

# Ultimate COVID-19 Detection Protocol Based on Saliva Sampling and qRT-PCR with Risk Probability Assessment

Joungha Won<sup>1,2</sup>, Hasan Hüseyin Kazan<sup>3</sup>, Jea Kwon<sup>2,4</sup>, Myungsun Park<sup>2</sup>, Mehmet Ali Ergun<sup>5</sup>, Sureyya Ozcan<sup>6</sup>, Byung Yoon Choi<sup>7</sup>, Won Do Heo<sup>1\*</sup> and C. Justin Lee<sup>2\*</sup>

<sup>1</sup>Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon 34141, <sup>2</sup>Center for Cognition and Sociality, Cognitive Glioscience Group, Institute for Basic Science, Daejeon 34126, Korea, <sup>3</sup>Department of Biological Sciences, Middle East Technical University, Ankara 06800, Turkey, <sup>4</sup>KU-KIST Graduate School of Converging Science and Technology, Korea University, Seoul 02841, Korea, <sup>5</sup>Department of Medical Genetics, Gazi University Faculty of Medicine, Ankara 06560, <sup>6</sup>Department of Chemistry, Middle East Technical University, Ankara 06800, Turkey, <sup>7</sup>Department of Otorhinolaryngology, Seoul National University Bundang Hospital, Seongnam 13620, Korea

In the era of COVID-19 outbreak, various efforts are undertaken to develop a quick, easy, inexpensive, and accurate way for diagnosis. Although many commercial diagnostic kits are available, detailed scientific evaluation is lacking, making the public vulnerable to fear of false-positive results. Moreover, current tissue sampling method from respiratory tract requires personal contact of medical staff with a potential asymptomatic SARS-CoV-2 carrier and calls for safe and less invasive sampling method. Here, we have developed a convenient detection protocol for SARS-CoV-2 based on a non-invasive saliva self-sampling method by extending our previous studies on development of a laboratory-safe and low-cost detection protocol based on qRT-PCR. We tested and compared various self-sampling methods of self-pharyngeal swab and self-saliva sampling from non-carrier volunteers. We found that the self-saliva sampling procedure gave expected negative results from all of the non-carrier volunteers within 2 hours, indicating cost-effectiveness, speed and reliability of the saliva-based method. For an automated assessment of the sampling quality and degree of positivity for COVID-19, we developed scalable formulae based on a logistic classification model using both cycle threshold and melting temperature from the qRT-PCR results. Our newly developed protocol will allow easy sampling and spatial-separation between patient and experimenter for guaranteed safety. Furthermore, our newly established risk assessment formula can be applied to a large-scale diagnosis in health institutions and agencies around the world.

**Key words:** Coronavirus, COVID-19, SARS-CoV-2, Saliva sampling, COVID-19 risk assessment

## INTRODUCTION

COVID-19, which is caused by the 2019 novel coronavirus

SARS-CoV-2, has emerged from Wuhan, China in 2019. SARS-CoV-2 belongs to the Coronaviridae RNA virus family that penetrates respiratory epithelial cells and causes fever, cough, myalgia or fatigue, other various symptoms, and even death [1]. More importantly, asymptomatic carriers of SARS-CoV-2 can be a silent source of uncontrolled infection [2]. It has been reported that the spike protein of SARS-CoV-2 has been mutated from SARS-CoV-2, and enabled it to increase the infectivity by contacting with respiratory epithelial cells [3]. WHO reports on October 10<sup>th</sup>, 2020, that over 3.6 million infections have been reported worldwide, with a mortality rate of three percent [4]. Although the rates

Submitted December 29, 2020, Revised January 5, 2021,  
Accepted January 6, 2021

\*To whom correspondence should be addressed.  
Won Do Heo, TEL: 82-42-350-2642, FAX: 82-42-350-2610  
e-mail: wondo@kaist.ac.kr  
C. Justin Lee, TEL: 82-42-878-9150, FAX: 82-42-878-9151  
e-mail: cjl@ibs.re.kr

of mortality are very various, it has been generally reported that the elderly, the immune-suppressed and those with other diseases are highly susceptible to SARS-CoV-2 infection and more likely to die [5]. What is more alarming is the fact that there is no sign of slowing down. Instead, there is a repeating pattern in several countries of a secondary outbreak, after a period of rapid initial outbreak. Therefore, tremendous research efforts on developing disinfection strategies and treatments to fight against COVID-19 are being actively conducted.

Another way to fight against COVID-19 is to develop a safe, fast and easy diagnostic method which allows to quickly determine the epidemiological surveillance of SARS-CoV-2 and execute preventive measures. The most popular diagnostic protocol for COVID-19 involves a step-by-step procedure starting from collection of tissue sample from a subject, isolation and extraction of viral RNA, and detection of SARS-CoV-2 specific genes [6]. Although various novel sampling and detection methods have been described, the traditional way of tissue sampling from respiratory tract and real-time gene expression analysis are widely used for the detection of SARS-CoV-2 in clinical settings [7]. There are two ways of tissue sampling from the upper respiratory tract: one from nasopharynx and the other from posterior oropharynx. Sampling through a nasal swab requires a deep insertion of a swab to nasopharynx inside the nose [8], which is a burdensome and uncomfortable process for the subject. Sampling from pharyngeal wall with a throat swab requires a step to press the tongue and scrape the tissue from the throat wall [9]. Even though a throat swab might be easier than a nasal swab, both methods should be performed by highly trained medical personnel. This requirement forces medical personnel to be constantly exposed to the infectious viruses. In fact, there have been numerous cases where medical personnel performing the tissue sampling inadvertently contracted COVID-19 [10]. In particular, a touching of the pharyngeal area could trigger pharyngeal reflex, resulting in nausea or cough from the subject and potentially increase the risk of transmitting the virus to nearby health care workers [10]. In fact, a risk of COVID-19 infection of the front-line health-care workers in the USA and the United Kingdom are ten times higher than normal population [11].

To protect medical staff from the danger of potential infection, we should be making additional efforts such as expanding non-contact health care, strengthening the protective gear for medical staff, and implementing self-tissue sampling method [11]. As a candidate self-tissue sampling, a saliva-based sampling has been previously used in the detection of Zika virus to ensure safe spatial separation between patient and medical staff and experimenters [6, 12, 13]. It has been already reported that SARS-CoV-2 virus can be

detected in saliva as well as in the respiratory tract in COVID-19 patients [14], and the detection sensitivity during the time course of disease progression is similar between saliva and nasopharyngeal swab samples [15-17]. Given that pharyngeal swab is a widely used method in clinics [7], we consider that saliva-sampling is equally practical to detect SARS-CoV-2 virus during the time course of disease progression. These raise a hope for a complete spatial separation of a subject and medical staff via self-saliva sampling.

In addition to the problem of accidental exposure to SARS-CoV-2 in the clinics, frequent false-positive results during the diagnosis remain as a serious challenge. Most of the popular diagnostic kits utilize DNA-amplification of SARS-CoV-2 genes by Real-Time Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR) to determine the presence or absence of SARS-CoV-2 genes or positive or negative infection for COVID-19 [6]. During a conventional diagnosis, an experimenter obtains the result of qRT-PCR reaction for each sample in the form of the value for the threshold cycle (CT value) of an amplification plot. A well-known pitfall of qRT-PCR method is that it can give a non-specific amplification product due to a primer-dimerization, leading to a false-positive result [9]. We have previously described in detail the primer-design guidelines to prevent non-specific amplification and false-positive reactions and reported ten validated primer sets for SARS-CoV-2 detection in SYBR-green based qRT-PCR and conventional PCR [9, 18]. However, even with highly optimized primer sets, qRT-PCR reactions can often lead to an inadvertent amplification due to unknown thermal reactions or random formation of a primer-dimer. To circumvent this issue, we have previously proposed to take advantage of the melting temperature,  $T_m$  value, which is calculated from each qRT-PCR reaction [18].  $T_m$  value for qRT-PCR reaction due to primer-dimerization should be different from the one due to a well-targeted reaction. An experimenter can evaluate in combination with CT and  $T_m$  values to determine the presence or absence of SARS-CoV-2 genes more accurately. Nevertheless, this process of evaluation is prone to errors arising from subjective opinions and personal biases of the experimenter. Therefore, there is a pressing need for developing an error-free and objective algorithm or formula composed of CT and  $T_m$  values to calculate the degree of the positivity or presence of SARS-CoV-2. Such algorithm or formula should be indispensable for large-scale testing facilities.

In a series of two previous papers, we have reported the laboratory-safe and low-cost SARS-CoV-2 detection protocol, which is composed of self-pharyngeal swab sampling procedure, a Trizol-based RNA extraction, cDNA reverse transcription, SYBR green-based qRT-PCR protocol or conventional PCR, and optimized primer sets for SARS-CoV-2 detection [9, 18]. In the current study,

we set out to further develop and optimize a saliva-based ultimate detection protocol, equipped with scalable formulae based on logistic classification model.

## MATERIALS AND METHODS

### *Volunteer recruitment*

The purpose of the sampling and procedure through a pharyngeal swab and saliva-based sampling from volunteers were approved by Seoul National University Hospital Institutional Review Board (IRBY-H-1807-197-966) and IRB guidance lines. Total of 7 volunteers (Volunteer A-G) participated in this experiment. Some of these volunteers have direct or indirectly contact with SARS-CoV-2 infected patients or visited a known COVID-19 outbreak area.

The saliva from two human subjects whom had been diagnosed as “COVID-19-positive subject” and “COVID-19-negative subject” in the Department of Medical Microbiology, Gazi University, Ankara, Turkey by a commercial TaqMan probe-based kit approved by Republic of Turkey Ministry of Health was also included in the study. The involvement of COVID-19-positive patient and –negative individual was allowed by Gazi University Clinical Research Committee 13.10.2020/678 and the self-saliva sampling was performed after obtaining signed consent form from the patient and the individual.

### *Self-pharyngeal swab and self-saliva sampling procedure*

Sampling was conducted strictly through self-collection procedure. Detailed procedure of self-pharyngeal swab is described in the previously published paper [9]. The saliva-based sampling procedure was modified from the previously described procedure [13]. Briefly, we asked volunteers to drool and spit at least 1ml saliva into sterile polypropylene medical container (Medical container, Catalog #:400025, SPL, Republic of Korea), after brushing the teeth and vigorously rinsing the mouth with tap water. And then we added 20 µg of Proteinase K solution (Proteinase K solution, 20 mg/ml, Catalog #:21560025-2, Bio-world, USA) for an inactivation of SARS-CoV-2. To further ensure an elimination of viral activity, we transferred 500 µl of saliva and Proteinase K sample into 500 µl Trizol (TRIzol™ Reagent, Catalog #: 15596026, Invitrogen™, USA), followed by mixing with pipetting vigorously up-and-down.

### *Total RNA extraction*

Detailed procedure for Trizol-based manual RNA extraction is described in previously published paper [18]. Briefly, each sample was incubated in Trizol for 5 minutes in room temperature, and

then 200 µl chloroform was added, mixing by inverting the tube 5 times, incubated for 3 minutes and centrifuged for 15 minutes at 12,000×g at 4°C. The clear upper aqueous layer which contained RNA was transferred to a new 1.5 ml tube and same volume of isopropanol was added. After incubating for 10 minutes on room temperature, gently mix by inverting 5 times was followed. The sample was centrifuged for 10 minutes at 12,000×g at 4°C. The supernatant was discarded and the remaining pellet was washed by 1 ml of 70% ethanol and centrifuged for 10 minutes at 7,500×g at 4°C. The sample was washed again with 70% ethanol and centrifuged for 10 minutes at 7,500×g at 4°C. The supernatant was discarded and the RNA pellet was air-dried for 5 minutes. To solubilize the RNA pellet, the pellet was re-suspended in 10 µl of RNase-free water.

For the RNA kit preparation, we used QIAamp® Viral RNA Mini (Catalog #: 52904, Qiagen, Germany) and modify the given protocol which is provided by the company. Each sample was incubated in Trizol for 5 minutes in room temperature, and then transferred to QIAamp Mini column and centrifuged at 6,000×g for 1 minute in room temperature. Then, we added 500 µl AW1 Buffer to the column and centrifuged at 6,000×g for 1 minute in room temperature. And then, we added 500 µl AW2 Buffer to the column and centrifuged at 20,000×g for 3 minutes in room temperature. Finally, we eluted with 20 µl of RNase-free water in a clean 1.5 ml microcentrifuge tube.

### *SARS-CoV-2 RNA as a positive control*

The positive control containing SARS-CoV-2 viral RNAs was obtained from the Korea Centers for Disease Control and Prevention (<http://www.cdc.go.kr/>). Detailed description of how SARS-CoV-2 viral RNA was prepared in a separate report [19]. Briefly, SARS-CoV-2 viral RNA was prepared by extracting total RNA from Vero cell line, which is originated from the kidney of African green monkey (*Cercopithecus aethiops*), infected with a viral clone, BetaCoV/Korea/KCDC03/2020 at MOI 0.05.

### *Reverse-transcription*

Extracted total RNA was converted to complementary DNA (cDNA) using SuperScript™ III First-Strand Synthesis System (Catalog #: 18080051; Invitrogen™, USA), following the manufacturer's recommended procedures with some modifications. Detailed procedure of reverse transcription is described in previously published paper [9].

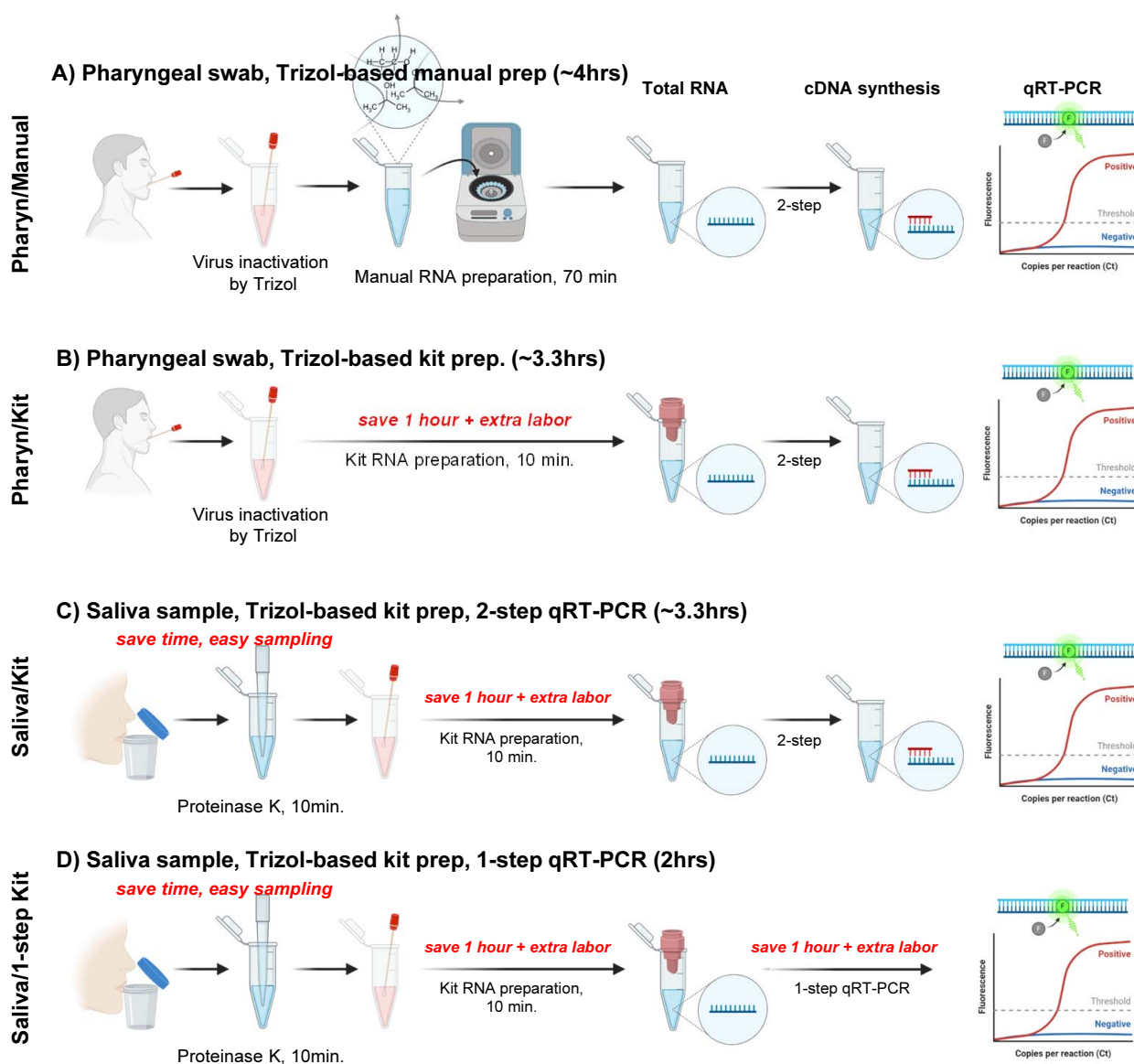
### *Real-time qPCR*

2X Power SYBR® Green PCR Master Mix (Catalog #: 4368577, Thermo Fisher Scientific, USA) was used. The thermal cycle con-

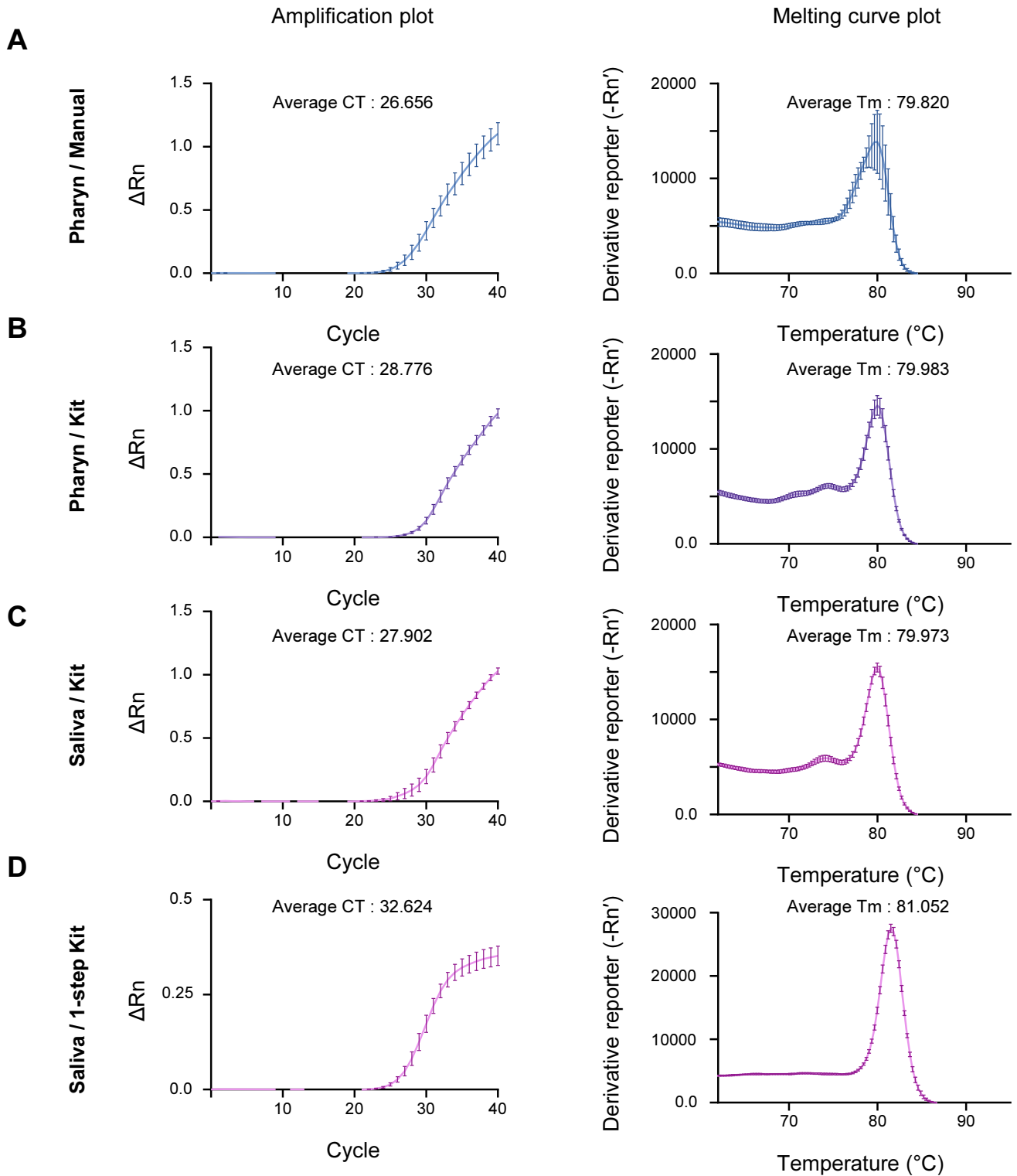
ditions were 50°C for 2 minutes, 95°C for 10 minutes, 40 cycles with 95°C for 15 second and 62°C for 1 minute, followed by melting curve stage at 95°C for 10 second and 60°C for 1 minutes in Quantstudio 1 Real-Time PCR system (Applied Biosystems, USA) and CFX96™ Real-Time PCR Detection System (Bio-Rad Laboratories, USA). Previously reported primer sets for SARS-CoV-2 detection and internal positive control are used [18]; SARS-CoV-2\_IBS\_E2, SARS-CoV-2\_IBS\_RdRP2, SARS-CoV-2\_IBS\_S2, SARS-CoV-2\_IBS\_N1, GAPDH primer sets. The final concentration of the primer mix was 500 nM.

### 1-step real-time qPCR

One Step TB Green® PrimeScript™ RT-PCR Kit (Catalog #: RR066A, Takara, Japan) was used. We followed the manufacturer's manual for preparing PCR mixture and thermal cycle conditions. In 20 µl total reaction, we used 2X One Step TB Green RT-PCR Buffer III 10 µl, TaKaRa Ex Taq HS 0.4 µl, PrimeScript RT enzyme Mix II 0.4 µl, PCR Forward/Reverse Primer (10 µM) 0.4 µl, each, ROX Reference Dye or Dye II (50 X) 0.4 µl, total RNA 2 µl, and RNase free water 6 µl. Thermal cycle was 42°C for 5 minutes, 95°C for 10 seconds, 40 cycles with 95°C for 5 second and 62°C for

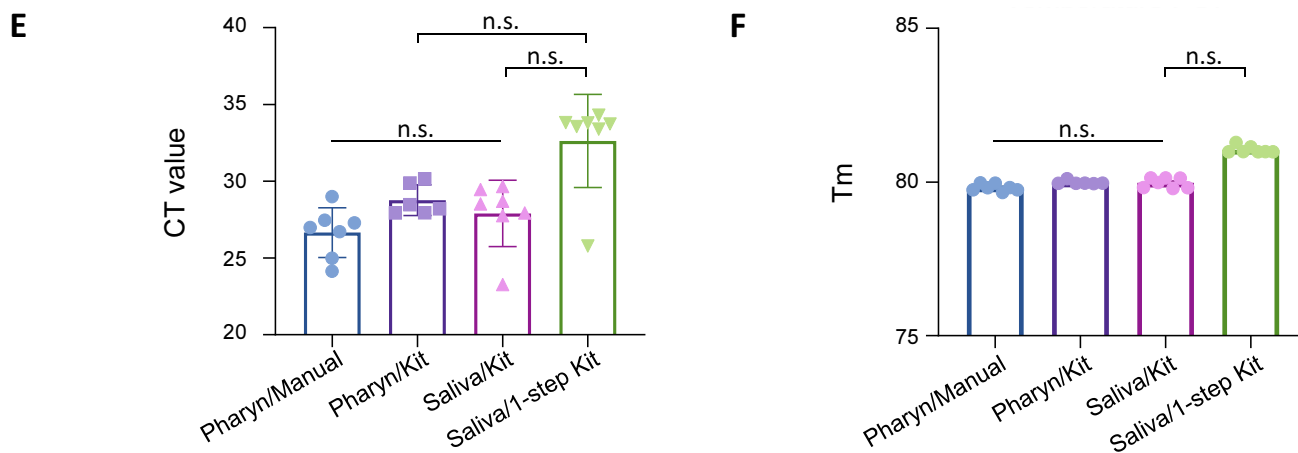


**Fig. 1.** Experimental scheme for safe, fast and easy human tissue sampling, total RNA extraction and qRT-PCR. (A and B) Samples are collected by self-pharyngeal swab and RNA are extracted by (A) Trizol-based manual RNA extraction method (Pharyn/Manual) or (B) Trizol and commercial RNA extraction kit (Pharyn/Kit). Trizol is used for inactivating viral infection. (C and D) Samples are collected by self-saliva sampling and RNA are extracted by Trizol and commercial RNA extraction kit. Trizol and proteinase K is used for inactivating viral infection. (C) cDNA synthesis and qRT-PCR are separately performed from saliva sample (Saliva/Kit) (D) cDNA synthesis and qRT-PCR are combined in one reaction (Saliva/1-step kit).



**Fig. 2.** Comparison of sampling quality between Pharyn/Manual, Pharyn/Kit, Saliva/Kit, and Saliva/1-step Kit by GAPDH qRT-PCR result. qRT-PCR result of amplification plot and melting curve plot from GAPDH primer set, from volunteer A-G, in (A) Pharyn/Manual, (B) Pharyn/Kit, (C) Saliva/Kit and (D) Saliva/1-step Kit prep sample. Average CT value and average Tm value are described in each figures. Reporter value (Rn) is calculated by normalized the fluorescence signal from SYBR green to reference dye, ROX. Peaks in the melting curve plot represent the melting temperature [20] of amplicons. Derivative reporter (-Rn) was calculated as the negative first derivative of Rn generated by the reporter during PCR amplification. All data are represented as the mean±S.E.M.





**Fig. 2.** Continued.

30 second. Other conditions were the same as described in Real-time qPCR.

## RESULTS

### *Self-tissue sampling protocols for SARS-CoV-2 detection*

To optimize the tissue sampling and RNA extraction procedure for enhanced safety, speed, and cost-effectiveness, we compared and contrasted four different procedures (Fig. 1A~C); 1) Pharyn/Manual: self-pharyngeal swab sampling and manual RNA preparation, 2) Pharyn/Kit: self-pharyngeal swab sampling and commercial kit-based 2-step RNA preparation, 3) Saliva/Kit: self-saliva sampling and commercial kit-based 2-step cDNA reverse transcription and qRT-PCR, and 4) Saliva/1-step Kit: self-saliva sampling and 1-step qRT-PCR with combined cDNA reverse transcription and qRT-PCR. Pharyn/Manual procedure utilized a manual Trizol-based RNA extraction, whereas other procedures utilized Trizol-based RNA extraction kit. The difference between Saliva/Kit and Saliva/1-step Kit was that Saliva/1-step Kit utilized 1-step qRT-PCR with combined cDNA reverse transcription and qRT-PCR, whereas Saliva/Kit utilized a 2-step procedure.

To assess the quality of tissue sampling, we performed qRT-PCR from 7 volunteers, who went through the two different tissue sampling methods and whose samples went through the four different procedures (Fig. 1A~D). We found that Pharyn/Manual, Pharyn/Kit, and Saliva/Kit sampling methods showed no significant difference in CT and Tm of amplicons in the human internal positive control, GAPDH (Fig. 2A~C, 2E, 2F), indicating that the self-saliva sampling has similar sensitivity as the self-pharyngeal sampling. In addition, we found that Saliva/1-step Kit sampling showed slightly

higher CT and Tm, compare to other procedures (Fig. 2D~F). This might be caused by lower efficiency of the 1-step qRT-PCR kit, which has a shorter time of cDNA reverse transcription, shorter extension time in the thermal cycle, and lower primer concentration. Nevertheless, the CT and Tm values were within the positive range of GAPDH (Fig. 2E, 2F), indicating that the self-saliva sampling even with 1-step Kit has similar sensitivity as the self-pharyngeal sampling or the self-saliva sampling with 2-step RNA Kit.

To compare and contrast the cost-effectiveness and speed among the four procedures, we estimated the total cost per each sample (Table 1) and duration of each procedure (Table 2). We found that the manual RNA extraction method was the most economical (~\$15), although it took longer duration (~4 hours) and more work load for the experimenter. Using the 2-step RNA prep kit slightly increased the cost, while reducing time to 3.3 hours and decreasing work load. Finally, with 1-step qRT-PCR kit and self-saliva sampling, the entire procedure took less than 2 hours and under \$14 (Table 1 and Table 2), making the Saliva/1-step Kit the fastest and most cost-effective.

### *Direct comparison of Pharyn/Manual, Pharyn/Kit, Saliva/Kit, and Saliva/1-step Kit procedures*

Right after the local outbreak of COVID-19 infections near Gwanghwamun area in Seoul, Republic of Korea, on August 15<sup>th</sup>, 2020, we performed the SARS-CoV-2 detection protocols for volunteers who have either unknowingly contacted SARS-CoV-2-positive patient or visited near Gwanghwamun area during the outbreak. Tissue sampling was conducted only on asymptomatic volunteers, after proper self-quarantine, following the governmental or institute's guidelines. All volunteers had no symptoms related

**Table 1.** Cost breakdown per volunteer sample

	Pharyn/Manual	Pharyn/Kit	Saliva/Kit	Saliva/1-step Kit
Swab or collection tube and etc.	\$0.50	\$0.70	\$0.70	\$0.70
Trizol	\$1.10	\$1.10	\$1.10	\$1.10
RNA preparation kit	-	\$7.70	\$7.70	\$7.70
Protease K	-	-	\$1.50	\$1.50
cDNA synthesis Kit	\$8.13	\$8.13	\$8.13	
SYBR master mix	\$3.25	\$3.25	\$3.25	
1-step-qRT-PCR kit				\$1.60
Primer set	\$0.10	\$0.10	\$0.10	\$0.10
Tubes and buffers	\$2.50	\$1.00	\$1.00	\$1.00
Total time	\$15.58	\$21.98	\$23.48	\$13.70

**Table 2.** Durations of each sampling, RNA extraction and qRT-PCR procedure

	Pharyn/Manual	Pharyn/Kit	Saliva/Kit	Saliva/1-step Kit
Self-pharyngeal swab collection	~10 min	~10 min	-	-
Self-saliva collection	-	-	~10 sec	~10 sec
RNA extraction	~70 min	~30 min	~40 min	~40 min
cDNA synthesis	60 min	60 min	60 min	5 min
qRT-PCR	100 min	100 min	100 min	65 min
Total time	~4 hrs	~3.3 hrs	~3.3 hrs	~2 hrs

to COVID-19.

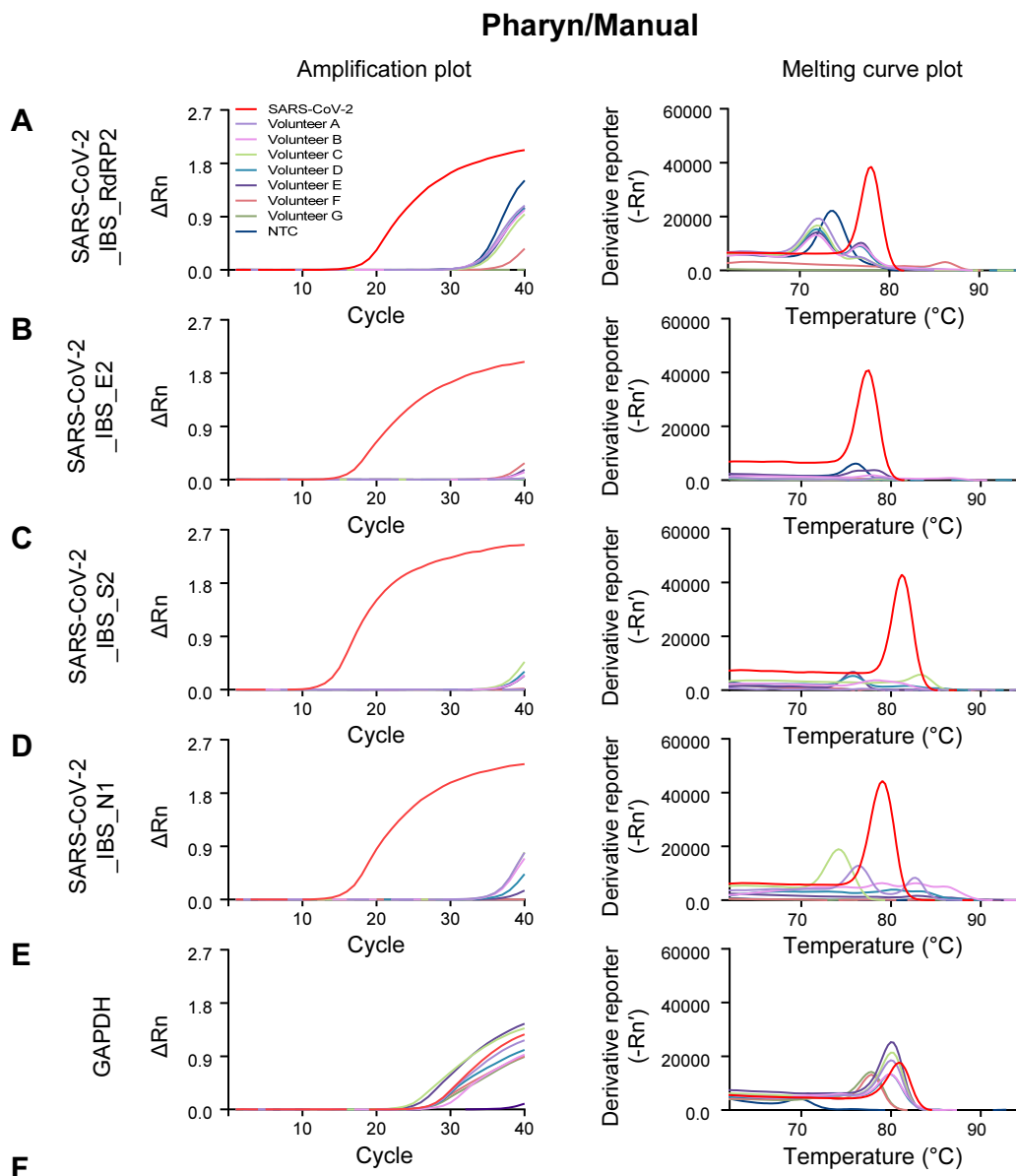
To directly compare the efficiency and sensitivity of Pharyn/Manual, Pharyn/Kit, Saliva/Kit, and Saliva/1-step Kit procedures, we firstly conducted self-tissue sampling and each diagnosis protocol for Volunteers A through G. We asked each volunteer to perform both self-pharyngeal swab sampling and self-saliva sampling. We then processed each sample through RNA extraction and qRT-PCR steps as depicted in Fig. 1A~D. For the qRT-PCR, we utilized four SARS-CoV-2-targeted SARS-CoV-2\_IBS\_RdRP2, SARS-CoV-2\_IBS\_E2, SARS-CoV-2\_IBS\_S2, SARS-CoV-2\_IBS\_N1, and one human gene-targeted GAPDH (for internal positive control) primer sets, which we have previously validated and reported [9, 18].

As expected, the SARS-CoV-2-targeting primer sets showed sufficient amplification in SARS-CoV-2 positive control (the red traces in Fig. 3A~D, Fig. 4A~D, Fig. 5A~D, and Fig. 6A~D), and the GAPDH primer set showed sufficient amplification in all of the volunteer samples in all four procedures (Fig. 3E, Fig. 4E, Fig. 5E, and Fig. 6E). Among the four SARS-CoV-2-targeted primer sets, the S2-targeted primer set, SARS-CoV-2\_IBS\_S2, showed CT>35 or undetermined (u.d.) in most of the volunteer samples (Fig. 3C and 3F; Fig. 4C and 4F; Fig. 5C and 5F; Fig. 6C and 6F), which can be interpreted as “negative” for SARS-CoV-2. This is based on the assumption that CT value of a single molecule (copy) amplification is near 35 [21]. In contrast, other SARS-CoV-2-targeted primer sets (SARS-CoV-2\_IBS\_RdRP2, SARS-CoV-2\_IBS\_E2, and

SARS-CoV-2\_IBS\_N1) sometimes showed CT<35 in some of the volunteer samples (Fig. 3~6; CT values in red), which could theoretically be interpreted as “positive” for SARS-CoV-2. However, Tm values for the corresponding melting curves showed a significant deviation of more than 0.3°C from the Tm value for SARS-CoV-2 positive control (Fig. 3~6; Tm values in red). In particular, Saliva/1-step Kit procedure group showed the most number of CT<35, which all turn out to be showing significant deviations of Tm from those of SARS-CoV-2 positive control (Fig. 6). These results indicate that the apparent CT<35 in some of the samples might be caused by an inappropriate amplification due to primer-dimerization [18]. Thus, considering both CT and Tm values, all of the samples in all four procedures gave “negative” for SARS-CoV-2 as expected. These results imply that considering both CT and Tm is practical in eliminating false positives. Taken together, these results validate the sufficient efficiency and sensitivity of the protocols with both self-pharyngeal and self-saliva sampling, even with the time-saving 1-step qRT-PCR kit.

***Development of a mathematical formula for automatic determination of sampling quality and risk assessment based on logistic classification model***

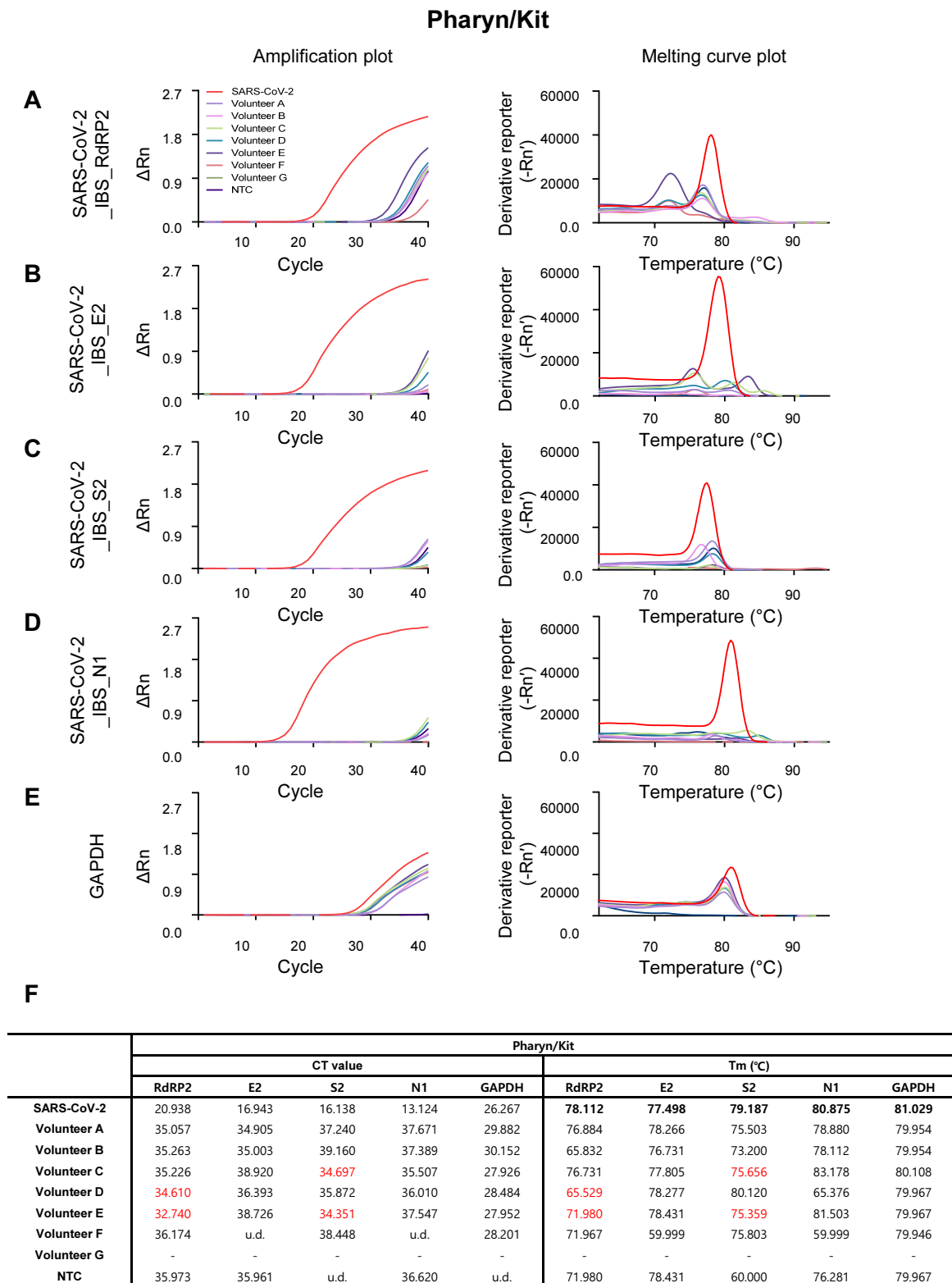
We targeted and detected four different genes in SARS-CoV-2 for COVID-19 detection. As previously described [9], we considered subject as COVID-19-positive if at least one of the four different SARS-CoV-2 genes is positive. Therefore, we considered



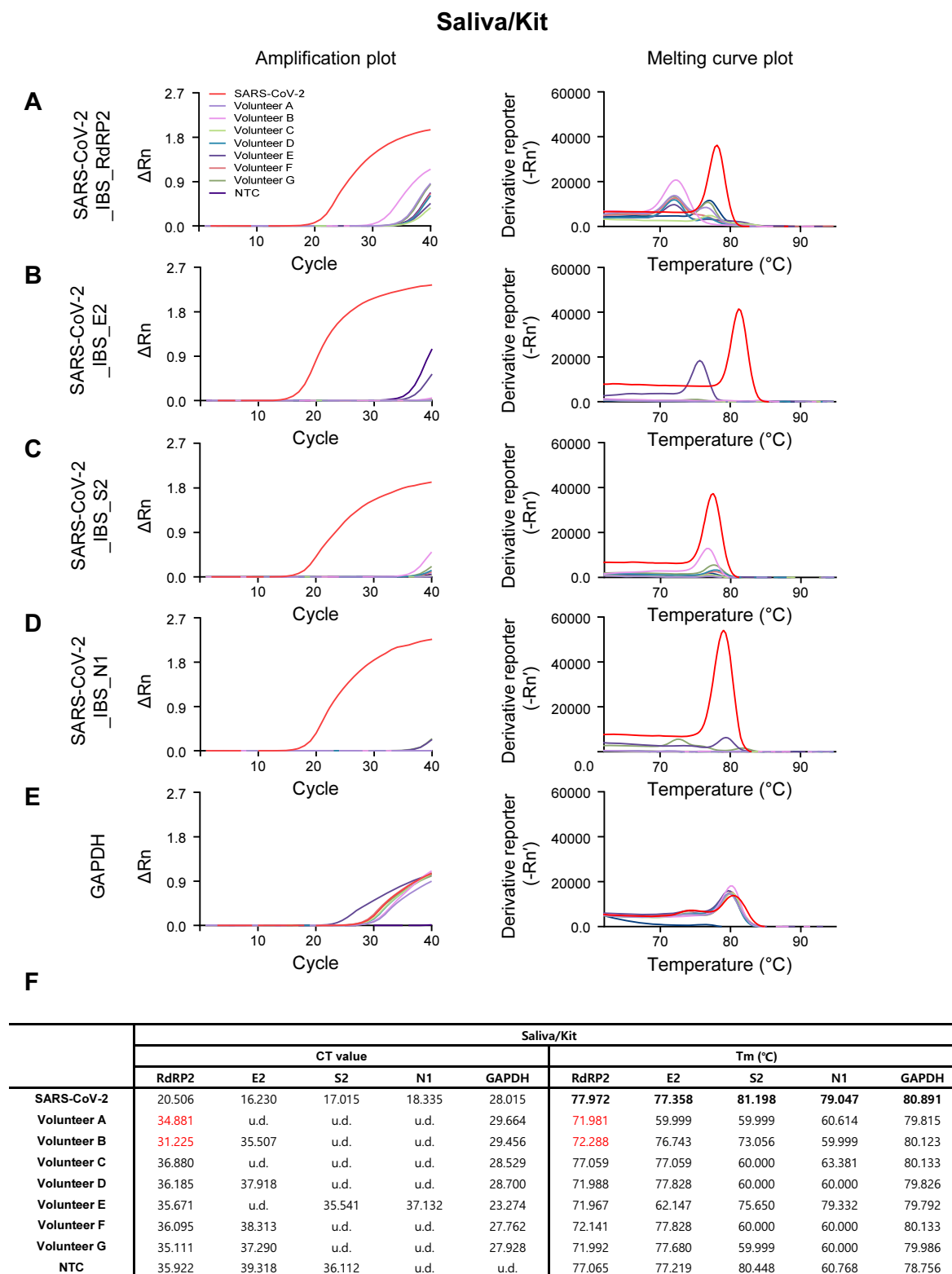
	Pharyn/Manual									
	CT value					Tm (°C)				
	RdRP2	E2	S2	N1	GAPDH	RdRP2	E2	S2	N1	GAPDH
SARS-CoV-2	17.974	14.848	13.848	15.988	26.956	<b>78.126</b>	<b>77.511</b>	<b>81.198</b>	<b>79.047</b>	<b>81.020</b>
Volunteer A	<b>34.796</b>	u.d.	u.d.	<b>34.473</b>	27.298	<b>72.042</b>	59.999	59.999	76.310	79.816
Volunteer B	<b>34.872</b>	37.818	36.908	<b>34.685</b>	28.996	<b>71.889</b>	77.682	77.986	<b>78.749</b>	79.663
Volunteer C	35.701	u.d.	35.913	<b>34.573</b>	24.145	76.767	59.999	83.017	<b>74.023</b>	79.968
Volunteer D	35.033	u.d.	36.670	35.882	27.465	71.889	59.999	75.704	80.126	79.821
Volunteer E	35.207	37.651	37.109	37.527	24.989	71.892	77.992	75.552	61.676	79.974
Volunteer F	37.782	u.d.	36.653	u.d.	26.708	66.589	73.627	77.911	62.000	79.747
Volunteer G	u.d.	u.d.	u.d.	u.d.	26.988	62.000	91.528	62.000	62.000	79.747
NTC	37.735	37.712	36.470	u.d.	u.d.	75.616	93.517	62.000	62.000	71.638

**Fig. 3.** Amplification and melting curve plot from Volunteer A~G's Pharyn/Manual samples. Amplification plot (left panel) and melting curve plot (right panel) from qRT-PCR of volunteer A~G's pharyn/manual samples. qRT-PCR was performed by using the previously reported SARS-CoV-2 primer sets; (A) SARS-CoV-2\_IBS\_RdRP2, (B) SARS-CoV-2\_IBS\_E2, (C) SARS-CoV-2\_IBS\_S2, (D) SARS-CoV-2\_IBS\_N1. (E) GAPDH primer set was used for IPC detection. Each color code represent individual volunteer result and SARS-CoV-2 amplification was represent as red color. (F) Ct value and Tm from volunteers A~G's Pharyn/Manual samples.



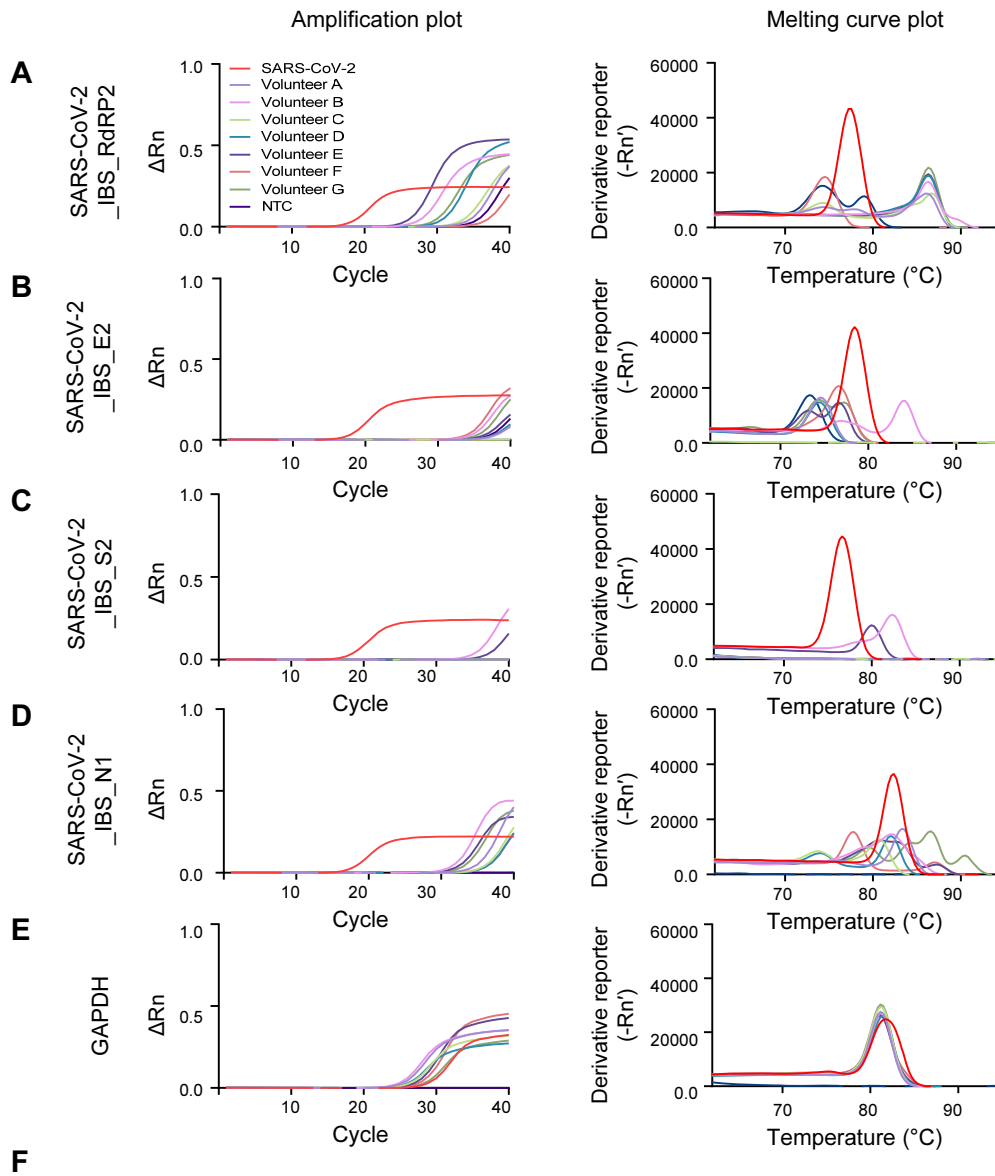


**Fig. 4.** Amplification and melting curve plot from Volunteer A~G's Pharyn/Kit samples. Amplification plot (left panel) and melting curve plot (right panel) from qRT-PCR of volunteer A~G's pharyn/kit samples. qRT-PCR was performed by using the previously reported SARS-CoV-2 primer sets; (A) SARS-CoV-2\_IBS\_RdRP2, (B) SARS-CoV-2\_IBS\_E2, (C) SARS-CoV-2\_IBS\_S2, (D) SARS-CoV-2\_IBS\_N1. (E) GAPDH primer set was used for IPC detection. Each color code represent individual volunteer result and SARS-CoV-2 amplification was represent as red color. (F) Ct value and Tm from volunteers A~G's Pharyn/Kit samples.



**Fig. 5.** Amplification and melting curve plot from Volunteer A~G's Saliva/Kit samples. Amplification plot (left panel) and melting curve plot (right panel) from qRT-PCR of volunteer A~G's saliva/kit samples. qRT-PCR was performed by using the previously reported SARS-CoV-2 primer sets; (A) SARS-CoV-2\_IBS\_RdRP2, (B) SARS-CoV-2\_IBS\_E2, (C) SARS-CoV-2\_IBS\_S2, (D) SARS-CoV-2\_IBS\_N1. (E) GAPDH primer set was used for IPC detection. Each color code represent individual volunteer result and SARS-CoV-2 amplification was represent as red color. (F) Ct value and Tm from volunteers A~G's Saliva/Kit samples.

### Saliva/1-step Kit



	Saliva/1-step Kit									
	CT value					Tm (°C)				
	RdRP2	E2	S2	N1	GAPDH	RdRP2	E2	S2	N1	GAPDH
SARS-CoV-2	18.594	18.237	18.517	18.071	30.446	77.369	78.427	76.614	82.354	81.891
Volunteer A	34.359	38.423	u.d.	34.614	34.309	86.130	74.651	60.000	83.411	80.985
Volunteer B	37.538	34.994	35.355	34.614	25.293	86.281	84.015	82.354	78.871	80.985
Volunteer C	33.470	u.d.	u.d.	35.885	33.793	86.734	61.510	60.000	80.692	80.985
Volunteer D	30.587	38.171	u.d.	36.248	33.806	86.292	74.355	60.000	82.061	80.989
Volunteer E	27.620	37.177	37.053	33.797	33.554	86.292	76.621	79.945	80.852	81.001
Volunteer F	36.769	34.373	u.d.	36.247	33.395	74.506	76.621	60.000	77.679	80.989
Volunteer G	29.723	35.917	u.d.	33.036	33.719	86.292	74.355	60.000	86.594	81.140
NTC	35.447	37.661	u.d.	u.d.	u.d.	75.555	78.274	59.999	78.123	59.999

**Fig. 6.** Amplification and melting curve plot from Volunteer A~G's Saliva/1-step Kit samples. Amplification plot (left panel) and melting curve plot (right panel) from qRT-PCR of volunteer A~G's saliva/1-step kit samples. To determine whether volunteers are infected to SARS-CoV-2, qRT-PCR was performed by using the previously reported SARS-CoV-2 primer sets; (A) SARS-CoV-2\_IBS\_RdRP2, (B) SARS-CoV-2\_IBS\_E2, (C) SARS-CoV-2\_IBS\_S2, (D) SARS-CoV-2\_IBS\_N1. (E) GAPDH primer set was used for IPC detection. Each color code represent individual volunteer result and SARS-CoV-2 amplification was represent as red color. (F) Ct value and Tm from volunteers A~G's Saliva/1-step Kit samples.

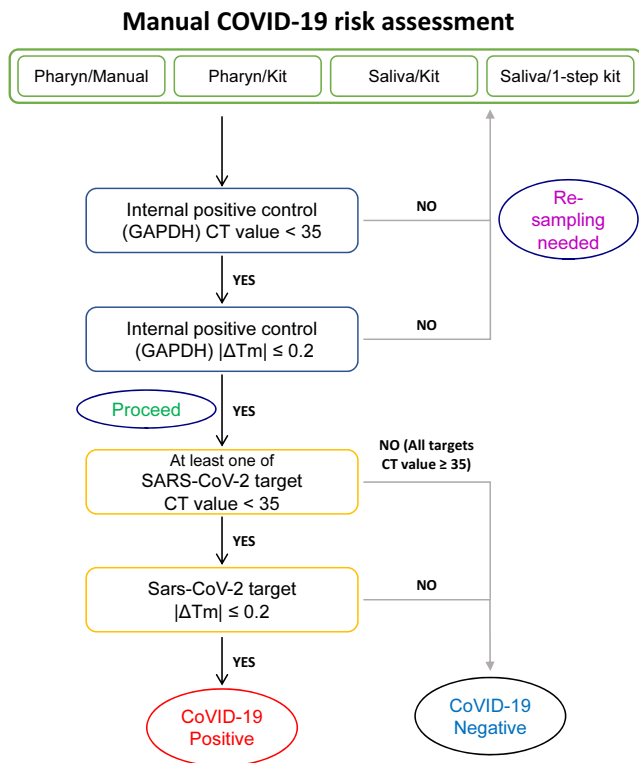


Fig. 7. Flow chart for manual SARS-CoV-2 risk assessment process.

subject as COVID-19-negative only when all four different SARS-CoV-2 genes were negative. This approach was optimized for ensuring and distinguishing the true-negative infection from potential positives.

During the processes of tissue sampling, RNA preparation and qRT-PCR, we realized that the manual determination of “positive” or “negative” by an experimenter requires some aspect of personal judgement (Fig. 7), which could allow biases and errors to enter. Especially in the case of Saliva/1-step Kit with many reactions showing CT<35 (Fig. 6), we identified multiple error-prone, decision-requiring steps (Fig. 7), in which personal biases and misjudgment could lead to false positives. To eliminate these possibilities, we set out to develop an error-free, unbiased mathematical formula to determine the “positive” or “negative” for SARS-CoV-2 based on the observed CT and Tm values.

For SARS-CoV-2 infection risk probability assessment and sampling quality assessment, we developed a formula for the probability  $\sigma$ , based on the logistic classification model [22]:

$$\sigma(x, a, b) = \frac{1}{1 + e^{a(x-b)}} \quad (\text{Equation 1})$$

(note,  $a$ : steepness,  $b$ : threshold).

In this equation, the probability,  $\sigma$  ranges from 0 to 1, where 0

represents “no risk” and 1 “high risk”. We developed a formula for CT value risk probability assessment ( $\sigma_{CT}$ ), based on Equation 1. We assumed that CT value of a single molecule (copy) amplification is  $\geq 35$  based on the previous work [21]. Thus, we setup the threshold value,  $b=35$ . We also adjusted the steepness value,  $a$  to give probabilities of 0.05 at CT 35+1 and 0.95 at CT 35-1. This estimation resulted in steepness value,  $b=3$ ;

$$\sigma_{CT}(CT, 3, 35) = \frac{1}{1 + e^{3(CT-35)}} \quad (\text{Equation 2}).$$

The graphical representation of  $\sigma_{CT}$  is shown in Fig. 8A.

We developed another formula for the sample quality and risk assessment using  $|\Delta Tm|$  ( $\sigma_{|\Delta Tm|}$ ), based on Equation 1. The threshold value of  $|\Delta Tm|$  was estimated to be 0.02 by the standard deviation of Tm among volunteer’s GAPDH results. And we adjusted the steepness to give probabilities of 0.05 at CT 0.2+0.1 and 0.95 at CT 0.2-0.1;

$$\sigma_{|\Delta Tm|}(|\Delta Tm|, 30, 0.2) = \frac{1}{1 + e^{30(|\Delta Tm| - 0.2)}} \quad (\text{Equation 3}).$$

The graphical representation of  $\sigma_{|\Delta Tm|}$  is shown in Fig. 8B.

We combined Equation 2 and Equation 3 into one formula, for SARS-CoV-2 infection risk ( $\sigma_{COVID19}$ ) probability assessment. Since  $\sigma_{CT}$  and  $\sigma_{|\Delta Tm|}$  are independent variables, we multiplied the two components:

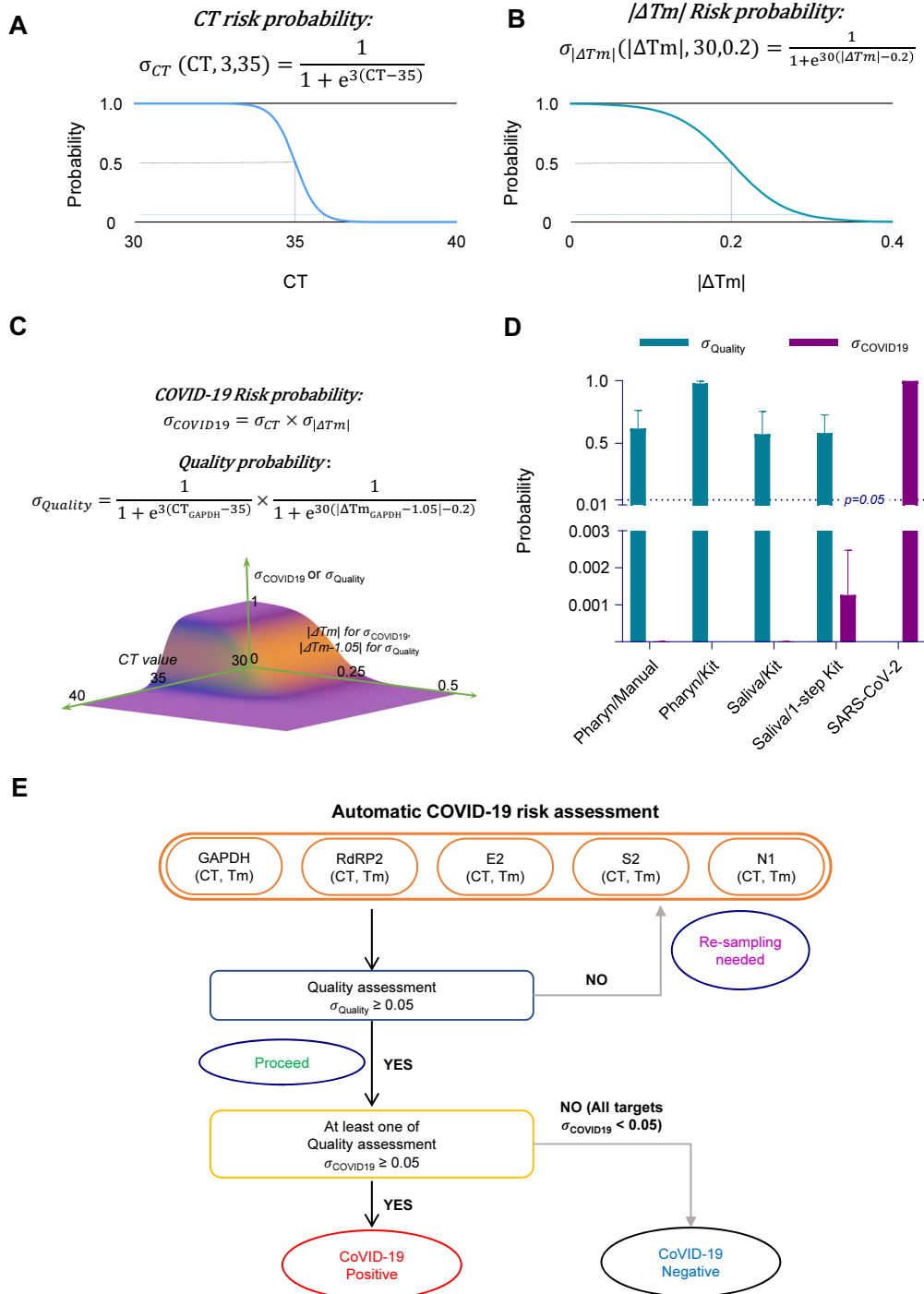
$$\sigma_{COVID19}(CT, |\Delta Tm|) = \sigma_{CT} \times \sigma_{|\Delta Tm|} = \frac{1}{1 + e^{3(CT-35)}} \times \frac{1}{1 + e^{30(|\Delta Tm| - 0.2)}} \quad (\text{Equation 4}).$$

After obtaining the probability value,  $\sigma_{COVID19}$  for each target gene for SARS-CoV-2, we selected the maximum value among the four genes as the final representative risk probability,  $\sigma_{COVID19}$  for each volunteer sample.

In case of the sampling quality assessment ( $\sigma_{Quality}$ ), only GAPDH was considered and we modified Equation 2 and 4. Since SARS-CoV-2 positive control RNA was originated from the Vero cell line, which shows a different Tm compared to human GAPDH, we corrected the  $\Delta Tm$  with a correction factor of 1.05:

$$\sigma_{Quality}(CT_{GAPDH}, |\Delta Tm|_{GAPDH}) = \sigma_{CT_{GAPDH}} \times \sigma_{|\Delta Tm|_{GAPDH}} = \frac{1}{1 + e^{3(CT_{GAPDH} - 35)}} \times \frac{1}{1 + e^{30(|\Delta Tm| - 1.05|_{GAPDH} - 0.5)}} \quad (\text{Equation 5}).$$

The correction factor of 1.05 for  $\Delta Tm$  of Vero cell GAPDH



**Fig. 8.** Risk probability and sampling quality assessment from CT value and ΔTm. (A) CT risk probability formula ( $\sigma_{CT}$ ) and probability graph for CT value. Dash line represent 0.5 probability at CT 35 (black), and 0.05 probability at CT 36 (blue). (B) Tm risk probability formula ( $\sigma_{|\Delta Tm|}$ ) and probability graph for |ΔTm|. Dash line represent 0.05 probability at 0.5 probability at 0.2 |ΔTm| (black), and 0.3 |ΔTm| (green). (C) SARS-CoV-2 risk probability formula ( $\sigma_{COVID19}$ ), quality probability formula ( $\sigma_{Quality}$ ) and probability graph for  $\sigma_{COVID19}$ , and  $\sigma_{Quality}$ . (D) Summary bar graph for  $\sigma_{Quality}$  (green) and  $\sigma_{COVID19}$  (purple) for Pharyn/Manual, Pharyn/Kit, Saliva/Kit, Saliva/1-step Kit, and SARS-CoV-2 positive controls.  $\sigma_{Quality}$  in SARS-CoV-2 is not included. Dotted line indicates p=0.05. Data are represented as the mean±S.E.M. (E) Flow chart for automatic SARS-CoV-2 risk assessment process.



**Table 3.** Sampling quality assessment from  $\sigma_{\text{Quality}}$  using internal positive control (GAPDH)

Quality assessment	Target/assessment	SARS-CoV-2	Volunteer A	Volunteer B	Volunteer C	Volunteer D	Volunteer E	Volunteer F	Volunteer G	NTC
Pharyn/Manual	GAPDH Assessment	0.998 Proceed	0.799 Proceed	0.039 Re-sampling	0.997 Proceed	0.822 Proceed	0.997 Proceed	0.334 Proceed	0.334 Proceed	0.000
Pharyn/Kit	GAPDH Assessment	0.998 Proceed	0.995 Proceed	0.995 Proceed	0.894 Proceed	0.996 Proceed	0.996 Proceed	0.993 Proceed	- -	0.000
Saliva/Kit	GAPDH Assessment	0.998 Proceed	0.995 Proceed	0.079 Proceed	0.060 Proceed	0.996 Proceed	0.989 Proceed	0.060 Proceed	0.839 Proceed	0.000
Saliva/1-step Kit	GAPDH Assessment	1.000 Proceed	0.749 Proceed	0.843 Proceed	0.821 Proceed	0.804 Proceed	0.759 Proceed	0.820 Proceed	0.048 Re-sampling	0.000

should be taken out in case of the positive control originated from the human samples. The graphical representation of  $\sigma_{\text{COVID19}}$  and  $\sigma_{\text{Quality}}$  is shown in Fig. 8C.

Based on those equations, we devised an automatic assessment algorithm, starting with the CT and Tm values for each volunteer sample (Fig. 8E). We generated an Excel worksheet (Supplementary Data 1) composed of input cells for CT and Tm values for each target gene of a volunteer sample. Once the CT and Tm values are entered, the embedded formulae automatically calculate  $\sigma_{\text{Quality}}$  to assess “Re-sampling needed” or “Proceed” and  $\sigma_{\text{COVID19}}$  to assess “COVID-19 Positive” or “COVID-19 Negative.”

To automatically assess the sampling quality,  $\sigma_{\text{Quality}}$  was calculated from the CT and Tm values of GAPDH qRT-PCR results from Volunteer A through G (Table 3). For the undetermined CT values (u.d.), each CT value was set to 40, which is the maximum cycle number in qRT-PCR. We found that the volunteer samples in Pharyn/Manual, Pharyn/Kit, Saliva/Kit, and Saliva/1-step Kit showed  $\sigma_{\text{Quality}} > 0.05$ , except 2 cases, indicating that the sampling quality was adequate to “proceed” for further analysis and that there is no need for re-sampling (Table 3). We asked for the re-sampling from volunteers who had showed  $\sigma_{\text{Quality}} < 0.05$  in assessment. Notably, the mean  $\sigma_{\text{Quality}}$  was well above 0.5 in all procedure groups (Fig. 8D).

After confirming the sampling quality, each qRT-PCR set was further processed for the risk assessment of COVID-19 by automatically calculating  $\sigma_{\text{CT}}$  (Table 4),  $\sigma_{|\Delta T_m|}$  (Table 5), and  $\sigma_{\text{COVID19}}$  (Table 6). We found that many samples showed  $\sigma_{\text{CT}} > 0.05$ : ten in Pharyn/Manual, eleven in Pharyn/Kit, six in Saliva/Kit, and fourteen in Saliva/1-step Kit (Table 4), likely caused by  $\text{CT} < 35$ . In contrast, most samples showed  $\sigma_{|\Delta T_m|} < 0.05$ , except for only one in Pharyn/Manual, one in Pharyn/Kit, two in Saliva/Kit, and one in Saliva/1-step Kit showed  $\sigma_{|\Delta T_m|} > 0.05$  (Table 5). Finally, none showed  $\sigma_{\text{COVID19}} > 0.05$ , indicating that all volunteers were “negative” for COVID-18 (Table 6), while the mean  $\sigma_{\text{COVID19}}$  was 0.9975 for

SARS-CoV-2 positive control (Table 6, Fig. 8D). Taken together, the automatic calculation of  $\sigma_{\text{Quality}}$ ,  $\sigma_{\text{CT}}$ ,  $\sigma_{|\Delta T_m|}$ , and  $\sigma_{\text{COVID19}}$ , based on CT and Tm values, allowed fast, unsupervised and unbiased assessment of sampling quality and risk of COVID-19.

#### **Application of automatic determination of sampling quality and risk assessment from COVID-19-positive and -negative subjects with Saliva/Kit**

As a proof-of-concept experiment, we tested Saliva/Kit from one SARS-CoV-2-positive and one SARS-CoV-2-negative subjects in the clinic (Fig. 9). We found that SARS-CoV-2 genes and GAPDH were detected in SARS-CoV-2-positive subject, although only GAPDH was detected in SARS-CoV-2-negative subject (Fig. 9A~E). SARS-CoV-2 positive subject showed 26.5 to 33.5 CT value for SARS-CoV-2 detection (Fig. 9F). Based on our estimation of the protocol efficiency in previous study [9], we estimated that 10~100 viral copies per each qRT-PCR reaction were present in the self-saliva sample. For automatic determination of sampling quality and risk assessment, we entered CT value and Tm value for each test in the Automatic COVID-19 risk assessment sheet in Supplementary Data 1 (Fig. 9F). We adjusted GAPDH CT correction value from 1.05 to 0.5 to correct for the different qRT-PCR machine what we used. This sheet automatically determined COVID-19-positive or -negative, which accurately corresponded to the clinical diagnosis. Therefore, we validated that our Saliva/Kit sampling method, SARS-CoV-2 detection protocol, and COVID-19 risk assessment algorithm work properly in clinical settings.

## **DISCUSSION**

In this study we have extended our previous studies on development of laboratory-safe detection protocol using pharyngeal swab sampling method and further optimization of primer sets for SARS-CoV-2 and internal positive control (GAPDH) to develop

**Table 4.** SARS-CoV-2 infection risk assessment from  $\sigma_{CT}$

Risk probability (CT)	Target	SARS-CoV-2	Volunteer A	Volunteer B	Volunteer C	Volunteer D	Volunteer E	Volunteer F	Volunteer G	NTC
Pharyngeal, Trizol-manual	RdRP2	1.000	0.648	0.595	0.109	0.475	0.350	0.000	0.000	0.000
	E2	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	S2	1.000	0.000	0.003	0.061	0.007	0.002	0.007	0.000	0.012
	N1	1.000	0.829	0.720	0.783	0.066	0.001	0.000	0.000	0.000
Pharyngeal, Trizol-Kit	RdRP2	1.000	0.457	0.312	0.337	0.763	0.999	0.029	-	0.350
	E2	1.000	0.571	0.498	0.000	0.015	0.000	0.000	-	0.053
	S2	1.000	0.001	0.000	0.713	0.068	0.875	0.000	-	0.000
	N1	1.000	0.000	0.001	0.179	0.046	0.000	0.000	-	0.008
Saliva, Trizol-Kit, 2-step qRT-PCR	RdRP2	1.000	0.588	1.000	0.004	0.028	0.118	0.036	0.418	0.059
	E2	1.000	0.000	0.179	0.000	0.000	0.000	0.000	0.001	0.000
	S2	1.000	0.000	0.000	0.000	0.000	0.165	0.000	0.000	0.034
Saliva, Trizol-Kit, 1-step qRT-PCR	N1	1.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000
	RdRP2	1.000	0.872	0.000	0.990	1.000	1.000	0.005	1.000	0.207
	E2	1.000	0.000	0.504	0.000	0.000	0.001	0.868	0.060	0.000
	S2	1.000	0.000	0.256	0.000	0.000	0.002	0.000	1.000	0.000
	N1	1.000	0.761	0.761	0.066	0.023	0.974	0.023	0.997	0.000

**Table 5.** SARS-CoV-2 infection risk assessment from  $\sigma_{|\Delta Tm|}$

Risk probability ( $ \Delta Tm $ )	Target	SARS-CoV-2	Volunteer A	Volunteer B	Volunteer C	Volunteer D	Volunteer E	Volunteer F	Volunteer G	NTC
Pharyngeal, Trizol-manual	RdRP2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	E2	0.998	0.000	0.705	0.000	0.000	0.000	0.000	0.000	0.000
	S2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N1	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Pharyngeal, Trizol-Kit	RdRP2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	-	0.000
	E2	0.998	0.000	0.000	0.039	0.000	0.000	0.000	-	0.000
	S2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	-	0.000
	N1	0.998	0.000	0.000	0.000	0.000	0.000	0.000	-	0.000
Saliva, Trizol-Kit, 2-step qRT-PCR	RdRP2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	E2	0.998	0.000	0.000	0.050	0.000	0.000	0.000	0.025	0.863
	S2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Saliva, Trizol-Kit, 1-step qRT-PCR	N1	0.998	0.000	0.000	0.000	0.000	0.073	0.000	0.000	0.000
	RdRP2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	E2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	S2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.804
	N1	0.998	0.000	0.000	0.000	0.058	0.000	0.000	0.000	0.000

the ultimate detection protocol using saliva-based sampling with automatic assessment of sample quality and risk for COVID-19. The advantages of our protocol over other previously reported protocols include; 1) the saliva-based self-sampling allows complete spatial separation and medical staff or experimenter, eliminating a possibility of cross-infection, 2) the low-cost and fast (less than 2 hour) protocol provides a flexibility of use in the same-day events, such as attendance in face-to-face conference meetings or admission to sports/performance, and 3) the automatic assessment of sampling quality and risk for COVID-19 allows unbiased, unsupervised, objective, and false-positive-free testing, suitable for

mass-scale, high-throughput testing facilities.

In some countries, such as the USA and the United Kingdom, the saliva-sampling has been recently approved as a national verified procedure [7, 23]. Saliva-based sampling provides an important advantage of complete spatial separation of medical staff and patients. The convenient feature of saliva-sampling procedure also enables shipping-based collection of samples, provided that the subject inactivates the viral particles of SARS-CoV-2 with Proteinase K at the site of collection before sending out via delivery services. The rest of the procedures of RNA extraction and qRT-PCR can be carried out at any Biosafety Level II Grade molecular

**Table 6.** SARS-CoV-2 infection risk assessment from  $\sigma_{\text{COVID19}}$

Risk probability	Target/assessment	SARS-CoV-2	Volunteer A	Volunteer B	Volunteer C	Volunteer D	Volunteer E	Volunteer F	Volunteer G	NTC
Pharyngeal, Trizol-manual	RdRP2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	E2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	S2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N1	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Assessment	Positive	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative
Pharyngeal, Trizol-Kit	RdRP2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	E2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	S2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N1	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Assessment	Positive	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative
Saliva, Trizol-Kit, 2-step qRT-PCR	RdRP2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	E2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	S2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N1	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Assessment	Positive	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative
Saliva, Trizol-Kit, 1-step qRT-PCR	RdRP2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	E2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	S2	0.998	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N1	0.998	0.000	0.034	0.000	0.001	0.000	0.000	0.000	0.000
	Assessment	Positive	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative

biology laboratories. We have used both Proteinase K and Trizol to inactivate the viral activity to ensure the complete inactivation. However, there are numerous reports using Protein K alone to inactivate SARS-CoV-2 [20, 24]. If Proteinase K alone is sufficient to inactivate the viral activity, then it should be used instead of Trizol because Trizol is known for its toxicity [25]. Future experiments are needed to confirm this key feature of Proteinase K to inactivate SARS-CoV-2.

Another advantage of using Proteinase K alone is that it allows detection of SARS-CoV-2 without RNA extraction [26, 27], which should further reduce the total procedure time to less than one hour. Although we have not tested this possibility, it needs future investigation. Taken together, our study suggests that self-saliva sampling is a reliable method to replace pharyngeal swab for SARS-CoV-2 detection protocol.

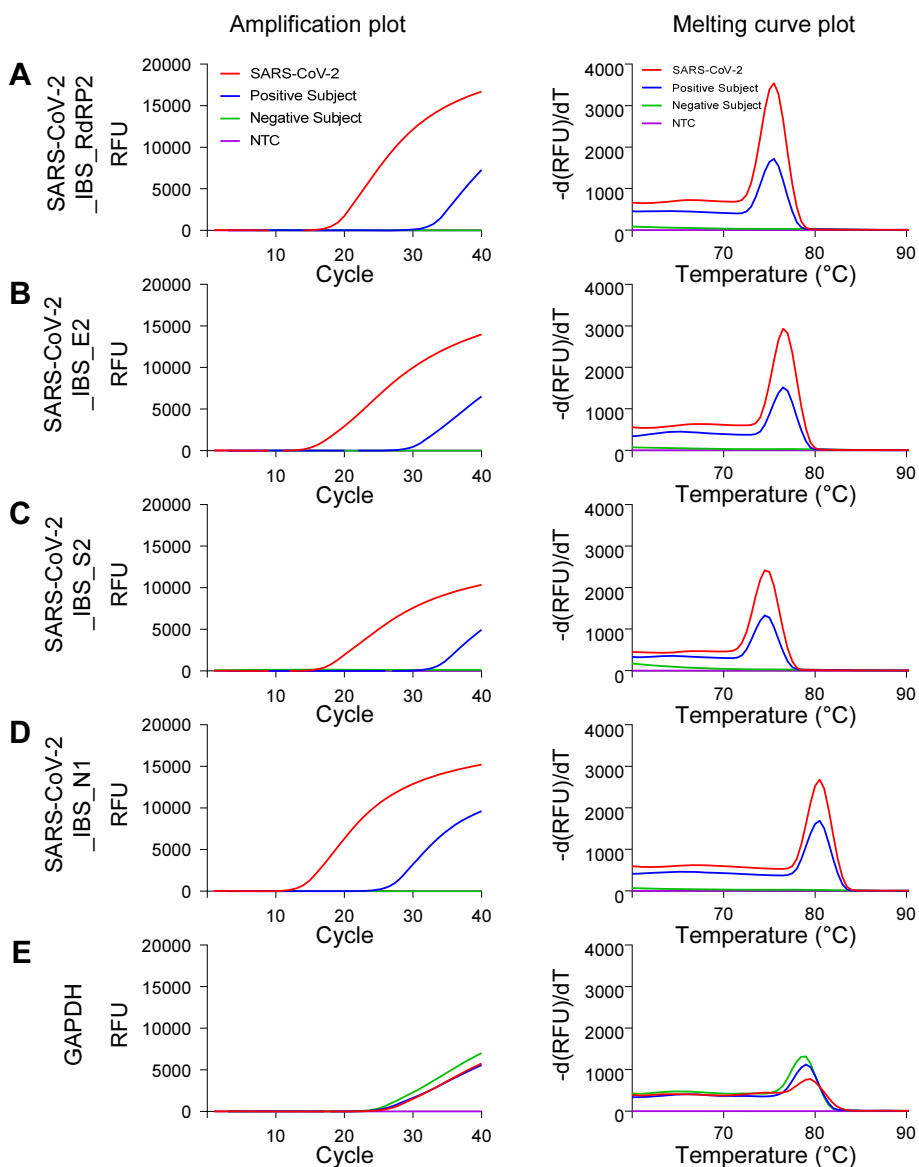
During the assessment of sampling quality, we used the correctional factor of 1.05 for the positive control obtained from the Vero cell line originated from the African green monkey [9, 18]. This correctional factor was necessary because the Tm value for GAPDH-targeted qRT-PCR showed an average deviation of 1.05°C between SARS-CoV-2 positive control and volunteer samples (Fig. 3F, 4F, 5F, and 6F). It is critical to remember that in case of positive controls obtained directly from human patients, this correctional factor should be taken out from the formula.

In addition to our focus on the sampling method, we have focused on developing the ways to minimize the occurrence of a

false-positive. To do this, we considered both CT and Tm values of the qRT-PCR results. In most of the currently available commercial detection kits, only the threshold of CT value is used as a basis for “positive” versus “negative”. However, we strongly suggest that such protocol imposes a danger of false-positive due to primer-dimerization-mediated amplification, which we have demonstrated that it has a different Tm value from the positive control [18]. To prevent this possibility, we developed the formulae for automatic risk assessment, which reflects both CT and Tm values. This formula utilizes the popular logistic classification model, which is a regression technique that can be performed when the dependent variable is to be bisected. We did not choose a two-dimensional logistic classification model because CT and Tm are independent variables. Instead, we multiplied the two probability functions for CT and Tm to obtain the combined risk probability for each targeted gene (Equation 4). The results of the automatic calculation and evaluation show that all of the volunteer samples have expected “negative” assessment. These promising results show the high potential and usefulness our model and formulae in large-scale testing.

We conducted this study only with one COVID-19 positive subject. Further research with more COVID-19 patients at a large-scale may provide more clinically optimized protocol.

### Saliva/Kit, Human Subjects



ID	Sampling Quality Assessment <sup>(3)</sup>					
	GAPDH CT	GAPDH Tm	Sampling Quality assessment	$\sigma_{CT\_GAPDH}$	$\sigma_{Tm\_GAPDH}$	$\sigma_{Quality}$
Positive Ctrl <sup>(2)</sup>	27.57	79.50	Proceed	1.000	1.000	1.000
Positive subject	26.89	79.00	Proceed	1.000	0.998	0.998
Negative subject	25.83	79.00	Proceed	1.000	0.998	0.998

ID	COVID19 Risk Assessment <sup>(4)</sup>																					
	RdRP2_CT	E2_CT	S2_CT	N1_CT	RdRP2_Tm	E2_Tm	S2_Tm	N1_Tm	COVID-19 risk assessment	$\sigma_{CT\_RdRP2}$	$\sigma_{CT\_E2}$	$\sigma_{CT\_S2}$	$\sigma_{CT\_N1}$	$\sigma_{Tm\_RdRP2}$	$\sigma_{Tm\_E2}$	$\sigma_{Tm\_S2}$	$\sigma_{Tm\_N1}$	$\sigma_{Consent\_RdRP2}$	$\sigma_{Consent\_E2}$	$\sigma_{Consent\_S2}$	$\sigma_{Consent\_N1}$	
Positive Ctrl <sup>(2)</sup>	18.315	15.820	17.563	13.734	75.500	76.500	74.500	80.500	Positive	1.000	1.000	1.000	1.000	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998
Positive subject	32.544	30.359	33.558	28.521	75.500	74.500	80.500	Positive	0.999	1.000	0.997	1.000	0.998	0.998	0.998	0.998	0.997	0.998	0.998	0.985	0.998	0.998
Negative subject	40.000	40.000	40.000	40.000	52.500	0.000	52.500	0.000	Negative	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

**Fig. 9.** Amplification and melting curve plot from Human Subject's Saliva/Kit samples. Amplification plot (left panel) and melting curve plot (right panel) from qRT-PCR of human subjects saliva/Kit samples. To determine whether volunteers are infected to SARS-CoV-2, qRT-PCR was performed by using the previously reported SARS-CoV-2 primer sets; (A) SARS-CoV-2\_IBS\_RdRP2, (B) SARS-CoV-2\_IBS\_E2, (C) SARS-CoV-2\_IBS\_S2, (D) SARS-CoV-2\_IBS\_N1. (E) GAPDH primer set was used for IPC detection. Each color code represent individual subject result and SARS-CoV-2 amplification was represent as red color. (F) Sampling quality assessment and COVID-19 risk assessment with Automatic COVID-19 risk assessment in Supplementary Data 1.

## ACKNOWLEDGEMENTS

This work was supported by financial support by Institute for Basic Science, Center for Cognition and Sociality (IBS-R001-D2) to C.J.L.

## REFERENCES

- Jiang F, Deng L, Zhang L, Cai Y, Cheung CW, Xia Z (2020) Review of the clinical characteristics of coronavirus disease 2019 (COVID-19). *J Gen Intern Med* 35:1545-1549.
- Gao Z, Xu Y, Sun C, Wang X, Guo Y, Qiu S, Ma K (2020) A systematic review of asymptomatic infections with COVID-19. *J Microbiol Immunol Infect*. doi: 10.1016/j.jmii.2020.05.001.
- Ou X, Liu Y, Lei X, Li P, Mi D, Ren L, Guo L, Guo R, Chen T, Hu J, Xiang Z, Mu Z, Chen X, Chen J, Hu K, Jin Q, Wang J, Qian Z (2020) Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat Commun* 11:1620.
- World Health Organization (2020) WHO coronavirus disease (COVID-19) dashboard [Internet]. World Health Organization, Geneva. Available from: [https://covid19.who.int/?gclid=CjwKCAjwID8BRAFEiwAnUoK1TvTluMRgHPrdYhLgrLexxW2npVTCJ4nfhfTPDuvf-L9iESDW6joqRoCW1AQAvD\\_BwE](https://covid19.who.int/?gclid=CjwKCAjwID8BRAFEiwAnUoK1TvTluMRgHPrdYhLgrLexxW2npVTCJ4nfhfTPDuvf-L9iESDW6joqRoCW1AQAvD_BwE).
- Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, Curtis HJ, Mehrkar A, Evans D, Inglesby P, Cockburn J, McDonald HI, MacKenna B, Tomlinson L, Douglas IJ, Rentsch CT, Mathur R, Wong AYS, Grieve R, Harrison D, Forbes H, Schultze A, Croker R, Parry J, Hester F, Harper S, Perera R, Evans SJW, Smeeth L, Goldacre B (2020) Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 584:430-436.
- Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, Bleicker T, Brünink S, Schneider J, Schmidt ML, Mulders DG, Haagmans BL, van der Veer B, van den Brink S, Wijsman L, Goderski G, Romette JL, Ellis J, Zambon M, Peiris M, Goossens H, Reusken C, Koopmans MP, Drosten C (2020) Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill* 25:2000045.
- World Health Organization (2020) Diagnostic testing for SARS-CoV-2 [Internet]. World Health Organization, Geneva. Available from: <https://www.who.int/publications/i/item/diagnostic-testing-for-sars-cov-2>.
- Marty FM, Chen K, Verrill KA (2020) How to obtain a nasopharyngeal swab specimen. *N Engl J Med* 382:e76.
- Won J, Lee S, Park M, Kim TY, Park MG, Choi BY, Kim D, Chang H, Kim VN, Lee CJ (2020) Development of a laboratory-safe and low-cost detection protocol for SARS-CoV-2 of the coronavirus disease 2019 (COVID-19). *Exp Neurobiol* 29:107-119.
- Malecki M, Lüsebrink J, Teves S, Wendel AF (2020) Pharynx gargle samples are suitable for SARS-CoV-2 diagnostic use and save personal protective equipment and swabs. *Infect Control Hosp Epidemiol*. doi: 10.1017/ice.2020.229.
- Nguyen LH, Drew DA, Graham MS, Joshi AD, Guo CG, Ma W, Mehta RS, Warner ET, Sikavi DR, Lo CH, Kwon S, Song M, Mucci LA, Stampfer MJ, Willett WC, Eliassen AH, Hart JE, Chavarro JE, Rich-Edwards JW, Davies R, Capdevila J, Lee KA, Lochlainn MN, Varsavsky T, Sudre CH, Cardoso MJ, Wolf J, Spector TD, Ourselin S, Steves CJ, Chan AT; COOronavirus Pandemic Epidemiology Consortium (2020) Risk of COVID-19 among front-line health-care workers and the general community: a prospective cohort study. *Lancet Public Health* 5:e475-e483.
- 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: a cross-sectional study. *Clin Microbiol Infect*. doi: S1198-743X(20)30278-0.
- Wyllie AL, Fournier J, Casanovas-Massana A, Campbell M, Tokuyama M, Vijayakumar P, Geng B, Muenker MC, Moore AJ, Vogels CBF, Petrone ME, Ott IM, Lu P, Venkataraman A, Lu-Culligan A, Klein J, Earnest R, Simonov M, Datta R, Handoko R, Naushad N, Sewanan LR, Valdez J, White EB, Lapidus S, Kalinich CC, Jiang X, Kim DJ, Kudo E, Linehan M, Mao T, Moriyama M, Oh JE, Park A, Silva J, Song E, Takahashi T, Taura M, Weizman OE, Wong P, Yang Y, Bermejo S, Odio C, Omer SB, Dela Cruz CS, Farhadian S, Martinello RA, Iwasaki A, Grubaugh ND, Ko AI (2020) Saliva is more sensitive for SARS-CoV-2 detection in COVID-19 patients than nasopharyngeal swabs. *medRxiv*. doi: 10.1101/2020.04.16.20067835.
- Yoon JG, Yoon J, Song JY, Yoon SY, Lim CS, Seong H, Noh JY, Cheong HJ, Kim WJ (2020) Clinical significance of a high SARS-CoV-2 viral load in the saliva. *J Korean Med* 35:e195.
- Iwasaki S, Fujisawa S, Nakakubo S, Kamada K, Yamashita Y, Fukumoto T, Sato K, Oguri S, Taki K, Senjo H, Sugita J, Haya-saka K, Konno S, Nishida M, Teshima T (2020) Comparison of SARS-CoV-2 detection in nasopharyngeal swab and saliva. *J Infect* 81:e145-e147.
- Sakanashi D, Asai N, Nakamura A, Miyazaki N, Kawamoto Y, Ohno T, Yamada A, Koita I, Suematsu H, Hagihara M, Shiota A, Kurumiya A, Sakata M, Kato S, Muramatsu Y, Koizumi Y, Kishino T, Ohashi W, Yamagishi Y, Mikamo H (2021)



- Comparative evaluation of nasopharyngeal swab and saliva specimens for the molecular detection of SARS-CoV-2 RNA in Japanese patients with COVID-19. *J Infect Chemother* 27:126-129.
17. Kim SE, Lee JY, Lee A, Kim S, Park KH, Jung SI, Kang SJ, Oh TH, Kim UJ, Lee SY, Kee SJ, Jang HC (2020) Viral load kinetics of SARS-CoV-2 infection in saliva in Korean patients: a prospective multi-center comparative study. *J Korean Med Sci* 35:e287.
  18. Park M, Won J, Choi BY, Lee CJ (2020) Optimization of primer sets and detection protocols for SARS-CoV-2 of coronavirus disease 2019 (COVID-19) using PCR and real-time PCR. *Exp Mol Med* 52:963-977.
  19. Kim D, Lee JY, Yang JS, Kim JW, Kim VN, Chang H (2020) The architecture of SARS-CoV-2 transcriptome. *Cell* 181:914-921.e10.
  20. Sentmanat M, Kouranova E, Cui X (2020) One-step RNA extraction for RT-qPCR detection of 2019-nCoV. *bioRxiv*. doi: 10.1101/2020.04.02.022384.
  21. Chu DKW, Pan Y, Cheng SMS, Hui KPY, Krishnan P, Liu Y, Ng DYM, Wan CKC, Yang P, Wang Q, Peiris M, Poon LLM (2020) Molecular diagnosis of a novel coronavirus (2019-nCoV) causing an outbreak of pneumonia. *Clin Chem* 66:549-555.
  22. Tolles J, Meurer WJ (2016) Logistic regression: relating patient characteristics to outcomes. *JAMA* 316:533-534.
  23. Department of Health and Social Care (2020) National technical validation process for manufacturers of SARS-CoV-2 (COVID-19) tests [Internet]. Department of Health and Social Care, London. Available from: <https://www.gov.uk/government/publications/assessment-and-procurement-of-coronavirus-covid-19-tests/coronavirus-covid-19-serology-and-viral-detection-testing-uk-procurement-overview>.
  24. Chu AW, Chan WM, Ip JD, Yip CC, Chan JF, Yuen KY, To KK (2020) Evaluation of simple nucleic acid extraction methods for the detection of SARS-CoV-2 in nasopharyngeal and saliva specimens during global shortage of extraction kits. *J Clin Virol* 129:104519.
  25. Blow JA, Dohm DJ, Negley DL, Mores CN (2004) Virus inactivation by nucleic acid extraction reagents. *J Virol Methods* 119:195-198.
  26. Lübke N, Senff T, Scherger S, Hauka S, Andrée M, Adams O, Timm J, Walker A (2020) Extraction-free SARS-CoV-2 detection by rapid RT-qPCR universal for all primary respiratory materials. *J Clin Virol* 130:104579.
  27. Lalli MA, Langmade SJ, Chen X, Fronick CC, Sawyer CS, Burcea LC, Wilkinson MN, Fulton RS, Heinz M, Buchser WJ, Head RD, Mitra RD, Milbrandt J (2020) Rapid and extraction-free detection of SARS-CoV-2 from saliva by colorimetric reverse-transcription loop-mediated isothermal amplification. *Clin Chem*. doi: 10.1093/clinchem/hvaa267.