Evaluation of a Shortened Qualitative Fit Test Method for Filtering Facpiece Respirators

Rhiannon Mogridge[1]*, Alison Bowry[1], Mike Clayton[1,2]

[1] Health and Safety Executive, Harpur Hill, Buxton, SK17 9JN, UK
[2] Change of address since completion of this work: 3M Centre, Cain Road, Bracknell, Berkshire, RG12 8HT, UK

* Corresponding author and E-mail: rhiannon.mogridge@hsl.gsi.gov.uk

ABSTRACT

The most commonly used type of respiratory protective equipment (RPE) in the healthcare sector in Great Britain (GB) is the disposable filtering facepiece. As a type of tight-fitting respirator, the Health and Safety Executive's guidance requires that fit testing be carried out as part of the selection process. The GB healthcare industry typically employs the qualitative fit testing method. If a shortened qualitative fit test protocol could be validated, its use could help to reduce overall time investment required for an effective fit testing programme.

A shortened qualitative fit test method (duration approximately 2 minutes) was compared against a generated aerosol quantitative fit test method using the approach given in the American National Standards Institute (ANSI) standard Z88.10-2010 Annex 2 *Criteria for Evaluating New Fit Tests Methods*.

Results show that the shortened qualitative fit test method failed to meet the evaluation criteria of ANSI Z88.10. The *test sensitivity*, which gives the probability of the shortened qualitative fit test method failing a mask that is a poor fit, was 0.70, which is less than the requirement of ≥ 0.95. The value of 0.70 means that 70% of poor fits would be expected to fail the shortened qualitative fit test method (true failures). Conversely, 30% of poor fits would be expected to pass the shortened qualitative fit test method (false passes). The shortened qualitative fit test method is therefore likely to produce a higher proportion of false passes, potentially leading to a reduction in protection than that expected from this type of RPE.

Keywords: Respiratory protective equipment (RPE), respirator fit test, qualitative fit test, quantitative fit test, Bitrex, PortaCount, filtering facepiece respirator

INTRODUCTION

The Health and Safety Executive (HSE) has approved codes of practice and guidance supporting the legal requirements governing the use of tight-fitting respiratory protective equipment (RPE) within Great Britain (GB). These include a requirement that each individual wearer of tight-fitting RPE has passed a fit test to establish that the facepiece is suitable for their face as part of the selection process, before they use the RPE for protection against workplace hazards (HSE, 2002; 2013a; 2013b).

HSE's guidance document on fit testing RPE, HSE OC 282/28 (HSE, 2012), gives guidance including information on fit testing methods for different types of RPE and recommends fit test pass criteria. Several fit testing methods may be used for fit testing filtering facepiece class 3 (FFP3) (European Standard EN149 2001+A1:2009 (CEN, 2009)) respirators. In practice, the methods which are almost exclusively used for fit testing FFP3 in GB are:

- A qualitative fit test (QLFT) using either
    - Bitrex
    - Or Saccharin, as the test agent
- A quantitative fit test (QNFT) using the TSI PortaCount. (a condensation nuclei counter (CNC) based instrument)

The FFP3 is the most common type of RPE used by GB health workers for protection against infectious agents such as influenza. Although several different fit testing methods can be used to fit test FFP3, for practical reasons the qualitative fit test method has become the healthcare sector's default fit testing method.

The HSE guidance OC 282/28 establishes a common exercise protocol for qualitative and quantitative fit tests. The protocol is given in Table I.

Table I. The HSE guidance OC 282/28 fit test exercises for QLFT and QNFT methods

Exercise	Duration (min)
Normal breathing	1
Deep breathing	1
Moving head side to side	1
Moving head up and down	1
Talking aloud	1
Bending at the waist	1
Normal breathing	1
Total exercise duration	**7**

Although the total exercise time is 7 minutes, the total time required to conduct a fit test is longer than just the exercise time and is dependent on the method used. A typical duration for a qualitative fit test is shown in Table II. If more detailed training is to be given to the wearer or a repeat fit test is necessary, the time taken will be longer.

Additionally, repeat fit testing is periodically required to ensure that the selected RPE continues to fit the wearer. This could place a significant additional time demand on the duty holder, which in turn can reduce compliance with regulations.

Table II. Typical time required to conduct a QNFT

Activity	Estimated time (min)
Explain the fit test to the wearer	2
Carry out the sensitivity test	1
Allow time for the wearer's palate to clear (during which the wearer is instructed on how to don and conduct a wearer seal check the RPE)	15
The wearer dons the RPE and conducts a wearer seal check	1
Conduct the fit test	7
Remove the RPE, explain and record the results	2
Total estimated fit test duration	**28**

Recent studies in the USA have looked at shortening the PortaCount fit testing protocol (Richardson *et al*, 2013; Richardson *et al*, 2014a; Richardson *et al*, 2014b). While some of this work is not directly relevant for qualitative fit testing, Richardson *et al's* research identified the key exercises in the fit test exercise protocol in which face seal leaks occurred, which are relevant regardless of the method of fit testing being used. The exercises and durations that they selected are given in Table III. Using this exercise protocol would reduce the time required to carry out a QNFT by 5 minutes.

Table III. Shortened fit test exercise protocol

Exercise	Sample duration (min)
Bending	0.5
Talking out loud	0.5
Moving head side to side	0.5
Moving head up and down	0.5
Total sample duration	**2**

Note: For the shortened quantitative fit test method, the actual total duration is 2 min 29 s due to the challenge sample, and challenge and mask purge times.

A shortened QLFT method based on the Bitrex taste test agent was previously explored by Nelson *et al* (2003). While fulfilling a number of the American National Standards Institute (ANSI) fit test comparison acceptance requirements (ANSI, 2010), the key requirement of the fit test method to identify an inadequate fit was not achieved - achieving a test sensitivity of 0.92 compared with the requirement of at least 0.95. Although the fit tests conducted using a seven 15-second exercise protocol gave comparable results to those achieved from a seven 60-second exercise protocol when analysed using a binary logistic regression analysis, the shortened QLFT method was not accepted by OSHA nor included in the ANSI fit test standard.

The effectiveness of the Bitrex QLFT relies on the wearer's taste and their response. This can vary by individuals; their taste response is not always immediate and there can be variation with how wearers define a positive taste (Frost *et al* 2014a). Effective application of Bitrex into the hood can also prove to be challenging within a short time frame. However, reducing the number of exercises and reducing the duration from 60 seconds to 30 seconds, rather than to 15 seconds as in the earlier study, may improve application of the Bitrex fit test method and the comparative results.

This paper explores the validation of a shortened QLFT method using the test exercise protocol proposed by Richardson *et al* (2014b) and is shown in Table III.

METHODS

Selection of Masks

The GB National Health Service (NHS) staff are supplied with their personal protective equipment (PPE) through the NHS Supply Chain. At HSE's request, the NHS Supply Chain identified the five most popular models of FFP3 based on the number purchased; samples of these FFP3s were obtained from the Sheffield NHS Foundation Trust and direct from the RPE manufacturers.

Volunteers

All work with human volunteers undertaken for this study was ethically cleared by the Sheffield University Ethics Committee. All volunteers were medically screened and gave informed consent before testing commenced. For their safety, the volunteers' heart rates were monitored using a Garmin Forerunner 305 heart rate monitor. The heart rate withdrawal criterion was:

$$Max\ heart\ rate = 185 - (Age \times 0.65)$$

Volunteers were clean shaven in the areas where the facepiece sealed against the face. Volunteers' anthropometric measurements, specifically face width (bizygomatic breadth) and face length (menton-sellion length), were taken and compared with the National Institute for Occupational Safety and Health (NIOSH) bivariate panel, as shown in Figure 1.

Fit Tests

The fit tests were carried out in sessions. During each test session, a volunteer donned a FFP3 and was fit tested using three different methods. As far as was possible, the fit of the mask was not disturbed during the session.

The methods of fit testing used were the generated aerosol QNFT method (as the reference), the shortened QLFT method based on the Bitrex taste test agent, and the QNFT PortaCount ambient particle counting method. This paper focusses on the comparison between the generated aerosol QNFT method and the shortened QLFT method. The QNFT PortaCount method was included as an additional part of the study to allow for a further comparison as a secondary reference.

Each volunteer attended up to five test sessions, and wore a different FFP3 for each session. The order in which volunteers wore the masks was randomised, as was the order of the fit test methods for each session. Regardless of the order of the fit test methods, the sensitivity test for the shortened QLFT method was always performed first, to allow the volunteer's palate to clear.

Volunteers were given assistance in fitting the FFP3 where necessary, to ensure that they were worn in accordance with the manufacturer's instructions. After carrying out a pre-use wearer seal check (also known as a user seal check or fit check), volunteers were asked to indicate whether, in their opinion, a good fit had been achieved; this opinion was recorded, but testing continued even if the volunteer reported a poor fit. After this point, care was taken by the researchers and volunteers not to alter the fit of the FFP3.

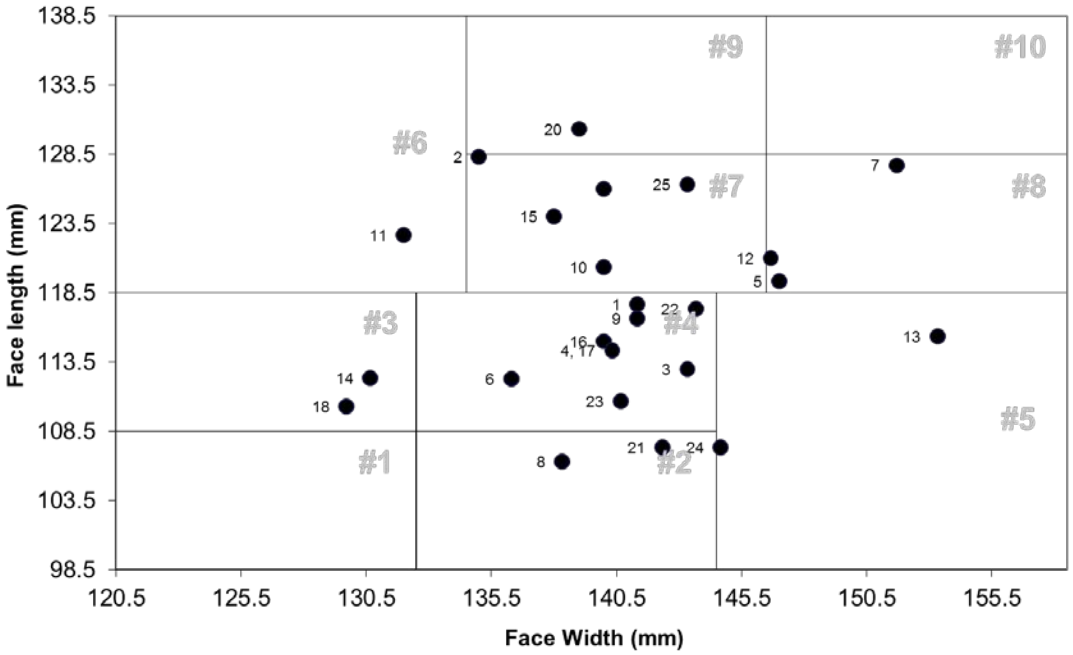

Figure 1. Distribution of volunteers' facial dimensions for 25 out of 26 volunteers. Volunteer 21 was not available when the facial measurements were taken.

All researchers were accredited fit testers under the GB Fit2Fit scheme (http://fit2fit.org) for both Bitrex and PortaCount fit test methods and experienced in measuring fit and performance using the generated aerosol method.

Probing of Masks

For each fit test, a new mask was prepared with two metal sampling ports using a TSI disposable probe kit 8025-N95 (TSI, Minnesota, USA). A plastic template was made for each model of mask so that the sampling ports would be in the same location on each sample of that model. A length of rigid tubing was used to extend from one of the metal sampling ports into the breathing zone of the wear (between the nose and mouth, <10 mm from the face) in accordance with the HSE guidance OC 282/28. The tubing was fitted with a multi-holed sampling ball, as described in EN 13274-1 (CEN, 2001). The second port was positioned as close to the centre of the mask as possible, and was blocked off using a rubber cap for most of the test session; this metal sampling port was used to carry out a taste reveal test at the end of the shortened QLFT, which is described in more detail later. Figure 2 shows an example mask with sampling ports.

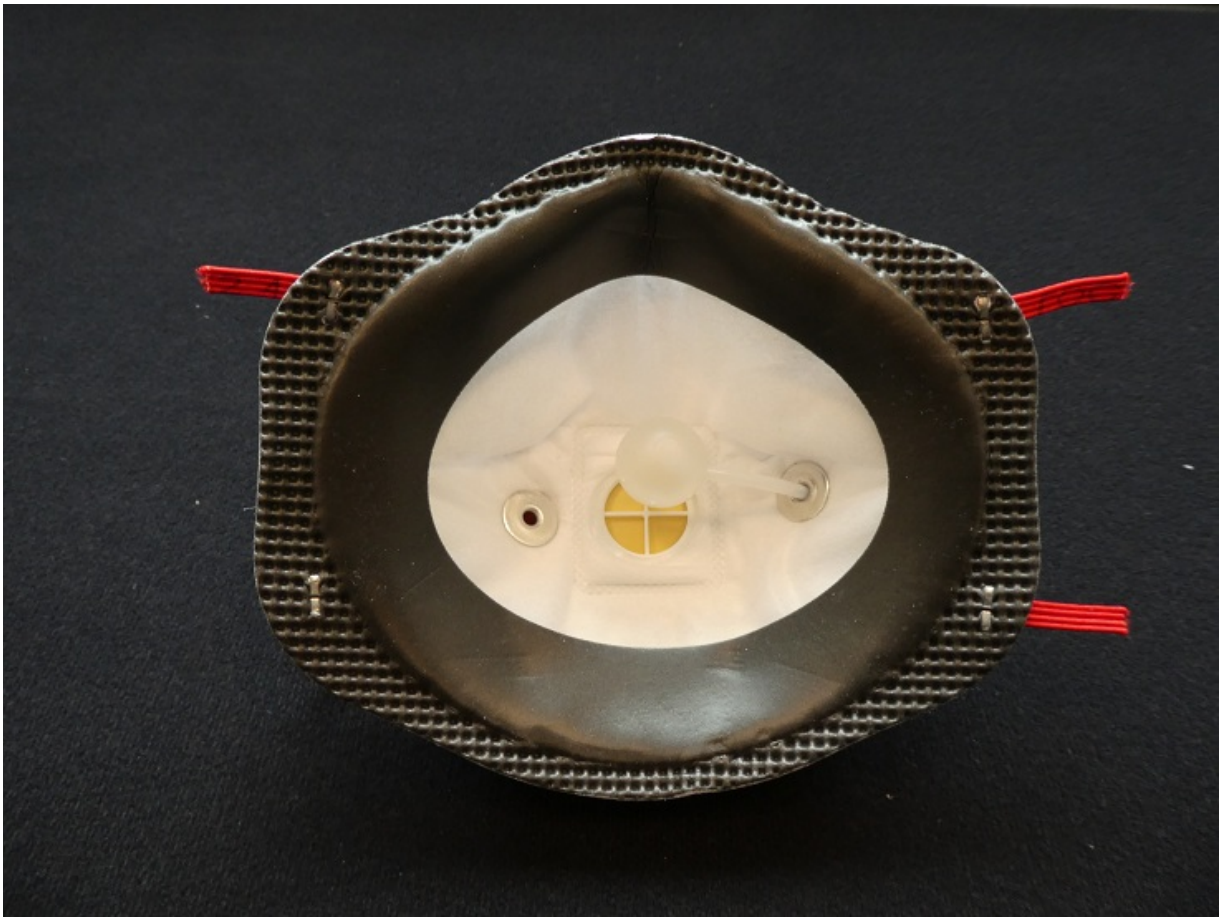

Figure 2. Example FFP3 mask with metal sampling ports in place.

Generated Aerosol QNFT Method

The generated aerosol QNFT method was based on the European standard method for measuring Total Inward Leakage (TIL) of RPE, as described in EN 13274-1 (CEN, 2001). For this fit testing method, a volunteer undertook the exercises in Table IV while wearing the RPE and walking on a treadmill in a test chamber. The test chamber was filled with a generated sodium chloride aerosol. Air samples were taken from inside and outside the FFP3, and the sodium concentration in each sample was measured using a Moore's low flow sodium flame photometer type 1250 (SFP, Dorset, UK). TIL measurements are near real-time and extremely sensitive; protection factors of >100,000 can be measured. The exercise protocol used for this test was based on EN 13274-1 (CEN, 2001), but with the sample times reduced to 1 minute to better match the fit testing protocol given in the HSE guidance OC 282/28.

Dry air was mixed with the in-mask sample to prevent the moisture in the wearer's exhaled breath from interacting with the NaCl in the sampling lines and thus affecting the measurement. The Moore's low flow sodium flame photometers draw a sample flow of 2 l/min, of which 1 l/min was drying air. After the measurements were taken, all in-mask measurements were multiplied by a correction factor of 2, to account for the use of the drying air.

Sodium chloride is retained in the lungs, resulting in a reduced concentration of sodium in the wearer's exhaled breath. A correction factor of 0.6 was used to account for lung retention (BSI, 1980) and the in-mask concentration was calculated using the formula:

$$In\text{-}mask\ concentration = \frac{Measured\ concentration}{0.6}$$

This correction factor is based on the assumption that 80% of the sodium chloride aerosol is absorbed in the lungs, and that the wearer spends half of the time inhaling, and half of the time exhaling:

$$C_m = (0.5 \times C_{in}) + (0.5 \times C_{ex})$$
$$C_{ex} = 0.2C_{in}$$
$$C_m = (0.5 \times C_{in}) + (0.5 \times 0.2C_{ex}) = 0.6C_{in}$$
$$C_{in} = \frac{C_m}{0.6}$$

Where C_m is the measured concentration, C_{in} is the concentration during the inhalation portion of the breath, and C_{ex} is the concentration during the exhalation portion of the breath.

The lung-retention correction factor was applied to all in-mask measurements. Once all the correction factors had been applied to the in-mask measurements, the fit factor was calculated as:

$$Fit\ Factor = \frac{C_o}{C_i}$$

Where C_i is the in-mask concentration (corrected for the drying air and lung retention), and C_o is the outside-mask concentration.

The fit factor was calculated for each of the five fit test exercises listed in Table IV. The exercise with the lowest fit factor, referred to from now on as the *minimum exercise fit factor*, was the fit factor used for that test in the analysis. If the minimum exercise fit factor was ≥100, the result was recorded as a pass, and if the minimum exercise fit factor was <100, the result was recorded as a fail. This is the standard fit test pass requirement in the UK, as described in HSE guidance OC 282/28 (HSE, 2012).

Table IV. Exercise protocol for the generated aerosol QNFT method

Exercise	Approximate duration (min)	Measurement duration (min)
Walking on treadmill	1.5	1
Walking on treadmill while moving head side to side	1.5	1
Walking on treadmill while moving head up and down	1.5	1
Walking on treadmill while talking aloud	1.5	1
Walking on treadmill	1.5	1
Total time	**7.5**	**5**

Shortened QLFT Method

The shortened QLFT method followed the HSE guidance OC 282/28, but using the exercise protocol given in Table III.

For the sensitivity test, the wearer was asked to wear a hood, into which Bitrex (denatonium benzoate) was atomised (sprayed) using a hand operated nebuliser. The wearer was asked to breathe through their mouth, and the number of sprays of Bitrex required before they could taste it was recorded.

Only volunteers who could taste the Bitrex in 10 sprays or fewer were included in this study. Volunteers with a higher taste threshold were excluded on the basis that it would be inappropriate to use a potentially less sensitive test method for them. If used in the workplace, the shortened QLFT test method would therefore be restricted to those with a taste threshold of 10 sprays or fewer.

After the sensitivity test, the volunteer had a drink of water, and was given at least 15 minutes for their palate to clear. For test sessions where the QLFT was not the first fit test method to be performed, the sensitivity test was still performed first, and the PortaCount and/or generated aerosol QNFT was carried out in between the sensitivity test and the shortened QLFT.

At the beginning of the shortened QLFT, 10 sprays of Bitrex were introduced into the hood worn by the volunteer. The volunteer performed the exercises shown in Table III whilst standing. Each exercise lasted for 30 seconds, and every 30 seconds an additional 5 sprays of Bitrex were introduced into the hood.

If the volunteer tasted the Bitrex during the test, the test was aborted and the result reported as a fail. If the volunteer did not taste Bitrex during the test, they were asked to carefully remove the rubber cap on the second sampling port, thereby introducing a deliberate leak into the FFP3 without altering the fit. They were asked if they could taste the Bitrex through the deliberate leak, and the result was recorded. Provided the wearer did not taste Bitrex during the test, but did taste Bitrex through the deliberate leak, the result was recorded as a pass.

PortaCount Fit Test Method

The PortaCount fit tests were carried out using a model 8038 PortaCount (TSI, Minnesota, USA), and followed the protocol given in the HSE guidance OC 282/28. The test exercise protocol employed is given in Table V. The ambient particle concentration was maintained within the range 10,000 – 20,000 particles/cc. As for the generated aerosol QNFT method the fit factors were calculated for each of the fit test exercises listed in Table V . The exercise with the lowest fit factor, (i.e. the *minimum exercise fit factor)*, was the fit factor used for that test in the analysis. If the minimum exercise fit factor was ≥100, the result was recorded as a pass, and if the minimum exercise fit factor was <100, the result was recorded as a fail.

Table V. Exercise protocol for the PortaCount fit test method

Exercise	Exercise duration (min)
Normal breathing	1
Deep breathing	1
Moving head side to side	1
Moving head up and down	1
Talking aloud	1
Bending at the waist	1
Normal breathing	1
Total exercise duration	**7**

Note: The actual total duration of the fit test was 9 min 48 s due to the challenge measurement and challenge and mask purge times.

ANSI Z88.10 Comparison of Fit Test Methods

The fit test results from the proposed new shortened QLFT method were compared to the fit test results from the reference generated aerosol QNFT method using the criteria given in ANSI Z88.10.

In the ANSI Z88.10 comparison, data pairs where the reference fit test result is within one coefficient of variation of the fit test pass level requirement are excluded from the analysis. This prevents borderline fits from having a major effect on the results. The ANSI Z88.10 comparison also excludes data pairs with a reference test fit factor of less than 5% of the pass level, as these are deemed to correlate too poorly to give meaningful value to the analysis.

Once all exclusions have been applied, the ANSI Z88.10 comparison requires a minimum of 100 data pairs for the analysis to be considered valid. Since the identification of poor fits is the primary purpose of the fit test, at least 50 of these valid data pairs must have failed the reference fit test method. The valid data pairs are then summarised in a contingency table, and the results used to calculate the test statistics. Table VI and Table VII show how the contingency table and test statistics are derived. The most important of the test statistics is the *test sensitivity*, which gives the probability that the new fit test method will correctly identify a poor fit.

Table VI. ANSI Z88.10 comparison 2 x 2 contingency table definition

	Failed reference fit test	**Passed reference fit test**
Passed new qualitative fit test	*A*	*B*
Failed new qualitative fit test	*C*	*D*

Consistency Checks

A comparison of fit test methods relies on the fit of the mask remaining the same on the wearer's face for all tests. Wearers were asked not to adjust the fit of the mask between fit tests, but movement of the mask cannot realistically be eliminated.

In order to account for large changes of the fit of the mask in between the three fit test methods, a consistency check procedure was added into each test session. Prior to, in between and after each fit test method, a consistency check was carried out using the PortaCount. The volunteer was asked to stand still and breathe normally whilst looking straight ahead. The fit factor was measured for a period of 10 s, and compared to the other consistency check fit factors for that session. If there was a factor of 10 or more difference between the consistency check fit factors, the fit was deemed to have changed substantially (a difference greater than a factor of 10), and the data pair was excluded from the analysis.

Table VII. Description of ANSI Z88.10 comparison test statistics

Test statistic	Calculation	Requirement / recommendation *
Test sensitivity	$\frac{C}{A+C}$	≥ 0.95
Predictive value of a pass	$\frac{B}{A+B}$	≥ 0.95
Test specificity	$\frac{B}{B+D}$	> 0.50
Predictive value of a fail	$\frac{C}{C+D}$	> 0.50
P_o	$\frac{B+C}{A+B+C+D}$	-
P_e	$\frac{(A+B)(B+D)+(C+D)(A+C)}{(A+B+C+D)^2}$	-
Kappa statistic	$\frac{P_o-P_e}{1-P_e}$	> 0.70
* ≥0.95 for the *Test Sensitivity* is *required*; all other test statistics are *recommended*.		

RESULTS AND DISCUSSION

Models of mask

The five masks most commonly used by the NHS, as identified by the NHS Supply Chain, were selected for use in the study. They are identified here as Models A, B, C, D and E. Table VIII summarises the design feature of each model of mask used. Initial results showed very low QNFT (generated aerosol and PortaCount) pass rates for Models B and D. In order to generate sufficient data for the evaluation, Model B was replaced with Model F. Model F was not on the list of masks provided by the NHS Supply Chain (though available via the NHS Supply Chain), but was identified by the researchers as a model of mask that was likely to give a balance of passes and failures when fit tested on a group of wearers. Because models B and D were basically the same mask, but with model B having an exhalation valve, only this model was changed-out.

Table VIII. Description of the models of mask used for this study

Model of mask	Description
Model A	Lightweight horizontal fold-flat mask with a nose clip, non-adjustable straps and no exhalation valve.
Model B	Horizontal fold-flat mask with a nose clip, non-adjustable straps, a foam nose strip and an exhalation valve.
Model C	Moulded mask with a nose clip, non-adjustable straps, a gasket-type face seal and no exhalation valve.
Model D	Horizontal fold-flat mask with a nose clip, non-adjustable straps, a foam nose strip (similar to Model B) and no exhalation valve.
Model E	Fold-flat mask with a nose clip, non-adjustable straps, a foam nose strip (different to Models B and D) and no exhalation valve.
Model F	Moulded mask with a nose clip, non-adjustable straps, a gasket-type (similar to Model C) face seal and an exhalation valve.

Volunteers

In total, 26 volunteers participated in the study: 11 females and 15 males. Of the 26 volunteers, 5 completed testing with all six models of mask, 19 completed the testing with 5 models of mask, 1 completed the testing with four models of mask, and 1 completed the testing with two models of mask. Some volunteers were experienced RPE wearers, while others were relatively inexperienced.

Consistency checks

Analysis showed that 13 of the test sessions resulted in a change in the consistency check fit factor of greater than a factor of 10. This was taken to be evidence that the fit of the mask had changed significantly, and so these data pairs were excluded from further analysis.

Comparison of the shortened QLFT results with the generated aerosol results

The generated aerosol fit test exclusion zone for FFP3 masks had previously been measured by Frost *et al* (2015), who determined that it should cover fit factors from 92 to 108 exclusive. Frost *et al*'s exclusion zone was deemed to be applicable to this study as well, as it was measured using the same inward leakage test equipment and aerosol generating equipment. When applied to the dataset, 6 data pairs were excluded as the TIL fit factor was too close to the pass level of 100.

In addition to the exclusion zone, one test session resulted in a generated aerosol minimum exercise fit factor of 3. This is less than the cut off at 5% of the pass level, and so was also excluded. After all exclusions had been applied, 111 test sessions remained. Table IX is the contingency table for the remaining data. The calculated test statistics can be found in Table X.

From Table X, it can be seen that the shortened QLFT method failed to meet the requirements of the ANSI Z88.10 comparison. The *test sensitivity*, which is the primary requirement of the ANSI Z88.10 comparison, was 0.70, which is much lower than the requirement of ≥ 0.95. This means that 70% of the time, a poor fit (as identified by the generated aerosol QNFT method) resulted in a fail with the qualitative method. Conversely, 30% of poor fits resulted in a pass with the shortened QLFT method.

Table IX. Contingency table for the generated aerosol reference and shortened QLFT

	Failed generated aerosol reference	**Passed generated aerosol reference**
Passed shortened QLFT	21 (*A*)	28 (*B*)
Failed shortened QLFT	49 (*C*)	13 (*D*)

Table X. ANSI Z88.10 comparison test statistics the generated aerosol reference and shortened QLFT

Test statistic	**Calculation**	**Recommended value***	**Calculated value**
Test sensitivity	$\frac{\boldsymbol{C}}{\boldsymbol{A}+\boldsymbol{C}}$	≥ 0.95	0.70
Predictive value of a pass	$\frac{\boldsymbol{B}}{\boldsymbol{A}+\boldsymbol{B}}$	≥ 0.95	0.57
Test specificity	$\frac{\boldsymbol{B}}{\boldsymbol{B}+\boldsymbol{D}}$	≥ 0.50	0.68
Predictive value of a fail	$\frac{\boldsymbol{C}}{\boldsymbol{C}+\boldsymbol{D}}$	≥ 0.50	0.79
P_o	$\frac{\boldsymbol{B}+\boldsymbol{C}}{\boldsymbol{A}+\boldsymbol{B}+\boldsymbol{C}+\boldsymbol{D}}$	-	0.69
P_e	$\frac{(\boldsymbol{A}+\boldsymbol{B})+(\boldsymbol{C}+\boldsymbol{D})(\boldsymbol{A}+\boldsymbol{C})}{(\boldsymbol{A}+\boldsymbol{B}+\boldsymbol{C}+\boldsymbol{D})^2}$	-	0.51
Kappa statistic	$\frac{\boldsymbol{P_o}-\boldsymbol{P_e}}{\boldsymbol{1}-\boldsymbol{P_e}}$	> 0.70	0.37
* ≥0.95 for the *Test Sensitivity* is *required*; all other test statistics are *recommended*.			

The *predictive value of a pass* was 0.57, compared to the recommendation of ≥0.95, so this recommendation was not met. This means that 57% of the time, a fit identified as being a pass by the shortened QLFT method was actually a good fit, as measured by the generated aerosol QNFT method. Conversely, 43% of fits identified as good fits by the shortened QLFT method were identified as poor fits by the generated aerosol QNFT method.

The *test specificity* was 0.68, compared to the recommendation of >0.50, so this recommendation was met. This means that 68% of the time, a good fit (as identified by generated aerosol QNFT method) was identified as a good fit by the shortened QLFT method. This recommendation relates to efficiency, rather than safety.

The *predictive value of a fail* was 0.79, compared to the recommendation of >0.50, so this recommendation was met. This means that 79% of the time, a fit that failed the shortened QLFT method was actually a poor fit (as identified by the generated aerosol QNFT method).

The *Kappa statistic* was 0.37, compared to the recommendation of >0.70, so this recommendation was not met. The *Kappa statistic* gives a measure of the overall agreement between the two fit testing methods. Positive values of the *Kappa statistic* indicate some degree of agreement between the methods; a value of 1 would be perfect agreement. In this case, there is a degree of agreement between the two test methods, but it falls well short of the recommended level of agreement of >0.7.

Comparison of the shortened QLFT results with the PortaCount results

The data set was further analysed by comparing the shortened QLFT data set with the PortaCount data set, with the latter being considered as the reference. Table XI and Table XII show the contingency table and calculated ANSI requirements.

Table XI. Contingency table for PortaCount as a reference and shortened QLFT

	Failed PortaCount fit test	**Passed PortaCount fit test**
Passed shortened QLFT	21 (*A*)	24 (*B*)
Failed shortened QLFT	48 (*C*)	10 (*D*)

Table XII. ANSI Z88.10 comparison test statistics for PortaCount as a reference and shortened QLFT

Test statistic	Calculation	Recommended value*	Calculated value
Test sensitivity	$\frac{C}{A+C}$	≥ 0.95	0.70
Predictive value of a pass	$\frac{B}{A+B}$	≥ 0.95	0.53
Test specificity	$\frac{B}{B+D}$	> 0.50	0.71
Predictive value of a fail	$\frac{C}{C+D}$	> 0.50	0.83
P_o	$\frac{B+C}{A+B+C+D}$	-	0.70
P_e	$\frac{(A+B)+(C+D)(A+C)}{(A+B+C+D)^2}$	-	0.52
Kappa statistic	$\frac{P_o-P_e}{1-P_e}$	> 0.70	0.37
* ≥0.95 for the *Test Sensitivity* is *required*; all other test statistics are *recommended*.			

By comparison with the generated aerosol QNFT as the reference, the overall statistics are very similar. At a pass level of 100, the *test sensitivity*, which is the primary requirement of the ANSI Z88.10 comparison, was 0.70, which is much lower than the requirement of ≥0.95. This improved slightly to 0.78 when a pass level of 70 for the QNFT PortaCount fit tests was employed (Frost *et al*, 2014a) The Kappa statistic also increased from 0.37 to 0.47, but again this is still short of the recommendation of >0.70.

Comparison of wearer seal checks with the generated aerosol QNFT reference method

During testing, all wearers performed a wearer seal check prior to their first fit test, and the result was recorded. The results of the wearer seal checks were compared to the result of the generated aerosol reference fit test, using the ANSI comparison. Data pairs were excluded (as with previous comparisons) if the minimum generated aerosol fit factor was less than 5 (5% of the pass level) or if the consistency checks between the wearer seal check and the generated aerosol reference fit test were more than a factor of 10 different.

Table XIII gives the contingency table for the fit check compared to the reference TIL fit test method, and Table XIV gives the ANSI comparison test statistics for the wearer seal check compared to the generated aerosol reference fit test method.

Table XIII. Contingency table for the generated aerosol reference and the wearer seal checks

	Failed generated aerosol reference	**Passed generated aerosol reference**
Passed wearer seal check	37 (*A*)	33 (*B*)
Failed wearer seal check	30 (*C*)	12 (*D*)

Table XIV. ANSI Z88.10 comparison test statistics for the generated aerosol reference and the wearer seal checks

Test statistic	**Calculation**	**Recommended value***	**Calculated value**
Test sensitivity	$\frac{C}{A+C}$	≥ 0.95	0.45
Predictive value of a pass	$\frac{B}{A+B}$	≥ 0.95	0.47
Test specificity	$\frac{B}{B+D}$	≥ 0.50	0.73
Predictive value of a fail	$\frac{C}{C+D}$	≥ 0.50	0.71
P_o	$\frac{B+C}{A+B+C+D}$	-	0.56
P_e	$\frac{(A+B)+(C+D)(A+C)}{(A+B+C+D)^2}$	-	0.48
Kappa statistic	$\frac{P_o-P_e}{1-P_e}$	> 0.70	0.17
* ≥0.95 for the *Test Sensitivity* is *required*; all other test statistics are *recommended*.			

The test sensitivity was 0.45, which is substantially lower than the requirement of ≥0.95. This means that wearers correctly identified that a fit was inadequate in only 45% of cases. Given that wearers failed to identify inadequate fits in more than half of the mask donnings, this further supports the conclusions from previous work (Frost *et al*, 2014b) that a wearer seal check is not an effective substitute for a fit test when using FFP3 masks.

DISCUSSION

The results of the comparison study demonstrate the shortened QLFT method assessed in this study falls short of the ANSI requirements by some margin for both the recommended and required test statistics. The values are also lower than those obtained by Nelson *et al* (2003). Since a different shortened fit test protocol was adopted, a reference fit test pass criterion of the minimum rather than the overall fit factor was used, a different reference method was employed, NaCl aerosol was used rather than a corn oil aerosol, and that the very nature of the qualitative fit test method is subjective, this is not surprising. Previous studies have found that even the full-length QLFT method may not meet the ANSI criteria, when compared against the generated aerosol QNFT method (Frost *et* al, 2014a; Coffey *et al*, 2002). Additionally, although FFP3 were used, there is still a degree of filter penetration present and measured during the generated aerosol QNFT method, and this does raise the question as to what the ideal fit test reference method is (Frost *et al,* 2014a).

To account for an element of filter penetration, Frost *et at* (2014a) explored the effect of a different pass level applied to the generated aerosol reference method and also to the PortaCount method; this resulted in improved test statistics. When the pass level of 70 was applied to the comparison of the shortened QLFT method and generated aerosol QNFT method, there was an increase in the test sensitivity from 0.70 to 0.78. This still fell short of the requirement of ≥ 0.95.

Although not the main aim of this study, we did explore what the test statistics would look like if the PortaCount fit test method was used as the reference fit test method. The overall statistics compared favourably with the generated aerosol reference for the overall rejection of the shortened QLFT. Although further data on the suitability of the PortaCount method as a reference is required, it does appear that there is scope for using the PortaCount as a reference fit test method over and above the existing allowance in ANSI Z88 A2.1, which permits the use of an accepted QNFT protocol as a reference for evaluating a modified QNFT protocol.

While at face value, reducing the time taken for a fit test seems totally sensible, there must be a careful balance struck. Although the test should not be unduly lengthy it is important that sufficient time, resources and commitment – both on the part of the RPE wearer and fit tester – are realised for a fit test to be effective. The push to reduce the fit test protocols, TSI fast fit and REDON, while appearing to be saving time, may in the end result in less time given in that all-important face to face training in respiratory donning provided by competent fit testers. There must be a balance between minimising downtime and retaining the essential added benefits of a fit test.

CONCLUSIONS

The The shortened qualitative fit test method failed to meet the requirements of the ANSI Z88.10 comparison. The *test sensitivity*, which is the primary test statistic, and which gives the probability of the shortened fit test method failing a mask that is an inadequate fit, was 0.70; this is less than the requirement of ≥0.95. The value of 0.70 means that only 70% of inadequate fits would fail the shortened QLFT method. Conversely, 30% of poor fits would be expected to pass the shortened qualitative fit test method, a proportion that is unacceptably high.

Acknowledgements

The authors would like to thank the members of the HSE volunteer pool for their willingness to undergo countless fit tests.

Disclaimer

This publication and the work it describes were funded by the Health and Safety Executive (HSE). Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect HSE policy. Throughout this paper, reference is made to a number of proprietary and trademarked products, such as Bitrex, TSI, PortaCount, etc. Their mention here does not constitute any endorsement by the authors or HSE.

REFERENCES

ANSI (2010) American National Standard (ANSI/AIHA Z88.10-2010) Respirator Fit Testing Methods ISB: 978-1-935082-24-8.

BSI (1980) BS6016:1980 Specification for filtering facepiece dust respirators. ISBN 0 580 11705 7. BSI London.

CEN (2001) EN 13274-1:2001 Respiratory protective devices – Methods of test – Part 1: Determination of inward leakage and total inward leakage. CEN, Brussels. ISBN 0 580 37080 1.

CEN (2009) EN 149:2001+A1:2009 Respiratory protective devices – Filtering half masks to protect against particles – Requirements, testing, marking. CEN, Brussels. ISBN 978 0580 62117 8.

Coffey CC, Lawrence RB, Zhuang Z, Campbell DL, Jensen PA and Myers WR. (2002) Comparison of five methods for fit-testing N95 filtering-facepiece respirators. *Appl. Occup. Environ. Hyg.* 17:723–730.

Fit2Fit RPE Fit Test Providers Accreditation Scheme. http://fit2fit.org/. Accessed 29 November 2017.

Frost S, Mogridge R, Roff M. (2014a) Fit of Filtering Facepiece Class 3 (FFP3) Respirators Part 1: A Comparison of Fit Test Methods. JISRP, Vol 31, Issue 1, pp. 57-72.

Frost S, Mogridge R, (2014b) Fit of Filtering Facepiece Class 3 (FFP3) Respirators Part 2: The Value of Fit-Checking and Simple Subjective Assessments of Fit as an Alternative to a Fit Test. JISRP, Vol 31, Issue 2, pp. 79-88.

Frost S, Mogridge R, Roff M. (2015) HSE Internal Report RR1029. Review of fit test pass criteria for Filtering Facepieces Class 3 (FFP3) Respirators. http://www.hse.gov.uk/research/rrhtm/rr1029.htm. Accessed 29 November 2017.

HSE (2002) L132 Control of lead at work (Third edition). Approved Code of Practice and guidance. HSE Books ISBN 9780717625659. http://www.hse.gov.uk/pubns/books/l132.htm. Accessed 8 January 2018

HSE (2012) Guidance: Fit testing of Respiratory Protective Equipment Facepieces HSE OC 282/28

(30/4/2012) http://www.hse.gov.uk/foi/internalops/ocs/200-299/282_28.pdf. Accessed 29 November 2017.

HSE (2013a) L143 Managing and working with asbestos. Control of Asbestos Regulations 2012. Approved Code of Practice and guidance. HSE Books ISBN 9780717666188. http://www.hse.gov.uk/pubns/books/l143.htm. Accessed 8 January 2018.

HSE (2013b) L5 The Control of Substances Hazardous to Health Regulations 2002. Approved Code of Practice and guidance. HSE Books ISBN 9780717665822. http://www.hse.gov.uk/pubns/books/l5.htm. Accessed 8 January 2018.

Nelson TJ, Janssen LL, Luinenburg MD *et al*. (2003) Determination of an abbreviated qualitative fit test using bitter aerosol. JISRP; 20: 102-9.

Richardson WA, Hofacre KC, Weed J, Holm R, Remiarz R. (2013) Evaluation of a faster fit testing method for full-facepiece respirators based on the TSI PortaCount®. JISRP, Vol. 30, Issue 2, pp. 116-128.

Richardson WA, Hofacre KC, Weed J, Holm R, Remiarz R. (2014a) Evaluation of a faster fit testing method for elastomeric half-mask respirators based on the TSI PortaCount®. JISRP, Vol. 31, Issue 1, pp. 9-22.

Richardson WA, Hofacre KC, Weed J, Holm R, Remiarz R. (2014b) Evaluation of a faster fit testing method of filtering facepiece respirators based on the TSI PortaCount®. JISRP, Vol. 31, Issue 1, pp. 43-56.