On the Breathability Measurement of Surgical Masks: Uncertainty, Repeatability, and Reproducibility Analysis

Juri Taborri, Beatrice Stocchi, Giuseppe Calabrò, and Stefano Rossi

***Abstract*— Wearing protective face mask has been demonstrated as one of the most effective COVID-19 prevention measures. The breathability measurement of surgical masks is normed by the UNI EN 14863:2019. This study aims at evaluating the uncertainty and reproducibility in the breathability (ΔP) measurement, as well as understanding the effects induced by the number and location of the measurement points. A total of 300 surgical masks were tested for breathability. For each mask, 15 measurement points have been considered. Results showed that more than 60% of the tested masks do not satisfy the standard requirements in terms of breathability, with the lower percentage of compliant mask associated with type II. Repeatability was found comparable to the instrument uncertainty related to the digital resolution of the device. A relative percentage variability related to the reproducibility up to 21.2% was found for the mask type II, thus allowing to affirm that different areas of the same mask can lead to different results in terms of breathability. The reproducibility error was found as the main source of uncertainty, leading to an underestimation or overestimation of the breathability value and, consequently, of the number of compliant masks. Finally, the number and location of measurement points represent a further source of uncertainty influencing the final value of the breathability also among the masks produced by the same manufacturer.**

***Index Terms*— Breathability, COVID-19, reproducibility, surgical masks, uncertainty.**

I. Introduction

TO DATE, the infection caused by the virus SARS-CoV-2, which has been defined by the World Health Organization (WHO) as pandemic since March 2020, has led to more than 230 million worldwide cases. Among them, a number of deaths close to 5 million have been certified [1].

Among the most important countermeasures adopted for avoiding the pandemic spread, social distancing and home confinement have been demonstrated as the most effective [2], [3]. Even if the introduction of recognized vaccine [4], to date, only 2.3% of people in low-income countries have received at least one dose [5]. These data allow affirming that the spread of the pandemic is not completely stopped and some prevention measures are still mandatory. In this context, the use of protective individual face masks still remains part of the prevention measure that can limit the spread of respiratory viral diseases [6], [7]. A face mask is a general name given to a mask able to cover the noise and mouth of the user; however, one specific classification of the mask can be conducted by considering the filtration efficiency level, subdividing them into three categories: the respirators, the surgical face masks, and the community masks. Respirators are defined as personal protective devices design to achieve a very efficient filtration of airborne particles (lower than 0.1 μm). Surgical masks are, instead, a medical device able to create a physical barrier between the mouth and nose and potential contaminants in the immediate environments (filter droplets greater than 3 μm) [8]. Conversely, the community masks are not classified as medical devices nor as personal protective equipment and they are made of a range of different materials, either woven or nonwoven layers. It is clear how the spread of the COVID-19 pandemic has completely changed the interest in the face masks among the population, leading to broader use of them. Thus, one of the first consequences of the pandemic spread was the increment in the demand of face masks, causing a severe shortage of surgical face masks, especially for healthcare workers who usually use them, and a collapse of the supply chain [9]. Considering this issue, several studies have been conducted both by higher institutions, aiming at proposing innovative methods for mask production and/or sanification for reuse [10] or to implement models to early estimate new cases of infections [11], and companies that decided to change their productive chain in order to increase the mask demand [12]. In addition, self-made masks made of cloths or other materials, addressed as community masks, have become popular even if they were not certificated.

To deal with this situation, the University of Tuscia, Viterbo, Italy, sets up the multidisciplinary group, named I-SUM, with the aim to provide a scientific support on face masks testing to companies that wanted to produce face masks [13]. I-SUM laboratory is equipped with several instrumentations for the characterization of face mask performance, both considering the respirators and the surgical masks. Among others, the following analyses have to be conducted: 1) bacterial filtration efficiency (BFE); 2) particle filtration efficiency (PFE);

Manuscript received October 7, 2021; revised December 30, 2021; accepted January 3, 2022. Date of publication January 13, 2022; date of current version February 25, 2022. This work was supported by the University of Tuscia, Viterbo, Italy, through the "Actions for research activities related to COVID-19 emergency" Rector under Grant 5445 (05.05.2020). The Associate Editor coordinating the review process was Dr. Rosenda Valdés Arencibia. *(Corresponding author: Juri Taborri.)*

The authors are with the Department of Economics, Engineering, Society and Business Organization, University of Tuscia, 01100 Viterbo, Italy (e-mail: juri.taborri@unitus.it).

Digital Object Identifier 10.1109/TIM.2022.3142754

3) microbial cleanliness (MC); 4) respiratory resistance; and 5) breathability. When considering surgical masks, the UNI EN 14863:2019 [14] depicts the specific analyses that have to be followed before the certification and the commercialization of the masks. They are BFE, breathability, splash resistance, MC, and biocompatibility.

Focusing on breathability, it represents a fundamental parameter to describe the user's comfort in wearing the mask. Since it has been demonstrated as the correct use of the face mask is strictly correlated with the user comfort [15], it results evident how the accurate and precise measurement of the breathability is mandatory. To date, only the study conducted by Tessarolo *et al.* [8] deals with the analysis of the variability in the breathability measurement by testing a sample size of 120 masks. However, in this study, no information regarding the uncertainty and the reproducibility of the breathability measurement is reported. To the best of the authors' knowledge, a wider investigation on the reproducibility and uncertainty of the above reported measure procedure as well as the assessment of the effects induced by the selection of the number and location of the measurement points are still lacking. From this perspective, this study aims at assessing the uncertainty, the repeatability and the reproducibility in the breathability measurement by testing 30 surgical mask models following a method equivalent to that specified in Annex C of the UNI EN 14863:2019 [14]. Specifically, ten samples for each model were tested for a total of 300 masks. In addition, the effects on the breathability value of the number and location of measurement points on the mask were explored. This article is the extended version of a conference proceeding already published in [16], in which the metrological characteristics of the breathability measurement were assessed on a reduced sample size, i.e., nine masks.

II. Surgical Mask Classification and Breathability Measurement

Surgical masks are face masks that provide a physical barrier to particulate material, meeting a certain level of protection from bacterial and particle filtration. The definition of surgical mask and its characteristics are specified in the UNI EN 14863:2019 "Medical Face Masks—Requirements and test methods" [14].

Three types of surgical masks, which are named types I, II, and IIR, are identified considering a combination of requirements in terms of breathability and BFE. Specifically, the mentioned standard requires that the mask-type classification must be conducted by considering four variables: 1) the BFE, which represents the efficiency of the mask material as a barrier to bacterial penetration and it is generally expressed in percentage; 2) breathability (ΔP), which is a measure of the air permeability of the mask and it is expressed as Pa/cm^2; 3) splash resistance pressure (SRP), which indicates that the mask is designed to be fluid resistant to splash and splatter of blood at a given pressure and it is expressed as kPa; and 4) MC, which represents the total number of viable microorganisms on the face mask and it is expressed in cfu/g. Table I reports the normed threshold values to meet for each surgical mask classification type. As observable in the table, only the type IIR is also characterized by splash resistance.

TABLE I
Surgical Mask Types and Classification Criteria

Variable	I	II	IIR
BFE (%)	≥95	≥98	≥98
ΔP (Pa/cm^2)	<40	<40	*<60*
SRP (kPa)	Not required	Not required	≥16.0
MC (cfu/g)	≤30	≤30	≤30

As reported in Section I, in this study, we only focused on the breathability measurement. The specifications for the measurement of the breathability related to a surgical mask are depicted in Annex C of the UNI EN 14863:2019 [14]. Such annex provides indications on both experimental setup and experimental protocol to be performed in order to gather the breathability value. More in detail, the experimental setup can be any sensor system that is composed of instruments that are able to: 1) measure a differential pressure; 2) provide compressed air through the measured surface of the mask; 3) control and adjust the flow rate of the compressed rate by means of a needle valve; and 4) constantly measure the airflow.

As regards the experimental protocol, the measurement must be conducted by imposing a constant airflow equal to 8 l/min and the measured surface of the mask must be equal to 4.9 cm^2. During the measurement, a perfect alignment between the flow rate and the examined area must be guaranteed. The protocol consists in the measurement of pressure difference in, at least, five different areas without overlaying. Finally, the breathability value has to be assumed as the average across ΔP obtained for each tested area. It is to be noticed that no information on the location of the measurement points is reported in the procedure, as well the maximum number of them.

III. Material and Methods

A. Experimental Setup and Protocol

The instrumentation selected for this study was the commercially available GBN701 face mask air flow resistance and differential pressure tester (GBPI, Guangzhou, China). Such a device meets all the requirements reported in Section II and it is in accordance with the UNI EN 14683:2019. The device is equipped with differential pressure sensor and a gas flowmeter. The control unit of the system is able to perform the breathability test by controlling the flow, measuring the differential pressure, and providing the value of pressure drop expressed in Pa/cm^2. The main specifications of the device are reported in Table II. The system was calibrated by the producer and no further calibrations have been performed during the measurements.

Before starting with the measurements, the operator must define the protocol parameters related to the flow rate and the area of the tested surface. Both parameters have been selected in compliance with the values in the UNI EN 14683:2019 and reported in Section II. In addition, the gas path was connected

TABLE II
SPECIFICATION OF THE USED INSTRUMENT

Specifications	Value
Gas source	Compressed air
Air flow range	1 – 10 l/min
Air flow accuracy	±1% FS
Pressure sensor range	500 Pa
Pressure sensor accuracy	±0.5% FS
Sample size	diameter 25 mm
Power supply	220 V, 50 Hz

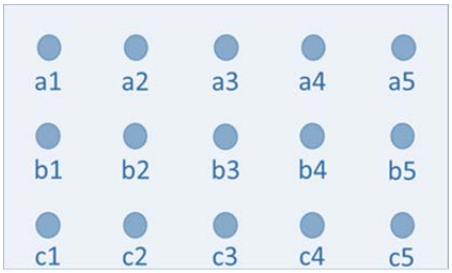

Fig. 1. 3 × 5 matrix of measurement points.

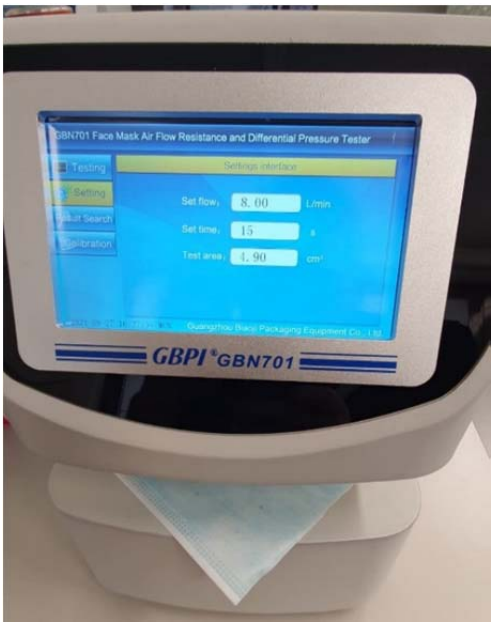

Fig. 2. Used device for the breathability measure with the display indicating the selected values for the protocol conduction and a mask inserted in it.

with the device and the mask has been prepared for the measurement procedure by removing the earloop in order to improve the handling and facilitate its placement on the tested area. Finally, each specimen has been conditioned according to the specifications reported in Annex C.3 of the standard. After conditioning, the operator has drawn 15 measurement points on the mask surface according to a matrix of three rows and five columns. It is worth underlining that the selected measurement points ideally represented the center of 15 different circular areas with a diameter of 25 ± 1 mm, according to the standard, without overlaying and guaranteeing that all the points in the area include mask material. Fig. 1 shows a schematic example of the measurement point selection.

After these preliminary operations, the operator put the gasket on the device in order to avoid air flow leak and then the mask in the device, adjusted the alignment between the center of the measured area and the center of the sensitive area of the pressure system embedded in the device, and clamped the mask in order to avoid leaks. During the measurement, the air flow passed through the examined area, the current pressure was detected by the pressure sensor embedded in the device, and the pressure difference downstream from the mask was computed and provided on the device display.

The experiments included a total of 30 surgical mask models, equally distributed among the three types. For each mask model, we tested ten replicates for an overall number of tested masks equal to 300. All the tested mask models are commercially available, produced by 30 different manufacturers and claimed by the manufacturer as compliant with the standard. Furthermore, the ten masks were a sample among a selected lot of 100 masks per each manufacturer. As regards the material, all the masks are made by two layers of nonwoven fabric spunbond polypropylene (68% of the total) and one intermediate layer of meltblown (32% of the total), with the exception of the fourth manufacturer of type I, the third and fourth manufacturers of type II, and the fifth manufacturer of type IIR who produced masks with three nonwoven fabric spunbond polypropylene layers.

For the repeatability analysis, the measurement was repeated five times per each point; such a procedure was applied on one mask per each company.

As a preliminary study, we evaluated the effect of air flow rate variation on the measure of breathability, by testing three masks for each type and measuring for each of them the breathability value on three points by setting three values of air flow rate, which were 7.9, 8.0, and 8.1 l/min. Such values have been chosen by considering the accuracy of the flowmeter reported in Table II (±1% FS, FS = 10 l/min).

Fig. 2 shows the used device with a surgical mask placed in it.

B. Data Processing and Analysis

Data gathered from all the tested surgical masks was collected and analyzed through customized spreadsheets. First, we evaluated the compliance of the breathability value (ΔP) with the requirements reported in the international standard by computing the average value across the 15 measurement points for each mask and comparing it with the acceptability threshold (t_a). Then, for each type of mask, we also computed the rate of compliant masks, expressed in percentage with respect to the total masks for each type.

1) Repeatability and Reproducibility: In order to evaluate the repeatability [17], [18] of the breathability measurement, the standard deviation (SD) among the five repetitions on the same point was calculated for all the points related to one mask per each company. The relative repeatability error (RRE) was obtained by dividing the SD with the average value and it is expressed as percentage for each mask. The maximum

value of RRE across the 15 points was assumed to assess the repeatability for each company. The higher the RRE, the lower the repeatability.

Within the aim to understand the reproducibility [17], [18] of the breathability measurement considering different areas of the same mask, we computed the SD among the 15 measurement points for each mask. Successively, the intramask relative percentage variability of ΔP was computed by dividing the SD with the average value and it is expressed as percentage for each mask. Considering the ten face masks for each company, the range of the relative variability within the same type of masks and same company was also calculated.

Finally, we considered the breathability variability among the ten different masks of the same manufacturer, computing the SD among the mean values of the breathability of the ten masks. Successively, the intermask relative percentage variability of ΔP was computed by dividing the SD with the average. For all the computations, the higher the relative percentage variability, the lower the reproducibility of the breathability measurement.

2) Uncertainty and Its Effect on the Mask Compliance: To assess the uncertainty [17], [18] associated with the breathability measurement, we considered the following sources of error: 1) repeatability, named u_{rt}, considering the ratio between the SD and the square root of the measurements, i.e., 5; 2) intramask reproducibility, named u_{rd}, considering the ratio between the SD and the square root of the measurement points, i.e., 15; 3) intermask reproducibility, named u^*_{rd}, considering the ratio between the SD and the square root of the number of measurements, i.e., 10; 4) instrumentation, named u_i computed by considering the digital resolution of the instrumentation, i.e., one decimal point for the breathability measurement, as a uniform distribution, thus following the equation $(a/\sqrt{3})$ where a is equal to 0.05 Pa/cm^2 [18]; 5) flowmeter accuracy, named u_F and computed by considering the results of the preliminary analysis, in which we found a maximum variation of the ΔP equal to ± 1 Pa/cm^2 and applying a uniform distribution (where a is equal to 1 Pa/cm^2); and 6) pressure sensor accuracy, named u_P and computed by considering the reported accuracy and applying a uniform distribution (where a is equal to 0.5 Pa/cm^2). Thus, for each mask, the combined uncertainty was calculated by using the following equation:

$$u_c = \sqrt{u_{rt}^2 + u_{rd}^2 + \left(u_{rd}^*\right)^2 + u_i^2 + u_F^2 + u_P^2}.$$

The maximum value of uncertainty across the ten masks of each company was considered and the effects of the uncertainty value in the comparison between each ΔP and t_a were evaluated individually for each mask. Finally, we computed the synthetic indices N that represents the number of compliant masks when considering or not the uncertainty on the final measurement of ΔP.

3) Effects of the Number and Location of Measurement Points on the Mask Compliance: As the last analysis, we wanted to understand the effects induced on the breathability value by the selection of the number and location of measurement points. In fact, it is worth highlighting that the UNI EN 14683:2019 defines sufficiently the computation of the ΔP as the average value across at least five different points. For this reason, we computed all the combinations of the 15 measurement points through the binominal coefficient, according to the equation

$$c = \frac{n!}{k!(n-k)!}$$

where n represents the number of measures, i.e., 15, and k is the subgroup of measures contained in each combination. All the combinations obtainable by varying k from 5 to 14 were considered. For each combination, ΔP was computed as the average value across the k considered measurement points and compared with the acceptability threshold. The percentage of combinations per each k (P_k), which allows to affirm the compliance of the mask, was evaluated. P_k can range from 0% to 100%, representing the absence and the totality of combinations associated with a compliant mask, respectively. This analysis has been performed on those masks for which the comparison between breathability value and acceptability threshold can lead to a disagreement when considering a different number of measuring points. Hypothesizing that masks belonging to the same type and the same company should be characterized by the same value of breathability, obtainable with the same number of measurement points, we assessed the number k^* that guarantees the same outcome of the comparison for all the masks of the same company, i.e., $P_k = 100\%$ or $P_k = 0\%$.

IV. Results and Discussion

A. Mask Compliance

All the surgical masks of two companies, i.e., one for type I and one for type II, were not considered in the study since the measured ΔP values reached the full scale of the instrumentation. The values of the breathability for all the remaining tested masks are reported in Fig. 3. By analyzing the figure, we can affirm that the acceptability criteria defined by the standards were met only by 27.8%, 25.6%, and 52.0% of the masks for type I, type II, and type IIR, respectively. Thus, considering the overall 280 tested masks, only 35.7%, i.e., 100 masks, is characterized by breathability values in compliance with the UNI EN 14683:2019. More specifically, the behavior of the masks within each tested company is reported in Table III. By analyzing the results reported in Table III, it is clear how the overall low percentage of compliant masks is due to the presence of some companies, for which all the ten tested masks did not meet the acceptability threshold. In fact, only three companies for type I and five for type IIR satisfy requirements for ΔP per each mask. Instead, the maximum percentage achieved by companies producing type II mask was 80%, with six companies characterized by the totality of noncompliant masks.

Although with a greater percentage, these findings confirm the ones reported in [8], in which 40% of the tested masks were found not in line with the requirements in terms of both breathability and bacterial filtration efficacy. It is also interesting to highlight that, also in [8], the worst percentages were associated with the type II masks. Since a low breathability,

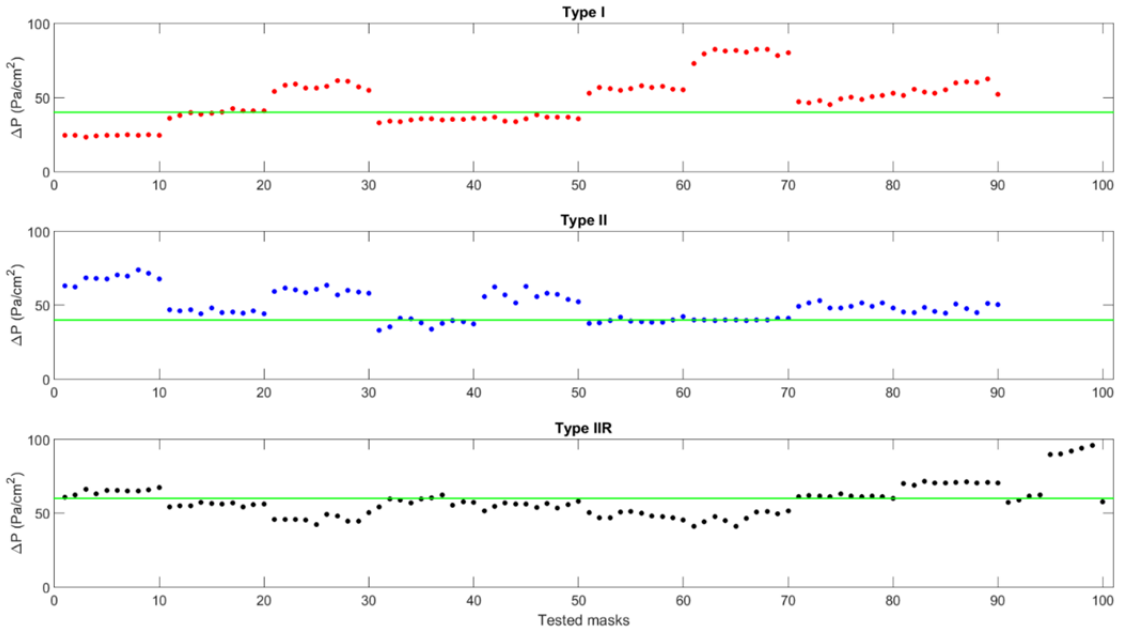

Fig. 3. Breathability values for all the tested masks. Markers represent the mean value of breathability across the 15 measurement points. Green lines indicate the acceptability threshold for the mask compliance.

TABLE III

RATE OF COMPLIANT MASKS FOR EACH TYPE AND EACH TESTED COMPANY. THE COMPANIES ARE DIFFERENT ACROSS THE TYPES OF MASK. #NUMBER INDICATES THE ID CODE SELECTED IN THIS STUDY TO IDENTIFY THE TEN DIFFERENT MANUFACTURERS. * INDICATES THAT FOR COMPANIES, SOME MASKS ARE CONSIDERED NONCOMPLIANT SINCE THEY OVERCOME THE FULL SCALE OF THE INSTRUMENT

	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
I	100%	50%	0%	100%	100%	0%	0%	0%	0%	0%*
II	0%	0%	0%	80%	0%	70%	80%	0%	0%	0%*
IIR	0%	100%	100%	80%	100%	100%	100%	10%	0%	30%

TABLE IV

RANGE OF RRE (%) ACROSS COMPANY BELONGING TO THE SAME MASK TYPE

	RRE range [%]
Type I	0.9 – 1.6%
Type II	1.0 – 1.5 %
Type IIR	0.7 – 1.4%

i.e., a greater value of ΔP, causes a negative impact on the user's comfort [19], these outcomes should be considered in order to avoid an incorrect use of the mask among population due to its discomfort [15], [20].

B. Repeatability and Reproducibility

The range of RRE values across the company related to the same mask type are reported in Table IV.

The results confirm the ones reported in the previous study [16] and permit to assess that the repeatability is robust enough to avoid the repetition of the breathability measurements more than one time on the same point.

Table V indicates the range of the intramask relative percentage variability within each company individually for each type of mask. Considering type I, a relative percentage variability up to 17.9% (SD = 10.4 Pa/cm^2) was found for company I#3, with the lowest value equal to 4.7% (SD = 1.2 Pa/cm^2) for company I#1. Moving to type II, a minimum value of 4.0% (SD = 1.8 Pa/cm^2) was observed for company II#8, whereas up to 21.2% (SD = 11.2 Pa/cm^2) was observed for company for II#9. As for type IIR, a relative percentage variability equal to 2.7% for company IIR#9 (SD = 1.9 Pa/cm^2) was found as the lowest value, with a maximum of 18.5% (SD = 8.9 Pa/cm^2) associated with the company IIR#7 and IIR#10. A comparable value of the relative percentage variability was also found in [8]. Table V also shows the results related to the intermask relative percentage variability. Considering type I, the relative percentage ranged from 2.1% (SD = 0.5 Pa/cm^2) to 7.0% (SD = 4.0 Pa/cm^2). As for type II, a maximum value of intermask relative percentage variability was found equal to 7.5% (SD = 2.8 Pa/cm^2), with a minimum value of 1.5% (SD = 0.6 Pa/cm^2), whereas a range equal to 1.0% (SD = 0.7 Pa/cm^2)–8.3% (SD = 3.9 Pa/cm^2) was associated with type IIR.

Generally, these findings allow affirming that the reproducibility error is not negligible, suggesting that different areas of the same mask can produce different values in terms of breathability. Such outcomes can be ascribed to the

TABLE V

RANGE OF THE INTRAMASK RELATIVE PERCENTAGE VARIABILITY (%) AND INTERMASK RELATIVE PERCENTAGE (%) VALUE FOR EACH EXAMINED COMPANY CONSIDERING INDIVIDUALLY THE MASK TYPE. THE COMPANIES ARE DIFFERENT ACROSS TYPES OF MASK

	Type	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
Intra-mask	I	4.7-8.1 %	7.2-15.8 %	6.2-17.9 %	7.4-11.0 %	9.9-15.7 %	6.6-13.5 %	6.8-12.5 %	5.8-10.1 %	7.9-12.5 %	-
	II	7.9-18.4 %	4.6-9.7 %	9.9-15.1 %	5.7-11.7 %	6.8-14.9 %	5.0-8.9 %	6.6-10.4 %	4.0-12.7 %	9.9-21.2 %	-
	IIR	5.9-10.0 %	4.8-10.9 %	5.7-14.0 %	7.8-10.0 %	5.6-9.0 %	8.0-16.7 %	9.9-18.5 %	4.7-8.3 %	6.2-18.5 %	2.7-5.2 %
Inter-mask	I	2.1%	4.7%	4.2%	2.9%	3.9%	2.7%	3.7%	4.9%	7.0%	-
	II	5.1%	2.6%	3.2%	7.5%	6.5%	4.1%	1.5%	3.6%	5.7%	-
	IIR	3.1%	1.8%	5.0%	4.1%	3.4%	4.3%	8.3%	1.3%	1.0%	1.9%

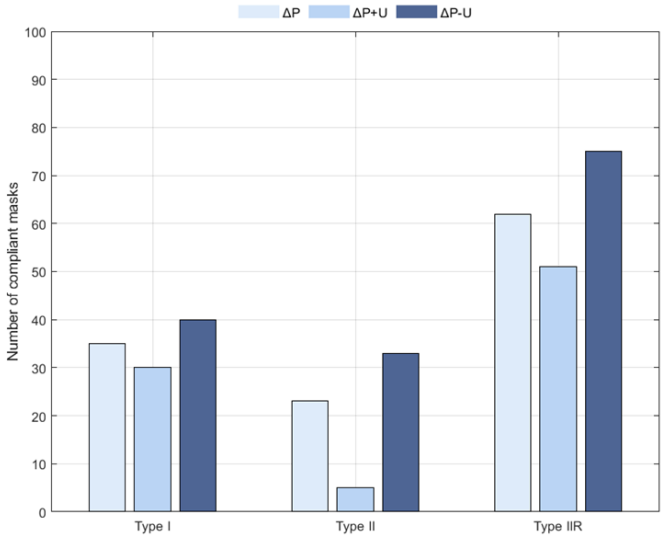

Fig. 4. Value of N, which is the number of compliant masks, when considering the uncertainty interval for each mask type.

heterogeneity of the material composing the mask, as also assessed in [21]. In addition, the found intramask repeatability should also be ascribed to a precompression of the measurement points, consequence of the pressure exerted by the gasket on the measurement point when measuring adjacent areas.

C. Uncertainty and Synthetic Index N

The found not-negligible values of the reproducibility error are also found as the main contribution of the uncertainty, whereas the repeatability error is comparable with the instrumentation uncertainty. More specifically, considering per each company both the highest value of the intramask relative percentage variability, the value of the intermask relative percentage variability, the contribution associated with the maximum repeatability, the instrument reading, the uncertainty related to the airflow meter, and the one associated with the pressure sensor, a combined uncertainty u_c ranged from 1.6 to 5.8 Pa/cm^2 for type I, from 2.0 to 6.9 Pa/cm^2 for type II, and from 1.8 to 5.4 Pa/cm^2 for type IIR was obtained.

Fig. 4 shows the obtained values of the synthetic index N, i.e., the number of masks per each type that satisfies the standard requirements when considering: 1) only the breathability value ΔP; 2) the breathability value increased by $U\Delta P + U$; and 3) the breathability value decreased by $U\Delta P - U$; where U is the expanded uncertainty corresponding to two times the combined uncertainty u_c.

Focusing on the figure, it is evident as the uncertainty can substantially change the outcome of the breathability test when data are compared with the acceptability threshold. Specifically, a decrement of compliant masks equal to 5%, 18%, and 8% was associated with types I, II, and IIR, respectively. Consequently, considering only the average value of ΔP can lead to both underestimate and overestimate the actual mask breathability value. These findings should be considered during the certification process of surgical masks; in fact, it appears evident as the information regarding the uncertainty level of the breathability cannot be neglected for a more truthful characterization of the mask comfort. The high value of uncertainty and its significant impact on the mask compliance is ascribed as the direct consequence of the above-discussed low reproducibility due to the selection of the tested areas on each mask and, consequently, to the heterogeneity of the materials composing the mask. This high level of uncertainty mainly related to the selected tested area confirms the necessity to repeat the measure in more points of the mask in order to decrease its influence in the final value of breathability. A deep analysis on the number of test repetitions to choose for each mask appears to be mandatory.

D. P_k Index: The Influence of Number and Location of Measurement Points on the Mask Compliance

Table VI shows those companies for which the breathability tests performed on the ten masks always give the same compliance results varying the number of measures and the location of the measurement points.

For the remaining companies, we evaluated the possibility to identify the minimum number of measurement points that permit to achieve the same compliance results for all the ten masks. The trends of P_k, i.e., the percentage of combinations that satisfy the standard requirement per each number k of measurement points, are reported in Fig. 5. From the analysis of the figure, it is clear how only for 5 out 12 companies was possible to identify a number k^* for which all the combinations of measurement points led to the same results in terms of mask compliance, specifically $k^* = 12$ was found for company I #5 and IIR #5, $k^* = 14$ for company IIR #1, $k^* = 9$ for company IIR #2, and $k^* = 6$ for company IIR #7. The results reported in Fig. 5 should also be discussed by considering the AQL requirements reported in the standards [14]. By assuming that the production lot is equal to 100 as the selected lot, the AQL4.0 level requires the test of 13 masks and only one

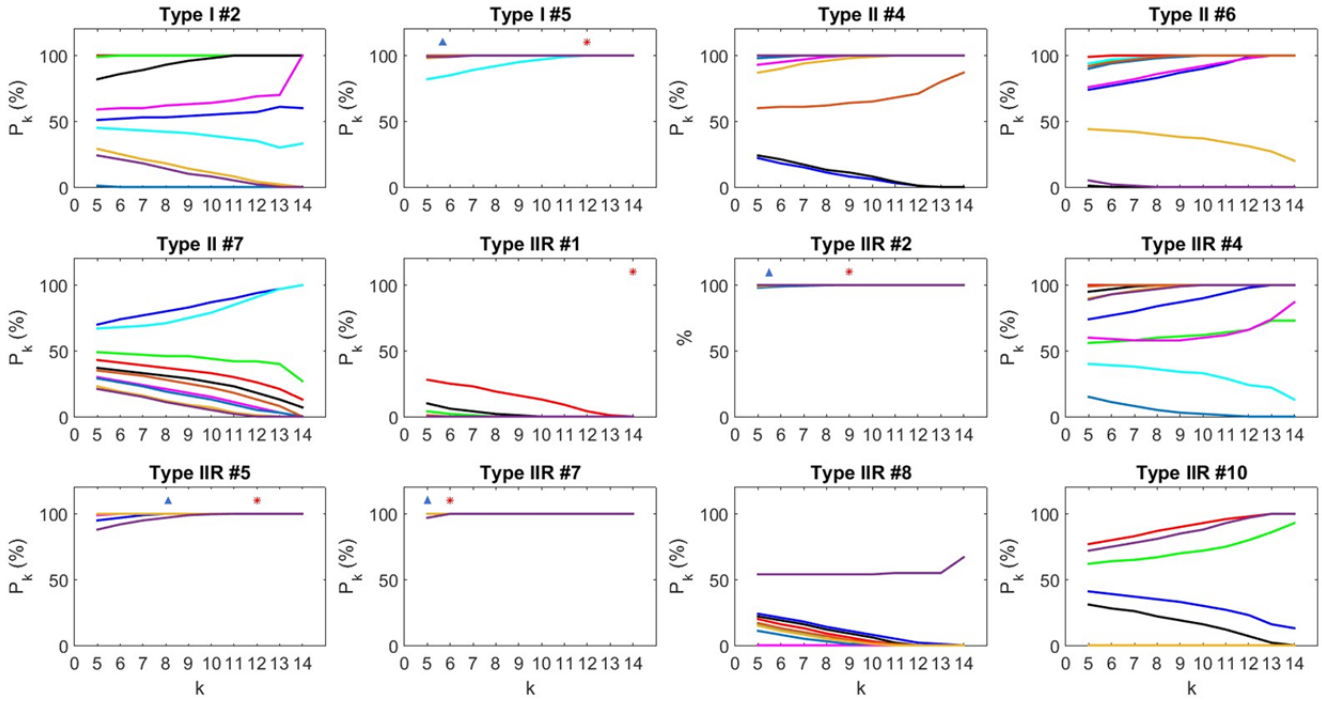

Fig. 5. Percentage (P_k) of combinations that satisfy the acceptability criterion according to the number of measurement points (k). In each graph, the behavior of the ten masks for the specific company is reported. The red * indicates number for which all the masks within the same company give the same results in terms of mask compliance. The blue triangles represent, instead, the value of k^* to overcome for avoiding noncompliance with the AQL of 4%.

TABLE VI

COMPANIES RELATED TO EACH MASK TYPE FOR WHICH THE COMPLIANCE RESULTS ARE INDEPENDENT ON THE SELECTION OF THE NUMBER AND THE LOCATION OF MEASUREMENT POINTS

Company	Compliance results
I #1	100% compliant masks
I #3	0% compliant masks
I #4	100% compliant masks
I #6	0% compliant masks
I #7	0% compliant masks
I #8	0% compliant masks
I #9	0% compliant masks
II #1	0% compliant masks
II #2	0% compliant masks
II #3	0% compliant masks
II #5	0% compliant masks
II #8	0% compliant masks
II #9	0% compliant masks
IIR #3	100% compliant masks
IIR #6	100% compliant masks
IIR #9	0% compliant masks

can be noncompliant. However, it should be noted that we tested the ten available masks for each manufacturer. Based on this approach, we can affirm that no value of k^* permits to meet the AQL requirement for the following masks: type I #2, type II #4, #6, and #7, and type IIR #4, #8, and #10 since at least two noncompliant masks were observed. Conversely, for the remaining manufacturers, we can assess that: 1) types I #5 and IIR #2 do not meet the AQL criterion if k^* is equal to 5; 2) type IIR #5 does not meet the AQL criterion if k^* is equal or lower than 7; and 3) type IIR #7 meets the AQL requirement for any value of k^*.

By looking at the graphs, it is easy to understand how the selection of the number and the location of the measurement points leads to a different classification of mask compliance also within the ten masks of the same company. Consequently, it is possible to incur in a high rate of false positives or false negatives when randomly selecting five mask areas since it is shown as different sets of five areas lead to different results in terms of mask compliance. In addition, in some cases, the P_k curve is characterized by a decreasing trend, indicating that the selection of a greater number of measurement points results in a greater probability to find a noncompliant mask, whereas a lower number generally leads to a positive test, i.e., satisfying the acceptability threshold. It implies that considering only five measurement points could facilitate the passing of thresholds reported in Table I and it could be the reason of the low number of masks that reaches the acceptability criteria defined by the standards despite as reported by the manufacturers.

Furthermore, the worst occurrence is for those companies in which it is impossible to find a common number of measurement points on which performing the breathability measurement on masks that should be nominally identical. It suggests that the selection of the five areas on the mask on which conducting the breathability test can give different compliance results since there is a not negligible variability on the mask manufacturing. Finally, some P_k curves are characterized by an increasing trend, reaching 100% with the

TABLE VII
NUMBER OF COMPANIES THAT SATISFY THE STANDARD REQUIREMENTS WITH ALL THE MASKS BASED ON THE SELECTED NUMBER k OF MEASUREMENT POINTS

k	Number of companies
5	4 (2 for type I, 2 for type IIR)
6	5 (2 for type I, 3 for type IIR)
7	5 (2 for type I, 3 for type IIR)
8	5 (2 for type I, 3 for type IIR)
9	6 (2 for type I, 4 for type IIR)
10	6 (2 for type I, 4 for type IIR)
11	6 (2 for type I, 4 for type IIR)
12	8 (3 for type I, 5 for type IIR)
13	8 (3 for type I, 5 for type IIR)
14	8 (3 for type I, 5 for type IIR)

increase of the measurement points. Specifically, Table VII shows the number of companies out of the tested 30 that satisfy the standard requirements with all the ten masks based on the number k of measurement points. These results confirm that the increment of the measurement points leads to a lower possibility to reach a wrong classification for the breathability of the masks. In addition, it is confirmed as no companies producing type II mask are associated with ten compliant masks regardless of the number of measurement points.

By summarizing, these outcomes suggest that the breathability measure is affected by a significant variability that can cause several disagreements on the mask compliance also when testing the same mask. For these reasons, a parameter related to reproducibility should be added to ΔP in the breathability-evaluation procedure in order to reduce the probability to experience misclassification in terms of compliance. Moreover, an increase in the number of tested areas should be considered to reduce the misleading of the final classification. Finally, it is useful to underline as the results of the present study could be exploited not only in the context of COVID-19 pandemic since the use of surgical masks is daily required by healthcare workers, without forgetting Eastern population in which the use of masks increased in the recent decades before COVID-19 for counteracting the pollutants and infectious agents [8].

E. Study Limitations

Three limitations should be considered when considering the above-reported outcomes.

First, the instrument used for the breathability measurement requires the use of compressed air to generate and maintain the airflow; however, the standard procedure reported in Annex C is based on the presence of a vacuum pump rather than a compressor. Nevertheless, it is worth noticing that most of the commercially available certified instruments for the measurement of breathability used compressed air, as well other research papers focused on this topic [8], [22].

A second limitation is related to the AQL computation; in fact, not knowing the effective number of production lots, we can only speculate on possible modifications on compliant masks due to the AQL level considering the production lot equal to the selected one, i.e., 100 samples. In fact, it should be considered that a greater number of production lots leads to a different value of the minimum sample size to test and, consequently, the number of allowed noncompliant masks.

Finally, even though the tested masks are all claimed as compliant with the standard by the manufacturer, we are not able to confirm the procedure applied to assess the conformity of these masks.

V. CONCLUSION

The demand for surgical mask is drastically increased due to the COVID-19 pandemic. Breathability is one of the parameters to test for the certification of a face mask as surgical. In this study, we tested 300 commercially available surgical masks in order to evaluate the repetability, the reproducibility and the uncertainty in the breathability measurement. In addition, the effects induced on breathability values by the selection of number and location of measurement points were assessed. Findings reveal as more than 60% of the tested masks do not satisfy the standard requirements. Although the repeatability error is comparable to the instrument uncertainty, the reproducibility related to the selected tested area affected the output of the breathability test since it is the main source of uncertainty of the measurement. Therefore, it emerges how the absence in the standards of information regarding the uncertainty and the number and location of the measurement points represents a nonnegligible error source, leading to a disagreement in breathability values also when examining masks nominally identical. Thus, a further parameter focused on the measurement reproducibility should be introduced in the standards in order to reduce misclassifications related to the breathability compliance of surgical masks.

ACKNOWLEDGMENT

This work has been carried out within the framework of the COVID-19 activities of the I-SUM/CINTEST group.

REFERENCES

[1] *Coronavirus Update (Live): 90, 190, 146 Cases and 1, 936, 857 Deaths From COVID-19 Virus Pandemic–Worldometer*. Accessed: Jan. 10, 2021. [Online]. Available: https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1?%22

[2] L. Matrajt and T. Leung, "Evaluating the effectiveness of social distancing interventions to delay or flatten the epidemic curve of coronavirus disease," *Emerg. Infectious Diseases*, vol. 26, no. 8, pp. 1740–1748, Aug. 2020.

[3] U. Glogowsky, E. Hansen, and S. Schächtele, "How effective are social distancing policies? Evidence on the fight against COVID-19," *PLoS ONE*, vol. 16, no. 9, Sep. 2021, Art. no. e0257363.

[4] K. Rawat, P. Kumari, and L. Saha, "COVID-19 vaccine: A recent update in pipeline vaccines, their design and development strategies," *Eur. J. Pharmacol.*, vol. 892, Feb. 2021, Art. no. 173751.

[5] E. Mathieu *et al.*, "A global database of COVID-19 vaccinations," *Nature Hum. Behav.*, vol. 5, no. 7, pp. 947–953, Jul. 2021.

[6] *Advice on the Use of Masks in the Community, During Home Care and in Healthcare Settings in the Context of the Novel Coronavirus (COVID-19) Outbreak*. Accessed: Sep. 29, 2021. [Online]. Available: https://www.who.int/publications/i/item/advice-on-the-use-of-masks-in-the-community-during-home-care-and-in-healthcare-settings-in-the-context-of-the-novel-coronavirus-(2019-ncov)-outbreak

[7] B. Wang, Y. Zhao, and C. L. P. Chen, "Hybrid transfer learning and broad learning system for wearing mask detection in the COVID-19 era," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021.
[8] F. Tessarolo *et al.*, "Testing surgical face masks in an emergency context: The experience of Italian laboratories during the COVID-19 pandemic crisis," *Int. J. Environ. Res. Public Health*, vol. 18, no. 4, p. 1462, 2021.
[9] D. Rondinone, A. B. Godfrey, and A. A. Khankar, "Challenges in the supply chain for personal protective equipment (PPE) during COVID-19," *J. Textile Apparel Technol. Manage.*, vol. 12, pp. 1–23, Mar. 2021.
[10] I. Armentano *et al.*, "Polymer materials for respiratory protection: Processing, end use, and testing methods," *ACS Appl. Polym. Mater.*, vol. 3, no. 2, pp. 531–548, Feb. 2021.
[11] R. R. Sharma, M. Kumar, S. Maheshwari, and K. P. Ray, "EVDHM-ARIMA-based time series forecasting model and its application for COVID-19 cases," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.
[12] G. Czifra and Z. Molnár, "COVID-19 and industry 4.0," Fac. Mater. Sci. Technol., Slovak Univ. Technol., Bratislava, Slovakia, Tech. Rep., 2020, doi: 10.2478/rput-2020-0005.
[13] U. of Tuscia. *I-SUM*. Accessed: Sep. 10, 2021. [Online]. Available: http://www.unitus.it/it/unitus/i-sum/articolo/progetto-i-sum
[14] *UNI EN 14683:2019*. Accessed: Jan. 10, 2021. [Online]. Available: http://store.uni.com/catalogo/uni-en-14683-2019-292195?josso_back_to=http://store.uni.com/josso-security-check.php&josso_cmd=login_optional&josso_partnerapp_host=store.uni.com
[15] M. Machida *et al.*, "Incorrect use of face masks during the current COVID-19 pandemic among the general public in Japan," *Int. J. Environ. Res. Public Health*, vol. 17, no. 18, p. 6484, Sep. 2020.
[16] J. Taborri, B. Stocchi, G. Calabro, and S. Rossi, "Repeatability and reproducibility in the breathability measurement of surgical masks," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, Jun. 2021, pp. 1–6.
[17] *International Vocabulary of Metrology–Basic and General Concepts and Associated Terms (VIM)*, Joint Committee Guides Metrol., Sevres, France, 2012.
[18] (2008). *Evaluation of Measurement Data-Guide to the Expression of Uncertainty in Measurement Évaluation des Données de Mesure-Guide Pour L'Expression de L'Incertitude de Mesure*. Accessed: Nov. 25, 2021. [Online]. Available: http://www.bipm.org
[19] K.-P. Lee, J. Yip, C.-W. Kan, J.-C. Chiou, and K.-F. Yung, "Reusable face masks as alternative for disposable medical masks: Factors that affect their wear-comfort," *Int. J. Environ. Res. Public Health*, vol. 17, no. 18, p. 6623, Sep. 2020.
[20] J. Khubchandani, D. Saiki, and J. Kandiah, "Masks, gloves, and the COVID-19 pandemic: Rapid assessment of public behaviors in the united states," *Epidemiologia*, vol. 1, no. 1, pp. 16–22, Nov. 2020.
[21] D. Coclite *et al.*, "Face mask use in the community for reducing the spread of COVID-19: A systematic review," *Frontiers Med.*, vol. 7, p. 1060, Jan. 2021.
[22] M. H. Bagheri *et al.*, "Filtration efficiency, breathability, and reusability of improvised materials for face masks," *Aerosol Sci. Technol.*, vol. 55, no. 7, pp. 817–827, 2021.

Juri Taborri received the master's degree (*cum laude*) in biomedical engineering and the Ph.D. degree in industrial and management engineering from the Sapienza University of Rome, Rome, Italy, in 2014 and 2017, respectively, with a thesis focused on novel solutions for human motion analyses performed through wearable sensors.

He is currently an Assistant Professor in mechanical and thermal measurements at the University of Tuscia, Viterbo, Italy. His research interests include biomechanical measurements, validation and metrological characterization of innovative sensor systems, electromyography and muscle synergies, and machine learning algorithms for human motion identification and control of robotic devices for rehabilitation.

Beatrice Stocchi received the bachelor's degree in industrial engineering from the University of Tuscia, Viterbo, Italy, in 2021, by presenting an experimental thesis on the evaluation of the metrological characteristics of breathability measurement of surgical masks, where she is currently pursuing the master's degree in mechanical engineering.

Giuseppe Calabrò received the Ph.D. degree in electrical sciences and engineering from the Mediterranea University of Reggio Calabria, Italy, in 2003.

He is currently a Full Professor of electrical engineering and magnetic fusion with the Department of Economics, Engineering, Society and Business Organization, University of Tuscia, Viterbo, Italy. He is actually the Rector's delegate for relations with enterprises and fundraising, the Vice-Head of the Ph.D. program in engineering and environment, and the Director of the CINTEST Research Center. From 2003 to 2017, he worked as a Researcher Scientist at the ENEA-Fusion Energy Department, Frascati, Italy. His present research collaborations include ENEA-EUROFUSION, CNR, INFN, ASIPP, EPFL-SPC, Max-Planck IPP, PPPL, and several industrial partners. His current research interests include fusion plasma equilibrium configurations modeling, computational electromagnetics, electromagnetic nondestructive evaluation, big data science, and artificial intelligence (AI).

Stefano Rossi graduated in mechanical engineering (Hons.) from the Sapienza University of Rome, Rome, Italy, in 2004. He received the Ph.D. degree from the University of Padova, Padua, Italy, in 2008. In 2008, he defended his Ph.D. thesis in "Measurements in Mechanical Engineering" at the University of Padova.

He is currently an Associate Professor and a Ph.D. Advisor at the Department of Economics, Engineering, Society and Business Organization, University of Tuscia, Viterbo, Italy. He is actually the Chair of the Engineering and Design Programs. He is also the Coordinator of the master's degree in mechanical engineering and the Head of the Industrial Measurements and Experimental Biomechanics Laboratory. He has more than ten years of experience in the design of experimental setup to investigate metrological aspects of sensors in the fields of experimental mechanics and biomechanics. His research is focused on the design of new robotic systems for the movement recovery and the individuation of indices to objectively measure human motor performance by means of inertial systems, optoelectronic systems, and electromyography.