RASSEGNE E ARTICOLI

Characteristics of patients affecting the duration of positivity at SARS-CoV-2: a cohort analysis of the first wave of epidemic in Italy

Caratteristiche dei pazienti che influiscono sulla durata della positività a SARS-CoV-2: un'analisi di coorte nella prima ondata epidemica in Italia

Lorenzo Milani,[1] Fabrizio Cigliano,[2,3] Alberto Catalano,[3] Alessandra Macciotta,[3] Marco Viola,[1] Valeria Caramello,[4] Giuseppe Costa,[3,5] Fulvio Ricceri,[3,5]* Carlotta Sacerdote[1]*

[1] Unit of Cancer Epidemiology, Città della Salute e della Scienza, University-Hospital, Turin (Italy)
[2] Department of Public Health and Pediatrics Sciences, University of Turin, Turin (Italy)
[3] Department of Clinical and Biological Sciences, University of Turin, Orbassano (Italy)
[4] Emergency Department and High Dependency Unit, San Luigi Gonzaga University Hospital, Orbassano (Italy)
[5] Unit of Epidemiology, Regional Health Service ASL TO3, Grugliasco (Italy)
* Joint last authors.
Corresponding author: Carlotta Sacerdote; carlotta.sacerdote@cpo.it

ABSTRACT

OBJECTIVES: to investigate the characteristics of patients affecting the duration of positivity test by RT-PCR in the population of Piedmont, a Region of North-West of Italy.

DESIGN: observational cohort study.

SETTING AND PARTICIPANTS: from the administrative database of the regional SARS-CoV-2 surveillance system, a cohort of all patients who tested positive by a RT-PCR assay to SARS-CoV-2 occurring from 22.02.2020 to 30.09.2020 in the Piedmont Region (N. 29,292) was obtained. The cohort has been linked to the hospital discharge database and to the vital statistics database.

MAIN OUTCOMES MEASURES: outcome of the study was the risk of non negativization, estimated by fitting Generalizing Estimating Equation model (GEE), a longitudinal model which consider for each subject several records collected on fixed time intervals 15, 30, 45 or 60+ days from the first positive test. Negativization was defined as the condition in which two consecutive samples taken from the patient at least 24 hours apart were negative for the presence of SARS-CoV-2.

RESULTS: the median duration of positive RT-PCR was 27 days. A higher median of days until positive persistence was observed in people over 80 (34 days, IQR 25-49), female (28 days, IQR 18-40), symptomatic patients (28 days, IQR 19-40), hospitalized people (32 days, IQR 21-44), patients with Charlson's index >0 (34 days, IQR 23-49), patients host of elderly nursing homes (37 days, IQR 25-51). In the GEE multivariable model, the variables associated to the non negativization at all times intervals were: older age (at 15th day: class 65+, OR 2.56, 95%CI 2.39-2.74), female gender (at 15th day: OR 1.12, 95%CI 1.06-1.18), and to be hospitalized for COVID-19 (at 15th day: OR 1.38, 95%CI 1.29-1.48). The presence of comorbidities and of symptoms were associate with the non negativization at 15th day (respectively, class 4+: OR 1.29, 95%CI 1.08-1.56 and symptoms: OR 1.20, 95%CI 1.13-1.27), but not at 45th day.

CONCLUSIONS: older age, female gender, presence of comorbidities and severity of disease (proxy hospitalization for COVID-19) were risk factors for non negativization at all times intervals. The presence of symptoms was a risk factors for the non negativization after 2 weeks from the first diagnosis and not at 45th day. Using a longitudinal model for the analysis of the dataset, it is possible to compare the weight of the variables included in the model at different times and correct an overestimation of the attributable risk after the first considered time interval.

WHAT IS ALREADY KNOWN

■ Older age, presence of symptoms, presence of comorbidities, and severity of disease are risk factors for positivity persistence after 2 weeks from the first diagnosis.

WHAT THIS PAPER ADDS

■ Older age, female gender, presence of comorbidities, and severity of disease are risk factors for positivity persistence not only after 2 weeks, but also after 30, 45, and 60 days from diagnosis.
■ When negativization is really late (>45 days after first positivity), presence of symptoms and presence of comorbidities are no longer associated with positivity persistence.

RIASSUNTO

OBIETTIVI: indagare le caratteristiche dei pazienti che influenzano la durata della positività al test RT-PCR nella popolazione del Piemonte.

DISEGNO: studio osservazionale di coorte.

SETTING E PARTECIPANTI: dal database amministrativo del sistema di sorveglianza regionale SARS-CoV-2, è stata ottenuta una coorte di tutti i pazienti residenti, risultati positivi al test RT-PCR per SARS-CoV-2 verificatisi dal 22.02.2020 al 30.09.2020 in Regione Piemonte (n. 29.292). La coorte è stata linkata al database delle dimissioni ospedaliere e al database dello stato in vita.

PRINCIPALI MISURE DI OUTCOME: risultato dello studio è il rischio di non negativizzazione, stimato mediante il *Generalizing Estimating Equation model* (GEE), un modello longitudinale che considera per ogni soggetto diversi record raccolti su intervalli di tempo fissi 15, 30, 45 o 60+ giorni dal pri-

mo test positivo. La negativizzazione è stata definita come la condizione in cui due campioni consecutivi, prelevati dal paziente ad almeno 24 ore di distanza, risultano negativi per la presenza di SARS-CoV-2.

RISULTATI: la durata mediana di positività al SARS-CoV-2 era di 27 giorni. È stata osservata una mediana più alta di giorni di positività in persone di età superiore a 80 anni (34 giorni, IQR 25-49), donne (28 giorni, IQR 18-40), pazienti sintomatici (28 giorni, IQR 19-40), persone ospedalizzate (32 giorni, IQR 21-44), pazienti con indice di Charlson >0 (34 giorni, IQR 23-49), pazienti ospiti di residenze sanitarie assistenziali (RSA) (37 giorni, IQR 25-51).

Nel modello multivariato GEE, le variabili associate alla non negativizzazione a tutti gli intervalli di tempo, sono state: età avanzata (al 15° giorno: classe 65+, OR 2,56, 95%CI 2,39-2,74), genere femminile (al 15° giorno: OR 1,12, 95%CI 1,06-1,18) e essere ricoverato per COVID-19 (al 15° gior-

no: OR 1.38, 95%CI 1.29-1.48). La presenza di comorbilità e di sintomi era associata alla non negativizzazione in 15a giornata (rispettivamente, classe 4+: OR 1.29, 95%CI 1.08-1.56 e sintomi: OR 1.20, 95%CI 1.13-1.27) ma non in 45a giornata.

CONCLUSIONI: l'età avanzata, il genere femminile, la presenza di comorbilità e la gravità della malattia (ricovero per COVID-19) sono fattori di rischio per la non negativizzazione a tutti gli intervalli di tempo. La presenza di sintomi è un fattore di rischio per la non negativizzazione dopo 2 settimane dalla prima diagnosi e non al 45° giorno. L'utilizzo di un modello longitudinale per l'analisi dei dati consente di confrontare il peso delle variabili incluse nel modello in tempi diversi e di correggere una sovrastima del rischio attribuibile nel tempo successivo al primo intervallo di tempo.

Parole chiave: SARS-CoV-2, COVID-19, tempo di eliminazione virale, tempo di negativizzazione, fattori di rischio

INTRODUCTION

The duration of isolation of severe acute respiratory syndrome Coronavirus 2 (SARS-CoV-2) infected patients is a crucial public health issue. In drafting the guidelines for the definition of the 'healed subject', in fact, a balance should be found among the need to reduce the possibility of further spread of the virus in the population, the imperative to preserve the personal rights of citizens and the opportunity to reduce the number of work days lost. Released from World Health Organization (WHO) in January 2020, the initial recommendation to confirm clearance of the virus, and thus allow discharge from isolation, required a patient to be clinically recovered and to have two negative Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) results on sequential samples taken at least 24 hours apart.[1] These recommendations were based on knowledge and experience with similar Coronaviruses, including those that caused the severe acute respiratory syndrome coronavirus 1 (SARS-CoV-1) and the Middle East Respiratory Syndrome Coronavirus (MERS-CoV) outbreaks quite recently.[2]

In February 2020, in Italy the first cases of SARS-CoV-2 infection were recorded.[3] Until September 2020, the management of the isolation of positive patients in Italy followed the first guidelines of the WHO, previously described. In Italy, at that time, it was the only way to acquire the status of 'clinically cured patient' (or simply 'cured patient') and thus put an end to the precautionary isolation of the positive subject, conducted under the supervision of the Hygiene Service and Public Health (SISP) of the Local Health Authority (ASL) competent for the territory.[4]

This study aims at investigating the characteristics of patients affecting the duration of positivity test by RT-PCR in the population of a Region of North-west of Italy, in the first 8 months of SARS-CoV-2 epidemics. This is in order to gain a better understanding of the characteristics of the patients that influence a longer positivity at RT-PCR test, crucial for both clinical and social reasons and to assist public health decisions to be made in the short time typical of this pandemic.

MATERIAL AND METHODS

STUDY DESIGN AND POPULATION

An observational cohort study built using administrative data from the Piedmont Region in Northwest Italy (about 4,400,000 inhabitants) was conducted. From the regional SARS-CoV-2 surveillance system, the cohort obtained included all resident patients aged 0-98 years old who tested positive by a RT-PCR assay on nasal and/or pharyngeal specimen or bronco-alveolar lavage (BAL) to SARS-CoV-2 occurring from 22.02.2020 (beginning of the epidemics) to 30.09.2020 (N. 29,292). The negativization was considered when patients have two negative RT-PCR results on sequential samples taken at least 24 hours apart. The cohort of SARS-CoV-2 infected patients has been linked to the hospital discharge database and to the vital statistics database by a unique anonymous identifier. Patients with formally incorrect anonymous identifier or patients still positive on 30.09.2021 were excluded from the cohort. The final cohort included 28,408 patients. (Figure 1)

The sequence of the swabs varies from subject to subject: the simplest series have a positive swab followed by two negative swabs. However, some patients presented complicated series of swabs with several consecutive positive RT-PCR exams or alternating positives and negatives before reaching the two negative swabs that close the sequence (Figure 2). Anyway, all the alive patients closed the sequence with the two negative swabs, because it was the only way for the patient to exit the quarantine period.
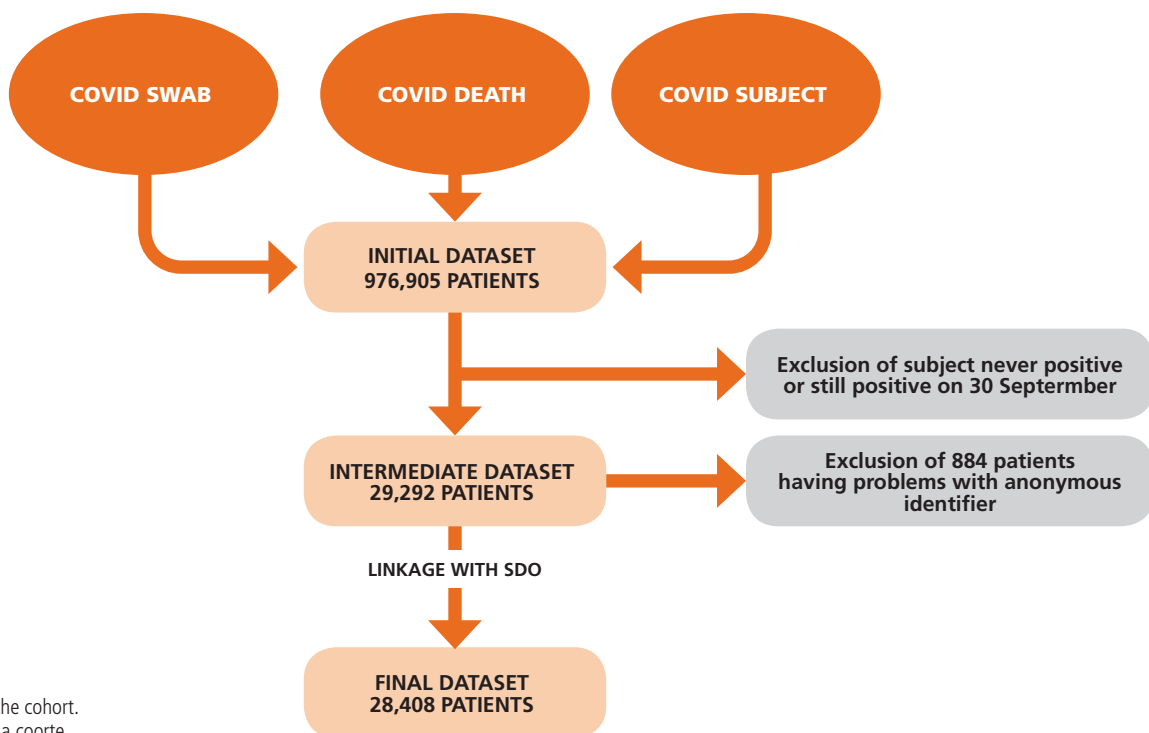
RASSEGNE E ARTICOLI

COVID SWAB

COVID DEATH

COVID SUBJECT

**INITIAL DATASET
976,905 PATIENTS**

Exclusion of subject never positive or still positive on 30 Septermber

**INTERMEDIATE DATASET
29,292 PATIENTS**

Exclusion of 884 patients having problems with anonymous identifier

LINKAGE WITH SDO

**FINAL DATASET
28,408 PATIENTS**

Figure 1. Flowchart of the cohort.
Figura 1. Flowchart della coorte.
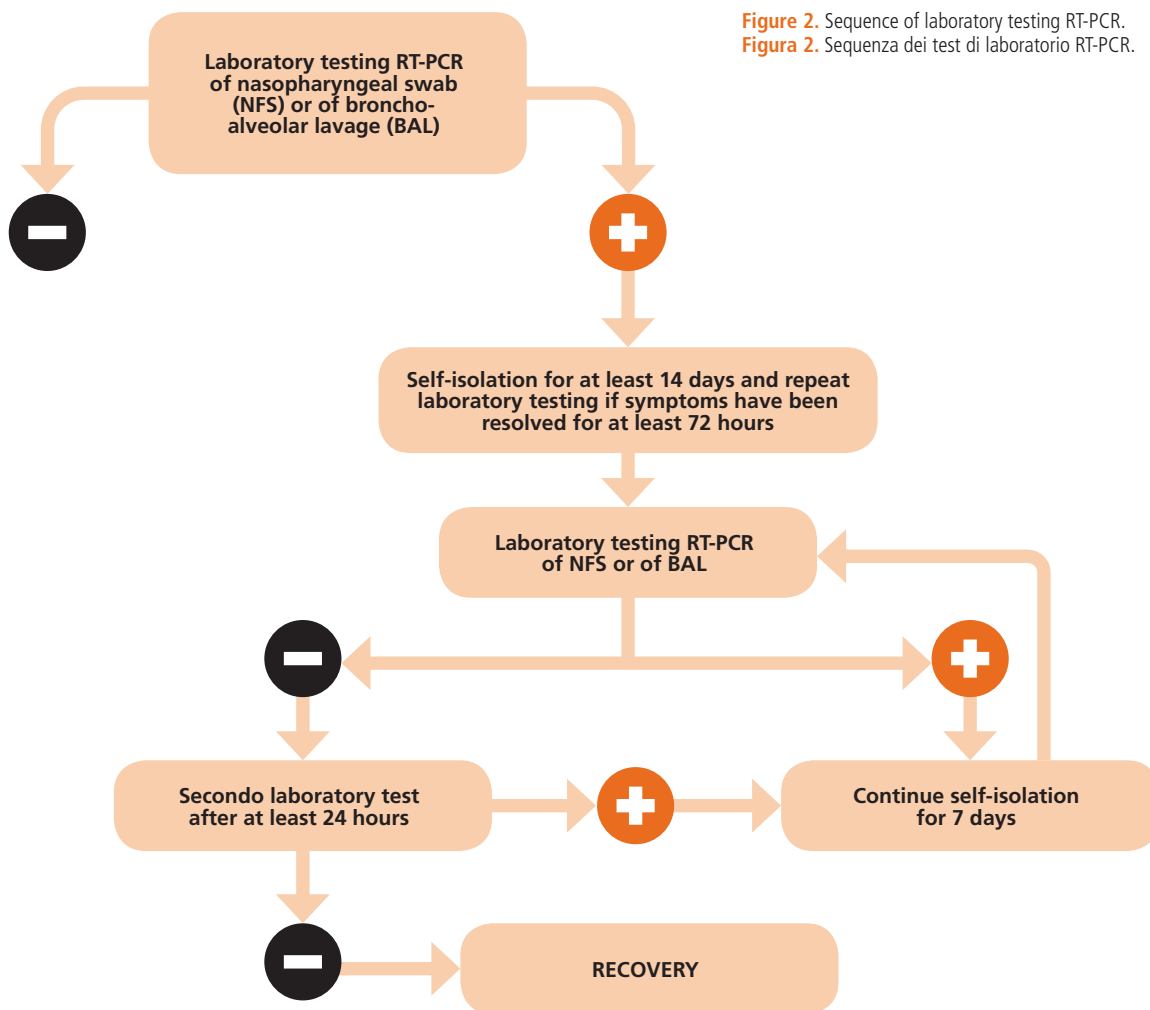
Figure 2. Sequence of laboratory testing RT-PCR.
Figura 2. Sequenza dei test di laboratorio RT-PCR.

Laboratory testing RT-PCR of nasopharyngeal swab (NFS) or of broncho-alveolar lavage (BAL)

Self-isolation for at least 14 days and repeat laboratory testing if symptoms have been resolved for at least 72 hours

Laboratory testing RT-PCR of NFS or of BAL

Secondo laboratory test after at least 24 hours

Continue self-isolation for 7 days

RECOVERY

VARIABLES

Epidemiological characteristics, clinical manifestation and the sequence of RT-PCR test results were collected. In particular, we considered as possible risk factors:

- **Age** stratified in three classes: 0-49, 50-64, 65+ years;
- **Hospitalization** having value 1 if patient had at least one hospitalization for COVID-19 (using a combination of ICD-9 codes of diagnosis and procedures to identify respiratory diseases in the hospital discharge database) during the sequence of positivity, 0 otherwise;
- **Hospitalization during positivity** having value 1 if subject was hospitalized for COVID-19 at the time of corresponding positive RT-PCR swab, 0 otherwise;
- **Symptoms** having value 1 for symptomatic patients (at least one symptom related to COVID-19 at the moment of diagnosis as reported in the infectious disease report form), 0 for asymptomatic patients;
- **Charlson's comorbidity index** stratified in four classes (score 0,1,2-3,4 or more);[5] this variable is created from hospital discharge diagnosis codes of hospitalizations in the last five years;
- **Number of performed swabs** having value 1 if patient has more than one swabs per week, 0 otherwise.
- **Nursing home host** having value 1 if patient was host of a elderly nursing home at the moment of first positive RT-PCR swab, 0 otherwise.

Despite national guidelines recommended to apply rigorous interval of time when RT-PCR tests are performed, in clinical practice they are not strictly respected for many reasons. Anyway, to set fixed times in which evaluate estimates, a variable was create to indicate how many days have passed from the first positive test. This variable assumes 15, 30, 45 and 60+ days values, grouping swabs with an approximation of ±7 days (for example, 30 value groups days from 23th to 37th). Median testing value was 9 days.

Response variable Y is created considering couples of RT-PCR results: if both are negative Y is equal 1 and the patients is discharged from the study, otherwise Y is equal 0 and patients isn't discharged. Considering the time between first positive test and the first negative test of the exit swabs couple, Y response was linked to the appropriate value of fixed days. Patients discharge when Y equal one. The variable death is enhanced for patients who died before negativization. Considering the day of death, the number of days the patient has been positive was calculate and death=1 was assigned in correspondence of the appropriate value of fixed days, otherwise death=0. These patients have Y equal zero always and they discharge when death is equal one.

STATISTICAL ANALYSIS

Descriptive analyses of the variables were expressed as absolute frequency and percentage and mean and standard deviation, as appropriate. Differences among groups were investigated Kruskal-Wallis test and chi² test.

The risks of positive persistence were estimated by fitting Generalizing Estimating Equation model (GEE), a longitudinal model which consider for each subject several records collected on fixed time intervals. This model is an extension of Quasi-Likelihood (QL) and Generalized Linear Model (GLM) for correlated outcomes. It allows n_i different for patients included in the study, where value n_i is not the same for each patient, so it adapts to the case of RT-PCR test where each subject has its own number of tests carried out. Such as GLM and QL models, it has some assumptions to be made before the estimates of parameters, in particular assumption upon correlation between subject outcomes. Results were expressed as Odds Ratios (ORs) and corresponding 95% confidence intervals (CIs). Two specifications of the Quasi Information Criterion (QIC and QICu) were calculated as an extension of Akaike Information Criterion for QL estimation models. The quasi-likelihood model criterion (QIC) was used to compare different correlation matrix, the quasi-likelihood model criterion under the independence model criterion (QICu) was used to select best set of variable given the correlation assumed.

Sensitivity analysis including the number of performed swabs in the model was conducted. This variable is built in order to account the swab's frequency: it is a binary variable which takes value 1 if subject has more than one swabs per week, 0 otherwise. Furthermore, a GEE model was fitted on 45th day, stratified for elderly nursing home host status and hospitalization.

All calculations were made using SAS version 9.4 software (SAS Institute, Cary, North Carolina).

RESULTS

28,408 subjects who tested positive by a polymerase chain reaction assay on nasal and/or pharyngeal specimen or BAL to SARS-CoV-2 occurring from 22.02.2020 to 30.09.2020 were identified. In this population, 40.4% were male and the mean age was 58.7±21.5 years (median 57). Most of patients had at least one symptom among those attributable to SARS-CoV-2 infection (56.2%, missing values 5.3%). Twenty-four percent of patients were hospitalized for COVID-19. The largest part of subjects considered (80.8%) had no pathologies according to Charlson's comorbidity index. The remaining 19.2%, was classified as follow: 7.3% class 1, 9.1% for class 2-3, and 2.3% for 4+.

As expected, higher mortality was present in older subjects (70.6% deaths in patients aged over 80 years) and in symptomatic patients (65.6%); the remaining percentage was equally divided in asymptomatic and patients with missing value for this variable. Most of the deaths occurred among

hospitalized subjects (77.8%) and among hosts of elderly nursing residence (32.2%) and surprisingly half of the deceased subjects had a Charlson's comorbidity index of 0. Finally, no gender differences in mortality rates were noticed (49% of male and 51% of women deceased) (Table 1).

Table 2 presents the analysis of the path of recovery for patients who reached two consecutive negative tests. Considering number of swabs, there were differences among groups for some variables (age, symptoms, hospitalization, and hosts of elderly nursing homes) while for other variables there were no differences among groups (gender and Charlson's comorbidity index). Classifying by age, median value for most classes was 3, while for class 80+ median is 4 (IQR 3-5). Number of swabs increased in oldest patients. Considering symptoms and hospitalization, a higher median number of swabs was identified for pa-

tients with symptoms (N. 4, IQR 3-6) and for hospitalized subjects (N. 4, IQR 3-6) rather than asymptomatic subjects (N. 3, IQR 3-5) and non-hospitalized subjects (N. 3, IQR 3-5): it could be noticed that people without symptoms and/or not hospitalized and/or nursing home hosts were tested less frequently. Furthermore, younger patients showed a lower median of days until recovery (N. 20, IQR 16-33), than people over 80 (N. 34, IQR 25-49). Female (N. 28, IQR 18-40) took one day longer than males until recovery (N. 27, IQR 18-39). Patients with symptoms had higher median of days until recovery (N. 28, IQR 19-40) than asymptomatic patients (N. 27, IQR 17-40), also hospitalized people had higher median of days until recovery (N. 32, IQR 21-44) than not hospitalized patients (N. 26, IQR 18-38) and hosts of elderly nursing homes (N. 37, IQR 25-51). Finally, considering Charlson's comorbidity index, an increase of median days until recovery was observed, with an increasing index: patients without comorbidity had 27 median days until recovery (IQR 18-38) while patients with Charlson's index not equal 0 had at least 34 median day (34 if index equal 1-3, IQR 22-47; 35 if equal 4+, IQR 23-49).

The characteristics of subjects by week of recovery or death are presented in table 3: in '15 days' column, for example, characteristics of patients who recovered or died within the first 15 days from first positive test were described. This table swows a higher mean age with an increasing of days of positivity: younger patients seem to recovery faster than older people. Considering two extreme periods (15 days and 60 days), the percentage of male at 15 days is higher than at 60 days, for male is less probably to have a sequence 60 days long. Approximately the same proportion of symptomatic patients was observed in the two selected moments, while for the hospitalization there is difference in percentage (at 15 days, percentage of hospitalized is higher than percentage at 60 days for at least one hospitalization during SARS-CoV-2 positivity). Finally, for Charlson's comorbidity index, there is a higher decrease in percentage for class 0, always comparing 15 days and 60 days: people without comorbidity seem to conclude their sequences earlier than 60 days (Table 3).

In Table 4, ORs from the GEE model at different recovery times are presented. All the variables at 15th days were associated to the non-negativization. In particular, the older age (class 65+, OR 2.56, 95%CI 2.39-2.74), the female gender (OR 1.12, 95%CI 1.06-1.18), the presence of symptoms (OR 1.20, 95%CI 1.13-1.27), a higher Charlson index (class 4+, OR 1.29, 95%CI 1.08-1.56), and to be hospitalized for COVID-19 at the 15th day (OR 1.38, 95%CI 1.29-1.48) were all risk factors for the non negativization after 2 weeks from the first diagnosis. Focusing on 45th day, there were four variables associated to the non-negativization: age (class 65+), gender, hospitaliz-

CHARACTER-ISTICS	RECOVERED PATIENTS*		DEATH		TOTAL	
	N.	%	N.	%	N.	%
COHORT	26,758		1,650		28,408	
AGE CLASSES (YEARS)						
0-14	406	1.52	0	0	406	1.42
15-49	8,418	31.46	18	1.09	8,436	29.70
50-64	7,134	26.66	95	5.76	7,229	25.45
65-79	4,427	16.54	372	22.55	4,799	16.89
80+	6,373	23.82	1,165	70.61	7,538	26.54
GENDER						
male	10,675	39.84	808	48.97	11,483	40.42
female	16,083	60.16	842	51.03	16,925	59.58
SYMPTOMS						
with	15,954	59.62	1,083	65.64	17,037	59.97
without	9,284	34.70	288	17.45	9,572	33.69
missing	1,520	5.68	279	16.91	1,799	6.34
HOSPITALIZATION						
yes	6,919	25.86	1,284	77.82	8,203	28.88
no	19,839	74.14	366	12.18	20,205	71.12
CHARLSON INDEX						
0	22,068	82.47	882	53.45	22,950	80.79
1	1,876	7.01	208	12.60	2,084	7.33
2-3	2,190	8.18	407	24.67	2,597	9.14
4+	624	2.34	153	9.27	777	2.74
NURSING HOME HOST						
yes	5,030	18.92	530	32.12	5,560	19.57
no	21,728	81.08	1,120	67.88	22,848	80.43

* Patients with two negative RT-PCR results on sequential samples taken at least 24 hours apart./ *Pazienti con 2 risultati RT-PCR negativi o in una sequenza di campioni raccolti almeno a 24 ore di distanza.*

Table 1. Demographic and clinical characteristics of the cohort of patients with at least a RT-PCR or BAL positive to SARS-CoV-2 in the Piedmont Region (22.02.2020-30.09.2020).
Tabella 1. Caratteristiche cliniche e demografiche della coorte di pazienti con almeno uun test RT-PCR o BAL positivo per SARS-CoV-2 in Piemonte (22.02.2020-30.09.2020).

RASSEGNE E ARTICOLI

CHARACTERISTICS	SWABS NUMBER UNTIL RECOVERY				DAYS UNTIL RECOVERY			
	MEAN	MEDIAN	MAXIMUM	P-VALUE**	MEAN	MEDIAN	MAXIMUM	P-VALUE**
COHORT RECOVERED	4.07	3	22		30.08	27	139	
AGE CLASSES								
0-14	3.92	3	12	<0.0001	23.15	20	74	<0.0001
15-49	4.07	3	20		26.38	23	137	
50-64	4.06	3	21		28.31	25	127	
65-79	4.13	3	18		31.77	29	119	
80+	4.23	4	22		36.81	34	139	
GENDER								
Male	4.17	3	22	0.06	29.53	27	139	<0.0001
Female	4.07	3	20		30.44	28	137	
SYMPTOMS								
With	4.22	4	21	<0.0001	30.28	28	139	<0.0001
Without	3.99	3	22		29.91	27	137	
Missing	3.70	3	10		29.01	27	100	
HOSPITALIZATION								
Hospitalized	4.44	4	21	<0.0001	32.50	32	139	<0.0001
Not Hospitalized	4.00	3	22		29.30	26	137	
CHIARLSON'S COMORBIDITY INDEX								
0	4.08	3	20	0.60	29.16	27	137	<0.0001
1	4.17	4	16		34.55	34	119	
2-3	4.30	4	22		34.39	34	139	
4+	4.56	4	20		36.36	35	121	
NURSING HOME HOST								
Yes	3.93	3	16	<0.01	41.33	36	182	<0.01
No	4.19	4	21		28.97	25	194	

* Kruskal-Wallis test for median / test di Kruskal-Wallis per la mediana

Table 2. Demographic and clinical characteristics of the cohort of patients with at least a RT-PCR or BAL positive to SARS-CoV-2 in the Piedmont Region (22.02.2020-30.09.2020), by number of performed tests and days from the first positive test and recovery (two consecutive negative tests).
Tabella 2. Caratteristiche cliniche e demografiche della coorte di pazienti con almeno un test RT-PCR o BAL positivo per SARS-CoV-2 in Piemonte (22.02.2020-30.09.2020), per numero di test eseguiti e giorni dal primo test positivo e guarigione (due test negativi consecutivi).

	15 DAYS		30 DAYS		45 DAYS		60 DAYS		P-VALUE
COHORT	11,153		9,604		5,067		2,583		
MEAN AGE	55.83 21.56		61.37 21.31		66.92 21.13		71.48 19.91		<0.001*
GENDER									
Male	4726	(42.37)	3840	(39.98)	2145	(42.33)	942	(36.47)	<0.001**
Female	6428	(57.63)	5764	(60.02)	3326	(65.64)	1641	(63.53)	
SYMPTOMS									
Yes	3894	(34.91)	3013	(31.37)	1664	(32.84)	1001	(38.75)	<0.0001**
No	6442	(57.76)	6009	(62.57)	3108	(61.34)	1477	(57.18)	
Missing	817	(7.33)	582	(6.06)	295	(5.82)	105	(4.07)	
HOSPITALIZION									
No	8361	(74.97)	6716	(69.93)	3484	(68.76)	1737	(67.25)	<0.0001**
Yes	2793	(25.04)	2888	(30.07)	1987	(39.21)	846	(32.75)	
CHARLSON									
0	9434	(84.59)	7789	(81.10)	4078	(80.84)	1846	(71.47)	<0.0001**
1	605	(5.42)	732	(7.62)	565	(11.15)	257	(9.95)	
2-3	858	(7.69)	833	(8.67)	636	(12.55)	367	(14.21)	
4+	256	(2.30)	250	(2.60)	192	(3.79)	113	(4.37)	

* Kruskal-Wallis test ** Chi2 test

Table 3. Demographic and clinical characteristics of the patients with at least a RT-PCR or BAL positive to SARS-CoV-2 in the Piedmont Region (22.02.2020-31.08.2020), by week of recovery (two consecutive negative tests).
Tabella 3. Caratteristiche cliniche e demografiche della coorte di pazienti con almeno un test RT-PCR o BAL positivo per SARS-CoV-2 in Piemonte (22.02.2020-30.09.2020), per settimana di guarigione (due test negativi consecutivi).

	15 DAYS	30 DAYS	45 DAYS
GENDER (F vs M)	1.12** (1.06-1.18)	1.04 (0.99-1.08)	1.09** (1.04-1.14)
SYMPTOMS (vs 0)	1.20** (1.13-1.27)	0.93** (0.89-0.97)	0.83** (0.79-0.87)
HOSPITALIZATION (vs 0)	1.38** (1.29-1.48)	1.35** (1.27-1.43)	1.67** (1.54-1.80)
AGE 50-64 YEARS (vs 0-49)	1.25** (1.17-1.33)	1.04 (0.99-1.10)	1.04 (0.98-1.11)
AGE 65+ YEARS (vs 0-49)	2.56** (2.39-2.74)	1.44** (1.36-1.51)	1.11** (1.05-1.17)
CHARLSON 1 (vs 0)	1.28** (1.14-1.44)	1.03 (0.95-1.11)	0.92 (0.85-1.00)
CHARLSON 2-3 (vs 0)	1.17** (1.06-1.30)	1.11** (1.03-1.20)	1.02 (0.94-1.10)
CHARLSON 4+ (vs 0)	1.29** (1.08-1.56)	1.12 (0.98-1.28)	1.01 (0.88-1.15)
QIC^	33,242.68	37,656.59	23,675.67
QICu^	33,242.55	37,662.66	23,684.73

* p-value < 0.05   ** p-value < 0.01
^Quasi-likelihood model criterion (QIC) and quasi-likelihood model criterion under the independence model criterion (QICu). / *Criterio del modello della quasi-verosimiglianza (QIC) e criterio del modello della quasi-verosimiglianza sotto il criterio del modello di indipendenza (QICu).*

Table 4. Generalized Estimating Equation (GEE) of patients with at least a RT-PCR or BAL positive to SARS-CoV-2in the Piedmont Region (22.02.2020-30.09.2020), by number of performed tests and days from the first positive test and recovery (two consecutive negative tests). Outcome: risk of positive persistence.
Tabella 4. Caratteristiche cliniche e demografiche della coorte di pazienti con almeno un test RT-PCR o BAL positivo per SARS-CoV-2 in Piemonte (22.02.2020-30.09.2020), per numero di test eseguiti e giorni dal primo test positivo e guarigione (due test negativi consecutivi). Outcome: rischio di persistenza alla positività.

ation, and symptoms. In particular, people from 50 to 64 years old have an increased risk of 11% comparing with class 0-49 (OR 1.11, 95%CI 1.05-1.17). Female patients have a risk of be still positive at 45th day 9% higher than male people (OR 1.09, 95%CI 1.04-1.14). Similarly, be hospitalized at 45th day increases the risk of be still positive (OR 1.67, CI. 1.54-1.80). Developing symptoms, instead, is a protective factor against be positive at 45th day: patients with symptoms have a 17% reduced risk compared to those with no symptoms (OR 1.17, 95%CI 1.54-1.80). The variable residence in an elderly nursing home was not included in the GEE model due to the important collinearity with age.

Since patients with symptomatic (versus paucis or asymptomatic) and inpatient (versus outpatient) SARS-CoV-2 infections were usually subjected to more extensive testing, a sensitivity analysis was performed to include in the model the non-homogeneous number of swabs performed according to the different conditions mentioned above. The results are available in Table S1 (see online supplementary materials). However, even when performing this analysis, the estimates did not change from the main analyzes. This proves that carrying out more than one test per week does not affect the estimates of other variables.

In Table S2, a GEE model at 45 days was presented stratified by elderly nursing home residence and by hospitalization. This analysis was included to exclude the possibility of an effect modification due to a possible different pattern of swabs execution in institutionalized patients. No effect modification was evident. As expected, largest confidence intervals that include the unit were seen, in particular for the groups with a lower numerosity.

DISCUSSION

During a period of epidemic spread of an infectious agent, such as the one experiencing for the SARS-CoV-2 pandemic, to know the duration of shedding of the agent is one of the determinants to define an appropriate period of isolation. In fact, this information is often used as a proxy of the possible infectivity of subjects who contract the infection.[6] Despite the importance of the topic, there are currently few studies and almost all based on Chinese population, investigating the issue of the factors that determine the persistence of swab positivity.[7-20] For these reasons, it was decided to explore what could be the factors contributing to the persistence of the swab positivity beyond the standard 2 weeks from first positive test in an Italian population.

A systematic review on the studies published until October 2020 found that viral spread, detected by the SARS-CoV-2 RT-PCR test on high respiratory tract samples, have a median duration of 16.8 days (95%CI 14.8-19.4) from disease onset (first appearance of symptoms or first positive RT-PCR test).[21] In the study population here considered, the median duration of positive RT-PCR was 27 days (calculated from first positive RT-PCR test for both asymptomatic and symptomatic patients). In Authors' knowledge, a higher median duration than 27 days was calculated only in another Italian population[14] and in a Chinese study.[7] Anyway, the discrepancies of duration among cohorts of patients could be attributed to several reasons regarding both the setting of the study (hospital or outpatients setting and the epidemics period) or the characteristics of the patients (median age and disease severity), but also the definition of duration of positivity test.

In fact, differently from the present analysis, almost all the previous studies were conducted on hospitalized patients,[7-13,15-17,19-20] and on the first four months of the epidemics.[7-17,19] The median age was lower in all studies (about 50 years old or less), except for the paper of Mancuso et al.[14] In all the studies, patients who tested negative for SARS-CoV-2 for two consecutive days in respiratory samples were considered to be clear of infection. Instead, the duration of the positive period was calculated in some studies as the time from the day of the first positive swab or from the hospital admission,[10-11,14,16,19] in

RASSEGNE E ARTICOLI

erlies. In fact, T-cell numbers and functions are affected by aging, leading to less controlled viral replication.[22] Several previous studies showed that older age is a risk factor for COVID-19 infection and mortality[3,23] then it is not surprising than could also affect the time of negativization.

Conversely, gender association with prolonged viral shedding is a controversial topic showing contradictory results in different studies. A longer positivity duration in female coherent with the present results has been calculated in a recent meta-analysis on 35 studies (viral shedding time in female 19.4, 95%CI 9.5-39.4, vs male 11.9, 95%CI 8.4-16.9) also if the difference between groups was not statistically significant.[21] This was an unexpected finding, because several studies have reported a significant difference in the rate of severe cases between adult females and males (42% vs 58%).[24] Usually, women exhibit lower infection and mortality rate for viral infectious disease,[25] however men seem to be more affected by seasonal viral infection (such as influenza), whereas postmenopausal women succumb more often to pandemic strains.[25] In women, elevated induction of inflammatory cytokines could be the reason of a higher morbidity and mortality, but also hormonal mechanisms related to estrogens that reduce the number of T lymphocytes could be involved.[25]

In previous studies, duration of positive RT-PCR was related to presence of symptoms: the duration is higher in symptomatic patients (19.7 days after the onset of symptoms, 95%CI 17.2-22.7) vs asymptomatic patients (10.9 days after the first positive RT-PCR test, 95%CI 8.3-14.3).[21] The presence of symptoms appears to be a risk factor of early non-negativization (within 14 days) also in the population here considered. Anyway, in this population, developing symptoms is a protective factor against being positive at 45th day in the adjusted model. The possible explanations of this result could be both:

■ a fast virus clearance in asymptomatic patients due to a higher immunity response;

■ a viral shedding who begun several days before the first positive RT-PCS test but ignored due to the absence of symptoms.

To confirm this result, was not due to a higher number of tests performed in symptomatic patients, also after the resolution of symptoms, a sensitivity analysis was performed to take into account the number of performed tests without any change in the main result.

The results obtained from the present cohort confirm the association between the delay in the negativization of the subjects' RT-PCR tests and the degree of severity of their disease, measured on the basis of their need for hospitalization. The association of this delay of negativization with the presence of comorbidities in the subject, on the other hand, was observed as significant only in the early phase (up to the 15th day). A plausible explanation for this event,

is that the presence of comorbidities manifests its effects on the duration of positivity of the RT-PCR test only in subjects who develop COVID of medium or moderate severity, and therefore their influence ceases in the short term. In contrast, in individuals who develop a more severe form of COVID, the delayed negativization may be more dependent on clinical variables that were not included in the present study model. This is in agreement with the fact that the probability of death, which is closely related to the presence of comorbidities, is not included in this study.

The evidence seems to go in the direction of those already collected not only for SARS-CoV-2, but also for other respiratory viruses such as Influenza and MERS, for which an alteration in the patient's immune status and the presence of significant comorbidities (such as diabetes mellitus) determine a prolonged viral spread, and a delayed negativization of the tests used to detect the presence of viruses in the host.[26-28] Influenza viral RNA can be detected in the respiratory tract for a prolonged period in patients with severe disease or immunosuppression after illness onset.[26]

Finally, it must be noticed that the detection of SARS-CoV-2 RNA from biological samples of the subjects does not necessarily mean that the subject is infectious and, therefore, capable of transmitting the virus to other people.[29] The risk of transmission of the infection, in fact, is also determined by other factors. Some are related to the virus, such as its residual ability to replicate. Others are related to the person carrying the virus, such as his social behaviour, his observance of anti-contagion regulations, the presence or absence of symptoms. Among the latter, coughing and sneezing, which are important sources of infectious droplets, are of particular importance.[30] Furthermore, it is now established that usually in a period between 5 and 10 days from the SARS-CoV-2 infection, the infected subject begins to gradually produce antibodies neutralizing the virus, and it is expected that the presence of these antibodies may reduce the likelihood of human-to-human transmission of the infection.[30] As seen, several patients whose symptoms have resolved still test positive for COVID-19 by RT-PCR for many weeks, and this complicates the development of guidelines to define individuals who are (likely) no longer infectious.[31]

The present study suffers from some limitations:

■ the beginning of the subjects' surveillance swab sequence coincides with the date of the first positive swab (i.e., the diagnostic swab of infection) rather than the date of onset of symptoms (which may be earlier than the date of the diagnostic swab);

■ the analysis was carried out using fixed and pre-established times that approximate the actual execution times of the swabs;

■ the moments when the control swabs were performed depend in the first instance on the date of resolution of the

subject's symptoms (the health authorities, in fact, do not consider it advisable to test subjects who are still symptomatic), but also on organizational and social-health needs, which was not possible to manage in the analysis phase. The possibility to be tested only when symptoms disappeared could affect the results on the higher probability to non negativization in symptomatic patients at 15 days. Anyway, in individuals who develop a more severe form of COVID-19, the presence of symptoms at diagnosis seems not be associated to probability of negativization at 30 and 45 days;

■ using administrative databases as a source of data, it is possible to carry out analysis on a very large number of topics but the provision of information relating to each case is meagre;

■ the variable 'symptoms' was assessed at the moment of confirmation of positivity at the RT-PCR swab; then, symptoms appearing later were not recorded. This could be the reason of the high mortality on asymptomatic patients. Furthermore, this variable has a high number of missing that could be not at random; therefore, it was decided to show in the descriptive tables missing as a separate category of the variable symptoms;

■ Ccomorbidities were assessed through the Charlson's index, then only comorbidities which have been diagnosed or reported in previous hospitalization were included; the presence of comorbidities could be underestimated.

■ positivity was considered only in symptoms in upper respiratory tract, but there is a general consensus on a prolonged viral shedding through feces.[32] Anyway the importance of this persistence on probability of transmission of the infection is still not confirmed;[32]

■ the Italian national guidelines for the definition of 'healed subject' were modified in October 2020. For this reason, it is not possible to make comparisons with the data referring to periods after September 2020.

Studies like the present one are important to support public health decisions with scientific evidence; for example, not recommending tailored timing of control swabs based on age, severity of SARS-CoV-2 disease or the presence of other concomitant pathologies. In fact, also if there is evidence of the presence of patient characteristics that are determinants of a late negativization, the median of duration of viral clearance are similar and do not justify the definition of personalized time of isolation. In this context, one of the strengths of using a longitudinal model for the analysis of the dataset is the possibility of comparing the weight of the variables included in the model at different times. This type of approach appears to be useful in correcting an overestimation of attributable risk in the time after the first, as appears in the results of the logistic regression.

REFERENCES

1. WHO. Laboratory testing of human suspected cases of novel coronavirus (nCoV) infection. Interim guidance, 10.01.2020. Available from: https://apps.who.int/iris/bitstream/handle/10665/330374/WHO-2019-nCoV-laboratory-2020.1-eng.pdf (last accessed:ed: 09.08.2021).
2. WHO. Laboratory assessment tool for laboratories implementing SARS-CoV-2 testing: https://www.who.int/publications/i/item/laboratory-assessment-tool-for-laboratories-implementing-covid-19-virus-testing (last accessed: 09.08.2021).
3. Onder G, Rezza G, Brusaferro S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. JAMA 2020;323(18):1775-76.
4. Mamo C, Dalmasso M, Pasqualini O, et al. Durata di isolamento dei casi COVID-19 nella prima ondata epidemica in un'ASL della Città Metropolitana di Torino. Boll Epidemiol Naz 2021;2(3):26-35.
5. Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. J Clin Epidemiol 1993;46(10):1075-79
6. Lau LL, Ip DK, Nishiura H, et al. Heterogeneity in viral shedding among individuals with medically attended influenza A virus infection. J Infect Dis 2013;207(8):1281-85
7. Zhou B, She J, Wang Y, Ma X. Duration of Viral Shedding of Discharged Patients With Severe COVID-19. Clin Infect Dis 2020;71(16):2240-42.
8. Zheng S, Fan J, Yu F, et al. Viral load dynamics and disease severity in patients infected with SARS-CoV-2 in Zhejiang province, China, January-March 2020: retrospective cohort study. BMJ 2020;369:m1443.
9. Liu Y, Yan LM, Wan L, et al. Viral dynamics in mild and severe cases of COVID-19. Lancet Infect Dis 2020;20(6):656-57.
10. Hu X, Xing Y, Jia J, et al. Factors associated with negative conversion of viral RNA in patients hospitalized with COVID-19. SCI Total Environ. 2020;728:138812.
11. Shi D, Wu W, Wang Q, et al. Clinical Characteristics and Factors Associated With Long-Term Viral Excretion in Patients With Severe Acute Respiratory Syndrome Corona-
virus 2 Infection: a Single-Center 28-Day Study. J Infect Dis 2020;222(6):910-18.
12. Xu K, Chen Y, Yuan J, et al. Factors Associated With Prolonged Viral RNA Shedding in Patients with Coronavirus Disease 2019 (COVID-19). Clin Infect Dis 2020;71(15):799-806.
13. Yan D, Liu XY, Zhu Y, Huang L, Dan B, Zhang G, Gao Y. Factors associated with prolonged viral shedding and impact of lopinavir/ritonavir treatment in hospitalised non-critically ill patients with SARS-CoV-2 infection. EurRespir J 2020;56(1):2000799.
14. Mancuso P, Venturelli F, Vicentini M, et al. Temporal profile and determinants of viral shedding and of viral clearance confirmation on nasopharyngeal swabs from SARS-CoV-2-positive subjects: a population-based prospective cohort study in Reggio Emilia, Italy. BMJ Open 2020;10(8):e040380.
15. Fu Y, Han P, Zhu R, et al. Risk factors for viral RNA shedding in COVID-19 patients. Eur Respir J 2020;56(1):2001190.
16. Yang Y, Hu X, Xiong L, Fet al. Clinical characteristics of hospitalized mild/moderate COVID-19 patients with a prolonged negative conversion time of SARS-CoV-2 nucleic acid detection. BMC Infect Dis 2021;21(1):141.
17. Huang R, Zhu L, Xue L, et al. Clinical and virological course of patients with coronavirus disease 2019 in Jiangsu province, China: a retrospective, multi-center cohort study. Virol J 2021;18(1):147.
18. Bennasrallah C, Zemni I, Dhouib W, et al. Factors associated with a prolonged negative conversion of viral RNA in patients with COVID-19. Int J Infect Dis 2021;105:463-69.
19. Chen X, Hu W, Yang M, et al. Risk factors for delayed viral clearance in COVID-19 patients. J Clin Hypertens (Greenwich) 2021;23(8):1483-89.
20. Hirai N, Nishioka Y, Sekine T, et al. Factors associated with viral clearance periods from patients with COVID-19: A retrospective observational cohort study. J Infect Chemother 2021;27(6):864-68.
21. Yan D, Zhang D, Chen C, et al. Characteristics of Viral Shedding Time in SARS-

CoV-2 Infections: A Systematic Review and Meta-Analysis. Front Public Health 2021;9:652842.

22. Goronzy JJ, Lee WW, Weyand CM. Aging and T-cell diversity. Exp Gerontol 2007;42(5):400-6.

23. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020; 395(10229):1054-62.

24. Guan W, Ni Z, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med 2020;382(18):1708-20.

25. Giefing-Kröll C, Berger P, Lepperdinger G, Grubeck-Loebenstein B. How sex and age affect immune responses, susceptibility to infections, and response to vaccination. Aging Cell 2015;14(3):309-21.

26. Hao S, Lian J, Lu Y, et al. Decreased B Cells on Admission Associated With Prolonged Viral RNA Shedding From the Respiratory Tract in Coronavirus Disease 2019: A Case-Control Study. J Infect Dis 2020;222(3):367-71.

27. Wang Y, Guo Q, Yan Z, et al. Factors Associated With Prolonged Viral Shedding in Patients With Avian Influenza A(H7N9) Virus Infection. J Infect Dis 2018; 217(11):1708-17.

28. Kojima N, Mores C, Farsai N, Klausner J. Duration of SARS-CoV-2 viral culture positivity among different specimen types. Clin Microbiol Infect 2021:27(10):1540-41.

29. Kissler SM, Fauver JR, Mack C, et al. Viral dynamics of acute SARS-CoV-2 infection and applications to diagnostic and public health strategies. PLoS Biol 2021;19(7):e3001333.

30. Peluso MJ, Deitchman AN, Torres L, et al. Long-term SARS-CoV-2-specific immune and inflammatory responses in individuals recovering from COVID-19 with and without post-acute symptoms. Cell Rep 2021;36(6):109518.

31. Morone g, Palomba A, Iosa M, et al. Incidence and Persistence of Viral Shedding in COVID-19 Post-acute Patients With Negativized Pharyngeal Swab: A Systematic Review. Front Med (Lausanne) 2020;7:562.

32. Yeo C, Kaushal S, Yeo D. Enteric involvement of coronaviruses: is faecal–oral transmission of SARS-CoV-2 possible? Lancet Gastroenterol Hepatol 2020;5:335-37.