des Plaines, IL), vortexed, and incubated at room temperature for 10 min. Samples from AMC with mucus or excessive viscosity were selectively digested as described above.

The CDC 2019 nCoV real-time reverse transcription-PCR (RT-PCR) diagnostic panel (26–28) was used to test all samples. Total nucleic acid extraction was conducted on the bioMérieux easyMAG or EMAG (bioMérieux Inc, Durham, NC) or MagNA Pure 96 system (Roche Diagnostics, Indianapolis, IN). Briefly, for all specimens tested using the easyMAG or EMAG, 110 μl of sample was added to 2 ml NucliSENS lysis buffer (bioMérieux) and extracted into 110 μl of eluate. For specimens extracted using the MagNA Pure 96, 100 μl specimen was added to 350 μl lysis buffer and eluted into 100 μl.

Testing was performed on ABI 7500FAST Dx instruments as described in the instructions for use (IFU) on the FDA website. All cycle threshold (C_T) values for viral RNA detection presented here are for the N1 gene.

Virus and viral RNA quantification. A SARS-CoV-2 isolate, USA-WA1/2020, was obtained from BEI Resources (NIH/ATCC, Manassas, VA) and amplified in Vero E6 cells, in accordance with biosafety level 3 (BSL-3) procedures. Viral isolate RNA was extracted and quantified by standard curve using RNA transcript synthesized to contain all SARS-CoV-2 N-gene targets in the CDC 2019-nCoV real-time RT-PCR diagnostic panel (Bio-Synthesis, Inc., Lewisville, TX).

Validation of specimen types for viral detection. Validation of all specimen types with the CDC 2019-nCoV real-time RT-PCR diagnostic panel included a limit of detection determination, followed by a clinical evaluation panel. To perform these experiments in a BSL-2 environment, quantified SARS-CoV-2 viral RNA, rather than infectious virus, was spiked into lysed negative specimens prior to extraction. Saliva was validated for use on the easyMAG and EMAG instruments as well as on the MagNA Pure 96 system (Roche Diagnostics, Indianapolis, IN). Combined saliva and nasal swab specimens were validated on the easyMAG and EMAG instruments only.

To investigate the possibility of maximizing virus detection without doubling the number of tests performed, mixed saliva and nasal swabs were validated as a sample type in the following manner: 110 μl of each specimen was added to the same 2-ml lysis buffer tube, eluted into 110 μl, and tested as one sample. To mimic situations in which saliva is positive and nasal swabs are negative or vice versa, and to confirm that there is no adverse impact on detection, limits of detection (LOD) were determined for spiked combined specimens.

LOD were assessed by first testing a range-finding dilution series, in triplicate, to estimate the LOD and then confirmed by spiking that concentration of SARS-CoV-2 viral RNA into 20 individual clinical specimens. The LOD was determined as that which gave at least 95% detection of the sample replicates.

To evaluate the detection of SARS-CoV-2 in clinical saliva samples, individual negative saliva specimens were spiked with quantified viral RNA at the following range of concentrations in triplicate: 2×, 4×, 6×, 8×, 10×, and 100× the established LOD. These evaluations were also performed with combined saliva and nasal swab specimens. Additionally, 19 previously positive nasal swabs in viral transport medium (VTM), with original C_T values ranging from 20 to 39, and three positive saliva specimens, with C_T values of 21 to 33, were individually reanalyzed and tested in combination with the other negative samples.

Saliva stabilizing experiments. To evaluate the stability of SARS-CoV-2 in saliva and to determine the necessity of stabilizing solutions for virus preservation in saliva, infectious SARS-CoV-2 virus was spiked into pooled negative saliva and sampled at multiple time points, with storage at multiple temperatures, with and without the addition of various stabilizing buffers. Individual saliva specimens were tested for SARS-CoV-2 by real-time RT-PCR before using in stability studies to verify their negative status. Table 1 summarizes the multiple parameters assessed in the stabilizing experiments.

SARS-CoV-2 isolate USA-WA1/2020, cultured and harvested as described above, was used in all stabilizing experiments. Stock viral RNA concentration was calculated at 2.2×10^7 RNA gene copies/μl as described above. Stock virus was diluted in Eagles minimal essential medium (EMEM) plus 2% fetal

bovine serum to working concentrations, and samples in each experiment were spiked with either high (1×10^6 RNA copies/ml specimen or 5,000 RNA copies/PCR) or low (2×10^5 RNA copies/ml specimen or 200 RNA copies/PCR) concentrations in duplicate.

Sampling of all specimens for the stability studies included lysis of 110 μl specimen into 2 ml NucliSENS lysis buffer, extraction on EMAG or easyMAG instruments into 110 μl of elution buffer, and analysis by the CDC real-time RT-PCR diagnostic panel. C_T values were averaged and plotted using GraphPad Prism software, version 8.3.0.

23andMe stabilization buffer. Saliva stabilization kits were kindly donated by 23andMe (Sunnyvale, CA). Stabilizing buffer was removed from the kits and pooled. A total of 32 tubes (15 ml polypropylene), each containing 1 ml saliva, were spiked with SARS-CoV-2 (16 high and 16 low virus concentrations), followed by the immediate addition of 1 ml stabilization buffer, and then mixed thoroughly. An additional 32 tubes were prepared containing 0.5 ml saliva spiked with high and low virus concentrations, followed by the immediate addition of 0.5 ml VTM, and then mixed thoroughly. All samples were then sampled and analyzed by real-time RT-PCR (day 0) by standard procedures, starting with a 110-μl transfer to lysis buffer tubes as described above. Duplicate tubes at each concentration were then stored at room temperature, 4°C, and −80°C, and samples were tested on each of days 1, 3, 5, and 7, except for those stored at −80°C, which were only analyzed on day seven.

AncestryDNA stabilization buffer. AncestryDNA saliva collection kits were kindly donated by the company (Lehi, UT) and used for additional saliva stability studies. In these experiments, the stability of SARS-CoV-2 was tested in saliva, VTM, saliva with VTM, and saliva with AncestryDNA (AcD) stabilization buffer. Briefly, tubes containing SARS-CoV-2-negative saliva or VTM were spiked with high and low virus concentrations (six replicates each), followed by the addition of stabilization buffer, mixing, and sampling as described above for the 23andMe experiment for day 0 analysis. The following volumes were used: saliva alone (1.5 ml), VTM alone (1.5 ml), saliva with VTM (0.5 ml saliva plus 1 ml VTM), and saliva with AcD (1 ml saliva plus 0.5 ml AcD buffer, per the manufacturer's instructions). Duplicate tubes of each concentration were prepared for each storage temperature and stored at room temperature, 4°C, and −80°C for up to 7 days. Stored, spiked samples were tested at appointed days when a 110-μl aliquot from each tube was lysed, extracted, and tested by real-time PCR at days 1, 3, and 7, except for those stored at −80°C, which were only sampled at day seven.

Stabilization with saliva and nasal swab combination. AncestryDNA saliva collection kits were used for stability studies in an experiment to investigate the stability of SARS-CoV-2 in saliva, nasal swab in VTM, and saliva with AcD stabilization buffer. In addition, to observe the effects of combining specimens on SARS-CoV-2 detection sensitivity, saliva alone and in combination with nasal swabs in VTM were lysed and analyzed together. Pooled negative saliva and pooled nasal swabs in VTM, all negative for SARS-CoV-2, were used in this experiment. A total of 18 tubes were prepared, 12 containing 2 ml saliva (alone) and 6 containing 2 ml NS in VTM. Four saliva tubes were spiked with high virus and four with low virus concentrations, and four remained unspiked. Two of each of these four (2 high, 2 low, and 2 no virus) were then mixed with 1 ml AcD stabilization buffer. Six NS tubes were also spiked with high, low, or no virus. Tubes were incubated at room temperature, and sampling was performed at three time points: days 0, 1.5, and 3. At each time point, 110 μl from each of the tubes was removed, lysed, extracted, and analyzed. In addition, 110 μl of each prepared saliva sample, with and without stabilization buffer, was lysed together with 110 μl of each prepared nasal swab sample, so that high, low, and negative saliva samples would be mixed with high, low, and negative nasal swab samples. Combined samples were then extracted as one specimen and analyzed.

Statistical analyses. For all statistical calculations, each C_T value used is the mean of the N1 and N2 C_T values for that positive specimen. In cases where a specimen type was not collected for an individual, that patient was then excluded from calculations of that specimen type's sensitivity. For the purpose of calculating sensitivities, true positives were determined as any C_T value of <45 for any specimen type. Sensitivities and associated confidence intervals (CI) were calculated using the R package pROC (29). Confidence intervals of sensitivities were calculated using 2,000 stratified bootstrap replicates. Two-sample Wilcoxon rank sum tests were used to compare mean ranks of C_T values between males and females for each specimen type. Wilcoxon rank sum tests were performed, and the results plotted, using the R package ggpubr (30). Conditional density plots were created using the R package vcd (31) with default settings. R, version 3.6.3 (32), was used for all preceding analyses.

RESULTS

Patient demographics and regional positivity rates. Specimens were collected from a total of 463 individuals at two collection sites in New York State during the early weeks of the SARS-CoV-2 pandemic. Samples were initially collected from 236 individuals at AMC. However, following initial testing of NPS, only 12 of 236 patients tested positive for SARS-CoV-2 RNA, a positivity rate too low (5%) to enable statistically meaningful comparisons (Fig. 1A). Specimen collection was immediately transferred to a location with known high positivity. During March 2020, New Rochelle (NR), NY, was one of the Northeast's earliest pandemic "hot spots," with a significantly higher positivity rate than the rest of the state. Specimens were collected from 227 patients in NR with demographics as shown in Table 2. Of the 227 patients tested at that site, 93 (41%) were positive for SARS-CoV-2, with 74 patients positive in all three sample types,

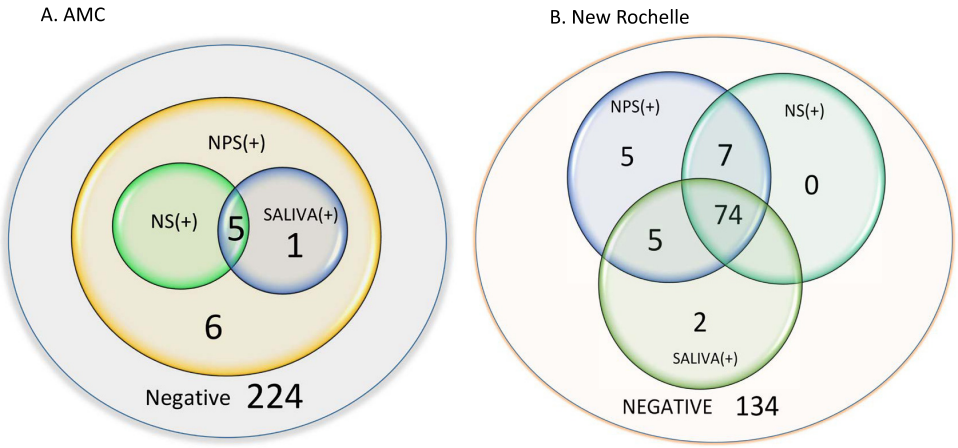

FIG 1 Comparison of positivity in different specimen types from the two collection sites. (A) Albany Medical Center ($n = 236$). (B) New Rochelle ($n = 227$).

7 positive in both NS and NPS, 5 positive in both saliva and NPS, and 5 positive in NPS alone. Interestingly, 2 patients were positive in saliva alone but none in NS alone, while NPS detected 91 of 93 positive patients (Fig. 1B). The distribution of viral load across age groups in the NR cohort, as indicated by C_T value, is displayed in Fig. 2 and provides a visual indication of differences seen across sample types and age groups.

Comparison of specimen sensitivities for SARS-CoV-2 detection. We sought to determine whether NS or saliva specimens could be used as a surrogate sample type in place of NPS for SARS-CoV-2 detection when using the CDC 2019 nCoV real-time RT-PCR diagnostic panel. Comparisons of grouped C_T values using a Kruskal-Wallis test with Dunn's multiple-comparison *post hoc* test revealed statistically significant differences between the three sample types: the mean C_T values for both NPS- and NS-positive samples were significantly lower ($P < 0.0001$) than that of positive saliva samples, and the mean C_T values for NPS and NS were not significantly different from one another (Fig. 3). When sensitivities and corresponding 95% confidence intervals for all specimen types were calculated, NPS and the combination of NS and saliva provided the highest sensitivities (97.8% and 94.6%, respectively) with overlapping 95% confidence intervals (Table 3). Both NS and saliva had lower, and identical, sensitivities and 95% confidence intervals (87.1% and 79.57 to 93.55 for both) than NPS or the combination of NS and saliva. The confidence intervals for NS and saliva overlapped that for the NS and saliva combination specimen type but not NPS specimen type (Table 3).

Validation of saliva and nasal swabs for clinical testing. Saliva and NS as individual specimen types and the combination of NS plus saliva were all validated for testing on the CDC 2019 nCoV real-time RT-PCR diagnostic panel after extraction on the EMAG and easyMAG. The LOD for saliva, NS, and NS plus saliva were all found to be 25 RNA

TABLE 2 Patient demographics

Parameter	Albany Medical Center cohort		New Rochelle cohort	
	All patients ($n = 236$)	Positive patients ($n = 12$)	All patients ($n = 227$)	Positive patients ($n = 93$)
Male gender, %	47	25	60	60
Age, yr				
0–5	4	0	0	0
6–18	13	0	2	1
19–45	133	6	149	63
46–65	68	6	59	21
>65	16	0	6	4
Unknown	2	0	11	4
Age range	3–105	27–61	14–77	14–71

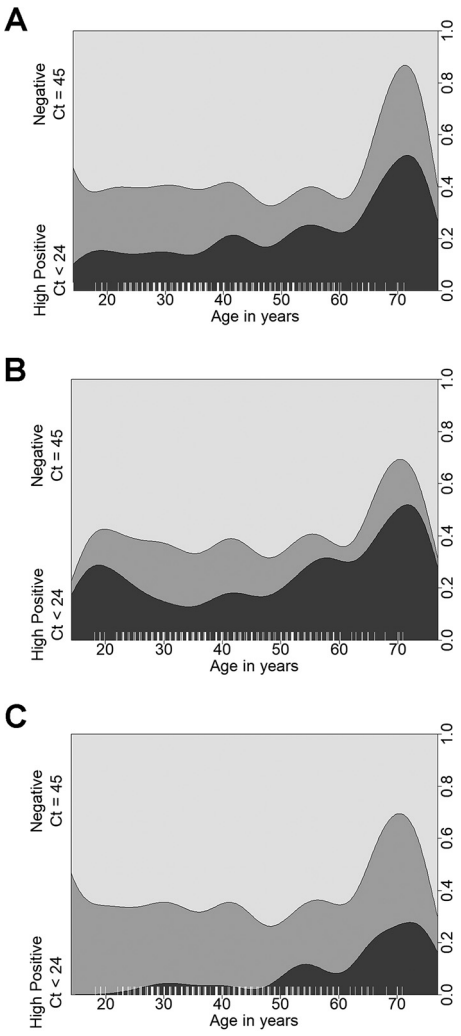

FIG 2 Distribution of negative, low-positive, and high-positive patients by age. Conditional density plots display the distribution of negative, low-, and high-positive samples by age. High-positive samples (darkest gray) showed C_T values of $<$24, low positives (medium gray) had $24 \leq C_T < 45$, and negatives (light gray) had C_T values of 45. Results are shown for NPS (A), NS (B), or saliva (C).

copies/PCR or 5 RNA copies/μl using both extraction platforms. Additionally, clinical evaluation studies of all specimen types with amended negative samples showed 100% accuracy (data not shown).

A total of 20 individual negative saliva specimens were separately lysed and spiked with viral RNA in an amount that was calculated to result in a concentration of 25 RNA copies/PCR. Next, 20 individual negative nasal swabs, in VTM, were added to the lysed saliva, and the combined sample was analyzed. For the reverse, 20 individual negative nasal swab specimens were separately lysed and spiked with 25 copies/PCR, and the same volume of 20 negative saliva samples was added to the lysed nasal swabs. These combined samples were then extracted and analyzed.

A clinical evaluation was also performed on samples of mixed NS and saliva specimens from positive patients. A total of 15 previously positive NS samples were retrieved and retested with and without the addition of saliva. All were positive when combined, and C_T values were not adversely affected (Fig. 4). Linear regression analysis indicated a high correlation between nasal swab only C_T values and nasal swab with saliva C_T values, even for samples with moderately weak values (C_T values between 30 and 35).

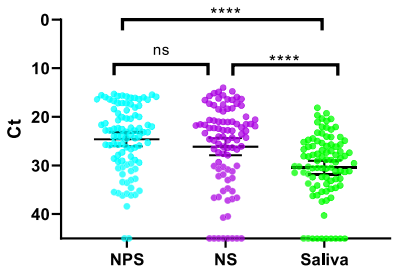

FIG 3 Comparison of SARS-CoV-2 detection by sample type and C_T value. Matched C_T values for NPS, NS, and saliva from all positive individuals collected at the ROC (n = 93) were compared using a Kruskal-Wallis test ($P < 0.0001$) with Dunn's multiple comparisons *post hoc* test, which were used to compare the mean ranks of C_T values between sample types. Bars represent means and 95% CI, and a C_T value of 45 was assigned to undetected specimens. Asterisks represent P values derived from Dunn's multiple-comparison test (ns, not significant; *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$; ****, $P \leq 0.0001$).

SARS-CoV-2 stability in saliva with stabilization solutions. A series of experiments were then performed in which virus-spiked saliva was supplemented with different stabilization buffers to examine the stability of SARS-CoV-2 in saliva and the utility of commercial stabilizing solutions. Samples were tested at high and low virus levels, held at different temperatures for various time periods (as described in Materials and Methods), and analyzed for changes in C_T values that would indicate a loss of detectable viral RNA.

An initial experiment compared the deterioration of virus in spiked saliva, stabilized by the addition of VTM, to that of spiked saliva with the stabilization buffer from 23andMe. Samples were stored at 4°C and room temperature and tested at days 0, 1, 3, 5, and 7. The SARS-CoV-2 virus in saliva was more stable in VTM than in 23andMe stabilizing solution, as indicated by C_T values over time at either temperature (Fig. 5). Greatest losses in detectable viral RNA were seen at day 7 in saliva with 23andMe stabilizing solution, particularly in samples that had been stored at room temperature, while those in VTM showed little to no change. A similar experiment comparing viral RNA detection in saliva supplemented with VTM or stabilization solution from Ancestry also showed no benefit from the addition of stabilization solution and again a significant loss of detectable virus at day 7 in samples that had stabilization solution added to them. Interestingly, in this experiment the virus in spiked saliva with no additives was just as stable as that in VTM, producing unchanged C_T values after 7 days of storage at either 4°C or room temperature (Fig. 6). The virus was also just as stable in NS at either 4°C or room temperature as in saliva over a 3-day test period, while again losing detectable virus in saliva if stabilization buffer was added (Fig. 7).

In a final series of experiments, we explored whether adding saliva and nasal swabs together and testing as a single specimen impacted the detection of the virus and whether stabilization solution being previously added to the saliva impacted detection when stored at room temperature for up to 3 days. Results are graphed in Fig. 8 and show that (i) the presence of a high virus level in saliva is not impacted by the addition of a high, low, or negatively spiked nasal swab, (ii) a low-level spiked saliva sample is not impacted by a low or negative spike nasal swab and the addition of a high-level

TABLE 3 Sensitivity of detection for NPS, NS, saliva, and NS plus saliva

Specimen type	Sensitivity (95% CI)
NPS	97.85 (94.62–100)
NS	87.10 (79.57–93.55)
Saliva	87.10 (79.57–93.55)
NS and saliva	94.62 (89.25–98.92)

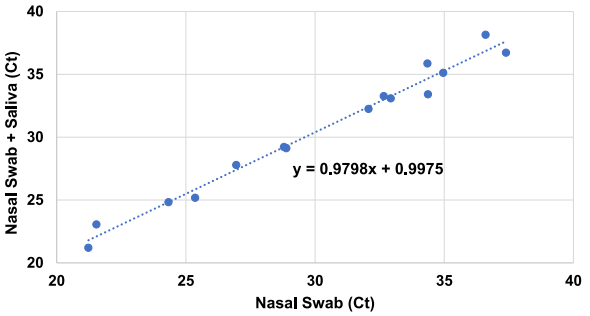

FIG 4 Addition of saliva does not interfere with C_T values of nasal swabs. Nasal swab specimens, previously positive for SARS-CoV-2, were retrieved from −80°C storage and retested by real-time RT-PCR. Negative saliva was extracted together with the same nasal swab specimens. Linear regression analysis was performed on both comparisons.

spiked nasal swab raises the virus level to high (detecting the high level of virus in the nasal swab), and (iii) the presence of saliva does not impact the detection of low- or high-level virus spiked into nasal swabs. All these findings are consistent across the 3 days that the sampling was performed. Interestingly, when the same sample combination experiments are performed using saliva supplemented with stabilization buffer, the presence of the solution does not affect the detection of the virus in nasal swabs; however, it does affect its detection in saliva (Fig. 8D, E, and F). With the exception of the detection of some high-level spiked virus in saliva at time point 0 (Fig. 8D), the addition of stabilization buffer effectively eliminated the virus detection in saliva at all other time points and concentrations (Fig. 8D and E), leaving detectable virus only in the NS.

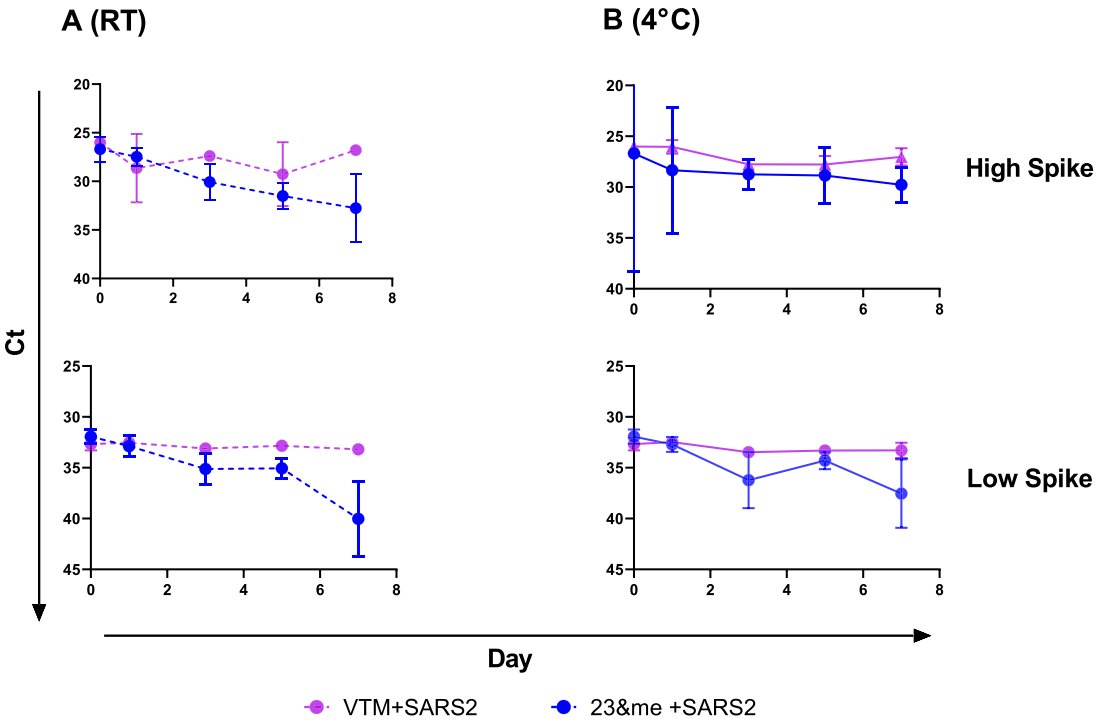

FIG 5 Virus stability in saliva samples mixed with 23andMe stabilization buffer or VTM. Mixtures of saliva plus 23andMe buffer (blue) or VTM (purple) and virus were held at room temperature (A, dashed lines) or 4°C (B, solid lines) and sampled at days 0, 1, 3, 5, and 7 in duplicate. Mean C_T values for N1 were plotted, and simple linear regression analysis was performed. The slope of all VTM curves did not deviate significantly from zero; the slope of the 23andMe low spike at 4°C did not deviate significantly from zero, and the slope of all other 23andMe plots significantly deviated from zero ($P < 0.02$).

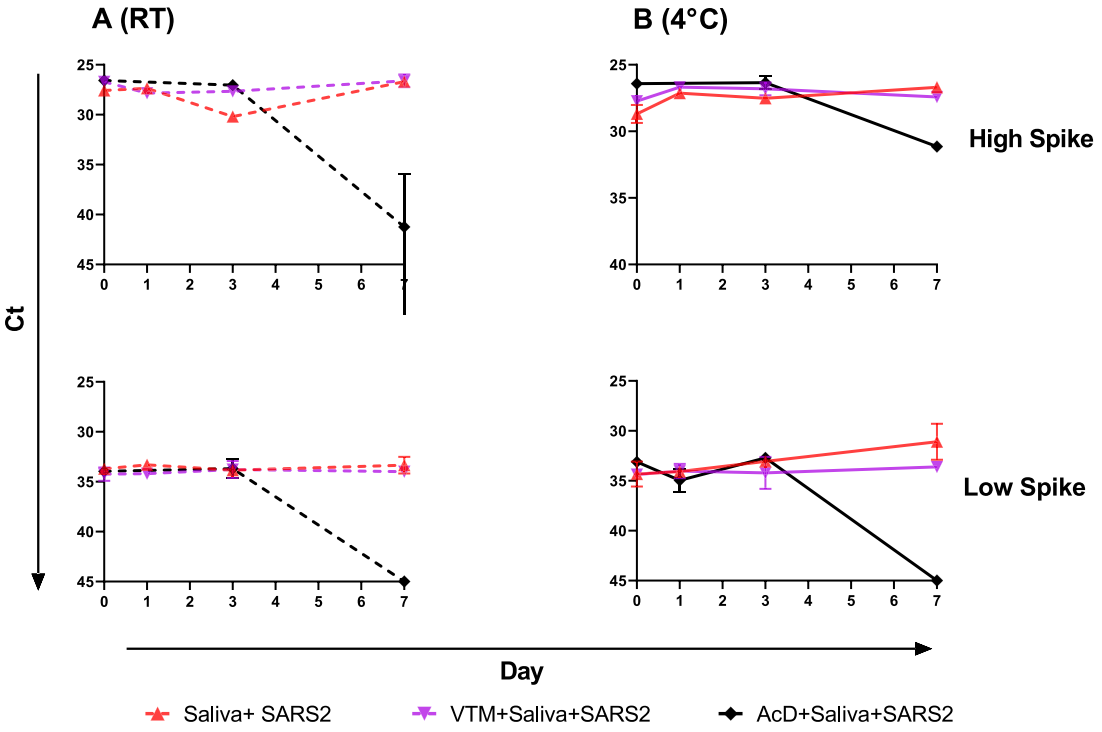

FIG 6 Virus stability in saliva alone, saliva with VTM, or saliva combined with AncestryDNA (AcD) stabilization buffer. Saliva alone (red), saliva plus VTM (purple), and saliva plus AcD buffer, all spiked with virus, were held at room temperature (A, dashed lines) or 4°C (B, solid lines) and sampled, in duplicate, at days 0, 1, 3, and 7. Mean C_T values were plotted, and simple linear regression analysis was performed. Slopes for neither saliva alone nor saliva plus VTM deviated from zero; all slopes for saliva plus AcD buffer significantly deviated from zero ($P < 0.02$).

DISCUSSION

As the COVID-19 pandemic continues, ongoing supply issues with collection devices and PPE at a national and international level, together with increasing demands for laboratory testing, has necessitated the exploration of alternative options for specimen types to maintain effective mitigation strategies. It will be difficult to find a consensus for the relative performance of different sample types across studies performed on different patient groups for a disease with such variable severity and presentations, across a wide range of age groups, using an array of collection devices and transport media, and with the additional variables of time postonset and time postsample collection.

COVID-19 has primarily affected adults and older persons. This is reflected in the populations of both cohorts tested in this study, as neither included large numbers of

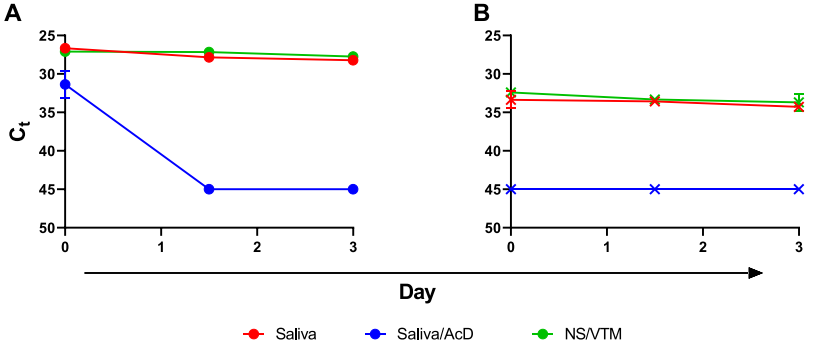

FIG 7 Virus stability over time. Viral RNA detection at both high (A) and low (B) spiked virus concentrations in saliva (red), saliva plus AcD buffer (blue), and NS in VTM (green). Samples were held at room temperature and collected, in duplicate, at 0, 36, and 72 h; mean C_T values are plotted.

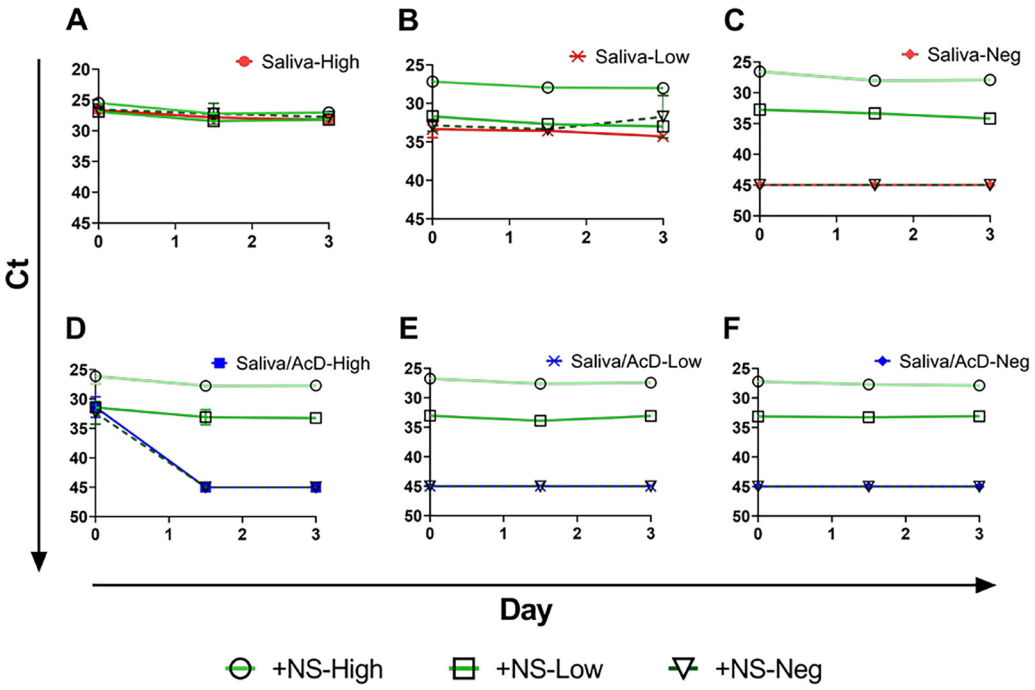

FIG 8 Effect of mixed saliva and NS specimens with high and low virus concentrations, with and without the addition of stabilization buffer to the saliva. (A, B, and C) High, low, and negative saliva mixed with high, low, and negative NS. (D, E, and F) High, low, and negative saliva with stabilization buffer mixed with high, low, and negative NS.

pediatric patient samples. The number of samples from the geriatric population (>65), however, was low in both cohorts and could be a result of the collections being performed at walk-through and drive-through clinics.

In the study reported here, the results show that in the outpatient group tested that was not severely ill, NPS specimens produced greatest sensitivity at approximately 98%, while NS or saliva alone provided sensitivities of approximately 87%. Whether the reduction in detection sensitivity that occurs with the use of these specimen types is acceptable has to be assessed in the settings in which they are being used, taking into account the more ready availability of NS swabs, the convenience of collection, the reduced discomfort of NS and saliva collection, and the potential for self-collection of these specimens and, thus, diminished risk of transmission to health care personnel. Notably, some other studies have not found sensitivity differences this great, while others have found them more extreme (13, 14, 16, 18, 19, 24, 25, 33). By combining the NS and saliva specimens in the laboratory, we could increase sensitivity to 95%, an additional 8% above NS or saliva alone; we believe this to be a substantial improvement and a beneficial option for sensitive diagnosis of SARS-CoV-2.

Of note, we did find that many of the submitted saliva specimens from NR were excessively mucoid, clearly containing material that had been coughed into the specimen and had to be digested, similar to the process used for sputum. Following this experience, we issued strict instructions not to cough or do anything that would result in excess mucus in the sample but to dribble or gently spit into the collection tube. In some other publications, the term "saliva" has been used loosely, and, on reading the specimen description carefully, has clearly included the deliberate collection of mucus. It is unclear from descriptions in most of the published papers where saliva has been studied, however, if they were truly saliva specimens or if they were a combination of saliva and sputum and what the relative sensitivity of the two specimen types is for detecting SARS-CoV-2. Saliva does present an attractive alternative to both NPS and NS specimens, because the only collection consumable needed is a tube. Notably, this sampling strategy has limitations: it may not be suitable for children, mentally

challenged patients, or severely ill patients. The collection of saliva is clearly not suitable for intubated patients, and the use of alternative sample types should be continued for those situations.

Saliva contains many antibodies and enzymes that could ultimately affect the stability of virus particles or viral RNA, which could impede detection. In our sample comparison from the NR cohort, significantly higher C_T values were observed in saliva samples than in NS or NPS. As a result, concerns were raised regarding viral stability in this sample type, and a series of experiments were undertaken to assess this and the possibility of improving stability with solutions from commercial saliva collection kits. Previous reports have shown viral RNA stability in saliva for 24 h at both room temperature and 4°C (25). In this study, however, we found surprising viral RNA stability in saliva alone for up to 7 days, even at room temperature, whereas the addition of stabilization solutions had adverse effects. It must be noted that these solutions are intended for stabilizing DNA, not RNA, and subsequent versions from commercial manufacturers specifically for RNA stabilization may have different results. However, at the time the work was done, neither company had equivalent saliva collection kits available with RNA stabilizing solutions, and it was their decision to have us proceed with the testing using the available stabilizers. It must also be noted that virus stability was only measured with regard to molecular detection. For investigators interested in studies on culturable infectious particles, the stability of virus infectivity may be different.

In symptomatic COVID-19 patients, SARS-CoV-2 viral RNA can be detected 1 to 2 days before symptom onset and several days postonset. Thereafter, viral loads decrease substantially; however, there are numerous reports of patients remaining RNA positive for prolonged periods of time and some individuals having intermittent bouts of repeated positivity. Whether these are reinfections, reactivations, or simply low levels of residual virus at the limits of detection of the assay is unclear. Moreover, we have much to learn about the behavior of viral loads in the large numbers of asymptomatic cases now being detected during mass screening of personnel and preprocedure patients. A few studies have recently become available on the relative sensitivities of various specimen types for SARS-CoV-2 detection in asymptomatic individuals and suggest that saliva can be considered an acceptable alternative sample type in these populations (34–37). The findings of our study, i.e., the combination of NS and saliva provides similar sensitivity to NPS while alleviating many of its limitations, should be carefully considered based on the setting and the study population being tested. Processes at laboratories must always be individually assessed for suitability at the site, with allowance for numerous aspects of pre- and postanalytical processes as well as the testing platforms and chemistries being utilized. It is quite feasible for this mixed sample type to be tested on a variety of platforms, including some of the sample-to-answer devices. However, as with all changes to processes, this would have to be assessed for suitability at the site and validated for the platform.

ACKNOWLEDGMENTS

We thank 23andMe and Ancestry for their generous provision of the saliva collection kits used in this work, participation in discussions regarding the experiments, and scientific openness regarding the findings. We also thank the staff of the Virology Laboratory at the Wadsworth Center for processing and testing the additional 1,500 specimens for the initial phase of this study at a time when testing services were already stretched to capacity. We thank Scott Wies, Peter Inglis, Akshay Vig, Kevin Flatley, and Druv Patel for assistance with specimen collection at the AMC site as well as the nurses and aides who staffed and collected specimens in the AMC tent. We also thank Alan Antenucci for coordinating retrieval of samples from the NR collection site and Bridget Anderson, Rebecca Goldberg, Dominick Chiumento, Richard Cook, and Ian Kerr, who coordinated the NR sample collection operation.

S.B.G., G.V.S., D.E., T.Y., D.A.C., A.C.W., A.K.C., and M.J.W. declare no conflict of interest. K.S. has served as a paid consultant for T2 Biosystems and Roche Diagnostics. K.S.G. receives research support from ThermoFisher for the evaluation of new assays for the diagnosis and characterization of viruses, including SARS-CoV-2. She also has a royalty-generating collaborative agreement with Zeptometrix.

This project was performed with internal funding from the NYSDOH and AMC.

REFERENCES

1. Patel R, Babady E, Theel ES, Storch GA, Pinsky BA, St George K, Smith TC, Bertuzzi S. 2020. Report from the American Society for Microbiology COVID-19 International Summit, 23 March 2020: value of diagnostic testing for SARS-CoV-2/COVID-19. mBio 11:e00722-20. https://doi.org/10.1128/mBio.00722-20.
2. Newton PN, Bond KC. 2020. COVID-19 and risks to the supply and quality of tests, drugs, and vaccines. Lancet Glob Health 8:e754–e755. https://doi.org/10.1016/S2214-109X(20)30136-4.
3. Pere H, Podglajen I, Wack M, Flamarion E, Mirault T, Goudot G, Hauw-Berlemont C, Le L, Caudron E, Carrabin S, Rodary J, Ribeyre T, Belec L, Veyer D. 2020. Nasal swab sampling for SARS-CoV-2: a convenient alternative in times of nasopharyngeal swab shortage. J Clin Microbiol 58:e00721-20. https://doi.org/10.1128/JCM.00721-20.
4. Rowan NJ, Laffey JG. 2020. Challenges and solutions for addressing critical shortage of supply chain for personal and protective equipment (PPE) arising from coronavirus disease (COVID19) pandemic–case study from the Republic of Ireland. Sci Total Environ 725:138532. https://doi.org/10.1016/j.scitotenv.2020.138532.
5. Vermeiren C, Marchand-Senecal X, Sheldrake E, Bulir D, Smieja M, Chong S, Forbes JD, Katz K. 2020. Comparison of Copan ESwab and FLOQSwab for COVID-19 diagnosis: working around a supply shortage. J Clin Microbiol 58:e00669-20. https://doi.org/10.1128/JCM.00669-20.
6. Wang X, Wu W, Song P, He J. 2020. An international comparison analysis of reserve and supply system for emergency medical supplies between China, the United States, Australia, and Canada. Biosci Trends 14:231–240. https://doi.org/10.5582/bst.2020.03093.
7. Boskoski I, Gallo C, Wallace MB, Costamagna G. 2020. COVID-19 pandemic and personal protective equipment shortage: protective efficacy comparing masks and scientific methods for respirator reuse. Gastrointest Endosc 92:519–523. https://doi.org/10.1016/j.gie.2020.04.048.
8. Perkins DJ, Villescas S, Wu TH, Muller T, Bradfute S, Hurwitz I, Cheng Q, Wilcox H, Weiss M, Bartlett C, Langsjoen J, Seidenberg P. 2020. COVID-19 global pandemic planning: decontamination and reuse processes for N95 respirators. Exp Biol Med 245:933–939. https://doi.org/10.1177/1535370220925768.
9. Rimmer A. 2020. Covid-19: experts question guidance to reuse PPE. BMJ 369:m1577. https://doi.org/10.1136/bmj.m1577.
10. Radbel J, Jagpal S, Roy J, Brooks A, Tischfield J, Sheldon M, Bixby C, Witt D, Gennaro ML, Horton DB, Barrett ES, Carson JL, Panettieri RA, Jr, Blaser MJ. 2020. Detection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is comparable in clinical samples preserved in saline or viral transport medium. J Mol Diagn 22:871–875. https://doi.org/10.1016/j.jmoldx.2020.04.209.
11. Rodino KG, Espy MJ, Buckwalter SP, Walchak RC, Germer JJ, Fernholz E, Boerger A, Schuetz AN, Yao JD, Binnicker MJ. 2020. Evaluation of saline, phosphate-buffered saline, and minimum essential medium as potential alternatives to viral transport media for SARS-CoV-2 testing. J Clin Microbiol 58:e00590-20. https://doi.org/10.1128/JCM.00590-20.
12. Rogers AA, Baumann RE, Borillo GA, Kagan RM, Batterman HJ, Galdzicka M, Marlowe EM. 2020. Evaluation of transport media and specimen transport conditions for the detection of SARS-CoV-2 using real time reverse transcription PCR. J Clin Microbiol 58:e00708-20. https://doi.org/10.1128/JCM.00708-20.
13. Alizargar J, Etemadi Sh M, Aghamohammadi M, Hatefi S. 2020. Saliva samples as an alternative for novel coronavirus (COVID-19) diagnosis. J Formos Med Assoc 119:1234–1235. https://doi.org/10.1016/j.jfma.2020.04.030.
14. Azzi L, Carcano G, Gianfagna F, Grossi P, Gasperina DD, Genoni A, Fasano M, Sessa F, Tettamanti L, Carinci F, Maurino V, Rossi A, Tagliabue A, Baj A. 2020. Saliva is a reliable tool to detect SARS-CoV-2. J Infect 81:e45–e50. https://doi.org/10.1016/j.jinf.2020.04.005.
15. LeBlanc JJ, Heinstein C, MacDonald J, Pettipas J, Hatchette TF, Patriquin G. 2020. A combined oropharyngeal/nares swab is a suitable alternative to nasopharyngeal swabs for the detection of SARS-CoV-2. J Clin Virol 128:104442. https://doi.org/10.1016/j.jcv.2020.104442.
16. Perchetti GA, Nalla AK, Huang ML, Zhu H, Wei Y, Stensland L, Loprieno MA, Jerome KR, Greninger AL. 2020. Validation of SARS-CoV-2 detection across multiple specimen types. J Clin Virol 128:104438. https://doi.org/10.1016/j.jcv.2020.104438.
17. Sullivan PS, Sailey C, Guest JL, Guarner J, Kelley C, Siegler AJ, Valentine-Graves M, Gravens L, Del Rio C, Sanchez TH. 2020. Detection of SARS-CoV-2 RNA and antibodies in diverse samples: protocol to validate the sufficiency of provider-observed, home-collected blood, saliva, and oropharyngeal samples. JMIR Public Health Surveill 6:e19054. https://doi.org/10.2196/19054.
18. Wang WK, Chen SY, Liu IJ, Chen YC, Chen HL, Yang CF, Chen PJ, Yeh SH, Kao CL, Huang LM, Hsueh PR, Wang JT, Sheng WH, Fang CT, Hung CC, Hsieh SM, Su CP, Chiang WC, Yang JY, Lin JH, Hsieh SC, Hu HP, Chiang YP, Wang JT, Yang PC, Chang SC, SARS Research Group of the National Taiwan University/National Taiwan University Hospital. 2004. Detection of SARS-associated coronavirus in throat wash and saliva in early diagnosis. Emerg Infect Dis 10:1213–1219. https://doi.org/10.3201/eid1007.031113.
19. Xie C, Jiang L, Huang G, Pu H, Gong B, Lin H, Ma S, Chen X, Long B, Si G, Yu H, Jiang L, Yang X, Shi Y, Yang Z. 2020. Comparison of different samples for 2019 novel coronavirus detection by nucleic acid amplification tests. Int J Infect Dis 93:264–267. https://doi.org/10.1016/j.ijid.2020.02.050.
20. Seaman CP, Tran LTT, Cowling BJ, Sullivan SG. 2019. Self-collected compared with professional-collected swabbing in the diagnosis of influenza in symptomatic individuals: a meta-analysis and assessment of validity. J Clin Virol 118:28–35. https://doi.org/10.1016/j.jcv.2019.07.010.
21. Wenham C, Gray ER, Keane CE, Donati M, Paolotti D, Pebody R, Fragaszy E, McKendry RA, Edmunds WJ. 2018. Self-swabbing for virological confirmation of influenza-like illness among an internet-based cohort in the UK during the 2014–2015 flu season: pilot study. J Med Internet Res 20:e71. https://doi.org/10.2196/jmir.9084.
22. To KK, Lu L, Yip CC, Poon RW, Fung AM, Cheng A, Lui DH, Ho DT, Hung IF, Chan KH, Yuen KY. 2017. Additional molecular testing of saliva specimens improves the detection of respiratory viruses. Emerg Microbes Infect 6:e49. https://doi.org/10.1038/emi.2017.35.
23. Pasomsub E, Watcharananan SP, Boonyawat K, Janchompoo P, Wongtabtim G, Suksuwan W, Sungkanuparph S, Phuphuakrat A. 2020. Saliva sample as a non-invasive specimen for the diagnosis of coronavirus disease-2019 (COVID-19): a cross-sectional study. Clin Microbiol Infect 27:285.e1–285.e4. https://doi.org/10.1016/j.cmi.2020.05.001.
24. To KK, Tsang OT, Chik-Yan Yip C, Chan KH, Wu TC, Chan JMC, Leung WS, Chik TS, Choi CY, Kandamby DH, Lung DC, Tam AR, Poon RW, Fung AY, Hung IF, Cheng VC, Chan JF, Yuen KY. 2020. Consistent detection of 2019 novel coronavirus in saliva. Clin Infect Dis 71:841–843. https://doi.org/10.1093/cid/ciaa149.
25. Babady NE, McMillen T, Jani K, Viale A, Robilotti EV, Aslam A, Diver M, Sokoli D, Mason G, Shah MK, Korenstein D, Kamboj M. 2021. Performance of severe acute respiratory syndrome coronavirus 2 real-time RT-PCR tests on oral rinses and saliva samples. J Mol Diagn 23:3–9. https://doi.org/10.1016/j.jmoldx.2020.10.018.
26. Centers for Disease Control and Prevention RVB, Division of Viral Diseases. 2020. CDC 2019-nCoV real-time RT-PCR diagnostic panel, acceptable alternative primer and probe sets. CDC, Atlanta, GA.
27. Centers for Disease Control and Prevention RVB, Division of Viral Diseases. 2020. Research use only 2019-novel coronavirus (2019-nCoV) real-time RT-PCR primers and probes. CDC, Atlanta, GA.
28. Centers for Disease Control and Prevention RVB, Division of Viral Diseases. 2020. CDC 2019-novel coronavirus (2019-nCoV) real-time RT-PCR diagnostic panel, instructions for use. CDC, Atlanta, GA.
29. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M. 2011. pROC: an open-source package for R and S+ to analyze and

compare ROC curves. BMC Bioinformatics 12:77. https://doi.org/10.1186/1471-2105-12-77.
30. Kassambara A. 2020. ggpubr: "ggplot2" based publication ready plots. R package, version 0.3.0.
31. Meyer D, Zeileis A, Hornik K. 2020. vcd: visualizing categorical data. R package, version 1.4–7.
32. R Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
33. Ott IM, Strine MS, Watkins AE, Boot M, Kalinich CC, Harden CA, Vogels CBF, Casanovas-Massana A, Moore AJ, Muenker MC, Nakahata M, Tokuyama M, Nelson A, Fournier J, Bermejo S, Campbell M, Datta R, Dela Cruz CS, Farhadian SF, Ko AI, Iwasaki A, Grubaugh ND, Wilen CB, Wyllie AL. 4 August 2020. Simply saliva: stability of SARS-CoV-2 detection negates the need for expensive collection devices. medRxiv https://doi.org/10.1101/2020.08.03.20165233.
34. Migueres M, Mengelle C, Dimeglio C, Didier A, Alvarez M, Delobel P, Mansuy JM, Izopet J. 2020. Saliva sampling for diagnosing SARS-CoV-2 infections in symptomatic patients and asymptomatic carriers. J Clin Virol 130:104580. https://doi.org/10.1016/j.jcv.2020.104580.
35. Teo AKJ, Choudhury Y, Tan IB, Cher CY, Chew SH, Wan ZY, Cheng LTE, Oon LLE, Tan MH, Chan KS, Hsu LY. 2021. Saliva is more sensitive than nasopharyngeal or nasal swabs for diagnosis of asymptomatic and mild COVID-19 infection. Sci Rep 11:3134. https://doi.org/10.1038/s41598-021-82787-z.
36. Tutuncu EE, Ozgur D, Karamese M. 26 January 2021. Saliva samples for detection of SARS-CoV-2 in mildly symptomatic and asymptomatic patients. J Med Virol https://doi.org/10.1002/jmv.26821.
37. Yokota I, Shane PY, Okada K, Unoki Y, Yang Y, Inao T, Sakamaki K, Iwasaki S, Hayasaka K, Sugita J, Nishida M, Fujisawa S, Teshima T. 25 September 2020. Mass screening of asymptomatic persons for SARS-CoV-2 using saliva. Clin Infect Dis https://doi.org/10.1093/cid/ciaa1388.