Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Coronavirus disease 2019 (COVID-19) is the respiratory disease first identified in humans at Wuhan, China, in December 2019, and was declared as ‘pandemic’ by the World Health Organization (WHO) on 11th March 2020. The use of personal protective equipment (PPE), quarantining, isolation, social distancing, respiratory and hand hygiene are further recommended to control the spread of disease.

Face masks are an essential component of the hygiene are further recommended to control the spread of infection.2 Face masks are an essential component of the World Health Organization (WHO) on 11th March 2020.1 The use of personal protective equipment (PPE), quarantining, isolation, social distancing, respiratory and hand hygiene are further recommended to control the spread of infection.2 Face masks are an essential component of the PPE and the popular varieties are respirator masks (N95/KN95, surgical, and fabric masks. N95 mask (US standard) usually consists of 3 layers- an outer and inner layer of melt-blown polypropylene, and melt-blown polypropylene filter material in the middle. KN95 (Chinese standard) mask consists of 4 layers- outermost and innermost layer of melt-blown polypropylene, along with 2 layers made respectively of melt-blown polypropylene filter material, and cotton fibers, in the middle.3 Both the respirators are designed with 95% ability to filter the particles in the environment using static electricity which requires tight-fitting, and hence might be uncomfortable.4 Surgical mask (medical/procedure mask) approved by the United States Food and Drug Administration (USFDA) consists of usually three layers- an outer and inner layer of blue, non-woven fabric polypropylene, with a filtering material in the middle made from ‘spin bond-melt blown spun bond’ technology.5 A fabric mask (Cloth mask) is usually 2-layered and tightly woven either from a single fabric of cotton or with other fabric material.3 Both surgical and fabric mask protects from large droplets and does not require tight-fitting.4

Air filtration efficiency of N95 is reported to be 10% higher compared to KN95 under the pristine conditions as the ‘filter’ layer of N95 is 8-fold thicker and the ‘fitting’ factor is 10-fold higher than KN95.3 Some of the KN95 masks are reported to perform similar to N95, although not approved by National Institute for Occupational Safety and Health (NIOSH, USA).5 Undoubtedly, respirators with an established level of air filtration ability achieve better filtration of airborne particles than surgical masks if used properly and continuously.6 Hence, these respirators are recommended for healthcare professionals who are in very close contact with patients and involved in aerosol-generating procedures (AGPs). A review suggested that
conventional surgical masks do not offer protection against high-risk AGPs. However, another review suggested that surgical mask offers a similar level of protection against viral respiratory infection during non-aerosol generating procedures as N95 respirators. Hence, this surgical mask is commonly recommended for all other healthcare professionals, who are not in very close contact with patients and not involved in AGPs. The air filtering efficiency of the cloth mask is generally lower than that of surgical masks and/or respirators and yet, better than no masks at all. However, another review reported that the efficiency can be improved when made up of two layers of different fabrics (cotton & chiffon, cotton & silk) or cotton quilt made with multiple layers. Therefore, this cloth mask is recommended by the WHO and United States Centers for Disease Control and Prevention (US, CDC) for the general public in non-medical settings, along with 6 feet of physical distance.

In this pandemic, the person has to breathe with the mask on, as well as communicate wearing it. A review of the literature suggests that wearing masks impairs communication in medical and non-medical healthcare workers. Self-perceived symptoms of increased vocal fatigue, vocal discomfort, and vocal effort in individuals who wear the mask for professional activities compared to essential activities have also been reported. Prevalence of voice problems can vary from mild to severe, in healthcare workers wearing a mask during this pandemic. Voice problems can be a result of the type of mask used, as these masks vary in terms of composition, filter, fitting type, and thickness (layers). Till date, a systematic review on the voice changes across types of masks is not available. In this context, this study will provide a comprehensive review of the available literature to the general population as well as various professionals regarding the effect of wearing different masks on healthy individuals’ voices. This can be useful in educating individuals regarding the selection of appropriate masks, without considerable aberrations of voice in order to have optimal communication.

**METHOD**

Study design: A systematic review was completed by following the ‘Preferred Reporting Items for Systematic Review and Meta-Analyses Protocols’ (PRISMA-P) (Figure 1).
Study Selection Criteria: Studies published in scientific journals (‘print’ as well as ‘pre-print’ versions) involving only original data that satisfied the following criteria were considered: studies with the objective of investigating voice measures in individuals wearing masks during COVID-19 pandemic, males and/or females aged 18 years and older wearing masks were considered. In addition, studies involving all types of research study designs were included. Effect of wearing the mask on voice, in terms of perceptual/ acoustic/ aerodynamic/ physiological parameters should have been the target outcome reported, and the studies should be published in English. The exclusion criteria involved eliminating case reports, case series, letters to editors, short communication, conference abstracts as they have been considered as the lowest level of evidence.17

Search Strategy: The literature search strategy followed the PICO Model, where P (Population of interest) is individuals wearing the face mask, I (Intervention) and C (Comparison) are not applicable, and O (Outcome) is voice measures. Cochrane collaboration18 and Preferred Reporting Items for Systematic review and Meta-Analyses (PRISMA) guidelines19 were followed to carry out the review. Electronic databases such as ‘Cochrane Central Register of Control Trials (CENTRAL), Cumulative Index to Nursing and Allied Health Literature (CINHAL), Embase, Ovid Medline, ProQuest Medical Library, PubMed/Medline, Scopus, and Web of Science (WOS)’ were used to search articles published between the year 1st January 2020 and 30th April 2021. A combination of keywords (Table 1) was used to identify the primary studies in the databases. These keywords were used to develop search strings as per the requirements of the databases using Boolean operators (AND, OR). The title, abstract, and full-text screening was done by two independent reviewers (SS and KS). A manual search of various journals did not provide any additional relevant studies (SS and KS). In total, ten journal articles reporting primary studies (8 published & 2 in pre-print) were included in the systematic review.

Data Extraction and Management: The data was synthesized from the 10 studies using the tabular form to provide descriptive summaries of selected studies to readers. The data was tabulated in terms of the name of the author(s), title/time of publication, type of mask with manufacturer details, study design, participants, gender, age-range, inclusion/exclusion criteria, questionnaire/instrument used, measures studied, and findings. The age across studies ranged from 18 to 69 years. As the necessary information required to answer the study questions had been published in the research articles, no additional communication with the authors was done.

Quality appraisal and level of evidence analysis of included studies: A methodological ‘quality appraisal’ of each study was carried out by developing a tool of 16 questions with a ‘yes/no’ response (Table 2). The tool was developed based on standard appraisal forms ‘critical review form-quantitative studies’,20 standard guidelines of ‘quality assessment tool for observational studies of cross-sectional, case-control & cohort research design’21 and also quantitative study appraisal, part of ‘mixed methods appraisal tool (MMAT) version 2018’.22 The response ‘Yes’ was rated as 1 and ‘No/Not reported’ was rated as 0. Studies that scored

| TABLE 1. Keywords Used to Identify the Primary Studies |
|-----------------------------|------------------|
| Keywords | Synonyms |
| #1 | “Effect” OR “influence” OR “impact” |
| #2 | “Face mask” OR “face cover” OR “face guard” OR “face shield” OR “respirator” OR “mask filter” |
| #3 | “Voice” OR “speech” OR “speaking” OR “talking” OR “communication” |
| #4 | “COVID-19” OR “2019 coronavirus disease” OR “coronavirus disease 2019” OR “2019 novel coronavirus” OR “2019-new coronavirus” OR “2019-nCoV” OR “coronavirus” OR “COVID” OR “corona” OR “SARS-CoV-2” |

| TABLE 2. The Methodological Quality Appraisal Tool |
|-----------------------------|------------------|
| Item | Questions |
| 1. | Was the research objective clearly stated? |
| 2. | Was the study population specified and defined? |
| 3. | Were all the participants recruited from the same population in the same period? |
| 4. | Were inclusion and exclusion criteria for the participants defined in the study? |
| 5. | Was the research design mentioned? |
| 6. | Was the sampling method mentioned? |
| 7. | Was the pilot study tested or reviewed before the actual study? |
| 8. | Was the sample size justification provided? |
| 9. | Were the outcome measures defined clearly? |
| 10. | Were ethical approval and informed consent obtained? |
| 11. | Was the order of ‘mask condition’ and ‘no mask condition’ randomized? |
| 12. | Were the outcome assessors (or raters) blinded to the participants’ exposures/interventions (‘mask condition’ and ‘no mask condition’)? |
| 13. | Was the loss to follow-up after baseline testing (mask condition /no mask condition) 20% or less? |
| 14. | Were the tools tested for their reliability and validity? |
| 15. | Was the statistical analysis appropriate to answer the research question? |
| 16. | Was there a mention of the settings under which the findings could be applied? |
TABLE 3. Quality Assessment of Primary Studies Included

| Study ID      | Items | Quality appraisal score (%) |
|---------------|-------|-----------------------------|
| Ribeiro et al 14 | 1 1 1 0 0 0 1 1 0 0 1 0 1 1 | 62.50 |
| Heider et al 16 | 1 1 1 1 0 0 1 1 1 1 0 1 0 1 | 68.75 |
| Corey et al 25  | 1 0 0 0 0 1 1 1 0 0 1 0 1 1 | 37.5 |
| Bottalico et al 26 | 1 1 1 1 0 0 1 1 1 1 0 1 0 1 | 68.75 |
| Cavallaro et al 27 | 1 1 1 0 0 0 1 1 1 1 0 1 0 1 | 56.25 |
| Magee et al 28  | 1 1 1 1 0 0 0 1 1 1 0 1 1 1 | 62.50 |
| Fiorella et al 29 | 1 1 1 1 0 0 0 1 1 1 0 1 0 1 | 56.25 |
| Nguyen et al 30  | 1 1 1 1 1 0 0 0 1 1 1 0 1 1 1 | 75.00 |
| Rahne et al 31  | 1 1 1 1 1 0 0 0 1 1 1 1 0 1 1 | 62.5 |
| Lin et al 32    | 1 1 1 1 1 0 0 0 1 1 1 1 0 1 | 56.25 |

Rating of 1 for ‘Yes’ and 0 for ‘No/Not reported’; Quality appraisal score (%) = Total score obtained for a study/ Total number of items (16), multiplied by 100; Weak: 0% - 33.9%; Moderate: 34% - 66.9%; Strong: 67% - 100%.

RESULTS

The methodological appraisal was carried out using the tool mentioned in (Table 3). The quality appraisal score (%) was used to categorize the studies; 7 studies were categorized as ‘moderate’ and 3 as ‘strong’ (Table 3).

As per the GRADE approach, 8 out of 10 studies were rated to have a ‘low’ level of evidence.14,15,25-27,29,31,32 Because, six studies had employed cross-sectional observational study design,14,15,26-27,29,32 whereas 2 studies with experimental designs25-31 did not randomize the order of ‘mask/no mask condition’ (item 11) and the outcome assessors (or raters) were not blinded to the participants ‘mask condition’ and ‘no mask condition’ (item 12). The remaining 2 out of ten studies were rated as ‘moderate’.28,30 Because, in 1 study, outcome assessors (or raters) were blinded to the participants ‘mask condition’ and ‘no mask condition’ (item 12) although observational study design was used,28 and in another study, the order of ‘mask/no mask condition’ was randomized (item 11).30

Table 4 provides the important identifying details of ten studies in terms of the title of the study, time of publication, country of study, types of masks studied, and its manufacturer. These studies have investigated the influence of wearing respirators (N95 & KN95), surgical, various fabric, and transparent masks on voice-related parameters. Out of ten studies, three studies were carried out in the USA, two each in Australia & Italy and one each in Brazil, China, and Germany. Out of ten primary studies, eight were already published and two were pre-print articles that investigated the effect of wearing a mask on various voice measures.

Table 5 provides the summary of ten included studies investigating the effect of wearing a mask on voice production. These studies have used few self-rating scales and various acoustic measures to identify and quantify the voice problem following wearing a mask during this pandemic. Out of ten studies, 2 studies14,15 have used previously validated self-reported questionnaires such as the Brazilian Portuguese version of vocal fatigue index (VFI),33 self-perception of vocal tract discomfort scale (VTDS)34 and Spanish validated Voice Handicap Index (VHI-10)35 to identify vocal fatigue, discomfort and voice handicap problems respectively. Same 2 studies14-15 have also used author-developed questionnaires such as self-perception of vocal effort-5 point Likert Scale (0 = Never, 5 = Always) and an anonymous self-perceived voice symptoms-23 item questionnaire survey to identify vocal effort, and perceived voice problems. The remaining eight studies have researched the acoustic measures7-32 and one has also investigated an aerodynamic measure along with.29

DISCUSSION

The current systematic review aimed to identify the available literature on changes in voice in individuals wearing various face masks. Total 10 studies suited the selection criteria and were included for the final review. The discussion is based on the changes in voice in terms of variables of self-reported voice changes, acoustic measures, and aerodynamic measures noted across studies.
Self-reported measures of voice

Self-reported measures of voice provide one’s perception about the existence of voice-related problems. A study,14 has exclusively investigated the self-reported voice measures in terms of vocal fatigue, discomfort, vocal effort, perceived voice problem, and compared between the WG group and EAG group. The WG group had reported significantly higher vocal fatigue index score, VTDS score (both frequency, intensity), and perceived vocal effort rating compared to the EAG group. In another study,15 self-reported voice problem was compared among three groups (Group I, II & III) based on working duration. Irrespective of group type, 21.56% (47) had significantly perceived mild voice problems, 11.10% (24) had severe voice problems.
In continuation to the same study,\textsuperscript{15} 58 (26.24\%) had an abnormal VHI\textsuperscript{35} score (>11) compared to ‘no mask condition’. VHI score was significantly affected with mask condition in terms of duration, mask types & use (single or simultaneous), different healthcare personnel and no. of working hours. VHI score was also more in individuals who wore the mask for 4–8 and 8–12 hours duration than those who used it only for 1–4 hours. VHI score was more in individuals who wore two masks simultaneously (self-filtering & surgical) than those who used a single mask. Nurses had more VHI than physicians, medical residents, and Speech-Language Pathologists (SLPs). Those working in intermediate and intensive care units had more scores than those working in general wards. Those healthcare personnel who worked 44 hours per week (Group I) had more VHI followed by those who worked 24 hours per week with 3 free days subsequently (fourth shift) (Group III), and 22 hours per week (Group II).

Overall, the perceptual measures of voice across reviewed studies suggested that the individuals who wore a mask for their professional activities, as well as essential activities, perceive significant symptoms of voice problems in terms of vocal fatigue, discomfort, and effort. Health-care professionals have a potential risk of developing voice disorders based on their mask characteristics (duration, type, and use) and working characteristics (no. of working hours & working place).

**Acoustic measures of voice**

Along with perceptual measures, acoustic measures are also commonly used in routine clinical voice assessment procedures. The acoustic analysis of voice provides information on the source and filter characteristics of the signal. Generally, various types of instruments are used to objectively quantify the voice problems in terms of the fundamental frequency, intensity, perturbation, and noise-related measures. Many studies carried out on fundamental frequency related measures ($f_0$ mean, Hz; $f_0$ CoV, %), intensity related measures (mean intensity, dB; Intensity prominence dB; p95 Intensity), and perturbation related measures (Jitter, %; Shimmer, %) reported no significant difference between ‘no mask’ and ‘mask’ conditions with surgical, N95/KN95, and cloth masks.\textsuperscript{27-30} In contrary to these findings, a study\textsuperscript{32} is in partial agreement with the above findings which has reported no significant difference between ‘no mask’ and ‘surgical mask’ conditions for $f_0$, measure and significant difference between ‘no mask’ and ‘surgical mask’ conditions present for SPL (dB), Jitter (%) and shimer (%). That is, the intensity was increased and perturbations were decreased compared to the ‘no mask’ condition. The authors have attributed the decreased frequency and amplitude perturbation to the increasing trend in fundamental frequency and significantly increased intensity. Also, in another study,\textsuperscript{20} 35\% of participants reported an increase in loudness and a 65\%, decrease in vocal loudness for the ‘mask condition’.

Overall, most of the acoustic parameters related to ‘fundamental frequency measures’ are unaffected by wearing a mask. But, wearing a mask might alter ‘intensity-related’ measures. Both frequency and amplitude perturbations are also unaffected or even reduced by wearing a mask. That is, the acoustic measures of ‘pitch’ as well as ‘voice quality’ correlates are not affected by wearing a mask. But the acoustic measures of ‘loudness’ correlate might be affected by wearing a mask. The changes in loudness measures can be attributed to the effects of the mask such as acoustic attenuation property of mask, difficulty in coordination of speech and breathing, altered feedback, vocal adjustment based on proximity along with social-distancing.\textsuperscript{46-31}

Mask acts as a low-pass acoustic filter for speech ie mask attenuates the intensity in the speech frequency ranges esp. at mid and high frequency, attenuating the speakers’ volume of the voice.\textsuperscript{38,46-51} The use of the mask leads to a pressure drop across the mask resulting in reduced airflow intake. This in turn results in difficulty in coordinating respiratory-laryngeal systems during speech production leading to reduced loudness.\textsuperscript{14-15} The occlusion effect of the mask might lead to altered self-auditory feedback and the user might tend to speak softly than normal\textsuperscript{47,23,30} or loudly than normal.\textsuperscript{32} The ‘vocal effort adjustment’ based on the proximity\textsuperscript{46} especially ‘6-feet social distancing’ between speaker and the listener which has to be followed in this pandemic might have led to the perceived decrease or increase in loudness. In general, a mask user has increased demand and overload to the glottis to compensate for the effects of a mask (acoustic attenuation, breathing difficulty, altered feedback) along with ‘social distancing’ on voice projection. But, those individuals who speak wearing the mask for a prolonged time, compensating these effects along with improper vocal adjustments might misuse or abuse their voice. Such users self-perceive the increased vocal effort, discomfort, fatigue, and also are at risk for developing voice disorders.

Studies carried out on noise-related measures such as CPPS and HNR reported no significant differences between ‘no mask’ and ‘mask’ conditions.\textsuperscript{27,28,32} On the contrary, another study is in partial agreement with the previous study, reporting no significant difference between ‘no mask’ and mask conditions for CPPS, but reporting a significant increase in HNR while wearing either a surgical mask or KN95 compared to ‘no mask’.\textsuperscript{30} The HNR is a more sensitive acoustic measure that quantifies the relative noise in the voice signal\textsuperscript{49} and it was found to be improved with surgical and KN95 masks\textsuperscript{40} This inners that though altered auditory feedback reduces the vocal monitoring ability leading to change in voice quality,\textsuperscript{50} wearing a mask facilitates optimal voice quality. That is, wearing a mask reduces vocal constriction and increases frontal resonance similar to covering the mouth with either hand,\textsuperscript{51} or using a semi-occluded ventilation mask,\textsuperscript{52} facilitating more effective voice production in individuals with dysphonia as well as in those without dysphonia. However, further research is warranted to
### TABLE 5.
Summary of Studies Included for the Systematic Review of the Effect of Wearing a Mask on Voice Measures

| Study ID | Study design/Participants (n/ gender, age range) | Inclusion/Exclusion criteria | Participants’ profession | Type of questionnaire/instrument used - task | Measures studied | Findings during ‘no mask’ condition | Findings during ‘mask’ condition | Level of evidence |
|----------|-------------------------------------------------|------------------------------|--------------------------|---------------------------------------------|-----------------|--------------------------------------|----------------------------------|-------------------|
| 14       | Cross-sectional; n = 488 (122 males, 346 females); 18 – 59 years | Individuals diagnosed with a voice disorder/ COVID-19, and non-residents of Brazil were excluded. | Wearing a mask for professional and essential activities during the pandemic (WG group); Wearing mask only to perform essential activities during the pandemic (EAG group). | VFI - Brazilian Portuguese version 13 Self-perception of VTDS - Brazilian Portuguese version 14 | VFI total mean score VTDS frequency sub-scale total mean score VTDS severity sub-scale total mean score | WG = 21.66 ± 8.99; EAG = 19.61 ± 7.96 | WG = 19.78; EAG = 4.98 | Low |
| 15       | Cross-sectional; n = 221 (54 males, 167 females); 18 – 59 years | Three groups of participants: Group I-44 hours/week shift with 8 hours daily; Group II-22 hours/week shift with 4 hours daily; Group III-Fourth shift modality with 24 hours on-duty followed by 3 days off. Not stated | Nurses, physicians, medical residents, physical therapists, speech-language pathologists (SLP), and nursing assistants | Self-Perceived Voice Symptoms - 23 item questionnaire survey Spanish validated VHI-10 16 | No. of participants (%) VHI score | 67.43% (147) no voice problem, 21.56% (47) mild and 11.10% (24) severe | 160 had normal score & 58 abnormal scores (>11) | Low |
| 25       | n = 1 | Not stated | Not stated | Speechreading recorded for 30 s | Sound level attenuation (dB) | - | Transparent – 8; Cloth – 6.36; N95 – 5.4; Surgical – 2.8; KN95 – 2.6; Peak attenuation (>1 kHz: N95 – 6; Surgical; KN95 – 4 | Low |
| 26       | n = 1, male | A male speaker with a standard American English dialect. | Not stated | ‘Consonant-Nucleus-Consonant (CNC)/word list of monosyllabic words with equal phonemic distribution recorded using compact disc’ 20 | Sound level attenuation (dB) | - | Fabric – 4.2; N95 – 2.9; Surgical – 2.3 | Low |
| 27       | n = 50 (20 males, 30 females); 26 – 69 years | Individuals with an ability to sustain vowels for at least 10 s were included. Individuals with a current or history of voice disorder and voice therapy taken, or with a history of respiratory infection in 2 weeks before recording, were excluded. | Not stated | ‘Praat’ software (version 6.1.16)-Sustained vowel /a/ at comfortable pitch and loudness. Mean pitch (Hz) Number of pulses Mean HNR Jitter (%) Shimmer (%) | Mean pitch (Hz) 185.52 ± 55.12 Number of pulses 575.00 ± 198.76 Mean HNR 20.92 ± 3.47 Jitter (%) 0.298 ± 0.124 Shimmer (%) 3.165 ± 1.572 | 183.52 ± 51.13 574.18 ± 157.88 20.91 ± 3.44 0.327 ± 0.134 3.34 ± 1.420 | Low |

(Continued)
| Study ID | Study design/ Participants (n/ gender, age range) | Inclusion/ Exclusion criteria | Participants’ profession | Type of questionnaire/ instrument used - task | Measures studied | Findings during ‘no mask’ condition | Findings during ‘mask’ condition | Level of evidence |
|----------|--------------------------------------------------|------------------------------|--------------------------|---------------------------------------------|----------------|-----------------------------------|---------------------------------|-----------------|
| 28       | n = 7 (4 males, 3 females); 21 – 39 years        | Individuals with English speaking ability and with no history of voice, cognition, neurological impairments were included. | Not stated | “Praat” 28 Sustained vowel /a/; Reading phonetically balanced ‘Grandfather Passage’ 28 text | Mean intensity (dB) 71.54±3.89 | Surgical: 71.73±4.34; N95: 71.85±4.31; Cloth: 72.26±2.78 | Surgical: 71.77±4.3; N95: 72.95±4.37; Cloth: 73.52±2.84 | Moderate |
|          |                                                  |                             |                          | 85 intensity                                | 72.66±3.76 | Surgical: 72.87±4.3; N95: 73.51±4.37; Cloth: 73.52±2.84 |                                     |                 |
|          |                                                  |                             |                          | Mean f0 frequency (Hz) 155.80±63.25         | 0.71±0.09 | Surgical: 0.77±0.08; N95: 0.65±0.08; Cloth: 0.65±0.06 |                                     |                 |
|          |                                                  |                             |                          | f0 CoV (%) 0.71±0.09                          | Surgical: 0.77±0.08; N95: 0.65±0.08; Cloth: 0.65±0.06 |                     |                                     |                 |
|          |                                                  |                             |                          | CPPS (dB) 19.52±2.74                          | Surgical: 19.16±1.87; N95: 19.99±2.19; Cloth: 19.34±2.1 |                     |                                     |                 |
|          |                                                  |                             |                          | HNR 20.30±3.66                                | Surgical: 18.11±3.26; N95: 21.88±3.77; Cloth: 21.37±2.16 |                     |                                     |                 |
| 29       | n = 60 (24 males, 36 females), 26 – 69 years     | Individuals with the ability to sustain a vowel for at least 10 s were included. Individuals with a current or history of voice disorder and voice therapy, or with a history of respiratory infection in 2 weeks before recording, were excluded. | Workers of the ENT Department of the Polyclinic Hospital, University of Bari “Aldo Moro”, Italy. | “Praat” 41 Sustained vowel /a/ at comfort-able pitch and loudness. | MPT(s) 25.58±5.79 for males, 20.64±3.97 for females | Surgical: 25.29±5.52 for males, 19.86±4.17 for females | Surgical: 24.78±1.82* N95: 23.59±4.09*; Cloth: 20.23±4.96 | Low |
|          |                                                  |                             |                          | Mean f0 frequency (Hz) 131.8±24.40 for males, 213.34±41.40 for females | 1.51±0.23 | Surgical: 1.55±0.16; N95: 1.64±0.5; Cloth: 1.51±0.24 | Surgical: 132.27±24.33 for males, 211.19±33.38 for females |       |
|          |                                                  |                             |                          | Mean Vocal intensity (dB) 68.54±6.72 for males, 70.07±5.88 for females | 0.32±0.06 | Surgical: 0.38±0.11; N95: 0.31±0.06; Cloth: 0.32±0.06 | Surgical: 68.67±6.48 for males, 68.07±6.14 for females |       |
|          |                                                  |                             |                          | No. of glottal pulses 429.08±111.91 for males, 644.47±136.54 for females | 0.38±0.26 for males, 0.30±0.12 for females | Surgical: 427.8±98.71 for males, 653.86±109.64 for females | Surgical: 426.79±98.74 for males, 652.75±109.66 for females |       |
|          |                                                  |                             |                          | Mean HNR (dB) 18.91±3.80 for males, 21.08±3.76 for females | 0.38±0.26 for males, 0.30±0.12 for females | Surgical: 19.46±4.08 for males, 20.87±3.46 for females | Surgical: 19.46±4.08 for males, 20.87±3.46 for females |       |
|          |                                                  |                             |                          | Jitter (%) 0.32±0.06 for males, 0.30±0.12 for females | 0.32±0.06 for males, 0.30±0.12 for females | Surgical: 0.36±0.16 for males, 0.34±0.12 for females | Surgical: 0.36±0.16 for males, 0.34±0.12 for females |       |
|          |                                                  |                             |                          | Shimmer (%) 4.46±1.22 for males, 2.98±1.22 for females | 4.46±1.22 for males, 2.98±1.22 for females | Surgical: 3.92±2.21 for males, 3.51±1.64 for females | Surgical: 3.92±2.21 for males, 3.51±1.64 for females |       |

(Continued)
| Study ID | Study design/ Participants (n/ gender, age range) | Inclusion/ Exclusion criteria | Participants’ profession | Type of questionnaire/ instrument used- task | Measures studied | Findings during ‘no mask’ condition | Findings during ‘mask’ condition | Level of evidence |
|----------|-------------------------------------------------|------------------------------|--------------------------|--------------------------------------------|-----------------|-----------------------------------|-----------------------------|------------------|
| 30       | Within-subject study design; n = 16 (4 males, 12 females) | Individuals speaking English and without smoking habits, voice, hearing problems were included. | Otolaryngologists, SLPs, and a Nurse | “Praat” [4] - Sustained vowel /a/ for at least 10s; Reading 3rd phrase of ‘CAPE-V’ [4] and 2nd and 3rd sentences of ‘Rainbow’ passage. [4] | Mean spectral level at low (0-1kHz) & high frequency (1-8kHz) | Low/High Spectral ratio (0-1kHz/1-8kHz) | HNR ratio | - | N95 -5.2 dB* | Surgical - 2dB * | Moderate |
| 31       | Prospective exploratory experimental, n = 2 (1 male, 1 female) | Not stated | Not stated | 20 sentences from ‘German matrix Oldenburg Sentence test’ (OLSA) [4] | Sound level attenuation (dB) | - | - | Surgical-65.8±7.2 | N95-66.7±7.5 | Low |
| 32       | n = 53 (25 males, 28 females); 20 – 68 years | Participants with normal voice, no history of voice/ articulation/ anatomical problems, voice therapy, and vocal tract surgery, were included. | Not stated | “MDVP” (Version 3.3.0, Kay Pentax) & ‘Praat’ (version 6.1.13)-Sustained vowel /a/ at comfortable pitch and loudness for at least 3 s for acoustic measures, and as long as possible for MPT measure. | f0 (Hz) | 130.51±21.98 – male; 224.46±27.09 - female | SPL(dB) | 72.58±4.13 – male; 70.00±4.04 - female | Jitter (%) | 1.14±0.72-male; 1.24±0.92-female | Surgical:0.75±0.63-male; 0.98±0.69-female | Low |
|          |                                  |                              |                          |                                             | Shimmer (%) | 4.44±2.83- male; 4.53±1.48 -female | NHR (dB) | 0.13±0.03-male and female | Surgical:3.30±3.11-male; 3.98±0.69-female | Surgical:13.00±21.03-male; 228.91±27.26 -female | Surgical:73.59±3.54-male; 71.37±4.08; Surgical:0.75±0.63-male; 0.98±0.69-female | Low |
|          |                                  |                              |                          |                                             | CPP (dB) | 12.43±4.01- male; 12.63±3.91-female | MPT(s) | 20.36±7.38- male;16.36±6.11-female | Surgical:19.88±7.69-male;17.21±7.39-female | Surgical:13.00±21.03-male; 228.91±27.26 -female | Surgical:73.59±3.54-male; 71.37±4.08; Surgical:0.75±0.63-male; 0.98±0.69-female | Low |

Note. VFI, Vocal Fatigue Index; VTDS, Vocal Tract Discomfort Scale; VHI, Voice Handicap Index; CAPE-V, Consensus Auditory-Perceptual Evaluation of Voice; MDVP, Multi-Dimensional Voice Program; f0 CoV, Fundamental frequency Coefficient of Variation; CPPS, Cepstral Peak Prominence Smoothed; HNR, Harmonic- Noise Ratio; MPT, Maximum Phonation Time; f0, Fundamental frequency; SPL, Sound Pressure Level; NHR, Noise-Harmonic Ratio; CPP, Cepstral Peak Prominence.
acoustically quantify and confirm the negative and/or positive effects of different masks on voice quality.

There have been few studies\textsuperscript{8,30} that unraveled the effect of wearing a mask on spectral measures—‘spectral tilt’ and ‘spectral ratio’. Among these, the first study\textsuperscript{28} reported significantly lower spectral tilt for surgical mask and N95 conditions compared to ‘no mask’ for phonetically balanced text ‘the Grandfather Passage’.\textsuperscript{30} However, the relatively altered spectral tilt of both masks compared to ‘no mask’ has been attributed to the filtering property of the mask which acts as a low pass filter, attenuating the high frequency sounds of the speech signals. In another study,\textsuperscript{30} ‘spectral ratio’ for the 3rd sentence of the rainbow passage text\textsuperscript{43} reading was significantly lowest for ‘no mask’, and higher for surgical mask, and highest for KN95. This implies that users increase vocal effort to compensate for the mask (surgical or N95), while the mask alters spectral energy in high-frequency region increasing spectral slope. Therefore, wearing a mask increases spectral measures (spectral tilt as well as slope) of the speech signal, more so in filters (N95/KN95) followed by surgical mask compared to ‘no mask’ condition.

Four studies have quantified the amount of acoustic attenuation across masks. The first study\textsuperscript{25} played 30 s recorded reading speech of human talker while wearing a mask, played from 2-meter distance of ‘listener’ position in a sound-treated room. The highest attenuation measured at the listener position has been reported for transparent masks followed by cloth masks with varying material and weave, N95, surgical mask, and KN95 in 2 to 16 kHz. Most masks had little effect below 1 kHz and the maximum peak attenuation was reported for N95 followed by KN95 and surgical mask. The second study\textsuperscript{26} used recorded CNC words\textsuperscript{16-37} played in the presence of speech-shaped white noise through head and torso simulator (HATS, 45BC KEMAR, GRAS, Holte, Denmark) wearing no mask and mask conditions in a sound booth. The highest sound attenuation was for fabric masks followed by N95, and surgical masks across the octave bands from 63 Hz to 16 kHz. High frequencies from 2–16 kHz were attenuated for all the masks. The third study\textsuperscript{30} recorded connected speech of participants wearing a mask in a practicing clinic with the ambient noise of 33.3 dBA. They reported the highest significant attenuation of mean spectral level amplitude for KN95 compared to surgical mask in 1-8 kHz. The fourth study,\textsuperscript{31} played recorded sentences with various types of background noise through “Dummy Head” (KU 100, Neumann, Berlin, Germany) under surgical and N95 conditions. They reported the average maximum 8 dB amplitude spectra reduction at frequencies above 1 kHz by N95 mask, and frequencies above 2 kHz by a surgical mask, for olnoise female noise; olnoise male noise; ISTS noise.\textsuperscript{35} But, maximum amplitude attenuation occurred at two frequencies with N95 mask and single frequency by a surgical mask. For speech along with white noise background, maximum amplitude reduction was reported for N95 mask compared to surgical mask.

The amount of amplitude attenuation depends on mask composition, filter, fitting, and thickness\textsuperscript{53,25-26,30-31} If the mask has a less porous material composition, it dampens the oscillation of the air particles and absorbs sound energy especially in the high-frequency range.\textsuperscript{26} Mask with higher filter characteristics (KN95) attenuated more spectral amplitude compared to mask with lower filter characteristics (surgical).\textsuperscript{30} Similarly, N95 attenuates amplitude at more frequencies compared to surgical masks,\textsuperscript{31} that is tightly fitting masks (N95) and tightly woven fabric masks attenuate more than loosely fitting masks (surgical). Likewise, a multilayered mask attenuates more compared to single-layered or any loosely fitted mask.\textsuperscript{25} The 3-layered surgical mask and double-layered loosely woven 100% cloth mask made up of cotton jersey or cotton plain material perform acoustically better compared to other opaque and transparent masks.\textsuperscript{25}

Considering maximum amplitude attenuation and the number of frequencies at which attenuation occurs in various speech bands (1-8 kHz) across studies reviewed, the amplitude attenuation was highest for transparent mask followed by cloth mask, N95, KN95, and surgical mask. Also, it can be inferred that though there is an inverse relationship between breathability and filtration efficiency during the breathing mechanism,\textsuperscript{34} the current review confirms that there is a direct relationship between breathability and acoustic performance of the mask during speech production. Hence, manufacturers could consider (re) designing the masks with reduced sound attenuation without compromising droplet blocking efficiency.

**Aerodynamic measures of voice**

The laryngeal aerodynamic analysis captures the coordinated nature of respiratory and laryngeal functions during speech production. Most commonly used laryngeal aerodynamic measures include Maximum Phonation Time (MPT in s), Counts Per Breath (CPB), Mean Air Flow Rate (MAFR in L/s), Estimated Sub-Glottic Pressure (ESGP in cmH\textsubscript{2}O), and Laryngeal Airway Resistance (LAR in cmH\textsubscript{2}O/L/s). 2studies\textsuperscript{29,32} have been carried out to check the effect of a surgical mask on MPT, a measure of the efficiency of the glottal closure, as well as the respiratory system.\textsuperscript{25} Both these studies\textsuperscript{29,32} have reported no significant difference for MPTs measured with the ‘surgical mask’ compared to ‘no mask’. This implies that wearing a surgical mask doesn’t significantly influence the efficiency of the respiratory-laryngeal system during speech production. This can be attributed to the loose-fitting design of the mask which facilitates the normal intake of air. Future researchers could also consider aerodynamic measures of voice using an accelerometer, and non-invasive physiological measures using electroglottography. These will provide a comprehensive understanding of respiratory and phonatory functions in terms of glottic coaptation for reduced respiratory system...
efficiency and glottic configuration across masks when compared to ‘no mask’ during speech production.

Limitations and further recommendations
The original articles reviewed here have studied commonly used, commercially available masks, and their effects on voice measures. The current review is limited to articles published in English and did not look into the possible literature in other languages. Also, the included studies had limitations in terms of sample size, age range, study designs, randomization, blinding, and voice outcomes studied, waning the quality of evidence offered. Therefore, future studies can be of more value when higher quality researches are conducted with a larger sample size, randomized control trials along with the reliability data of perceptual and/or instrumental measures used. This would help in better understanding the properties of the masks and their effects on various measures of voice production. Further, with the increase in the transmission of the COVID-19 infection, even ‘double masking’ has been recommended, and more studies on such masking conditions possibly affecting voice production can be taken up as an objective by future studies.

CONCLUSION
In this pandemic, everyone is advised to wear an appropriate mask to protect oneself and others from the viral infection. Hence, it is essential to understand the impact of wearing the mask on voice production. The current systematic review suggested that wearing mask results in vocal fatigue, discomfort, and perceived voice problems. Further, few acoustic measures of voice confirmed the attenuation of signal amplitude at speech frequencies by transparent mask followed by cloth mask, N95, KN95, and surgical mask, leading to the increased vocal effort.

Considering the influence of masks on voice measures, the surgical mask is better compared to other masks for effective voice production. Although the level of protection from respiratory threat offered by a surgical mask is limited, it is still advisable for everyone (including health professionals) when not in very close contact with the patient and not involved in AGPs. Further, in a direct teaching, offline classroom scenario, ‘surgical mask’ can reduce the vocal load of teachers, smoothen the teacher-student interaction, and facilitate student learning. Although wearing a particular mask is an individual choice, users should be aware of the risk of developing voice problems with a mask. Vocal healthcare strategies such as avoiding speaking in background noise, using a microphone along with the mask, and using augmentative alternative communication (AAC) should also be considered for optimal voice production.

REFERENCES
1. World Health Organization. WHO Director-General’s opening remarks at the media briefing on COVID-19 - 11. 2020. Available at: https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020. Accessed March 11, 2020.
2. World Health Organization. Rational use of personal protective equipment for coronavirus disease (COVID-19) and considerations during severe shortages: interim guidance. World Health Organization; 2020.
3. Zangmeister CD, Radney JG, Vicenzi EP, et al. Filtration efficiencies of nanoscale aerosol by cloth mask materials used to slow the spread of SARS-CoV-2. ACS Nano. 2020;14:9188–9200.
4. Bench press. The difference between N95s, surgical masks and cloth masks. 2020. Available at: https://mgriblog.org/2020/04/29/the-difference-between-n95s-surgical-masks-and-cloth-masks/. Accessed on 29 April.
5. Yin W, Cheng D, Patel SH, et al. KN95 and N95 respirators retain filtration efficiency despite a loss of dipole charge during decontamination. ACS Appl Mater Interfaces. 2020;12:54473–55480.
6. Dau NQ, Peled H, Lau H, et al. Why N95 should be the standard for all COVID-19 inpatient care.
7. Sobti A, Fathi M, Mohiktar MA, et al. Aerosol generating procedures in trauma and orthopaedics in the era of the Covid-19 pandemic: What do we know? Surgeon. 2021;19:e42–e48.
8. Bartoszek JJ, Farooqi MA, Alhazzani W, et al. Medical masks vs N95 respirators for preventing COVID-19 in healthcare workers: a systematic review and meta-analysis of randomized trials. Influenza Other Respir Viruses. 2020;14:365–373.
9. Mondal A, Das A, Goswami RP. Utility of cloth masks in preventing respiratory infections: a systematic review. medRxiv; 2020.
10. Sharma SK, Mishra M, Mudgal SK. Efficacy of cloth face mask in prevention of novel coronavirus infection transmission: a systematic review and meta-analysis. J Educ Health Promot. 2020;9:1–8.
11. Chughtai AA, Seale H, Macintyre CR. Effectiveness of cloth masks for protection against severe acute respiratory syndrome coronavirus 2. Emerg Infect Dis. 2020;26:e1–e5.
12. Santos M, Torres D, Cardoso PC, et al. Are cloth masks a substitute to medical masks in reducing transmission and contamination? A systematic review. Braz Oral Res. 2020;34:1–17.
13. Marler H, Ditton A. “I’m smiling back at you”: Exploring the impact of mask wearing on communication in healthcare. Int J Lang Commun Disord. 2020;56:205–214.
14. Ribeiro VV, Dassie-Leite AP, Pereira EC, et al. Effect of wearing a face mask on vocal self-perception during a pandemic. J Voice. 2020. Article in Press.
15. Heider CA, Álvarez ML, Fuentes-López E, et al. Prevalence of voice disorders in healthcare workers in the universal masking COVID-19 era. Laryngoscope. 2021;131:E1227–E1233.
16. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015;349:1–25.
17. American Speech-Language-Hearing Association. Evidence-based practice in communication disorders: an introduction. Rockville, USA: ASHA; 2004. Available at: http://www.asha.org/policy/TR2004-00001/.
18. Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions. John Wiley & Sons; 2011.
19. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009;6:e1000100.
20. Law M, Stewart D, Letts L, et al. Guidelines for critical review of qualitative studies. McMaster University occupational therapy evidence-based practice research Group; 1998:1–9.
21. NHLBI N. Quality assessment tool for observational cohort and cross-sectional studies. Available from:Study Quality Assessment Tools | NHLBI, NIH; 2017.
22. Hong QN, Pluye P, Fabregues S, et al. Mixed methods appraisal tool (MMAT), version 2018. Registration of Copr. 2018; 1148552:10.
23. Saab MM, Landers M, Hegarty J. Males’ awareness of benign testicular disorders: an integrative review. Am J Mens Health. 2018;12:556–566.
24. GRADE Working Group. Grading quality of evidence and strength of recommendations. BMJ. 2004;328:1490.
25. Corey RM, Jones U, Singer AC. Acoustic effects of medical, cloth, and transparent face masks on speech signals. J Acoust Soc Am. 2020;148:2371–2375.
26. Botelho P, Murgia S, Puglisi GE, et al. Effect of masks on speech intelligibility in auralized classrooms. J Acoust Soc Am. 2020;148:2878–2884.
27. Cavallaro G, Di Nicola V, Quaranta N, et al. Acoustic voice analysis in the COVID-19 era. Acta Otorrinolaringol Ital. 2021;41:1.
28. Magee M, Lewis C, Noffs G, et al. Effects of face masks on acoustic analysis and speech perception: Implications for peri-pandemic protocols. J Acoust Soc Am. 2020;148:3562–3568.
29. Fiorella ML, Cavallaro G, Di Nicola V, et al. Voice Differences When Wearing and Not Wearing a Surgical Mask. J Voice. 2021. Article in Press.
30. Nguyen DD, McCabe P, Thomas D, et al. Acoustic voice characteristics with and without wearing a facemask. Sci Rep. 2021;11:1–11.
31. Rahne T, Fröhlich L, Plonke S, et al. Influence of face surgical and N95 face masks on speech perception and listening effort in noise. Plos One. 2021;16:1–11.
32. Lin Y, Cheng L, Wang Q, et al. Effects of medical masks on voice assessment during the COVID-19 pandemic. J Voice. 2021. Article in Press.
33. Zambon F, Moretti F, Ribeiro VV, et al. Vocal tract discomfort in adult cochlear implant users: its relationship to self-reported voice disorders. J Voice. 2011;25:482–488.
34. Zambon F, Moretti F, Ribeiro VV, et al. Vocal fatigue index: validation and cut-off values of the brazilian version. J Voice. 2020. Article in Press.
35. Rodríguez G, Zambon F, Mathieson L, et al. Vocal tract discomfort in teachers: its relationship to self-reported voice disorders. J Voice. 2013;27:473–480.
36. Nuñez-Batalia F, Corte-Santos P, Señaris-González B, et al. Adaptación y validación del índice de incapacidad vocal (VHI-30) y su versión abreviada (VHI-10) al español. Acta Otorrinolaringol Es. 2007;58:386–392.
37. Lehiste I, Peterson GE. Linguistic considerations in the study of speech intelligibility. J Acoust Soc Am. 1959;31:280–286.
38. Peterson GE, Lehiste I. Revised CNC lists for auditory tests. J Speech Hear Disord. 1962;27:62–70.
39. Auditory Potential LL. Minimum Speech Test Battery (MSTB) for adult cochlear implant users. Goodyear, AZ: Auditory Potential LLC; 2011.
40. Boersma P, Praat, a system for doing phonetics by computer. Glot Int. 2001;5:341–345.