Swabbing Thoroughly People for COVID-19 Positivity. Insights on the Main Bio-analytical and Microbiology Bias and Concerns

Salvatore Chirumbolo[1] ⬤

Abstract

Thorough swabbing is becoming an increasing approach to fight COVID-19 transmission, particularly among asymptomatic subjects, who are thought to represent the majority of potentially-contacting people. Particularly in the current management of COVID-19 emergency, the 3T approach, i.e., testing, tracing and treating, is felt as particularly crucial to fight COVID-19 pandemic. Aside from the time-consuming, cost expensive and the many burdensome issues associated with a thorough swabbing, adopting easy-to-make criteria such as "drive-thru-swab" may exacerbate the burden of critical biases and pre-analytical errors, which may impair the analytical reliability of these tests. This manuscript addresses some major points about.

Current Status

The pro-active recommendation from World Health Organization (WHO) to use the window opportunity during lockdown and fight against the SARS-CoV-2 pandemic, included the burdensome planning to thoroughly swabbing any subject. This approach is still adopted in the form called 3T, i.e., testing, tracing and treating. The urgency to trace up asymptomatic infected people and address SARS-CoV-2 outbreak is currently a matter of debate, despite subject's symptomatology is the best warning. This is carried out in order to prevent any infectious contact and slow down viral spread, a concern particularly felt in Italy to date [1]. The huge availability in reverse transcription-quantitative polymerase chain reaction (RT-qPCR) arrays, settled to monitor COVID-19 positivity by swabbing the highest amount of asymptomatic subjects, may show a decidedly high rate of failure, due to the many related bias and analytical errors [2–4]. A widespread consideration about COVID-19 led experts to consider asymptomatic subjects, who is believed to represent the majority of individuals, as potentially bearing SARS-CoV-2, if only keeping close social meetings on. during their daily life. This justified the burdensome workflow associated with a thorough swabbing process on the general population.

Very recently, people felt a sort of harassment merged with fear, due to the paroxysmal exposition of images and videos showing pandemic effects, with deaths, sick people and hospitals and probably even caregivers performing swabbing to any motorist on the road, the so-called "drive-thru-swab" or DTS, aroused some concern. Physicians exist suggesting a house-to-house swabbing, in order to thoroughly mapping the highest number of resident people for COVID-19 positivity [5, 6]. DTS appears as a rapid and straightforward approach to collect the highest amount of swabs, but is usually performed in a not standardized environment (open air instead of a lab), often with hasty operators to prevent traffic and in a context particularly enriched in airborne pollutants, such as engine emission exhausts [7, 8].

A paroxysmal seeking the virus is affecting the correctness by which these crucial tests must be performed, particularly if swabbing is carried out in an open air, highly polluted space and without a fully warranted aseptic process. Moreover, the huge need for swabs to probe citizens for COVID-19 positivity is causing warnings about the possible shortage in the availability of safer swab kits, endowed with virus inactivating buffers and preservatives. Pre-analytical errors are more frequent as much with the hasty employment of swabbing, particularly in a DTS context. In order to make easier and safer COVID-19 testing procedures, FDA recommended that people doing a test be supplied with the proper personal protective equipment. This must include protective masks, gowns, gloves, face shields to be worn and be enabled to conduct efficiently their own swab, a procedure that should

✉ Salvatore Chirumbolo
salvatore.chirumbolo@univr.it

[1] Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Verona, Italy

prevent swabbing shallowly in the nose cavity and carelessly in the throat (pharyngeal swabbing) [2].

DTS has the disadvantage to be performed in cumbersome circumstances such as traffic, high polluted environment with coal dust and diesel engine emissions alongside with the need to swab and collect the highest amount of samples very rapidly. Interestingly, engine exhausts with gases and particulate matters as major emissions are particularly able to promote and exacerbate pulmonary sickness caused by viral pathology. In a past report Hahon and colleagues showed that CD1 white Swiss mice undergoing breathing of 2 mg/m^3 of either diesel engine emission (DEE), coal dust or other pollutants as particulate matters for 1, 3 or 6 months, exhibited pulmonary damage (96.5% with diesel exhausts) respect to controls with filtered air (61.2%), just following 3 months exposure. Moreover, a higher influenza virus growth and an increased haemoagglutinin-antibody levels following 6 months exposure to particulate matters were observed [9]. Airborne viruses growth is particularly susceptible to DEE and particulate matter ≤10 μm diameter (PM_{10}), as reported by Harrod et al. for respiratory sincitial virus in C57Bl/6 ice exposed to 1000 μg/m^3 DEE or 30 μg/m^3 PM_{10} for 6 h/day in a week: a dose–response inflammation rapid enhancement occurred in the lungs [10]. Interestingly, preliminary studies recently reported by Setti e al., showed very recently that SARS-CoV-2 viral RNA can be detected in airborne pollutants, such as PM_{10} [11]. Despite the correct observation that these issues are commonly present aside from any DTS, the existence of these pre-analytical factors may hamper the reliability of swabbing and the subsequent RT-qPCR.

Micro- and nano-particulate matter from engine dust and industrial waste in a high urbanized environment may contain micro-organisms, whose genomic fragments may bias the outcome of any drive-thru swabbing [12]. Although further airborne pathogens do not seem to interfere with coronavirus RT-qPCR test, due to its high analytical specificity [13], engine-derived nitrogen dioxide (NO_2) may cause the formation of 8-nitroguanosine in the viral genome, i.e., in viral RNA [14], thus contributing in causing concerns about the reliability of DTS for a rapid COVID-19 diagnosis in the asymptomatic population under thorough screening. These considerations may be crucial for the DTS approach and ask for further investigation.

Technical Issues About COVID-19 Emergency RT-qPCR Testing and Related Bias

The ability to reduce the many pre-analytical biases with DTS depends also on the COVID-19 prevalence in the population and on the risk to be infected by SARS-CoV-2 and spread the virus in the human community. Therefore, any evaluation of the actual risk to be infected from SARS-CoV-2 may adjust the evaluation on how much biased may be the DTS approach in a test context. The prevalence of certain pathology in the population affects significantly the percentage of false positives and false negatives in a test.

DTS serves to reduce and crucially dampen the big concern caused by crowding healthcares with a huge number of RT-qPCR demands. The COVID-19 emergency is asking for a thorough mapping of SARS-CoV-2 positive subjects, in order to test, trace and treat people from COVID-19. In this paroxysmal policy, tests are lesser and lesser warranting the highest quality standards of their performance, if carried out in a hasty fashion. An excellent SARS-CoV-2 RT-qPCR should have the highest sensitivity, in order to reduce the amount of false negative subjects, who could infect people trusting to be negative when they are not. To evaluate the different performances of SARS-CoV-2 RT-qPCR in emergency (hospitals, high-demand circumstances, and so on), where the request for a rapid RT-qPCR assay is particularly burdensome, a Confusion Matrix (SPSS v 24.0) was calculated. In the field of artificial intelligence or computing calculation, a confusion matrix, which can be also called the "wrong classification table", or "error matrix" serves to provide a representation of the accuracy of a statistical classification. In the matrix, each column represents the predicted values, while each row represents the real values. This contributes in calculating the accuracy and fitness of predictive values with real or factual data.

Then, a Matthews correlation coefficient (MCC) [15] was reported on true positive and true negative COVID-19 outputs from publicly available data of the Italian Ministry of Health (https://www.salute.gov.it) and our local University Hospital in Verona (Italy), The MCC (or Φ coefficient) is used to assess the quality of the binary classification (e.g., predicted/real) and is related to the chi square (X^2 distribution). It is a kind of measure of the confusion matrix. MCC has values from 0 (no correlation, low quality) to 1 (maximal correlation, highest quality).

According to the evaluation held with data of SARS-COV-2 positive subjects in Italy collected on April 1st 2020, the COVID-19 RT-qPCR test might reach a sensitivity of 76.96% (MCC = 0.0808). This was quite far from the analytical sensitivity reported elsewhere [16], if considering that the whole bulk of positive subjects here investigated included hospitalized symptomatic subjects, patients in intensive care units (ICUs) and COVID-19-caused deaths, though this evaluation accounted on a 53.90% precision (*P*) and 54.26% accuracy (*A*) (F1 score = 0.6340). On the contrary, if elaborating data from the total amount of positive subjects only including SARS-CoV-2 positive asymptomatic and symptomatic subjects, but not hospitalized individuals, usually quarantined in their homes, we would reach an 84.87% sensitivity (MCC = 0.1798), with (*P* = 66.25%,

$A = 63.95\%$, F1 score $= 0.7441$) [9]. Briefly speaking, SARS-CoV-2 RT-qPCR is much more sensitive in less compulsive and anxious situations, most probably because of bias in the swabbing procedure.

Furthermore, this might suggest that major diagnostic criteria coming from hospitalization parameters are less sensitive than RT-qPCR for COVID-19 in capturing SARS-CoV-2 infected people. This appears quite comprehensible, if many interstitial pneumonia may be associated with COVID-19, without having full certainty about its ethiopathogenetic cause [17]. Furthermore, when considering the overall prevalence of the tested and confirmed SARS-CoV-2 positive subjects at the RT-qPCR, both hospitalized and not, the predictive positive value (PPV), i.e., the probability that subjects with a positive screening test truly have the disease, is 42% (42.0639%), a value that would indicate that the great bulk of asymptomatic people might include a great proportion of subjects who are only "purported" to be COVID-19 infected.

This is the situation within which any swabbing procedure and RT-qPCR test may be set. The analytical performance of such tests is another issue to be fully addressed, in order to highlight the actual effectiveness of RT-qPCR for SARS-CoV-2 in fulfilling emergency needs.

The RT-qPCR Performance

A recent study from Hallmaier-Wacker et al., has shown that the investigation about a microbial population in swab samples is affected by the different choice in storage buffer and extraction kits rather than in amplification methods [2]. As regarding buffers, some criticism has been reported for the use of phosphate balanced buffers (PBS) [2]. As this kind of buffer stabilizes only certain types of cells, the microbial profile may be altered in a subsequent RT-qPCR assay [18]. Past reports showed that sodium phosphate in the PCR transport medium inhibited Taq polymerase [19]. Various reagents, including RNA-later, in swab performance, may have a crucial role in ensuring the reliability of a RT-qPCR test for COVID-19 diagnosis [2].

Anyway, the official test performance usually takes into account patterns from the report by Corman et al. [20], but these patterns are not always taken into full consideration. The commonest expert operator who would like to perform an in-home RT-qPCR for COVID-19, may go through a typical "producer instruction", for example the Diatheva® detection kit (ref D-MBK0091) or the Genesig® Real Time PCR assay from Primedesign™ Ltd, practically simple to be performed, once endowed with a proper thermocycler (for example Applied Biosystem® 7500 Real-Time PCR System, Bio-Rad CFX Connect™ Real-Time PCR Detection System, or Roche® LightCycler 480 II), and a suited

laboratory equipment. The operator is aided in performing the test in an easy, cost effective way, thanks to the very simple instructions provided by the producer, likewise doing a simple enzyme-linked immune assay (EIA) test, but this high feasibility in performing a molecular test may generate the illusive belief that even swabbing is easy to perform Many pre-analytical errors, able to disturb the analytical test performance, occur during the specimen collection and sampling [2]. The existence of these "easy-to-use" bench methodologies is rapidly spreading, due the urgent need to screen the widest number of purported SARS-CoV-2 positive and asymptomatic subjects and hence mapping the epidemiology of a social community to improve further political strategies. However, pre-analytical errors carried over from DTS with different types of reagents are quite impossible to be targeted, if kit instructions do not specify the actual chemical composition of one-step RT-qPCR master mix, COVID-19 primers and probe mix, buffers for resuspension and preparation, control RNA, and so on.

COVID-19 Prevalence and Infection Risk in the Test Management: Some Comment

The majority of bias associated with a huge deal of RT-qPCR on COVID-19 occurs via the thorough need to screen asymptomatic subjects purported to bear potential infectivity. As indicated above, the amount of biases is closely associated with the prevalence of the tested pathology.

Therefore, how many asymptomatic individuals are truly positive to the SARS-CoV-2 presence? Recent reports have elucidated that the average duration from the onset of any symptom to death in Chinese people is 17.8 days (95% credible interval [$_{95}$CrI] $= 16.9$–19.2), with a crude lethality ratio of 3.67% ([$_{95}$CrI] $= 3.56$–3.80), but that these values, being adjusted onto different ages, amounted to a lethality rate of 0.32% ([$_{95}$CrI] $= 0.27$–0.38) for people ≤ 60 years and 6.4% ([$_{95}$CrI] $= 5.7$–7.2) for people ≥ 60 years and even 13.4% ([$_{95}$CrI] $= 11.2$–15.9) in people aged more than 80 years [21]. The odds ratio (OR) allows researcher to use statistics (or probabilities) in order to evaluate the strength of an association between two variables, so to express a qualitative decision on a certain trend, for example calculating how much people died for COVID-19 "because" of COVID-19. The OR is therefore a measure of the probability rate between two variables to evaluate their statistical correlation, indicating no correlation. If the OR value is equal to 1, then it means that the exposure odds in healthy people is equal to the exposure odds in patients, i.e., the risk factor does not affect the appearance of the disease. If the OR value is greater than 1, the risk factor may be implicated in the appearance of the disease; if the OR value is less than 1 the

risk factor is showing actually a defense mechanism against the disease.

In Italy, during the period January 1st–March 31st 2019 a total number of 185.967 people died for reasons very probably far from SARS-CoV-2 infection, as COVID-19 occurred in 2020 (data from the Government Institute of Statistics ISTAT, https://demo.istat.it/index.html). Taking into consideration the bulk of death in the same period in 2020 (data from the Italian Ministry of Health), the OR value for deaths caused by COVID-19 is 2.341364 ($CI_{95} = 2.295$–2.389), so, with OR > 1.0, deaths are more likely to occur in SARS-CoV-2 confirmed positive subjects by swabbing them. Interestingly, among people diagnosed as SARS-CoV-2 positive by the swab RT-qPCR test, an OR > 1 (OR = 13.421005, $CI_{95} = 13.174$–13.672) indicates that swab tests are more likely to be positive in symptomatic subjects (calculated on data from Ministry of Health April 7th 2020, https://www.salute.gov.it/portale/nuovocoronavirus/dettaglioNotizieNuovoCoronavirus.jsp?lingua=italiano&menu=notizie&p=dalministero&id=4417).

Mortality distribution was very similar to the aforementioned data on Chinese population (14.2%, [$_{95}$CrI] = 13.6–15.4, for people aged ≥81 years) [21]. The highly debated initial WHO decision to swab only symptomatic, sick people was correct whereas instead of running a policy aimed at thoroughly swabbing anyone without any apparent symptom of COVID-19, a best practical approach might be to swab thoroughly people aged more than 61 years old and possibly isolate them, even from intra-family contacts [22–24].

This proper evaluation about the dynamics with which viral spreading is occurring within the population, causing symptomatic, sick and hospitalized patients, is particularly crucial to forecast the numerous biases occurring during rapid and easy-to-perform swabbing approaches also on asymptomatic people.

Conclusions

Some fundamental issues have to be further debated in the scientific community regarding RT-qPCR tests by thoroughly swabbing any people in a social community. First, pre-analytical bias could frequently occur in those circumstances, such as DTS, where good laboratory practices are dismissed because of hasty process, environmental pollution mainly due to engine exhausts, but also for chemical reagents used in the test, scarcely equipped performances and poorly warranted portable mini-labs. Second, poor handling procedures in circumstances with low prevalence of COVID-19, may greatly affect the analytical performance of a RT-qPCR test, particularly if carried on asymptomatic subjects. Expanding the debate on the best attitude to be held

in performing swabs is crucial for COVID-19 diagnosis and for addressing the pandemic.

Compliance with Ethical Standards

Conflict of interest The author states he has no conflict of interest.

References

1. Paterlini M (2020) Covid-19: over 300 Italian doctors and scientists call for more testing. BMJ 368:m1274
2. Hallmaier-Wacker LK, Lueert S, Roos C, Knauf S (2018) The impact of storage buffer, DNA extraction method, and polymerase on microbial analysis. Sci Rep 8(1):6292
3. Mackay IM, Arden KE, Nitsche A (2002) Real-time PCR in virology. Nucl Acids Res 30(6):1292–1305
4. Kralik P, Ricchi M (2017) A basic guide to real time PCR in microbial diagnostics: definitions, parameters, and everything. Front Microbiol 8:108
5. Nundy.S, Patel KK (2020) Self-service diagnosis of COVID-19 ready for prime time? JAMA Network (JAMA Health Forum) https://jamanetwork.com/channels/health-forum/fullarticle/2763264. Accessed 16 Mar 2020
6. Minni A, Ralli M, Candelori F, Cialente F, Ercoli L, Parlapiano C, Greco A, De Vincentiis M (2020) Lessons learned from COVID-19 pandemic in Italy. Bosn J Basic Med Sci. https://doi.org/10.17305/bjbms.2020.4847
7. Lenner M (1967) Nitrogen dioxide in exhaust emission from motor vehicles. Atmos Environ (1967) 21(1):37–43
8. Greim H (2019) Diesel engine emissions: are they no longer tolerable? Arch Toxicol 93(9):2483–2490
9. Hahon N, Booth JA, Green F, Lewis TR (1985) Influenza virus infection in mice after exposure to coal dust and diesel engine emissions. Environ Res 37(1):44–60
10. Harrod KS, Jaramillo RJ, Rosenberger CL, Wang SZ, Berger JA, McDonald JD, Reed MD (2003) Increased susceptibility to RSV infection by exposure to inhaled diesel engine emissions. Am J Respir Cell Mol Biol 28(4):451–463
11. Setti L, Passarini F, De Gennaro G et al (2020) SARS-Cov-2 RNA found on particulate matter of bergamo in Northern Italy: first evidence. Environ Res 188:109754. https://doi.org/10.1016/j.envres.2020.109754
12. Liu H, Zhang X, Zhang H, Yao X, Zhou M, Wang J, He Z, Zhang H, Lou L, Mao W, Zheng P, Hu B (2018) Effect of air pollution on the total bacteria and pathogenic bacteria in different sizes of particulate matter. Environ Pollut 233:483–493
13. Pfefferle S, Reucher S, Nörz D, Lütgehetmann M (2020) Evaluation of a quantitative RT-PCR assay for the detection of the emerging coronavirus SARS-CoV-2 using a high throughput system. Euro Surveill 25(9):200152. https://doi.org/10.2807/1560-7917.ES.2020.25.9.2000152
14. Masuda M, Nishino H, Ohshima H (2002) Formation of 8-nitroguanosine in cellular RNA as a biomarker of exposure to reactive nitrogen species. Chem Biol Interact 139(2):187–197
15. Chicco D, Jurman G (2020) The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 21(1):6
16. Younes N, Al-Sadeq DW, Al-Jighefee H, Younes S, Jamal OA, Daas HI, Yassine HM, Nasrallah GK (2020) Challenges in laboratory diagnosis of the novel coronavirus SARS-CoV-2. Viruses 12(6):E582

17. Wuyts WA, Cavazza A, Rossi G, Bonella F, Sverzellati N, Spagnolo P (2014) Differential diagnosis of usual interstitial pneumonia: when is it truly idiopathic? Eur Respir Rev 23(133):308–319

18. Stulberg E, Fravel D, Proctor LM, Murray DM, LoTempio J, Chrisey L, Garland J, Goodwin K, Graber J, Harris MC, Jackson S, Mishkind M, Porterfield DM, Records A (2016) An assessment of US microbiome research. Nat Microbiol 11(1):15015

19. Johnson SR, Martin DH, Cammarata C, Morse SA (1995) Alterations in sample preparation increase sensitivity of PCR assay for diagnosis of chancroid. J Clin Microbiol 33(4):1036–1038

20. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DKW, Bleicker T, Brünink S, Schneider J, Schmidt ML, Mulders DGJC, 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 MPG, Drosten C (2020) Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill 25(3):2000045. https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045

21. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, Cuomo-Dannenburg G, Thompson H, Walker PGT, Fu H, Dighe A, Griffin JT, Baguelin M, Bhatia S, Boonyasiri A, Cori A, Cucunubá Z, FitzJohn R, Gaythorpe K, Green W, Hamlet A, Hinsley W, Laydon D, NedjatiGilani G, Riley S, van Elsland S, Volz E, Wang H, Wang Y, Xi X, Donnelly CA, Ghani AC, Ferguson NM (2020) Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infect Dis 20:669–677

22. Al-Tawfiq JA (2020) Asymptomatic coronavirus infection: MERS-CoV and SARS-CoV-2 (COVID-19). Travel Med Infect Dis 27:101608

23. Lu S, Lin J, Zhang Z, Xiao L, Jiang Z, Chen J, Hu C, Luo S (2020) Alert for non-respiratory symptoms of coronavirus disease 2019 (COVID-19) patients in epidemic period: a case report of familial cluster with three asymptomatic COVID-19 patients. J Med Virol. https://doi.org/10.1002/jmv.25776

24. 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. https://doi.org/10.5607/en20009