Surgical mask filter and fit performance

Tara Oberg, MS, and Lisa M. Brosseau, ScD
Minneapolis, Minnesota

Background: Surgical masks have been used since the early 1900s to minimize infection of surgical wounds from wearer-generated bacteria. There is ongoing debate, however, whether surgical masks can meet the expectations of respiratory protection devices. The goal of this study was to evaluate the filter performance and facial fit of a sample of surgical masks.

Methods: Filter penetration was measured for at least 3 replicates of 9 surgical masks using monodisperse latex sphere aerosols (0.895, 2.0, and 3.1 μm) at 6 L/min and 0.075-μm sodium chloride particles at 84 L/min. Facial fit was measured on 20 subjects for the 5 masks with lowest particle penetration, using both qualitative and quantitative fit tests.

Results: Masks typically used in dental settings collected particles with significantly lower efficiency than those typically used in hospital settings. All subjects failed the unassisted qualitative fit test on the first exercise (normal breathing). Eighteen subjects failed the assisted qualitative fit tests; 60% failed on the first exercise. Quantitative fit factors ranged from 2.5 to 9.6.

Conclusion: None of these surgical masks exhibited adequate filter performance and facial fit characteristics to be considered respiratory protection devices. (Am J Infect Control 2008;36:276-82.)

Surgical masks have been in widespread use since the early 1900s to help prevent infection of surgical wounds from staff-generated nasal and oral bacteria.\(^1\,^2\) Today, surgical masks vary widely in style and intended application and can be found in a broad range of hospital and health care settings. In some health care settings, applications have evolved from prevention of patient wound infection to prevention of employee exposures. There is ongoing debate, however, about the use of surgical masks as respiratory protection devices.

The Food and Drug Administration (FDA) oversees the sale and marketing of medical devices, including surgical masks, which may be known as procedure masks, dental masks, and laser masks as well as masks used in surgery settings. FDA recommends that manufacturers demonstrate surgical mask performance in 4 areas: fluid resistance, filter efficiency, differential pressure, and flammability.\(^3\) Two types of filter efficiency tests are recommended: (1) particulate filtration efficiency (PFE) using a nonneutralized aerosol of 0.1-μm latex spheres at a challenge velocity between 0.5 and 25 cm/s (approximately 8 to 380 L/min for a 9-cm radius mask)\(^4\,^5\) and (2) bacterial filtration efficiency (BFE) using a nonneutralized 3 ± 0.3-μm Staphylococcus aureus aerosol and a flow rate of 28.3 L/min.\(^6\,^7\) The FDA requires no minimum level of filter performance. The Centers for Disease Control and Prevention (CDC) publishes guidelines on the use of surgical masks in health care settings.\(^8\)

The National Institute for Occupational Safety and Health (NIOSH) regulates the testing and certification of respiratory protection equipment. The NIOSH tests filters for the effects of loading (particle burden), temperature, and relative humidity and requires a minimum filtration efficiency of 95%, 99%, or 99.97% using neutralized 0.075-μm count median diameter (CMD) solid aerosols at 85 L/min.\(^10\) Neutralized aerosols are more penetrating than charged ones and yield more consistent results. The CMD represents the median size of the aerosol when it is assessed by number concentration.

Certification tests also evaluate effects of oil aerosols for filter designations of N (not resistant to oil), R (somewhat resistant to oil), and P (strongly resistant to oil). NIOSH evaluates the fit performance of some respiratory protective devices using human panels with specified facial dimensions. Certification of filtering face-piece respirators, however, does not currently include an assessment of fit performance.

The Occupational Safety and Health Administration (OSHA) regulates the selection and use of respirators in a workplace.\(^11\) Facilities are required to have a respirator program that includes individual medical evaluation, training, and fit testing. OSHA has designated assigned protection factors that indicate to employers how well respirators in a particular class will reduce exposure to airborne contaminants. A fit test is then used to evaluate the fit of a respirator on an individual.
Fit testing involves assessment of leakage detected by odor or taste (qualitative) or by comparison of particle concentrations inside and outside the face piece (quantitative). The latter is referred to as an individual's fit factor, which must be equal to or greater than the assigned protection factor multiplied by a safety factor. In the case of a filtering facepiece respirator, an individual's fit factor must be greater than 100 (assigned protection factor = 10; safety factor = 10).\textsuperscript{11} The shared regulatory approach of NIOSH and OSHA to respiratory protection recognizes the 2 most important aspects of respiratory protection: providing known filtration efficiency while also ensuring the proper use and selection of devices, which includes initial and ongoing individual fit. The first goal of this study was to evaluate surgical mask filter efficiency, using NIOSH and OSHA tests, and compare results with reported BFE and PFE. The second goal was to measure the individual fit of surgical masks on volunteers.

METHODS

After consultation with local infection control professionals, we selected 9 surgical masks representative of those used in hospital and dental settings. The masks included a range of types (surgical, laser, procedure), models (cup, flat, duckbill), and fastenings (1 and 2 straps, ear loops). Masks were purchased from local or on-line suppliers. Results of PFE and BFE tests and FDA approval status are reported if available (Table 1). The fit testing portion of this project was approved by the University of Minnesota Institutional Review Board.

We first evaluated filtration performance using monodisperse latex sphere and sodium chloride aerosols. The fit was characterized for fit with volunteers, using both qualitative and quantitative fit tests.

Filter performance

Surgical masks were challenged using 3 sizes of monodisperse latex spheres (0.895, 2.0, and 3.1 μm) at a flow rate of 6 L/min. These particle sizes were selected to approximate the range of the Bitrex aerosol (Macfarlan Smith, United Kingdom) used in qualitative fit tests to ensure that particles used for fit testing could be captured by filter media (geometric mean, 2.4 μm; geometric standard deviation, 1.4).\textsuperscript{12} The challenge flow approximates a resting human breathing rate. Three or 4 replicates of each surgical mask were tested at each experimental condition.

Surgical masks were formed to simulate their as-worn shape and sealed to a metal plate mounted in a filter test apparatus similar to that described elsewhere.\textsuperscript{13} In some cases, a screen was used to ensure mask shape was maintained throughout testing. The aerosol was generated using a nebulizer (Inspiron 002305-A; Inter-tech Resources Inc., Lincolnshire, IL) containing a solution of filtered deionized water and monodisperse polystyrene latex spheres (Duke Scientific, Palo Alto, CA). The aerosol was charged neutralized using a Kr-85 source and diluted with dried high-efficiency particulate air–filtered air. The challenge aerosol concentration was approximately $2 \times 10^5$ particles/m$^3$.

Measurements of particle number concentration were made with a direct reading, light scattering photometer (APS Model 3321; TSI Inc., St. Paul, MN). Percent penetration was calculated by dividing the average downstream concentration by the average upstream concentration and multiplying by 100.

A 2-way analysis of variance was used to identify the effect of surgical mask type and particle size and their interaction on aerosol penetration. Multiple comparison procedures were performed, where appropriate, to examine the underlying causes of significance.

Three replicates of each surgical mask were also challenged with a neutralized 0.075-μm sodium chloride aerosol at 84 L/min (representative of a high work rate), following the NIOSH N-series respirator certification requirements.\textsuperscript{19} This test was performed with an Automated Filter Tester (AFT Model 8150; TSI Inc.), which generates a salt aerosol by liquid atomization, measures aerosol concentrations upstream and downstream of the filter with a light scattering photometer, and reports percent penetration.

Facial fit

Surgical masks with latex aerosol penetration less than 0.6% (at any particle size) were then evaluated for facial fit. This cutoff represents the point at which masks could be divided into 2 separate performance groups. Twenty subjects (10 male and 10 female) ranging in age from 19 to 57 years were recruited by e-mail and posted flyers. Respondents were screened by telephone; those with facial hair, a fear of closed-in places, or symptoms or history of lung illness or injury were excluded. Subjects were not screened for previous use of masks or respirators. Qualifying subjects were scheduled for testing.

Subjects were screened again at the start of a fit test, and written consent was obtained. Each subject wore a single randomly assigned surgical mask and performed 2 qualitative tests (1 unassisted and 1 assisted) followed by 2 quantitative tests (1 unassisted and 1 assisted). A total of 20 paired qualitative fit tests and 20 paired quantitative fit tests were performed with 2 donning protocols (4 tests for each of the 5 mask models). Every surgical mask was tested by 2 male and 2 female subjects. Fit test methods followed procedures recommended by OSHA.\textsuperscript{11}
Facial fit was first evaluated following OSHA’s Bitrex (Denatonium Benzoate) Solution Aerosol Qualitative Fit Test Protocol, which uses a taste threshold approach.¹¹,¹² Screening was initially performed without a surgical mask to ensure the subject could taste the test solution. Immediately after the screening, the subject donned a surgical mask without assistance. No donning instructions were included with the purchased masks. Subjects viewed packaging materials, which showed a diagram of a person wearing a mask.

The test aerosol was then nebulized from solution into a hood placed over the subject’s head. While standing, the subject was asked to perform a series of 1-minute exercises: normal breathing, deep breathing, turning head side to side, moving head up and down, reading from a prepared text, jogging in place, and normal breathing.

The subject then removed the hood and redonned the surgical mask with assistance, after which the test procedure was repeated. Donning assistance consisted of mask positioning on face, tie/strap placement (eg, base of neck and crown of head for 2 ties or straps), and nosepiece conformance to bridge of nose. A test was failed if the subject was able to detect the taste of Bitrex at any point during the procedure.

For quantitative fit tests, the subject was asked to don, without assistance, a new surgical mask with a center probe. The quantitative fit factor was determined with a Portacount Plus device (TSI Inc.), which relies on ambient particles detected by a condensation nuclei counter and laser photometer to measure concentrations outside and inside the mask.

Subjects performed the same 7, 1-minute exercises as well as a 15-second grimace (repeatedly smiling and frowning), following computer screen instructions. Software prompts subjects at each new exercise and calculates individual exercise and overall fit factors. After completing the quantitative fit test without assistance, a subject redonned the surgical mask with researcher assistance and repeated the fit test procedure.

The mean quantitative fit factors were compared using a 3-way analysis of variance with repeated measures. Statistical tests were used to identify significant between and within subject effects. A Bonferroni multiple comparisons test was performed when means were statistically significant.

RESULTS

Filter performance

Latex sphere challenge tests. The 9 masks exhibited a wide range of particle penetration (0%-84%) over the 3 particle sizes. Percent penetration generally decreased with increasing particle size for all surgical masks, but the degree of change was not consistent across masks. Average penetration was 16% (standard deviation [SD], 28%). 15% (SD, 26%), and 11% (SD, 2%) for the 0.895-, 2.0-, and 3.1-µm particles, respectively. Masks used in dental clinics (A, B, and C) showed significantly higher average penetration across all particle sizes (6%-75%) when compared with those used in hospital settings (D through I) (0.02%-0.7%) (Fig 1).

Sodium chloride—NIOSH—challenge tests. The 9 masks exhibited a similarly wide range of penetration of the smaller sodium chloride aerosol particles at 84 L/min (4%-90%). As expected because of the higher challenge flow, penetration for all filters was greater at the NIOSH test conditions, although the degree of change was not consistent (Table 2). Dental masks again showed significantly higher penetration (53%-90%) than hospital masks (4%-37%).
Facial fit

**Qualitative fit test.** All subjects failed the qualitative fit test on the first exercise (normal breathing) when masks were donned without assistance. All but 2 male subjects (wearing mask F and mask H) failed the qualitative fit test after receiving assistance. Sixty percent of subjects failed on the first exercise.

**Quantitative fit test.** Quantitative fit factors varied significantly with mask type ($P < .024$). Average quantitative fit factors ranged from 2.5 to 6.9 for unassisted donning and from 2.8 to 9.6 for assisted donning (Table 5). The average overall fit factor for masks donned without assistance (4.4; SD, 0.9) was less than the average overall fit factor for masks donned with assistance (5.7; SD, 0.8), but this relationship was not consistent for all mask types.

Protocol (unassisted vs assisted) had a significant effect on fit factor ($P < .0012$); however, there was a significant interaction between donning protocol and mask type ($P < .002$). Three of the masks had almost equal fit factors using the 2 donning protocols (D, F, and I). Increases in the fit factors for masks G and H from the unassisted to the assisted protocol accounted for the significant effect of protocol.

Sex had no effect on mask fit factor ($P < .54$), and mean fit factors for male and female subjects were similar for either donning protocol (unassisted: 4.5 ± 0.6 for males; 4.4 ± 0.6 for females; assisted: 6.1 ± 0.6 for males; 5.3 ± 0.6 for females).

**DISCUSSION**

Our experimental tests showed filter efficiencies ranging from 20% to 99% in the latex sphere tests and from 10% to 90% in the sodium chloride tests. Other investigators have found a similar broad range of filter efficiencies for surgical masks.\(^{13-18}\)

The filter tests required by the FDA are much less stringent than the NIOSH tests. Most of the experimental mask filters were reported to have bacterial or particle filter efficiency greater than 96% (Table 1). The smallest differences between NIOSH test results from this study and manufacturer-reported filter efficiencies were found for mask D with a 94% NIOSH test efficiency and a >99% BFE and PFE. The largest difference was found for mask C with 23% NIOSH test efficiency compared with 95% BFE. Higher filter efficiencies in the BFE and PFE tests may be the result of using nonneutralized aerosols, polydisperse aerosols (BFE test), and a range of flows (PFE test).

For exposures to infectious respiratory organisms, we are most concerned with the size and concentrations of aerosols generated during normal breathing, talking, coughing, and sneezing. One study found that particles from healthy subjects ranged from 0.09 to 3 μm particles. Concentrations ranged from 100 to 350 particles/L during normal breathing and 150 to 2000 particles/L during talking or coughing.\(^{19}\) Another showed concentrations ranging from 14 to greater than 3000 particles/L and an average particle size of 0.52 μm.\(^{20}\)

The filters of most of these surgical masks will allow a large majority of wearer-generated particles to penetrate and will collect only a small percentage of airborne particles generated by infectious patients. Even when equipped with filters demonstrating relatively high collection efficiency, 10% to 40% of particles
Table 2. Filter penetration

<table>
<thead>
<tr>
<th>Mask</th>
<th>NIOSH test (84 L/min)</th>
<th>Latex sphere tests (6 L/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.3 μm NaCl</td>
<td>0.8 μm</td>
</tr>
<tr>
<td>A</td>
<td>90.2 (0.43)</td>
<td>81.7 (1.51)</td>
</tr>
<tr>
<td>B</td>
<td>52.9 (2.4)</td>
<td>14.7 (4.36)</td>
</tr>
<tr>
<td>C</td>
<td>77.2 (1.2)</td>
<td>53.6 (2.31)</td>
</tr>
<tr>
<td>D</td>
<td>5.98 (0.3)</td>
<td>0.10 (0.02)</td>
</tr>
<tr>
<td>E</td>
<td>37.4 (0.4)</td>
<td>1.55 (0.31)</td>
</tr>
<tr>
<td>F</td>
<td>28.9 (0.7)</td>
<td>0.53 (0.12)</td>
</tr>
<tr>
<td>G</td>
<td>10.44 (0.8)</td>
<td>0.11 (0.06)</td>
</tr>
<tr>
<td>H</td>
<td>3.96 (0.2)</td>
<td>0.04 (0.02)</td>
</tr>
<tr>
<td>I</td>
<td>31.6 (1.1)</td>
<td>0.51 (0.11)</td>
</tr>
</tbody>
</table>

NOTE: Data are in percentage, mean and standard error.

will penetrate the face seal as a result of poor fit. For aerosols containing organisms with a low infectious dose (e.g., tuberculosis), this level of face seal leakage would not prevent a potentially infectious exposure during even a brief encounter with a patient generating copious amounts of aerosol.21-23 The poor performance of dental masks is of particular concern, given dental surgeons’ close proximity to patients and the high aerosol concentrations generated by dental procedures.24

A limited number of previous studies of surgical masks as personal protective devices also show that surgical masks are not equivalent to respirators. Two studies simulating inward leakage compared surgical masks sealed and unsealed on a mannequin face and found fit factors from 3 to 5 (18% to 32% face-piece leakage) for 1.8-μm dioctyl phthalate particles and 22-μm fungal spores.15,25 A significantly higher prevalence of antibodies for several respiratory tract viruses (influenza A and B and respiratory syncytial virus) was found in 50 dental surgeons compared with 50 controls. However, no significant difference in antibody prevalence was found between surgeons wearing masks versus those who occasionally or never wore surgical masks.26 Surgical masks were not effective at reducing internal deposition of Technetium-99 meta-stable, an aerosolized radiopharmaceutic, in nuclear medicine personnel and did not significantly reduce latex particle inhalation in a study of 20 health care workers.27,28 A surgical and laser mask showed similar fit factors of 3.0 (95% CI: 1.8-4.2) and 3.8 (95% CI: 2.9-4.6), much less than the fit factor of 102.6 (95% CI: 41.2-164) for a FFP2 respirator.29 A recent study found a geometric mean fit factor of 2.6 (geometric SD, 1.6) for 6 surgical masks, using a TSI Portacount device with an N95 Companion.30

There are few studies that compare the clinical efficacy of respirators versus surgical masks. In one Toronto hospital, all attending health care workers reported to be wearing “respirators” contracted severe acute respiratory syndrome (SARS) during a patient intubation.31 Closer examination reveals that employees were wearing surgical masks, not respirators. Another study found that nurses in a Toronto hospital not consistently wearing either a filtering face-piece N95 respirator or a surgical mask had 4 times the risk of contracting SARS as those consistently wearing respirators or surgical masks.32 Fit testing is not required and was infrequently employed in Canada during the SARS outbreaks.33

Our data illustrate how important it is to preventing inward leakage of particles. Qualitative fit tests are considered valid measures of personal protection for respirators that must achieve a fit factor of 100 (used in atmospheres less than 10 times the permissible exposure limit). Although 2 subjects were able to pass the qualitative fit test on 2 different surgical masks when assisted with fit, we believe these results occurred because of temporary taste desensitization. Our quantitative fit test results support this conclusion. None of the test surgical masks attained an individual fit factor of 100, the minimum level expected for a half-mask filtering face-piece respirator. Assistance with fit made no difference in the degree of fit.

Our qualitative fit test results also illustrate the importance of surgical mask design. Mask D showed the second highest filter efficiency but the lowest fit factor. One of the reasons for poor fit may be the ear loop design, which limits adjustability of fit. To be effective in reducing wearer’s exposure to airborne substances, a respiratory protection device needs to have sufficient fit as well as high filtration efficiency.

This study has several limitations. A relatively small number of masks were included, although our selections covered a range of styles and uses to be representative of the variety of commercially available surgical masks. The number of fit test subjects was also small, and no effort was made to obtain a sample
representative of the full range of facial shapes. Our data, however, are consistent for all 20 subjects, suggesting that most people will experience similar low levels of fit. Our quantitative fit tests employed a TSI Portacount device without a N95 companion. The latter instrument selects a narrower range of particles to minimize the contribution of filter penetration to the measurement of fit. We sought to minimize this by including in the fit test experiments only surgical masks with filter penetration less than 0.6% at all 3 test particle sizes. In addition, our findings are similar to those of a recent study using the N95 companion.

N95 filters must have less than 5% penetration for an aerosol with a mass median aerodynamic diameter of 0.3 microns. Half-mask respirators (including those with N95 filters) must have a fit factor (outside/inside particle concentration) of at least 100 to provide the wearer with a protection factor of 10. None of the surgical masks we tested met both of these performance criteria. Although 1 filter showed an average penetration of less than 5% in the NIOSH test, our experiments did not assess the full set of test conditions required for respirator certification. None of the surgical masks we evaluated met the filter or fit performance criteria for respiratory protection devices.

We conclude that surgical masks do not offer protection comparable with that of respiratory protective devices (and are not certified by NIOSH as such). Our measurements of inward leakage led us to infer that outward leakage will also occur while wearing a surgical mask. The FDA should evaluate the use of surgical masks for their intended purpose of preventing wound infection. If health care institutions continue to expend resources on surgical masks for both purposes, we strongly urge the FDA to employ a more robust regulatory approach to their approval. In health care settings in which both wound infection prevention and respiratory protection are needed, use of surgical N95 respirators, which are both FDA-certified as surgical masks and NIOSH-certified as N95 respirators, should be considered.

In the United States today, 29 CFR §1910.134 requires the use of NIOSH-approved respirators for protection against inhalation hazards. Our data show that surgical masks do not meet the filtration performance criteria for NIOSH-approved half-mask respirators and that test subjects were not able to pass a fit test as specified by the OSHA. It is therefore recommended that NIOSH-certified respirators, not surgical masks, be used to reduce employee exposure to airborne infectious organisms.

References