MATERIALS AND METHODS

Clinical study design

Study design

This was a diagnostic accuracy study of subjects consecutively recruited at a single institution (ASST dei Sette Laghi, Varese, Italy) when they underwent the nasopharyngeal swab procedure for the diagnosis of COVID-19. The recruitment period spanned from April 15th to May 3rd, 2020.

This study was approved by the local Ethical Committee (n° 68/2020) and registered on ClinicalTrials.gov (NCT04357327). This study adhered to the STARD-15 Guidelines.

Study recruitment

The subjects who were recruited were individuals who underwent the nasopharyngeal swab procedure for the diagnosis of SARS-CoV-2 infection. Inclusion criteria were COVID-19 inpatients who were scheduled for the nasopharyngeal swab, individuals with symptoms suggestive of the disease (i.e., cough, fever and/or dyspnea) or healthcare workers and administrative staff. A consecutive recruitment approach was adopted in three independent medical areas in our hospital: the COVID-19 wards (with the exclusion of patients admitted to the Intensive Care Unit), the Emergency Room and the area for healthcare workers.

This recruitment strategy was adopted in order to avoid spectrum bias and obtain samples from a sufficient number of individuals affected by the disease (*i.e.*, COVID-19 inpatients), at high risk of disease (*i.e.,* patients admitted to the Emergency Room) and low risk of disease (healthcare workers and administrative staff), respectively. The subjects were recruited for the RST when undergoing their scheduled nasopharyngeal swab, maintaining thus the consecutive recruitment and avoiding selection bias.
A signed informed consent was conditional on the participation to the study. There were no exclusion criteria.

Sample and data collection

Each recruited subject provided a salivary sample of about 1 mL at the same time as their nasopharyngeal swab in the morning. The collection method was the same for every patient, *i.e.*, the drooling technique. After this procedure, a nasopharyngeal swab was performed following the established clinical practice.

An anamnestic form was filled in by the clinician and this included data about the subject's age, sex, or the presence of symptoms suggestive for COVID-19, *i.e.*, fever, cough and dyspnea. In the case of COVID-19 inpatients, the severity of the disease was classified according to 3 degrees: mild, moderate, and severe. Symptomatic patients were considered affected by a mild disease when they could be adequately treated at home, by a moderate disease when they were hospitalized but able to breathe without respiratory support, by a severe disease when they required oxygen therapy.

Laboratory procedures

Rapid Salivary Test preparation procedure

The Lateral Flow Assay (LFA) was designed and performed using the Universal Lateral Flow Assay kit (Abcam, cat# ab270537), following the manufacturer's instructions. A customized sandwich LFA was designed, by combining Ulfa-Tag and GOLD conjugation technologies with an immunochromatography test performed on universal nitrocellulose LFA strips (*Fig.1* in the maintext). The rabbit polyclonal SARS-CoV-2 Spike antibody (αSpike; ProSci, cat# 3525) was used to detect the presence of the virus in the saliva samples, as both "capture" and "detection" antibody.

Briefly, antibody conjugation reactions were performed first. The capture αSpike antibody (αSpike-Ulfa) was conjugated with the Ulfa-tag using the Lightning-Link® Ulfa-Tag Conjugation kit, following kit instructions. Similarly, the detection αSpike antibody (αSpike-GOLD) was conjugated with 40 nm gold particles using the InnovaCoat® GOLD (20 OD) Conjugation kit. Then, the αSpike-Ulfa antibody was diluted in 1X Universal Running Buffer (URB) added with 0.1% bovine serum albumin and protease inhibitors (termed here "LFA RB") to the final concentration of 100 µg/ml. The αSpike-GOLD antibody was also diluted in LFA RB to the final concentration of 6 OD (corresponding to 6 µg/ml). The 40 nm InnovaCoat® GOLD-Biotin standard (GOLD-biotin), which represents the internal validity control of the test, was also diluted in LFA RB to the final concentration of 1 OD. Eventually, the saliva samples were also diluted (80 µl saliva in 60 µl LFA RB).

All diluted reagents were then mixed as follows: 5 µl diluted αSpike-Ulfa + 5 µl diluted αSpike-GOLD + 5 µl diluted GOLD-biotin + 75 µl diluted saliva. The mixture was incubated for 5 minutes at room temperature and then 80 µL were loaded on the sample pad. When the saliva appeared excessively viscous, 40 µL of LFA RB were added on the sample pad to facilitate the run.

At the end of the run, the presence of both "Control-line" (C-line) and "Test-line" (T-line) was evaluated. After 10 minutes, three independent observers (LA, AB and AG) attributed a number according to a scoring card provided with the commercial kit. The test was considered as positive when at least 2 of the 3 observers stated a value >= 1.
The procedure followed the same steps in all cases. If the control line failed to appear, the event was recorded. In this case, the test was repeated again if possible. Otherwise, the test was considered failed and the patient excluded from the final analysis.

Nasopharyngeal swab analysis

The sample was subjected to viral thermal inactivation for 1 minute at 90 °C. RNA extraction from the nasopharyngeal swab was performed with the Abbott mSample Preparation System (Promega corporation) and an automated extraction system (Extraction m2000SP, Abbott Molecular). The extracted RNA was amplified with GeneFinder™ COVID-19 Plus RealAmp PCR kit (ELITechGroup), a one-step rRT-PCR system targeting SARS-CoV-2 RdRp, E, and N genes.

All recruited subjects underwent the nasopharyngeal swab, thus Partial verification bias was absent. The salivary collection and nasopharyngeal swab procedures were performed at the same moment to avoid disease progression bias. Being the RST and the nasopharyngeal swab analyzed by independent blinded clinicians, information bias was avoided when interpreting the rRT-PCR results of the swabs. The index test was completely independent of the reference test, so to avoid the Incorporation bias. Inconclusive results were recorded.

Salivary rRT-PCR

The salivary samples were subjected to virus inactivation for 1 minute at 90°C, and then the molecular assays were performed. One hundred and forty µl of each salivary sample was subjected to RNA extraction by QIAmp Viral RNA mini kit (Qiagen) and eluted in 60 µl. One step rRT-PCR was performed using Luna® Universal qPCR Master Mix (New England BioLab) from 5 µl of extracted RNA. Forward (5'-ACCTTCCCAGGTAACAAACCA-3') and reverse (5'-TTACCTTTCGGTCACACCCG-3') primers targeting the 5'UTR region of SARS-CoV-2 were used. All samples were run in replicates and compared with progressive dilutions of a synthetic SARS-CoV-2 RNA control (Twist Bioscence) on QuantStudio 5 Real-Time PCR System (ThermoFisher Scientific). The quantification of the viral load was recorded as the number of viral copies/µL. Inconclusive results were recorded.

This molecular analysis was blinded as regards the RST and the nasopharyngeal swab. Differential verification bias was avoided because all included patients underwent both the index and reference tests.

Statistical analysis

We had neither previous information on the accuracy of the RST procedure, nor on the prevalence of positive nasopharyngeal swabs (reference standard) in the study population, for standard sample size computation. We considered that a sample size of 110 with a prevalence of 30% of positive nasopharyngeal swabs would suffice to estimate a sensitivity of around 0.75 with a precision of 0.15 for the 95% confidence interval and a power of 0.8. The final sample size was then fixed at 120 anticipating 10 % of technical RST procedure failures.
The demographic and clinical features of the eligible participants were summarized by using mean and standard proportion, or absolute and relative frequencies, for continuous and discrete variables, respectively. For descriptive purposes only, the same statistical analyses were reported also in positive and negative nasopharyngeal swab subjects.

In case of technical failure of the RST without the possibility to repeat the procedure, the participants were excluded from further analyses. In the other subjects, the sensitivity and specificity for the RST (index test) when compared to the nasopharyngeal swab (reference test) were estimated, and their 95% confidence interval was calculated from the exact binomial distribution. These analyses were also replicated according to the setting of recruitment and on the basis of the presence of COVID-19 symptoms when performing the swab. We characterized true positives (TP), false positives (FP), as well as false negatives (FN) and true negatives (TN), in terms of prevalence of positive salivary rRT-PCR results and quantitative viral load distribution (number of copies/µL). We tested the null association of no difference in viral load distribution across TP, FP, FN and TN using a Kruskal-Wallis test.

In addition, we assessed the sensitivity and specificity of the RST when compared to salivary rRT-PCR. Finally, we reported Cohen's kappa and 95% confidence interval as a measure of the agreement across observers in the definition of a positive (value ≥ 1) vs negative (value=0) RST sample.

RESULTS

Fig. S1: Study flow diagram (STARD-15).
One hundred and twenty-five subjects were found to be potentially eligible. One of these subjects was unable to sign the informed consent form, while 2 COVID-19 patients were unable to provide the salivary sample. A total number of 122 patients were recruited in this study.

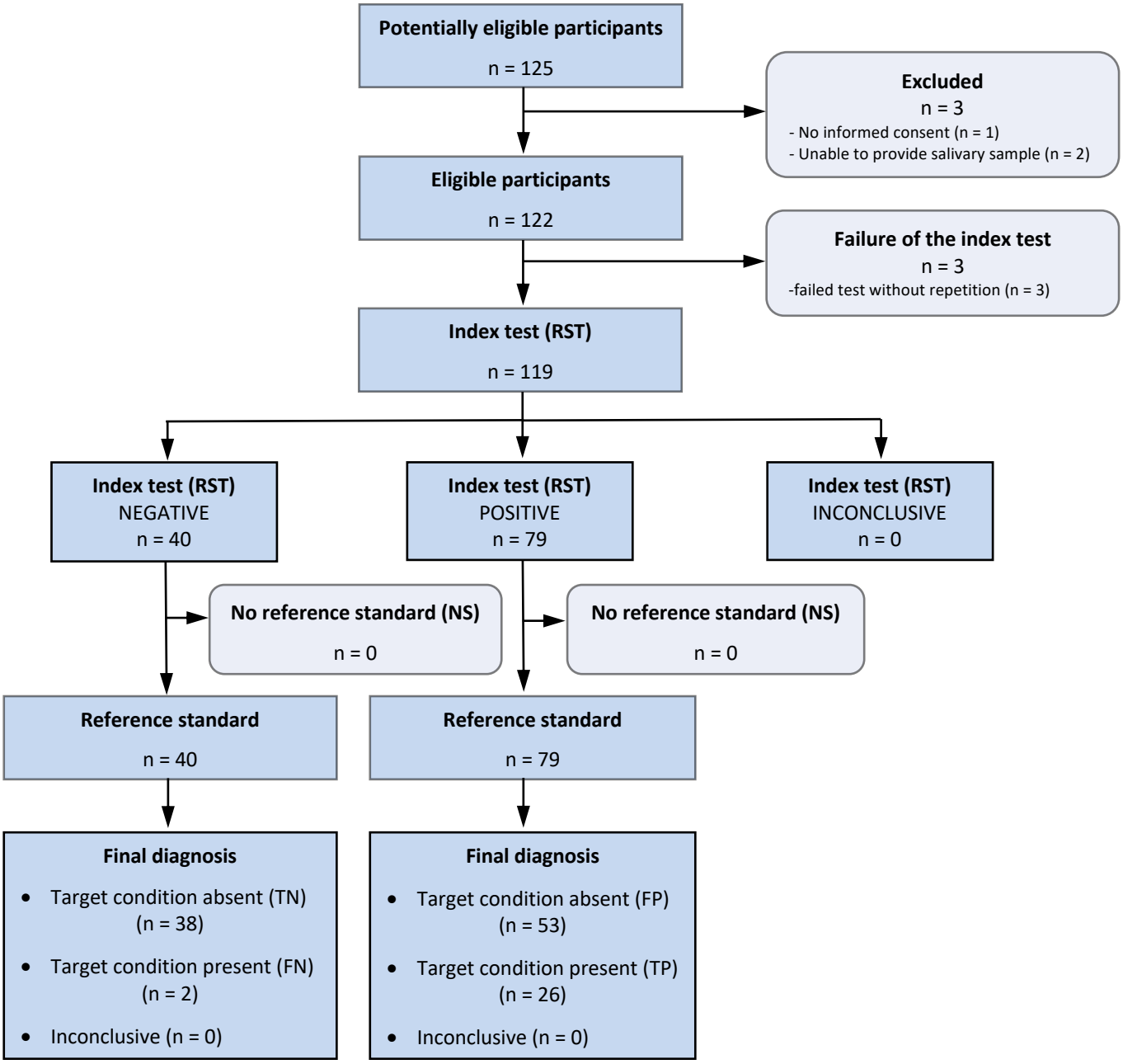

Table S1: Demographic and clinical characteristics of the study sample, by positivity to the nasopharyngeal swab

Variable	All subjects	Nasopharyngeal swab	
		Positive	Negative
N	122	29	93
Age, mean (SD)	53.5 (19.8)	68.4 (18.5)	48.8 (17.9)
Men, n (%)	40 (32.8)	12 (41.4)	28 (30.1)
Women, n (%)	82 (67.2)	17 (58.6)	65 (69.9)
Setting of the nasopharyngeal swab procedure, n (%)			
COVID-19 hospitalized patients	38 (31.2)	25 (86.2)	13 (14.0)
ER patients	21 (17.2)	2 (6.9)	19 (20.4)
Healthcare workers	63 (51.6)	2 (6.9)	61 (65.6)
COVID-19 symptoms at the time of nasopharyngeal swab procedure, n (%)	42 (34.4)	24 (82.8)	18 (19.4)
COVID-19 diagnosis, n (%)	35 (28.7)	29 (100)	6 (6.5)
COVID-19 severity at the time of nasopharyngeal swab procedure, n (%)*			
Asymptomatic	5 (14.3)	5 (17.2)	0 (0.0)
Mild disease	12 (34.3)	11 (37.9)	1 (16.7)
Moderate disease	12 (34.3)	8 (27.6)	4 (66.7)
Severe disease	5 (14.3)	5 (17.2)	0 (0.0)

*: Among the n=35 subjects with COVID-19 diagnosis. Severity not available for 1 subject (with negative nasopharyngeal swab)
SD= Standard Deviation

Fig. S2: ***Discrepancies between the salivary rRT-PCR and the nasopharyngeal swab.***

a) Prevalence of positive salivary rRT-PCR analyses in True Positives (TP), False Positives (FP), False Negatives (FN) and True Negatives (TN) who underwent the Rapid Salivary Test in comparison with the nasopharyngeal swab. The majority of FP subjects were positive also with salivary rRT-PCR, confirming that the virus was really present in their saliva.

b) The Box Plot highlights that there were not any differences between TP, FP and TN subjects regarding the viral load detected with salivary rRT-PCR. This result could be explained by the fact that a non-negligible percentage of the individuals classified as FP had the virus in their saliva and had a negative nasopharyngeal swab

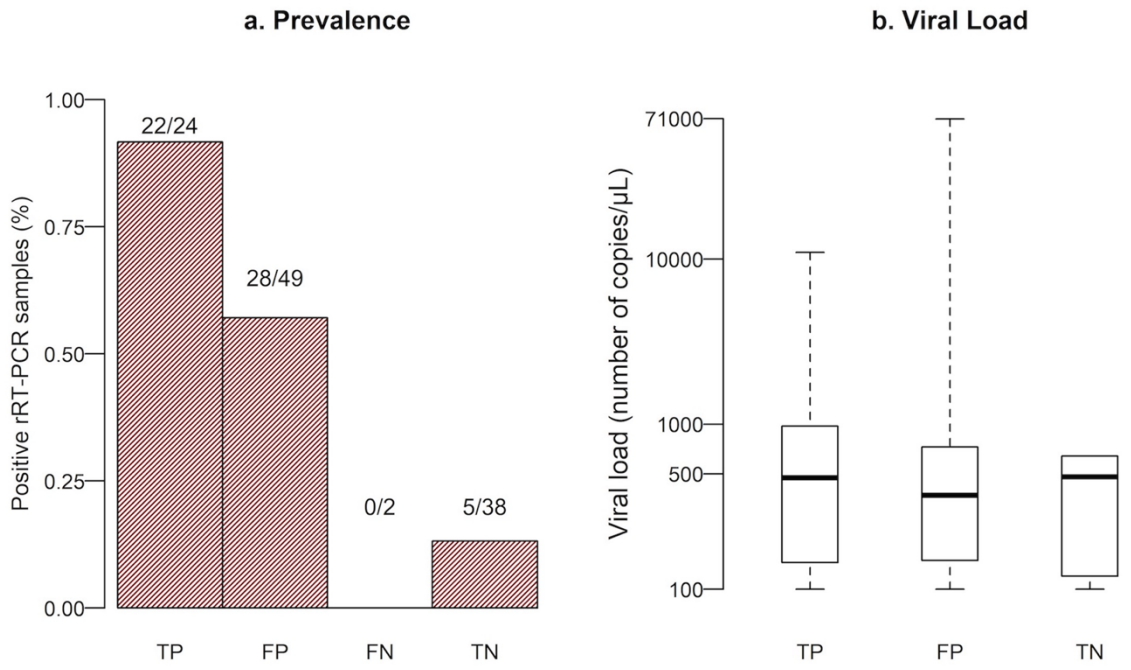

Fig. S3: Amplicons (250 bp) obtained after PCR with primers targeting 5'UTR region. a) Agarose gel electrophoresis; b) Capillary electrophoresis; c) Sanger method sequencing. BLAST sequences show 100% identity with SARS-CoV-2 5'UTR region.

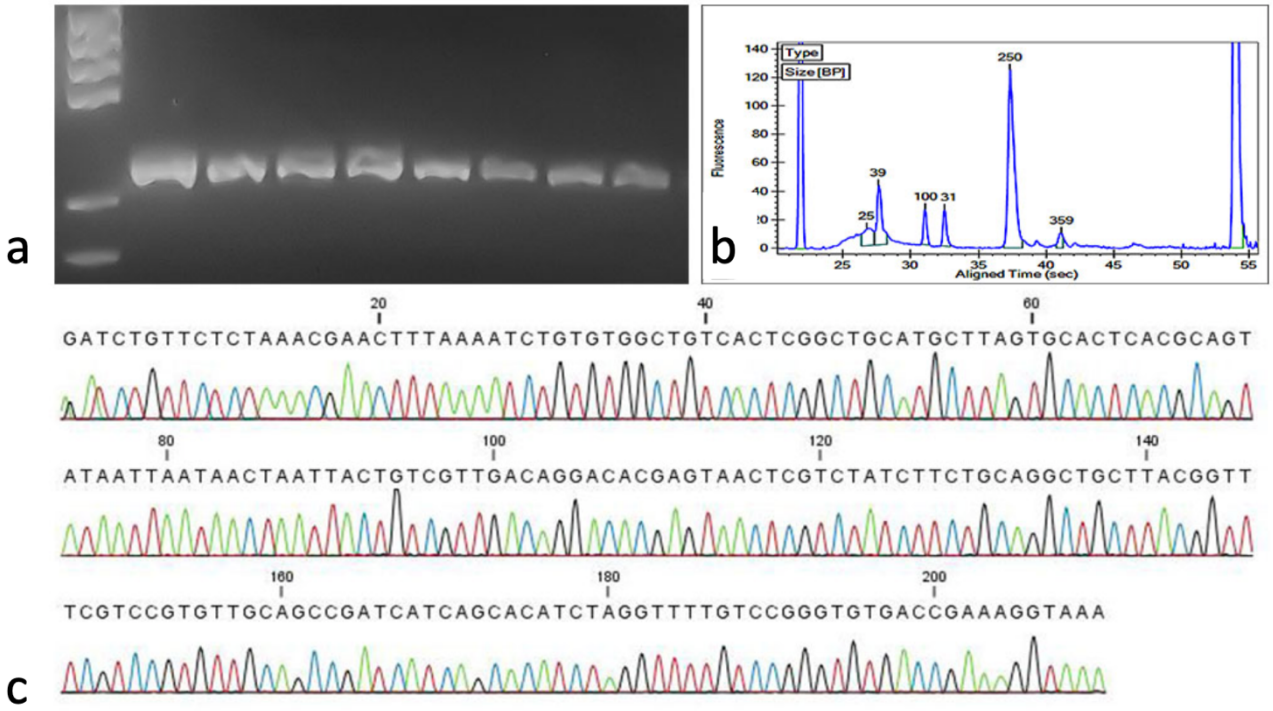

Table S2: Cross-controls were done in order to confirm the obtained results. In particular, five positive and five negative samples both on saliva and nasopharyngeal swab, together with 11 samples with discordant results were chosen. For these samples the salivary specimens were tested with the commercial kit and the nasopharyngeal swabs with the home-made one. In case of discordant results sequencing of 5'UTR region was done.

ID	RST	NPS (diagnostic kit)	NPS (5'UTR)	Salivary rRT-PCR (diagnostic kit)	Salivary rRT-PCR (5'UTR)	Sequence (5'UTR)
12	pos	pos	pos	pos	pos	
8	pos	pos	pos	neg	pos	SARS-CoV-2
4	pos	pos	pos	weak pos	pos	
9	pos	pos	pos	weak pos	weak pos	
13	pos	pos	pos	weak pos	pos	
3	neg	neg	neg	neg	neg	
26	neg	neg	neg	neg	neg	
88	neg	neg	neg	neg	neg	
6	neg	neg	neg	neg	neg	
67	pos	neg	neg	neg	neg	
75	pos	neg	neg	neg	neg	
21	pos	neg	pos	neg	neg	SARS-CoV-2
15	pos	neg	neg	pos	neg	
2	pos	neg	neg	neg	pos	SARS-CoV-2
51	pos	neg	neg	neg	weak pos	SARS-CoV-2
39	pos	neg	neg	neg	pos	SARS-CoV-2
10	neg	neg	neg	neg	weak pos	SARS-CoV-2
16	neg	neg	neg	neg	weak pos	SARS-CoV-2
68	pos	neg	neg	neg	pos	SARS-CoV-2
96	pos	neg	neg	neg	pos	SARS-CoV-2
72	pos	neg	neg	neg	pos	SARS-CoV-2
98	pos	neg	neg	neg	weak pos	SARS-CoV-2
97	pos	neg	deb pos	neg	weak pos	SARS-CoV-2
87	pos	neg	neg	neg	weak pos	SARS-CoV-2

Table S3a: Prevalence of positive rRT-PCR in the salivary sample in true positives, false positives, false negatives and true negatives, and distribution of the viral load.

	N	Positive rRT-PCR on salivary sample			p-value^
		n (%)	Viral Load (copies/µL)		
			Mean	Median (25°-75° Pct)	
True Positive	24	22 (91.7)	1858,4	472.5 (145.0-975.0)	0,64
False Positive	49	28 (57.1)	3398,6	371.0 (149.3-726.8)	
False Negative	2	0 (0.0)	-	-	
True Negative	38	5 (13.1)	396,4	480.0 (120.0-640.5)	

True positive=positive RST and positive nasopharyngeal swab
False positive=positive RST and negative nasopharyngeal swab
False negative=negative RST and positive nasopharyngeal swab
True negative=negative RST and negative nasopharyngeal swab
^: Kruskal-Wallis test
Pct=percentile

Table S3b: Prevalence of positive rRT-PCR in the salivary sample by presence of COVID-19 symptoms at the time of nasopharyngeal swab procedure,

	N	Positive rRT-PCR on salivary sample			p-value^
		n (%)	Viral Load (copies/µL)		
			Mean	Median (25°-75° Pct)	
COVID-19 symptoms at the time of nasopharyngeal swab procedure					
Any symptom	39	28 (71.8)	1645,1	195.0 (120.0-1200.0)	0,65
No symptoms	77	29 (37.7)	3398,2	480.0 (170.0-682.0)	

^: Kruskal-Wallis test
Pct=percentile