

CRANFIELD UNIVERSITY

Karen M. Long

Masking the Problem: Human Errors in Qualitative Respiratory Protective
Equipment Fit Testing

Cranfield Defence and Security
PhD in Defence and Security

PhD

Academic Year: 2018 - 2022

Supervisor: Dr Nathalie Mai
Associate Supervisor: Dr Michael Williams

April 2022

CRANFIELD UNIVERSITY

Cranfield Defence and Security
PhD in Defence and Security

PhD

Academic Year 2018 - 2022

Karen M. Long

Masking the Problem: Human Errors in Qualitative Respiratory Protective
Equipment Fit Testing

Supervisor: Dr Nathalie Mai

Associate Supervisor: Dr Michael Williams

April 2022

© Cranfield University 2022. All rights reserved. No part of this
publication may be reproduced without the written permission of the
copyright owner.

ABSTRACT

Respiratory protective devices are designed to protect wearers from inhalation hazards in the workplace. The performance of tight-fitting devices relies on a good face-to-facepiece seal, and this is tested through a prescribed procedure known as a ‘fit test’ which provides an assessment of face seal leakage using either a quantitative or qualitative detection method. Numerous anecdotal reports exist about human errors in workplace application of the qualitative fit test method, but published research is limited to laboratory-based studies.

This thesis provides the first investigation of real-world human errors in qualitative fit testing and proposes measures to improve the reliability of this widely used fit test method. Through an extensive literature review, the study examines the multi-faceted body of current research and identifies a major research gap in the knowledge pertaining to human errors. An assessment of the prevalence and sources of human errors in qualitative fit testing is conducted through secondary analysis of data acquired during Fit2Fit practical assessments. Findings provide information about the types of errors made by real-world fit testers, trends over time, and differences between Fit2Fit accredited and non-accredited fit testers.

Through a theoretical analysis of the fundamental method principles, this thesis postulates a set of ‘critical errors’ which have the potential to undermine the reliability of the test. An empirical assessment of the impact of critical errors is conducted through a series of aerosol concentration studies and live shortened fit tests. When these impacts are applied to secondary data of errors made by real-world fit testers, this thesis offers the first estimate of a workplace (Beta) error value for this fit test method. Through examination of current control measures, this thesis highlights gaps in current resources, and proposes a set of engineering and administrative solutions which aim to address critical errors and improve workplace safety.

Keywords:

Facepiece, face seal, leakage, taste test, hood test, Bitrex.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisory team for their abiding support and mentorship. Particularly to Prof Ken McNaught whose sunny outlook on the world of statistics helped me to become a better researcher and critical thinker. To Prof Emma Sparks for her ‘outside the box’ thinking which helped me to remember the bigger picture. To my supervisor, Dr Nathalie Mai, whose constantly cheerful, friendly, and helpful advice has been invaluable to me throughout this research. Nathalie has allowed me the freedom to pursue my own research interests, provided steady support in pursuit of good science, and always offered honest feedback – *merci mille fois Nat!* Particular thanks go to my assistant supervisor Dr Mike Williams who shares my enduring passion for improving respiratory protection and helped me to start on this research journey. Mike has been my sense-checker when hours of late-night writing needed a critical eye and my counsel in times of stress; I truly could not have done this without him.

I wish to acknowledge the invaluable feedback and suggestions from peers and participants at various conferences and seminars. I would particularly like to thank members of the International Society for Respiratory Protection for their learned questions which helped me to think critically and reflect on my own research. Thank you also to members of the BSIF Fit2Fit board who allowed me unfettered access to the archive of Fit2Fit assessment papers for my secondary research and put up with my piles of paperwork for months!

Most importantly, I would like to thank my family for their patience, support, and love without which this thesis would not have been possible. To my parents, for my innate drive and the promise of being ‘so proud’ which has helped me through the tough times. To my husband Simon, and children Zara and Alexander, thank you for allowing me the time and space to pursue my dreams; your support means the world to me and none of this would be worth it without you. To my sister Claire, thank you for being my best friend, business partner and sounding board in good times and bad. You truly are my rock. Last but not least, thank you to my dog Bonnie for being my faithful companion on long walks and for being the best listener when no one else is around.

TABLE OF CONTENTS

ABSTRACT	i
ACKNOWLEDGEMENTS	ii
LIST OF FIGURES	vi
LIST OF TABLES	vii
LIST OF EQUATIONS.....	ix
LIST OF ABBREVIATIONS	x
1 Introduction	1
1.1 Background.....	1
1.2 Aims and objectives.....	7
1.3 Thesis structure.....	8
1.4 References	12
2 Human Errors in qualitative fit testing: a literature review	17
2.1 Abstract.....	17
2.2 Background.....	18
2.3 Method.....	21
2.4 Results	23
2.4.1 Statistical errors inherent in the QLFT protocol.....	23
2.4.2 Errors arising from the subjective nature of the QLFT method	28
2.4.3 Human errors made in the real-world application of the QLFT method.....	32
2.5 Discussion.....	34
2.6 Conclusion.....	37
2.7 References	38
3 Human errors in qualitative fit testing: a study of real-world fit testers.....	45
3.1 Abstract.....	45
3.2 Background.....	46
3.3 Method.....	48
3.3.1 Data collection.....	48
3.3.2 Scoring and assessment criteria.....	50
3.4 Results	51
3.5 Discussion.....	57
3.6 Conclusion.....	61
3.7 References	62
4 Human errors in qualitative fit testing: a theoretical study of the impact of human errors on fit test outcome	64
4.1 Abstract.....	64
4.2 Background.....	65
4.3 Method.....	67
4.3.1 Selection of evaluation criteria	67
4.3.2 Theoretical assessment of impact	68

4.4 Results	71
4.5 Discussion.....	76
4.6 Conclusion.....	80
4.7 References	81
5 Human errors in qualitative fit testing: a practical study of the impact of human errors on fit test outcome.....	84
5.1 Abstract.....	84
5.2 Background.....	85
5.3 Method.....	87
5.3.1 Set up of control and error scenarios	87
5.3.2 Aerosol mass concentration studies.....	90
5.3.3 Shortened fit test studies.....	92
5.4 Results	96
5.4.1 Aerosol mass concentration studies.....	96
5.4.2 Shortened fit test studies.....	97
5.5 Discussion.....	102
5.6 Conclusion.....	107
5.7 References	108
6 Human errors in qualitative fit testing: an investigation of current controls.....	112
6.1 Abstract.....	112
6.2 Background.....	113
6.3 Method.....	115
6.3.1 Assessment of current administrative controls.....	115
6.3.2 Assessment of current engineering controls.....	117
6.4 Results	119
6.4.1 Current administrative controls	119
6.4.2 Current engineering controls	122
6.5 Discussion.....	128
6.6 Conclusion.....	131
6.7 References	132
7 Conclusions and summary.....	135
7.1 Summary of research findings.....	135
7.2 Contribution to knowledge	139
7.3 Recommendations	141
7.3.1 Recommendations for new administrative controls	141
7.3.2 Recommendations for new engineering controls	150
7.4 Research impact and dissemination.....	153
7.5 Limitations.....	154
7.5.1 Real-world data	154
7.5.2 Choice of the TSI Portacount quantitative reference method	155
7.6 Suggestions for future research	158
7.7 References	160

8 Appendices	163
8.1 Appendices A-N	163
Appendix A Keyword combinations for themes	163
Appendix B Fit2Fit Criteria referencing	164
Appendix C Sources of error during Fit2Fit QLFT practical assessments, 2009 – 2019	167
Appendix D Theoretical impacts of errors for Fit2Fit QLFT practical assessment criteria	169
Appendix E Theoretical impacts of errors applied to results of real-world study....	174
Appendix F Anatomy of BS ISO 16975-3 qualitative fit test nebuliser.....	176
Appendix G Informed consent form.....	177
Appendix H Results of aerosol mass concentration studies	179
Appendix I Results of shortened fit test studies	180
Appendix J Number of applications and timings for control scenario and error scenarios Fit test solution under-administered and Fit test timing incorrect.....	181
Appendix K Critical errors identified against key principles of the validated QLFT method	182
Appendix L Search terms and key phrases for critical errors identified against key principles of the validated QLFT method.....	183
Appendix M Current administrative and engineering controls available to qualitative fit testers.....	184
Appendix N Analysis of engineering controls in QLFT equipment	185
8.2 References	186

LIST OF FIGURES

Figure 2-1 Beta values reported in articles included in this study against a required fit factor (RFF) of 100, listed in chronological order	26
Figure 2-2 Research gaps identified in three themes pivotal to an understanding of errors in qualitative fit testing	34
Figure 3-1 Schematic of the process leading to Fit2Fit accreditation	48
Figure 3-2 Frequency of errors recorded in Fit2Fit QLFT practical assessment papers, 2009 – 2019.....	52
Figure 4-1 Process for theorising and judging the impact of human errors in QLFT for Fit2Fit criterion 8d	69
Figure 4-2 Distribution of error scenarios for QLFT key principles 1), 2), 3) or none..	71
Figure 5-1 Set up of control and error scenarios for nebuliser function tests	88
Figure 5-2 Procedure for aerosol mass concentration studies	91
Figure 5-3 Location of the sample port and adjustable leak port	92
Figure 5-4 Process for shortened fit tests	94
Figure 6-1 Gaps in administrative controls between QLFT equipment manufacturers' user instructions and INDG479, Fit2Fit QLFT Companion and Fit2Fit Approved Training content	122
Figure 6-2 Labelling / colouring of nebulisers identified as a control in mitigating the critical error <i>Fit test nebuliser filled with sensitivity solution</i>	125
Figure 6-3 Engineering controls identified as assisting in proper nebuliser filling.....	125
Figure 6-4 Presence / absence of controls for critical errors in QLFT	126

LIST OF TABLES

Table 3-1 Descriptive statistics by outcome for Fit2Fit QLFT practical assessments during the period 2009 – 2019	51
Table 3-2 Descriptive statistics for Fit2Fit QLFT practical assessment papers, 2009 – 2019.....	53
Table 3-3 Descriptive statistics by accreditation status for Fit2Fit QLFT practical assessments, 2009 – 2019	53
Table 3-4 Frequency of errors by section recorded in Fit2Fit QLFT practical assessment papers, 2009 – 2019	54
Table 3-5 Top ten sources of error recorded in Fit2Fit QLFT practical assessment papers, 2009 – 2019.....	55
Table 3-6 Descriptive statistics by Fit2Fit assessor for Fit2Fit QLFT practical assessments, 2009 – 2019	56
Table 4-1 Fit2Fit criteria with two or more theorised critical error scenarios	72
Table 4-2 Critical errors identified against the key principles of the QLFT method	73
Table 4-3 Critical errors applied to total errors made by real-world fit testers during Fit2Fit QLFT practical assessments [8]	74
Table 4-4 Critical errors applied to fit tests observed during Fit2Fit QLFT practical assessments	75
Table 5-1 Conditions for control and error scenarios for assessment of practical impact of equipment errors in QLFT	87
Table 5-2 Results of aerosol mass concentration tests for control and error scenarios..	96
Table 5-3 Summary of results for quantitative and qualitative fit tests for control and error scenarios.....	98
Table 5-4 Error scenarios with the highest impact on the ability of wearers to detect a known leak in QLFT	100
Table 6-1 UK QLFT equipment (kit) providers	119
Table 6-2 Administrative controls identified in QLFT guidance, training, and equipment user instructions	120
Table 6-3 Analysis of the control of critical errors in QLFT through administrative resources	121
Table 6-4 Engineering controls in QLFT kits	123
Table 6-5 Control of critical errors through engineering resources provided in QLFT kits	124

Table 7-1 Gaps identified in current administrative controls and proposed new controls	
.....	142
Table 7-2 Gaps identified in current engineering controls and proposed new controls	151

LIST OF EQUATIONS

Equation 7-1 Fit factor calculation	155
Equation 7-2 Overall quantitative fit factor calculation for the HSE INDG479 protocol based on seven test exercises	156

LIST OF ABBREVIATIONS

AIHA	American Industrial Hygiene Association
ANSI	American National Standards Institute
APC	Ambient Particle Counting
BS	British Standard
BSIF	British Safety Industry Federation
CI	Confidence Interval
CNP	Controlled Negative Pressure
COPD	Chronic Obstructive Pulmonary Disease
COSHH	Control of Substances Hazardous to Health
FF	Fit Factor
FFP	Filtering Facepiece
HSE	Health and Safety Executive
ID	Identification
INDG	Industry Guidance
ISO	International Organisation for Standards
MN	Minnesota
NIOSH	The National Institute for Occupational Safety and Health
OEM	Other Equipment Manufacturer
OSHA	Occupational Safety and Health Administration
PCR	Polymerase Chain Reaction
PM	Particulate Matter
PPE	Personal Protective Equipment
QLFT	Qualitative Fit Test
QNFF	Quantitative Fit Factor
QNFT	Quantitative Fit Test
QR	Quick Response

RFF	Required Fit Factor
RPD	Respiratory Protective Device
RPE	Respiratory Protective Equipment
SARS	Severe Acute Respiratory Syndrome
<i>SD</i>	Standard Deviation
SPSS	Statistical Programme for Social Sciences
TIL	Total Inward Leakage
UK	United Kingdom
USA	United States of America

1 Introduction

1.1 Background

Occupational exposure to airborne hazardous substances is of particular concern in workplace risk assessments where the links with respiratory diseases are well known [1][2]. A recent Health and Safety Executive report has estimated that 12,000 lung disease deaths and 17,000 self-reported breathing problems each year are associated with past exposure to hazardous substances in workplace settings [3][4]. Although the list is extensive, the most prevalent diseases were reported to include Chronic Obstructive Pulmonary Disease (COPD), lung cancer (both asbestos and non-asbestos-related), mesothelioma, silicosis, and asthma. Since many of these diseases have a long latency, their effects on worker health are often only apparent many years or decades after original exposure to an inhalation hazard. A wide variety of different substances carry the potential to cause harm by inhalation, these include dusts, mists, gases, fumes, micro-organisms, and airborne pathogens [5]. Although some respiratory hazards (e.g., asbestos) are now banned from use, their widespread historical application means that considerable risks still exist, particularly in the construction sector [6]. Beyond construction, inhalation hazards are present from a wide variety of sources across a multitude of diverse applications including nuclear power generation, mining, quarrying, agriculture, firefighting, chemical and petrochemical processing. In the healthcare sector, the SARS-CoV-2 pandemic has highlighted the risks of a respiratory hazard in the form of an infectious airborne pathogen; a risk which has had far reaching consequences.

Control measures which are designed to eliminate or reduce the risks of respiratory hazards at source are known to be the most effective method of mitigation, but their use is limited by both cost and practicality [5]. Millions of workers therefore rely on respiratory protective devices (RPD) to provide protection against the respiratory hazards they encounter at work. The ability of RPD to provide the wearer with breathable air is achieved either by filtration of the immediate atmosphere (termed ‘respirators’) or through a supply of clean air (termed ‘breathing apparatus’) [7]. The interface between the wearer and the RPD, known as the ‘facepiece’, can be designed to be loose-fitting (such as in powered hoods or helmets) or tight-fitting (such as in disposable and reusable

half and full-face devices) [7]. Whilst the protection afforded by loose-fitting facepieces relies on powered air to prevent ingress of hazardous atmospheres, the protection afforded by tight-fitting facepieces relies upon an effective facial seal [7]. This facial seal must accommodate the normal movements associated with workplace activities whilst maintaining a suitable and effective contact between the wearer's skin and the facepiece of the RPD. Since humans display innumerable differences in facial characteristics, it would be unlikely for one model of facepiece to accommodate all individual face shapes and sizes [8]. The ability to assess the adequacy of the facial seal of tight-fitting RPD facepieces through a 'fit test' is therefore critical to assuring the ability of these devices to be suitable in protecting the wearer. Such is the vital nature of the interface between the wearer and the facial seal of the RPD, that fit testing is a mandatory requirement for all tight-fitting facepieces in use in the UK workplace [5]. Fit testing was first introduced in the UK into the Approved Codes of Practice supporting the Control of Asbestos at Work Approved Code of Practice in 1999, and later expanded to all workplace settings through inclusion in the Control of Substances Hazardous to Health Regulations in 2002 [9][5]. Whilst in the US, fit testing has been a mandatory requirement since 1969, other national regulatory bodies have been slower to act, with only a handful of countries formally adopting the International Standard and setting any local recommendations for fit testing of RPD [10][11][12].

Fit testing has a wide and colourful history with a multitude of different methodologies having been proposed. Some of the earliest attempts at assessing the facial fit of tight-fitting facepieces were made by the United States Bureau of Mines. An information circular issued in 1940 detailed the testing schedules for the approval of respiratory protective equipment which included both laboratory tests and 'Man tests' [13]. Guidance for conducting 'Man tests' provided the following instructions for dispersoid (dust, fume, mist and fog) respirators:

'The dispersoid respirator is worn while coal dust is blown into the face. If no leakage occurs there is a sharp line of demarcation at the point of contact of the facepiece with the face'.

Whilst unsophisticated, this method set the stage for later quantitative and qualitative fit test methods where face seal leakage was assessed through visualisation, sensory response, or electronic detection.

Quantitative fit test (QNFT) methods provide an objective measure of fit termed a ‘fit factor’ which is derived as the ratio of concentration, or pressure, of a challenge agent outside and inside of a tight-fitting facepiece [8][12]. Early quantitative methods were developed in response to the requirements laid down by the US Atomic Energy Commission Respirator Committee to assess the performance of respirators against radioactive contaminants [14]. One such method utilised the fluorescent properties of uranine to quantify facepiece penetration to a lower detection limit of 0.05% [15]. Later QNFT methods used forward light-scattering photometers to measure inward leakage of a detectable challenge agent such as a polydisperse di (2-ethylhexyl) phthalate [16]. The first automated QNFT system was described in 1986 for use by the US military [17]. This system employed a computer-controlled forward light-scattering photometer detection system, which the authors reported as providing a ‘*substantial time and cost saving*’. A later miniaturised continuous flow Condensation Nucleus Counting device was developed by the US Army Chemical Research, Development and Engineering Center in conjunction with TSI Inc. (St. Paul, MN) [18]. This device was the forerunner to the TSI Portacount™, the most widely used ambient particle counting QNFT device today [19].

Whilst QNFT methods are often highly sophisticated, these procedures are not without their limitations, and studies which have attempted to compare the resulting fit factors give rise to much debate in this field [20][21]. Possible sampling errors in QNFT methods have been extensively studied and have been shown to arise from streamlining effects, leak site location, particle size distribution and lung deposition [22][23]. Nevertheless, QNFT methods are often described as ‘gold standard’ procedures and are considered ‘objective’ methods, since determination of leakage derives from instrumentation or other non-human response. Conversely, Qualitative fit test (QLFT) methods are considered ‘subjective’ methods, where determination of leakage relies upon the wearer’s sensory response to a ‘challenge agent’.

Early QLFT methods used challenge agents that were selected on the basis that they would cause visible irritation to the wearer if leakage occurred. The ‘irritant smoke’

method used the wearer's involuntary response to hydrogen chloride fumes which were produced (along with tin compounds) by a reaction of stannic chloride (SnCl_4) with moisture in the air [24]. A 2005 study by Campbell et al reported on the findings of a US labour market survey which indicated that the irritant smoke method was the most common fit testing practice amongst US private sector employers at that time [25]. This method fell out of use however following indications of health concerns including skin and eye irritation, one of which was reported as being serious enough to require hospital treatment [26].

The isoamyl acetate ('banana oil') odour test was the first QLFT method to use the wearer's voluntary sensory response [27]. This method was recognised in UK fit testing guidance up to 2019, but a lack of commercial availability and the practicalities of odour preparation in workplace settings later saw this method withdrawn from guidance [28]. Other more imaginative QLFT methods included the use of photosensitive paper which offered a visual indication of leakage, and biological monitoring using chlorofluorocarbon 113 (CFC-113, 1,1,2 trichloro-1,2,2 trifluoroethane) [29][30]. Although inexpensive to perform, these methods failed to gain commercial attraction since the practicalities of challenge-agent preparation and detection meant that their use was restricted to laboratory settings.

Of all the QLFT methods proposed, only two remain in widespread use today: the BitrexTM (denatonium benzoate) method and the sodium saccharin method [8]. Commonly referred to as the 'Bitrex' and 'saccharin' QLFT methods, both methods utilise a challenge agent comprising a strong-tasting substance and rely on the wearer's taste response to identify leakage. The saccharin QLFT method was proposed by the Minnesota Mining and Manufacturing Company, submitted to the US Occupational Safety and Health Administration (OSHA) for inclusion in the Lead Standard in 1982 and evaluated for use in 1984 [31][32]. The Bitrex method, evaluated in 1995, was developed to address the possible drawbacks of the saccharin method which was found to elicit a sweet, 'tolerable' response, leading some researchers to suggest that leakage may go unreported [33][34]. With its characteristically unpleasant taste at parts per million level and low toxicity, Bitrex was considered a robust candidate as a challenge agent in QLFT [35]. Despite the concerns raised about saccharin detection, both methods were

commercialised by respiratory manufacturers for fit testing of their tight-fitting filtering facepiece and reusable half masks, and soon became the first choice QLFT methods [36].

The current UK protocols for the Bitrex and saccharin QLFT methods, along with two QNFT methods, are described in the Health and Safety Executive (HSE) INDG479 guidance document on fit testing [8]. The procedure for both QLFT methods is based upon a subtle principle which assumes that a wearer will be able to identify a poor-fitting facepiece through detection of leakage of a strong-tasting substance (the challenge agent). The procedure follows a strict protocol which requires the wearer's taste threshold first to be screened against a diluted version of the challenge agent - a procedure known as the 'sensitivity test'. The results of the sensitivity test determine the amount of a more concentrated version which is used to challenge the facepiece during the second part of the procedure known as the 'fit test'. Both parts of this procedure rely on the wearer's ability and willingness to correctly identify and report taste – a subject which has been widely discussed by researchers in this field [34].

The saccharin QLFT method was developed almost 40 years ago, and little has changed in the design of the rudimentary equipment or methodology in the intervening period [31][32]. QLFT equipment requires manual operation and management, and thus the method relies heavily upon the skills of the human operator. During both the sensitivity and fit test procedures, aerosolisation of the challenge agent is achieved (according to the Venturi effect) through actuation of a hand-operated jet nebuliser, and containment / management of the aerosol is achieved through use of a test hood [37]. Proper use of this equipment and application of the test protocol are fundamental to the principles of the QLFT method and errors made by human operators introduce the potential for those principles to be undermined. Any process that involves a human operator has the potential to be affected by human errors, and errors which result in an unreliable fit testing outcome are of serious concern for the millions of workers that rely on the protection afforded by tight-fitting RPD. The worst outcome of human errors in QLFT is the potential for a wearer with an ill-fitting RPD to pass a fit test erroneously leading to a false pass result and assignment of a leaking RPD for use in the workplace. A false pass fit test result offers the wearer false confidence that a facepiece is safe to use in the workplace, when in truth it is not, potentially putting their health or life at risk.

Coupled with the absence of a requirement for formal training or certification of fit testers, there have been numerous anecdotal reports of the effects of human errors in workplace QLFT [38][39][25]. Although commentators have expressed concerns about the lack of competency of some fit test operators, there have been no published studies which have attempted to provide an understanding of human errors in QLFT to date [39]. Errors in laboratory QLFT have been extensively studied, but all have focused on the errors inherent in the method (the Alpha and Beta errors) rather than in the practical application of the method [40][41][42]. Errors arising from the subjective nature of the method have been widely discussed by numerous researchers, but none has examined the role of the fit test operator in managing the QLFT process [32][33]. Existing research thereby fails to provide an understanding of human errors in workplace QLFT and hence fails to address the potential for false passes and assignment of poor fitting RPD to wearers. The importance of research in understanding human errors in QLFT cannot be underestimated; the protection afforded by tight-fitting RPD facepieces is dependent upon an effective facial seal and human errors which lead to a poor-fitting facepiece being used in the workplace may have very grave consequences indeed.

1.2 Aims and objectives

This thesis presents a series of original studies which address human errors in QLFT through generation and analysis of data that is unique to this work. Data is generated through secondary analysis of Fit2Fit practical assessment papers, and through a series of studies that examine the theoretical and practical impacts of human errors. Building upon the background of existing research, the aim of this thesis is to better understand human errors in QLFT and relate the findings to impact on fit test outcome. The key motivation behind this aim is to improve the safety of workplace QLFT, and this overarching driver sets the rationale for discussions and suggestions for real-world change. The core research question that informs this thesis is summarised below:

Core research question:

How do human errors affect the outcome of a qualitative fit test and hence the potential for false pass results?

This thesis extends the boundaries of current knowledge through investigation of human errors in QLFT through a series of research objectives, summarised below.

Objective 1: To evaluate the current understanding of sources of error in QLFT through a review of current literature, and thus to identify research gaps.

Objective 2: To investigate the prevalence and sources of human errors amongst real-world QLFT operators.

Objective 3: To understand and assess the theoretical impacts of human errors upon QLFT outcome and hence the potential for false passes.

Objective 4: To evaluate the practical impacts of human errors upon QLFT outcome and hence the potential for false passes.

Objective 5: To assess the current control measures in QLFT and identify gaps to inform future improvements in design and methodology aimed at mitigating the effects of human errors

These objectives each contribute towards the overall aim and provide a cohesive study of human errors in QLFT thus offering a significant intellectual contribution in this field.

1.3 Thesis structure

This thesis is structured in ‘paper format’, whereby its intellectual contribution is presented through a series of distinct chapters written in the form of self-contained journal papers. Each chapter addresses one of the five research objectives detailed in section 1.2 and contains all aspects of a typical journal paper including background / rationale, method, results, and discussion sections. Collectively, these chapters describe a single programme of research which answers the core research question and effectuates the overall aim of this thesis. This thesis ends with a final conclusions chapter which brings together the major findings from each paper, and proposes new guidance, strategies, and solutions for tackling human errors in QLFT. A summary of each of the core chapters of this thesis is provided below.

Chapter 2: Human errors in qualitative fit testing: a literature review

This chapter addresses the first research objective that guides this study namely, to evaluate the current understanding of sources of error in QLFT, and to identify research gaps. The chapter begins by describing the methodology used for the study and details the processes used in identifying, reviewing, and synthesising the key literature in this field. The chapter presents a thematic analysis of the results arranged according to the three themes identified: statistical errors inherent in the method, errors arising from the subjective nature of the QLFT method, and human errors in real-world application of the QLFT method. Gaps in current understanding are identified and the results are synthesised in the discussion section which includes implications of findings in terms of real-world use of the method.

Chapter 3: Human errors in qualitative fit testing: a study of real-world fit testers

This chapter is linked to the second research objective and provides an analysis of the prevalence and sources of human errors in QLFT. The chapter opens with a description of the methodology which details the evaluation of human errors through secondary analysis of practical assessment papers gathered through the British Safety Industry

Federation (BSIF) RPE Fit Test Providers ‘Fit2Fit’ Accreditation Scheme (Fit2Fit scheme). An overview of the Fit2Fit scheme procedure for assessing and awarding Fit2Fit accreditation to competent fit testers is outlined along with a schematic flowchart to provide clarity of the process. The results are analysed and presented according to both prevalence and source of error with reference to the Fit2Fit scheme syllabus and practical assessment criteria. Further analysis reveals the trends over time and the differences between accredited and non-accredited fit testers in terms of the number of human errors observed. The chapter concludes with a discussion about the possible limitations of this research along with the implications of the observed errors in terms of the reliability of the QLFT method.

Chapter 4: Human errors in qualitative fit testing: a theoretical study of the impact of human errors on fit test outcome

This chapter addresses the third research objective which is to form a theoretical understanding and assessment of the impact of human errors on QLFT outcome. The paper commences by identifying the key principles of the QLFT method and how they underpin the fundamental theory of challenge aerosol generation, containment, and detection. A theoretical analysis of the impacts of human errors on these key principles is conducted for each separate stage and action of the fit testing process. Assessment of impact is then used to filter human errors, revealing those which are critical to the outcome of the fit test, and hence have the potential to lead to a false pass result. These errors, termed ‘critical errors’ form the basis of the discussion section which addresses the theoretical impact of human errors and links back to the errors observed amongst real-world fit testers in Chapter 3.

Chapter 5: Human errors in qualitative fit testing: a practical study of the impact of human errors on fit test outcome

This Chapter addresses the fourth research objective by assessing the practical impact of the ‘critical errors’ previously identified in Chapter 4. The practical impact of human errors is researched through studies which assess their effect on both challenge aerosol

concentration and fit test outcome in the presence of a known leak. The paper begins by describing the methodology involved in conceptualising and configuring the ‘control scenario’ and ‘error scenarios’ which aim to simulate typical errors made by real-world fit testers. The methodology adopted for the aerosol concentration study is then explained along with the procedure for size-segregated measurement using an optical aerosol monitor instrument. This section ends with an explanation of the method used for assessing the impact of errors on fit test outcome in the presence of a known leak using shortened QLFT against an ambient particle counting QNFT reference method. The findings of these two studies are reported and synthesised to provide a view on the impact of human errors on fit test outcome. A workplace Beta error is estimated to indicate the potential for a false pass result for human errors in the application of QLFT equipment.

Chapter 6: Human errors in qualitative fit testing: an investigation of current controls

This chapter addresses the final research objective namely, to assess the current control measures in QLFT and identify gaps to inform improvements in design and procedure. The paper begins by describing the methodology used to evaluate the current controls available to mitigate human errors in QLFT. Controls are classed as being either ‘administrative’ or ‘engineering’ in nature, and this paper continues by explaining the limits of evaluation and by describing the methods used to assess each type of control. The results are presented according to both type and source of error in order to reveal gaps in the current control mechanisms in QLFT. The discussion section provides an in-depth analysis of the real-world impacts of the gaps revealed and present suggestions for practical and administrative improvements which aim to reduce human errors in QLFT.

Chapter 7: Conclusions and summary

The final chapter of this thesis brings together the results from each of the separate papers and discusses their overall contribution in terms of real-world impact. Useful for practitioners, manufacturers, and policymakers, this chapter discusses the implications of the findings in terms of the reliability of the QLFT method and proposes new strategies

which aim to assist mitigation of the most critical human errors. This chapter presents recommendations for new guidance for fit test operators, and improvements in the design of QLFT equipment which together provide solutions for the gaps identified. The chapter concludes with a discussion of the limitations of this work and summarises the intellectual contribution of this thesis along with a list of presentations conducted throughout the registration period.

1.4 References

1. Nelson DI., Concha-Barrientos M., Driscoll T., Steenland K., Fingerhut M., Punnett L., et al. The global burden of selected occupational diseases and injury risks: methodology and summary. *American Journal of Industrial Medicine*. 2005; 48(6): 400–418. Available at: DOI:10.1002/ajim.20211
2. Driscoll T., Nelson DI., Steenland K., Leigh J., Concha-Barrientos M., Fingerhut M., et al. The global burden of non-malignant respiratory disease due to occupational airborne exposures. *American Journal of Industrial Medicine*. 2005; 48(6): 432–445. Available at: DOI:10.1002/ajim.20210
3. Health and Safety Executive. Occupational Lung Disease statistics in Great Britain. 2020. Available at: <https://www.hse.gov.uk/statistics>
4. Health and Safety Executive. Health and safety at work Summary statistics for Great Britain. 2021. Available at: <https://www.hse.gov.uk/statistics/overall/hssh2021.pdf>
5. Health and Safety Executive. L5 Control of substances hazardous to health. The Control of Substances Hazardous to Health Regulations 2002 (as amended). Approved Code of Practice and guidance L5 (Fifth edition). HSE Books; 2005. Available at: <http://www.hse.gov.uk/pubns/priced/l5.pdf>
6. Health and Safety Executive. Asbestos health and safety. Available at: <https://www.hse.gov.uk/asbestos/> (Accessed: 19 December 2021)
7. Health and Safety Executive. HSG53 Respiratory protective equipment at work. A practical guide. (Fourth edition). HSE Books; 2013. Available at: <http://www.hse.gov.uk/pubns/priced/hsg53.pdf>
8. Health and Safety Executive. INDG479 Guidance on respiratory protective equipment (RPE) fit testing. HSE Books; 2019. Available at: <https://www.hse.gov.uk/pubns/indg479.pdf>

9. Health and Safety Commission. Work with asbestos insulation, asbestos coating and asbestos insulating board. Control of Asbestos at Work Regulations 1987; Approved Code of Practice. 1999.
10. American National Standards Institute. ANSI Z88.2-1992: Respiratory Protection. 1992. Available at:
<https://law.resource.org/pub/us/cfr/ibr/002/ansi.z88.2.1992.pdf>
11. Murray A. UK best practices on RPD fit testing (Fit2Fit programme). National Fit Testing Programmes. 18th May 2021. International Society for Respiratory Protection; Available at: <https://www.isrp.com/minutes>
12. British Standards Institution. BS ISO 16975-3:2017 Respiratory protective devices. Selection, use and maintenance. Part 3: Fit-testing procedures. London: British Standards Institution; 2017.
13. Shrenk H. United States Department of the Interior, Bureau of Mines. Information Circular 7130. Testing Respiratory Protective Equipment for Approval. 1940. Available at: <https://ntrl.ntis.gov/NTRL/>
14. Campbell DL., Coffey CC., Lenhart S. Respiratory protection as a function of respirator fitting characteristics and fit-test accuracy. American Industrial Hygiene Association Journal. 2001; 62(1): 36–44. Available at:
DOI:10.1080/15298660108984607
15. Burgess W., Silverman L., Stein F. A new technique for evaluating respirator performance. American Industrial Hygiene Association Journal. 1961; 22(6): 422–429. Available at: DOI:10.1080/00028896109343432
16. Revoir W. Correlation of results of qualitative respirator facepiece fitting tests using isoamyl acetate vapor as test agent with results of quantitative respirator facepiece fitting tests using polydisperse diethylhexyl phthalate aerosol as test agent. Journal of the International Society for Respiratory Protection. 1983; 1(4): 42–59.
17. Brletich W., Allyson CR., Hughes FP. Development of an automated system for respirator quantitative fit testing. Journal of the International Society for Respiratory Protection. 1986; 4(4): 1–6.

18. Laye R. Evaluation of a miniaturized condensation nucleus counter for measurement of respirator fit factor. *Journal of the International Society for Respiratory Protection*. 1987; 5(3): 1–8.
19. Biermann AH., da ROZA RA., Sawyer SR., Foote KL., McCormack C., Sackett CR. Evaluation of Portacount for determining respirator fit factors. Part I: executive summary and Portacount manual. *Journal of the International Society for Respiratory Protection*. 1991; 9(1): 10–19.
20. Janssen LL., Luinenburg MD., Mullins HE., Nelson TJ. Comparison of three commercially available fit-test methods. *American Industrial Hygiene Association Journal*. 2002; 63(6): 762–767. Available at: DOI:10.1080/15428110208984767
21. Frost S., Mogridge R., Roff M. RR1029 Review of fit test pass criteria for filtering facepieces Class 3 (FFP3) respirators. Health and Safety Executive; 2015. Available at: <https://www.hse.gov.uk/research/rrpdf/rr1029.pdf>
22. Fairchild CI., Stampfer JF. Particle concentration in exhaled breath. *American Industrial Hygiene Association Journal*. 1987; 48(11): 948–949. Available at: DOI:10.1080/15298668791385868
23. Biermann AH., da Roza RA., Sawyer SR., Foote KL., McCormack C., Sackett CR. Evaluation of Portacount for determining respirator fit factors. Part II: laboratory breathing machine tests of stability and comparison with an aerosol photometer. *Journal of the International Society for Respiratory Protection*. 1991; 2: 26–41.
24. Marsh JL. Evaluation of irritant smoke qualitative fitting test for respirators. *American Industrial Hygiene Association Journal*. 1984; 45(4): 245–249. Available at: DOI:10.1080/15298668491399730
25. Campbell DL., Doney B., Groce D., Greskevitch M., Syamlal G., Coffey CC., et al. Respirator fit testing practices in the U.S. *Journal of the International Society for Respiratory Protection*. 2005; 22(1): 11–16.
26. Tharr D., Lenhart SW., Burroughs GE. Occupational health risks associated with the use of irritant smoke for qualitative fit testing of respirators. *Applied Occupational*

and Environmental Hygiene. 1993; 8(9): 745–751. Available at:
DOI:10.1080/1047322X.1993.10388191

27. Hardis KE., Cadena CA., Carlson GJ., da Roza RA., Held BJ. Correlation of qualitative and quantitative results from testing respirator fit. American Industrial Hygiene Association Journal. 1983; 44(2): 78–87. Available at:
DOI:10.1080/15298668391404419

28. Health and Safety Executive. OC 282/28 Fit testing of respiratory protective equipment facepieces. HSE Books; 2012.

29. Cohen HJ. Determining and validating the adequacy of air-purifying respirators used in industry. Part II: development of a new respirator fit test. Journal of the International Society for Respiratory Protection. 1984; 2(3): 305–312.

30. Decker JB., Crutchfield CD. Feasibility of a new method to determine respirator performance by biological monitoring of exhaled air. Journal of the International Society for Respiratory Protection. 1993; 1: 6–19.

31. Occupational Safety and Health Administration. Minnesota Mining and Manufacturing Company post hearing data submission on qualitative fit testing protocols under the lead standard before OSHA. Docket No. H-049A. 1981.

32. Marsh JL. Evaluation of saccharin qualitative fitting test for respirators. American Industrial Hygiene Association Journal. 1984; 45(6): 371–376. Available at:
DOI:10.1080/15298668491399965

33. Mullins H., Danisch S., Johnston A. Development of a new qualitative test for fit testing respirators. American Industrial Hygiene Association Journal. 1995; 56(11): 1068–1073. Available at: DOI:10.1080/15428119591016278

34. McKay RT., Davies E. Capability of respirator wearers to detect aerosolized qualitative fit test agents (sweetener and Bitrex) with known fixed leaks. Applied Occupational and Environmental Hygiene. 2000; 15(6): 479–484. Available at:
DOI:10.1080/104732200301269

35. Macfarlan Smith Ltd. Bitrex. Available at: <https://www.bitrex.com/> (Accessed: 6 May 2021)
36. 3M. Qualitative fit test apparatus FT30 user instructions. Available at: <https://multimedia.3m.com/mws/media/1464265O/3m-qualitative-fit-test-apparatus-ft-30-user-instruction.pdf> (Accessed: 16 October 2021)
37. Carvalho TC., McConville JT. The function and performance of aqueous aerosol devices for inhalation therapy. *Journal of Pharmacy and Pharmacology*. 2016; Available at: DOI:10.1111/jphp.12541
38. McKay RT. Respiratory Protection Newsletter. Respiratory Protection Newsletter. 2016. Available at: www.fit2fit.org/resources (Accessed: 16 February 2021)
39. Report of the stakeholders meeting on respiratory protective equipment fit testing. 22nd April 2005. Health and Safety Laboratory; Available at: <https://www.hse.gov.uk/asbestos/assets/docs/stakeholdermeeting220405.pdf>
40. Coffey CC., Lawrence RB., Zhuang Z., Campbell DL., Jensen PA., Myers WR. Comparison of five methods for fit-testing N95 filtering-facepiece respirators. *Applied Occupational and Environmental Hygiene*. 2002; 17(10): 723–730. Available at: DOI:10.1080/10473220290107002
41. Campbell DL., Coffey CC., Jensen P., Zhuang Z. Reducing respirator fit test errors: a multi-donning approach. *Journal of Occupational and Environmental Hygiene*. 2005; 2(8): 391–399. Available at: DOI:10.1080/15459620500182174
42. Coffey CC., Lawrence RB., Zhuang Z., Duling MG., Campbell DL. Errors associated with three methods of assessing respirator fit. *Journal of Occupational and Environmental Hygiene*. 2006; 3(1): 44–52. Available at: DOI:10.1080/15459620500455398

2 Human Errors in qualitative fit testing: a literature review

2.1 Abstract

Fit Testing is a mandatory requirement for all types of tight-fitting respiratory protective device facepieces worn in the UK workplace. Qualitative fit test methods based upon bitter or sweet taste appear relatively simple to use but are known to have limitations due to the subjective nature of the method. Concerns raised amongst fit testing stakeholders point to a lack of competency amongst some fit testers, raising questions about human errors in the real-world application of the method and hence the potential for assigning false pass results. Since respiratory protective devices are used in the workplace to protect against respiratory hazards, assigning a ‘pass’ fit test result in error can lead to serious health consequences for the wearer. This article presents a review of the published literature on errors in QLFT and identifies three key themes which are critical to an understanding of this area: statistical errors inherent in the method, errors arising from the subjective nature of the QLFT method, and human errors in real-world application of the QLFT method. Identification of research gaps reveals that much work is needed in better understanding human errors in the application of the method in real-world settings if false passes are to be mitigated.

2.2 Background

Many workers wear respiratory protective devices (RPD) incorporating tight-fitting facepieces to reduce their exposure to hazardous substances in the workplace. Tight-fitting facepieces are those whose performance relies on a good face-to-facepiece contact and examples include those used in filtering facepiece (disposable) respirators, and reusable half or full-face devices [1]. It is a well-established principle that all respirators leak, and that leakage into tight-fitting facepieces is derived from three sources: 1) via the filter media or components, 2) via the exhalation valve(s) if fitted, and 3) via the face seal [2][3][4]. Leakage via the filter media, components and exhalation valves is evaluated during performance testing according to generated aerosol (usually sodium chloride) Total Inward Leakage (TIL) measurements in line with established European Standards and test methods [3][5][6]. Since facial characteristics vary from wearer to wearer, leakage via the face seal cannot be established during TIL testing and must be evaluated separately for each individual wearer-to-facepiece combination. Fit testing is a method used to evaluate face seal leakage, and provides an assessment of whether a specific make, model, and size of facepiece is suitable for an individual wearer and hence has the potential to provide adequate protection against a respiratory hazard. The importance of proper fit testing as part of an effective RPD program has long been acknowledged, and numerous studies have demonstrated the benefit of fit testing in increasing the workplace protection afforded by RPD [7][8][9][10]. Current UK guidance on fit testing is provided in the Health and Safety Executive (HSE) document INDG479 and further information about ‘best-practice’ application is described in the British Safety Industry Federation (BSIF) Fit2Fit Companion series [11][12][13][14].

Fit test methods can broadly be divided into two types: quantitative fit test (QNFT) methods and qualitative fit test (QLFT) methods. QNFT methods provide a direct numerical assessment of leakage known as a quantitative fit factor (QNFF) which is derived as the ratio of challenge agent concentration, or pressure, outside and inside of a facepiece [15][16][17]. For filtering and reusable half facepieces the minimum QNFF that must be achieved in each part of the fit test regime is 100 [11]. QLFT methods in contrast do not provide a numerical result, but instead provide a simple pass / fail outcome based upon the wearer’s detection of a strong tasting or smelling challenge agent. The only

QLFT methods currently included in the HSE INDG479 protocol are those based on a bitter (denatonium benzoate, BitrexTM) or sweet (sodium saccharin) tasting challenge agent and are known widely as the ‘Bitrex’ and ‘saccharin’ methods [18][12].

In general, QLFT methods are viewed as being simpler and less expensive than QNFT methods but are known to be more subjective and are limited in application to reusable and filtering half mask facepieces [11]. In addition, QLFT methods place higher onus on the competency of the fit tester, with requirements to use and manage manual equipment to ensure that an adequate concentration of challenge agent is delivered and maintained [19][20]. Due to their relatively low cost and ease of use, QLFT methods are widely employed as the method of choice in settings such as healthcare where filtering facepieces are most frequently used. During the first wave of the SARS-CoV-2 pandemic, it was estimated that approximately one million fit tests were carried out in the UK, the vast majority in healthcare using the QLFT method [21].

A large part of the perceived simplicity of the QLFT method is the straightforward pass/fail nature of the outcome. If the wearer is unable to detect the challenge agent during the fit test, then leakage via the face seal is assumed to be below the critical level ($\leq 1\%$, $QNFF \geq 100$) and a ‘pass’ result is recorded. If, on the other hand, the wearer is able to detect the challenge agent at any time during the fit test then it is assumed that leakage is above the critical level and a ‘fail’ result is recorded. The binary nature of this result gives rise to two types of errors: false positive errors (false fails), and false negative errors (false passes) [22]. A false fail occurs when a wearer with a well-fitting facepiece (leakage $\leq 1\%$, $QNFF \geq 100$) indicates detection of the challenge agent during the fit test and an incorrect fail result is recorded. Whilst false fails are inconvenient, troublesome, and wasteful of resources, they are usually assumed not to present an unsafe fit test result for the wearer since they ultimately lead to a repeated fit test with the same or a new model of respirator. Conversely, a false pass result occurs when a wearer with a poor-fitting facepiece (leakage $> 1\%$, $QNFF < 100$) does not indicate detection of the challenge agent during the fit test and an incorrect pass result is recorded. In this situation, the facepiece is erroneously deemed to provide a suitable fit for the wearer, potentially leading to serious consequences for their immediate and long-term health [23].

Since the consequences of erroneous fit test results can be health critical, research in this area is crucial in improving the wearer's safety. This review describes the current understanding of sources of error in QLFT and identifies research gaps. This is a complex field involving both intrinsic and extrinsic factors; intrinsic factors include statistical errors inherent in the validated fit test method, and extrinsic factors include errors arising from the subjective nature of leak detection and human errors in application of the method. This review presents a discussion of those studies pivotal to an understanding of sources of error in the QLFT method. The research gaps identified in this review will inform future studies which aim to improve the reliability of the QLFT method by reducing errors and hence safeguarding workers' health.

2.3 Method

A protocol was developed to establish the objectives of this literature review, the search strategy, inclusion / exclusion criteria, and the method for data analysis. An initial scoping study was conducted through Scopus to understand the size of the literature and determine the key themes in this subject area. Results of the scoping study were used to define the primary objective of the literature review as: the identification, assessment, and discussion of evidence from pivotal studies of errors in qualitative respirator fit testing. The scoping study identified three major themes in the literature on errors in QLFT:

- 1) statistical errors inherent in the QLFT protocol,
- 2) errors arising from the subjective nature of the QLFT method,
- 3) human errors in the real-world application of the QLFT method.

Using the articles identified in the scoping study, a set of keywords and phrases were formulated and tested using iterative methods. This method yielded a total of sixteen keyword combinations: four relating to theme 1, six relating to theme 2, and six relating to theme 3. Appendix A provides details of the keyword combinations associated with each theme. Following identification of the relevant keyword combinations, criteria were developed for the inclusion and exclusion of evidence. Studies were included in the analysis if they met all of the following criteria:

1. The study reported use of the qualitative fit test method for testing filtering facepieces or reusable half mask facepieces.
2. The study reported the use of methods and materials described in any of the following Standards / guidance documents: Health and Safety Executive guidance HSE282/28 (studies conducted pre-March 2019) [24], Health and Safety Executive guidance HSE INDG479 (studies conducted post-March 2019) [11], British Standard BS ISO 16975-3 [25], American National Standards Institute ANSI Z88.10 [26].
3. The fit test timings, exercises and procedure were correctly reported according to the relevant Standards / guidance document cited, unless a deviation was introduced specifically for the purposes of the study.

4. The study was conducted using Bitrex or saccharin as the challenge agent.
5. The report was written in English.

The keyword combinations identified in Appendix A were used to perform an electronic search of four academic databases: Scopus, Web of Science, PubMed and ScienceDirect. Authors, sub-authors and institutions were investigated further using the ‘Author ID’ and ‘Affiliations’ tools available in Scopus. The review was not limited according to impact factor or date of publication since the number of results returned was of manageable size, yielding a total of 563 hits upon initial query. Detailed records were kept relating to search dates, keywords and number of results returned. Citation data and abstracts were catalogued in alphabetical order of author surname using Mendeley Desktop software (version 1.19.8). Removal of duplicates revealed a total of 188 unique entries which were filed for investigation. The title and abstract of each entry were examined, and those that did not meet the inclusion criteria were excluded. The main body of this review was conducted during the period February 2018 – November 2018 with continuous development up to December 2021 using alerts and citation feeds set up for key search phrases through Scopus.

2.4 Results

The selection process resulted in a final set of 56 entries which met all of the inclusion criteria and hence were considered pivotal to an understanding of the current knowledge in this field. These entries were studied through examination of the full text and detailed notes were kept according to the key themes identified. A narrative approach is used to report findings for the three themes identified in this study, and a meta-analytic approach is used to review results which report a Beta error value. An additional inclusion criterion for these results required the reported Beta value to include a 95% confidence interval (CI) statistic.

This results section provides a review of the pivotal studies relating to the understating of errors in qualitative fit testing through the three themes identified: 1) statistical errors inherent in the QLFT protocol, 2) errors arising from the subjective nature of the QLFT method, and 3) human errors made in the real-world application of the QLFT method. The first of these three themes are the method statistical errors; those associated with the sensitivity and efficacy of QLFT as a fit testing method.

2.4.1 Statistical errors inherent in the QLFT protocol

In statistical terms, there are two types of errors associated with fit test methods: Alpha errors and Beta errors. The Alpha error is the probability that a wearer with a well-fitting facepiece will fail the fit test (a false fail result), and the Beta error is the probability that a wearer with a poor-fitting facepiece will pass the fit test (a false pass result). Various descriptors have been used by researchers when reporting these values in articles, including the Greek symbols ' α ' and ' β ' and the terms 'false positive error' and 'false negative error' [23][27]. In this section describing the pivotal research on statistical errors, the terms 'Alpha error' and 'Beta error' are used in place of other terms except when quoting directly.

The Health and Safety Executive INDG479 guidance for fit testers refers to criteria in British Standard BS ISO 16975-3, which recommends the American National Standard ANSI/AIHA Z88.10-2010 as providing a suitable evaluation protocol and criteria for the

statistical analysis of new fit test methods [25][26]. The ANSI/AIHA Z88.10–2010 protocol (the ‘ANSI protocol’) evaluates any new fit test method through analysis of 100 sequential paired fit tests against a reference QNFT method. These paired fit tests must include 50 tests that result in a reference method fit factor that is greater than 5% and less than the required fit factor (RFF) value. Data collected from the paired fit tests are analysed using a 2×2 contingency table which allows reporting of the following:

- *Test Specificity* (probability that a wearer with an acceptable fit will pass the fit test),
- *Test Sensitivity* (probability that the test will correctly identify an inadequate fit),
- *Beta error* (probability of a false pass, defined as $1 - \text{Test Sensitivity}$)
- *Alpha error* (probability of a false fail, defined as $1 - \text{Test Specificity}$) [26].

The ANSI protocol criteria stipulate that the Test Specificity should be > 0.50 , Test Sensitivity must be ≥ 0.95 , and hence the calculated Beta error must therefore be ≤ 0.05 . Of these criteria, the Beta error is deemed by The National Institute for Occupational Safety and Health (NIOSH) to be the most important statistic from the standpoint of worker health, and is the key value quoted by researchers when investigating existing or new fit test methods [26]. The Alpha error by contrast is rarely reported by researchers, a subject which will be discussed later in this section.

Much of the body of research work involving QLFT methods is based upon the two original validation studies of the saccharin and Bitrex methods [28][29]. The Bitrex validation study is frequently cited by current researchers and the original methodology is the basis for extant Standards and guidance [30][31][32]. Although these two studies are regarded as seminal works in this field, both contain limitations which raise questions regarding their suitability as reference works today. The original saccharin validation study, conducted using half mask respirators with high efficiency filters, reported a point estimate for the Beta error at 0.0, 95% CI [0.0 – 0.05], a value which falls within the ANSI criteria. Although this study included a large number of paired fit tests ($N=144$), the methodology was specifically designed to evaluate whether the new test method was effective in rejecting fit tests with a required fit factor (RFF) of 10 and was therefore biased towards extremely low fit factors and hence poor-fitting respirators. When the authors re-analysed the data against an RFF of 100 (today’s standard for high-efficiency

filtering and reusable half mask respirators), the point estimate for the Beta error was reported to be 0.03, 95% CI [0.0 – 0.08] [28]. The authors used a one-sided confidence interval in their estimate of the Beta error in this study. This method appears to be appropriate in this context since the upper limit of the Beta error is of the most interest in safety terms where higher values indicate a greater probability of a false pass and hence are of greatest concern.

The original Bitrex QLFT validation study was conducted alongside a re-evaluation of the saccharin protocol and again used half mask respirators with high efficiency filters. This study reported a point estimate for the Beta error of 0.04, 95% CI [± 0.05] for both methods against an RFF of 100 [29]. The estimated two-sided confidence interval reported in this study however suggests a lower confidence limit of -0.01, a value which is impossible since the Beta error is a measure of proportion and must therefore lie in the range 0 to 1. Notes provided in the article reveal that this discrepancy arises as a result of the author's use of the Normal approximation for the confidence interval. This approximation has been shown to be unreliable when either the sample size is small, or the proportion is close to 0 or 1 as in this case [33]. A better approximation may have been obtained through use of an alternative method such as the Agresti-Coull interval or Wilson interval which use a relocated centre estimate to improve the approximation of the confidence interval. In addition, and unlike the original saccharin validation, there was no attempt to randomise the order of the reference / test method, and fail results were forced using respirators that were deliberately mis-sized '*by one or two sizes*'. Such practice introduces the potential for wearer-bias since failures would have been identifiable from comfort alone. Moreover, the authors tested both the saccharin and the Bitrex methods sequentially in a single donning (in randomised order) after the reference QNFT method, leaving only a 1-minute gap between tests '*to allow any taste from the first test to diminish*'. The current protocol following detection of taste would require the wearer's palate to be cleared completely and the sensitivity test repeated in order to re-establish the wearer's taste threshold [12]. This raises questions about the methodology of this study with potential for the taste to linger between the sequential tests, leading to unreliable reporting of fit test outcomes.

Later studies conducted with the saccharin and Bitrex QLFT methods have reported higher Beta values than the original studies, all of which fail to meet the ANSI criteria Beta error value of ≤ 0.05 . Figure 2-1 provides an overview of the Beta values reported by studies using the Bitrex or saccharin QLFT methods against an RFF of 100. This evidence appears to suggest that the estimate for the Beta error has increased over time potentially indicative of a less robust application of the original method.

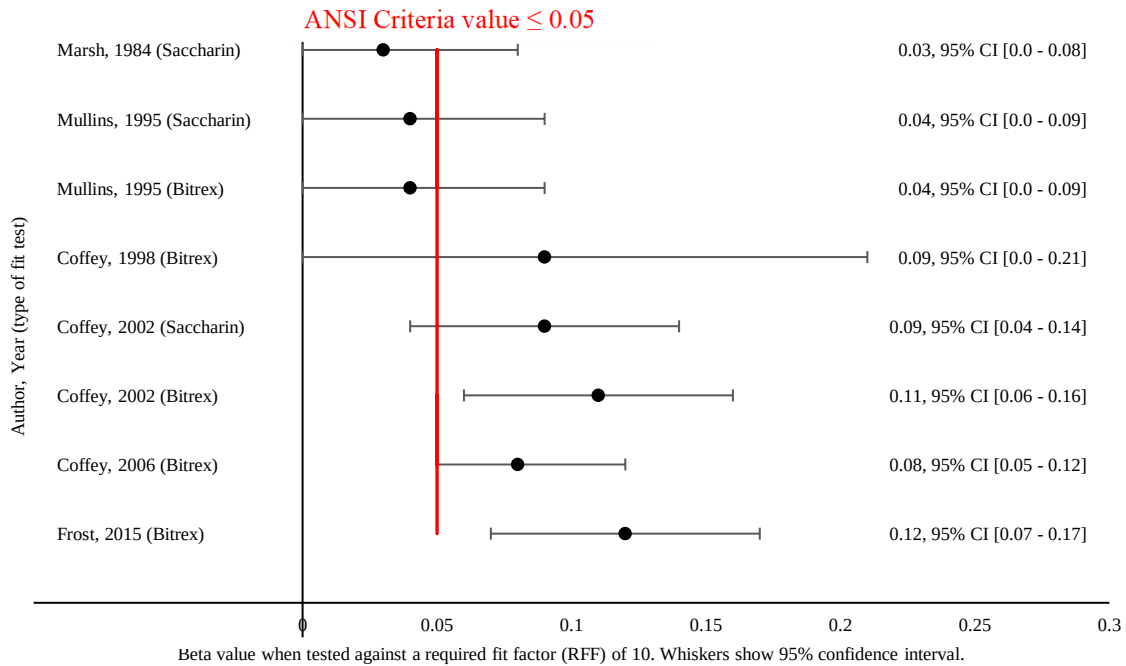


Figure 2-1 Beta values reported in articles included in this study against a required fit factor (RFF) of 100, listed in chronological order

Notes: ANSI Criteria value ≤ 0.05 is the acceptable Beta error for a new fit test method as defined by American National Standards Institute ANSI/AIHA Z88.10-2010 [26], CI = Confidence interval.

A 1998 study to evaluate the use of the Bitrex QLFT method for fit testing N95 filtering facepieces reported a Beta error of 0.09 95% CI [0.00 – 0.21] [34]. A later study by Coffey et al [22] compared the use of five different QLFT and QNFT methods for fit testing of N95 filtering facepieces and reported a Beta error of 0.11, 95% CI [± 0.05] for the Bitrex method and 0.09, 95% CI [± 0.05] for the saccharin method. Coffey et al [35] also re-evaluated the use of the Bitrex method in simulated workplace tests. The study was designed to assess the ability of the method to identify respirators that did not provide

adequate workplace protection, and results suggested a slightly improved Beta error of 0.08, 95% CI [0.05 – 0.12] when calculated across all classes of respirators. The authors recognised apparent discrepancies between the Beta errors reported by this and previous studies and suggested a possible effect due to the different types of respirators used in the studies. Indeed, several studies have suggested that Beta errors may vary with the type or class of respirator studied [36]. Coffey et al [35] reported that Beta errors were consistently higher when tests were conducted using reusable half mask respirators (Beta error 0.14, 95% CI [0.08 – 0.20]) than when compared to studies using filtering facepiece respirators (Beta error 0.05, 95% CI [0.03 – 0.10]). The authors suggested a possible mechanism for the improved Beta values reported with filtering facepiece respirators involving differences in the inhalation surface area (and hence differences in filter penetration rates) [35].

Several authors have suggested that the choice of QNFT reference method appears to affect the Test Sensitivity (and hence Beta error). To evaluate the comparability of reference methods, Janssen et al [37] tested the QLFT (Bitrex) method against two QNFT methods (ambient particle counting and controlled negative pressure) cited as possible reference-method alternatives to the usual generated aerosol QNFT method. Data reported from this study indicate that the alternative QNFT reference methods produced Test Sensitivity (and hence Beta) values which were not in agreement with each other and did not meet the ANSI criteria. These results led the authors to recommend the continued use of the generated aerosol QNFT method instead of the alternative ambient particle counting and controlled negative pressure methods for fit testing validation studies. The issue of the reference method was also highlighted by a study which reported on work conducted by the Health and Safety Laboratory in evaluating the pass level for high-efficiency filtering facepiece (FFP3) respirators. The study reported a Beta error of 0.12, 95% CI [$\pm$ 0.05] for the Bitrex QLFT method, a value significantly higher than the ANSI criteria value [38]. The authors reported the limitations of using the sodium chloride generated aerosol Total Inward Leakage (TIL) method as the reference method in validation studies for FFP3 respirators where filter-penetrating particles can never be eradicated entirely, and hence fit factors will be biased low.

Although not a directly assigned value according to the ANSI criteria, several researchers have reported on the Alpha errors associated with the QLFT method, noting values that are far higher than the recommended value of < 0.50 [22][35]. The Alpha error in QLFT is defined as '1 – Test Specificity', where Test Specificity is the fraction of wearers with acceptable fits that pass the new fit test method [26]. Whilst Alpha errors are widely considered to be inconvenient and wasteful of resources, they are often regarded as innocuous and are therefore largely ignored by researchers in favour of the more safety-critical Beta errors [23]. This assumption however was challenged by Coffey et al [35] in their study into the Alpha and Beta errors associated with common fit test methods. This study discussed the secondary effects of high Alpha errors and considered a possible impact on the Beta error associated with the QLFT method [35]. The authors argue that high Alpha errors can lead to an increase in the Beta error because, in line with current fit testing guidance, a failed fit test is permitted to be re-attempted up to twice. These repeated fit tests, using the same respirator-wearer combination, effectively allow another opportunity for a Beta error to be made creating a compound effect. This argument provides further weight to a previous study which suggested a model to describe the incorrect assignment of respirators, which predicts assignment errors of up to 20% [27].

When considering sources of error in QLFT, statistical values provide a key laboratory-based error rate, and these have been extensively studied and reported by numerous researchers in this field. Beyond these intrinsic errors, this review reveals two extrinsic factors which appear to be pivotal to an understanding of sources of errors in QLFT. The first of these are the errors that may arise due to the subjective nature of the QLFT method and the human detection system.

2.4.2 Errors arising from the subjective nature of the QLFT method

The problems of subjectivity associated with the QLFT method were first highlighted in the original evaluation study of the saccharin method which was conducted in 1984 [28]. The importance of the subjective nature of the method was not underestimated by the author who acknowledged that identification of a poor fitting facepiece rests upon the human detection system:

“The test relies on the subjective response of the wearer as to whether or not the test agent has penetrated into the interior of the facepiece. If the wearer is unable to detect any penetration, then it is assumed that there is a satisfactory fit between the facepiece and the wearer’s face.” [28]

The challenges of managing the subjective nature of the QLFT method have since been highlighted by numerous researchers. One study reported that wearers exposed to the saccharin challenge agent elicited *“no grimace or withdrawal”* and some wearers took several seconds to characterise the saccharin taste, compared to wearers exposed to the Bitrex taste agent, who *“elicited a strong, objectively observable withdrawal response.”* [39]. Indeed, the development of the Bitrex method aimed to provide an alternative to the ‘tolerable’ response reported for the saccharin method at that time [29]. The apparently greater adverse taste reaction to Bitrex may be explained by a comment by Mennela et al [40] in their overview of the research on Bitrex taste, where *“Bitrex taste is thought to have evolved as a deterrent against ingesting toxic substances”*. A more recent study described the problems that wearers had experienced in positively identifying taste during QLFT [38]. The study employed statistical methods to handle borderline results, considering them separately as ‘passes’ or ‘fails’, underlining the uncertainty and potential for error in these results. Further research in this area is needed to better understand whether errors can arise from wearers’ ability to elicit the desired taste response and how this may affect fit test outcome.

In addition to errors which may arise due to mischaracterisation of taste, some researchers have raised concerns about the ability of the method to properly determine the wearer’s taste threshold. The original QLFT methodology and current fit testing standards include a threshold screening test using a ‘sensitivity solution’ (a diluted version of the fit test solution) [41][11]. This threshold screening (or ‘sensitivity’) test is used to assess the wearer’s ability to detect the challenge agent at a concentration equivalent to a 1% face seal leakage, and hence screen out wearers with low taste response. Some researchers however have raised concerns that the sensitivity test may not adequately assure accurate detection of the test agent. This is supported by a number of studies which have reported that a proportion of wearers incorrectly identify the taste of a water placebo when substituted for Bitrex or saccharin challenge agents [36]. Indeed, in the original evaluation

of the new Bitrex QLFT method, the authors used an additional pre-qualification step to the sensitivity test which included the use of a distilled water placebo control as part of the threshold screening test [29]. Since these wearers were not permitted to continue to the fit test, the reported Beta error (0.04 [95% CI]) may be lower than would have been if these wearers had been allowed to continue. This provides a possible explanation for the higher Beta errors reported by later studies that did not include a placebo as a pre-qualification step. A number of researchers have since used a distilled water placebo as a control in their methodology and have reported on the proportion of wearers that incorrectly responded to the control. In their 2012 study, Mitchell et al [42] used a double-blinded trial with three different sensitivity solutions (commercial saccharin solution, homemade saccharin solution, and a distilled water placebo) and reported that 16.7% of participants ($n = 48$) incorrectly characterised the taste of the water placebo as saccharin. A more recent study by Fakherpour et al [43] in 2019 reported on a study of QLFT for 62 healthcare students in Iran. This single-blind, placebo (distilled water) controlled trial reported that 29% of wearers incorrectly characterised the placebo.

An additional problem with taste detection is that humans naturally have a range of taste detection threshold limits. The Bitrex and saccharin QLFT methods aim to accommodate wearers' different taste threshold values by grouping them into four sensitivity bandings. These sensitivity bandings are grouped according to the number of 'applications' of sensitivity solution required to elicit a response [12].

<i>Band 1</i> (1 – 10 applications):	High sensitivity
<i>Band 2</i> (11 – 20 applications):	Medium sensitivity
<i>Band 3</i> (21 – 30 applications):	Low sensitivity
<i>Band 4</i> (> 30 applications):	Sensitivity below level required for this method

One study reported that wearers indicated taste detection of saccharin after a mean of 2.38 applications of a commercially available sensitivity solution, and taste detection of Bitrex after 1.35 applications of a more concentrated version [39]. Sensitivity bandings are used to dictate the number of applications of fit test solution (challenge agent) which should be used during the second stage of the fit test. A more recent study reported on statistical evaluation of outcomes for wearers with different taste threshold values as part of an

investigation into pass criteria for FFP3 respirators [38]. This study reported an improved statistical Test Sensitivity when wearers with a sensitivity level > 5 (lower sensitivity to Bitrex) were removed, suggesting “*the Bitrex fit test may be giving more reliable results where test subjects have a more typical sensitivity to Bitrex*” [38]. There is a paucity of research in this area of subject-response and the overall contribution to errors in QLFT is not well understood. Mennella et al [40] described the wide variety in human and animal taste response, commenting that “*taste is experienced through a ‘sensory window’ that changes with age and experience and is partially defined by genetics*”. In a recent study using QLFT to evaluate novel respirators during the shortage in supply of RPD in the SARS-Cov-2 pandemic, researchers at Imperial College London emphasised that “*this psychophysical method has inherent limitations in that some will not be able to taste the solution ever and others may think they are tasting it when they are not*” [44].

This section has addressed the issues of taste from a wearer-positive position: where the wearer is believed to be truthful in their ability and willingness to detect and report taste. However, the subjective nature of the QLFT method introduces an additional source of error - a wearer that uses their wilful influence to ‘cheat’ the outcome of the test. A study of QLFT conducted amongst healthcare workers (HCWs) in Australia summarised this effect in their study by suggesting that, “*it is also possible that some HCWs with increased anxiety may intentionally or unintentionally fail the fit test (indicating a leak) during a pandemic*” [45]. The problem with a fit test based upon subjective taste is highlighted by Mennella et al [40] who describe the problems associated with taste perception which can only indirectly be measured through observation or response of the human subject.

Although research gaps remain in areas of taste characterisation and a possible placebo effect, errors arising from the subjective nature of QLFT have been broadly documented and are widely acknowledged. Current Standards and guidance recommend that factors which can affect the wearer’s taste should be ‘properly managed’ by the fit tester to ensure reliable outcomes [12]. Management of the wearer’s taste palate is one of the many tasks assigned to a qualitative fit test operator; a role which is critical in application of this highly manual method. Human operator errors in real-world application form the second of the extrinsic factors in understanding errors in QLFT.

2.4.3 Human errors made in the real-world application of the QLFT method

The QLFT method is a manual procedure which relies upon the ability of a fit tester to properly understand, apply, and manage a prescribed protocol. A number of sources have questioned whether the skills and knowledge of some fit testers is adequate to enable them to carry out their function satisfactorily [46]. Development of the BSIF RPE Fit Test Providers ‘Fit2Fit’ Accreditation Scheme (Fit2Fit scheme) acknowledged that a lack of competency exists in the UK workplace, with the potential to cause serious harm to the health of the wearer [47]. Indeed, the BS ISO Standard 16975-3 highlights that the reliability of the validated QLFT method depends upon the proper application of the accepted test protocol by the fit tester.

“any variation from the procedure specified below can invalidate the results, including changes in solution concentration, how the bulb is squeezed, the number of squeezes, and the size of the fit-test hood” [25].

Although there are numerous comments in the literature which suggest serious concerns, there has been little or no attempt to assess how well the QLFT method is applied in real-world settings by real-world fit testers. Coffey et al [35] attempted to predict the probability of workplace errors through development of a theoretical simulated workplace protection program model. This model suggested a predicted false pass rate of 19% for a simulated program which allowed three re-attempts using the same respirator model [35]. A 2018 study commented on the role of the fit tester in workplace fit testing programs in the United States, suggesting a severe lack of formal training and experience, a situation echoed by other commentators [48].

UK researchers have described how pressure has been placed on fit testers to take short cuts to minimise workers’ downtime and stress the importance of a properly conducted fit test [49]. This is further highlighted by evidence presented at a fit testing stakeholders’ meeting hosted by the Health and Safety Executive which heard reports that some fit testers deviate from guidance and apply fit testing methods inconsistently [50]. A former chair of the ANSI/AIHA Z88.10-2010 committee on respirator standards commented that,

“a common problem with respirator fit testing is that poorly fitting respirators may pass when the operator doesn’t administer the test correctly. This is particularly true for qualitative fit testing, but also applies to quantitative fit testing” [51].

These statements are supported through critical evaluation of the methodology section of studies which have included the use of the QLFT method. A recent review of the implementation of QLFT in dentistry during the SARS-CoV-2 pandemic highlighted several errors in application, including a failure to remove both bungs from the nebuliser and use of a full-face shield underneath the fit testing hood, both contravening current UK guidance [52]. Another recent review of fit testing, published in the British Dental Journal included a fit test exercise regime which differed from the published HSE guidance [53]. Inconsistencies in methodology between manufacturer’s user instructions and guidance adds further potential problems for real-world fit testers, including varying guidance about the required nebuliser fill volume and variation in the time required to clear the wearer’s palate [25][12].

Perhaps the biggest indicator of real-world application problems comes from the published results of assessments undertaken through the Fit2Fit scheme. This voluntary scheme provides assessments and subsequently awards the ‘Fit2Fit accreditation’ to fit testers who demonstrate their skills and knowledge against a syllabus which references HSE INDG479 [54]. Figures published for the year to February 2020, show that 30% of fit testers that put themselves forward for assessment failed to meet the required standard in their practical examination, giving the clearest indication yet of a lack of competency amongst real-world fit testers in UK workplace settings [55].

2.5 Discussion

Results of this literature review indicate a disparity in research efforts across the three themes identified as being pivotal to an understanding of errors in QLFT. Figure 2-2 provides an overview of the research gaps revealed through this literature review under the three themes.

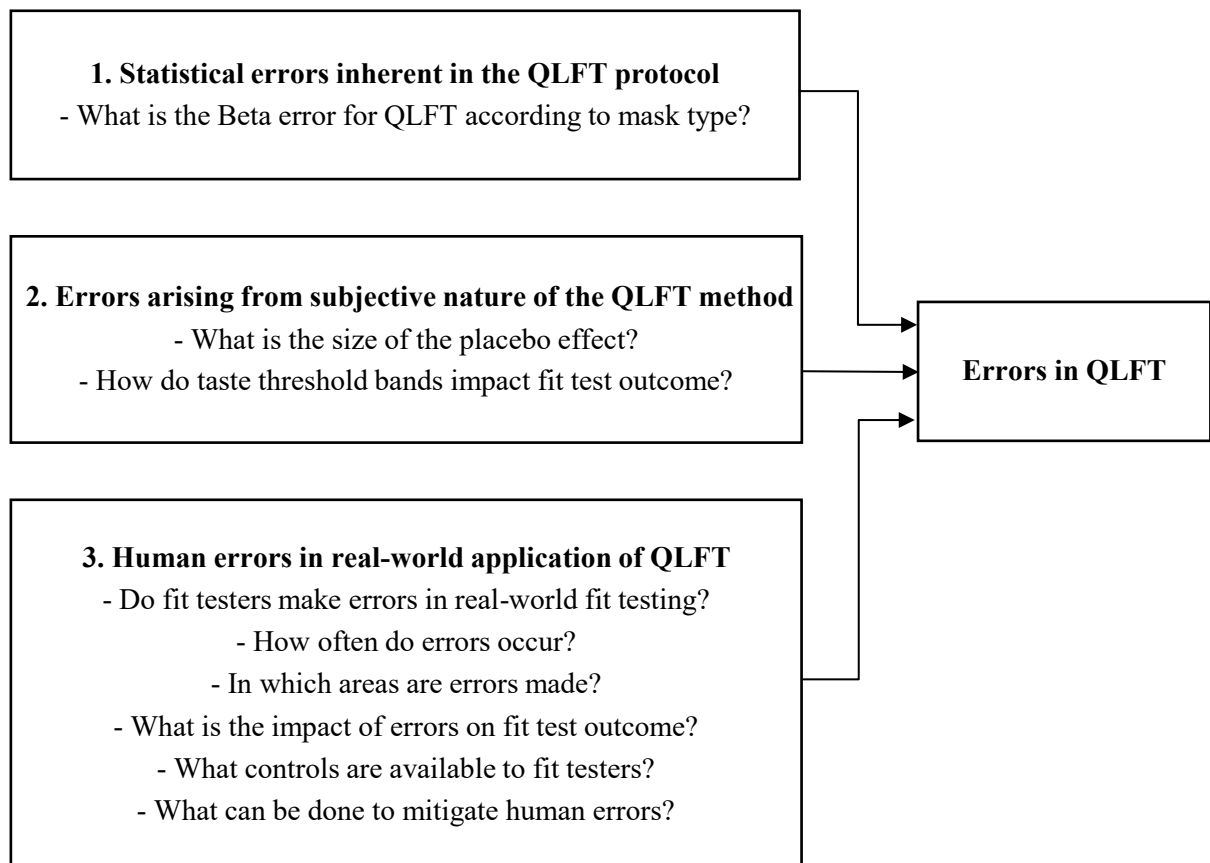


Figure 2-2 Research gaps identified in three themes pivotal to an understanding of errors in qualitative fit testing

This review indicates that statistical errors in QLFT have been extensively studied by numerous researchers, with the ANSI criteria providing a solid basis for comparison and evaluation. Whilst there appears to be agreement amongst researchers that Beta errors are the most significant statistical error arising from fit testing in terms of safety, the studies reported here do not agree on a definitive value for the Beta error for the saccharin and Bitrex QLFT methods, instead suggesting a range of 0.03 – 0.12. The studies analysed in this review suggest that reasons for this range may include differences in the design of

research, types and classes of respirators used, and the choice of QNFT reference method used. This complicates the ability of researchers to compare results from separate validation studies and to assess whether the method meets the ANSI criteria under varying conditions of study. The Beta error is reported along with other statistics during evaluation or validation of new fit test methods; a crucial process when assessing the safety and efficacy of any new method. It is of concern then, that this review indicates that the Bitrex and saccharin QLFT methods fail to meet the ANSI criteria in many recent studies and particularly when used for half mask classes of respirators. Further research is necessary to better understand these differences in order to assure wearer safety across all applicable respirator classes.

The results of this review show that possible errors arising from the subjective nature of the QLFT method have been widely discussed in the literature, but several research gaps remain. Given the importance of taste detection to the integrity of the QLFT method, it is perhaps surprising that the placebo effect reported by a number of researchers has not been further investigated. QLFT taste methods are entirely dependent on the ability of the wearer to properly detect the challenge agent aerosol. Incorrect characterisation of taste either at sensitivity threshold testing, or during fit testing, has the potential therefore to lead to under-detection of face seal leakage and hence produce false pass results. In real-world terms, the safety-first benefits of including an additional placebo-testing step in QLFT methodology would far outweigh any minimal additional time or financial burden that this may incur. In terms of real-world impact, very few researchers have investigated the impact on fit test outcomes for wearers that fall into different sensitivity bandings, or how these are affected by external factors such as smoking. Wearers that have a strong desire to pass a fit test, perhaps due to pressures of time, money or peer pressure can do so by choosing not to report a detectable taste, leading to a false pass result. Further research including surveys of human taste detection in QLFT, and wearer attitudes is needed to better understand possible factors and provide solutions to these problems.

A wealth of research literature is devoted to evaluation of the QLFT method in laboratory settings, but it is the gaps in research in real-world settings that are most apparent. This review has identified very few studies that have examined the application of QLFT in workplace settings, and none that have attempted to investigate whether real-world fit

testers make errors in applying the validated protocol and how these may impact fit test outcome. Human errors in real-world application of the QLFT method appear to be poorly understood, with no formal studies in this area. Evidence about the sources, prevalence, and impacts of human errors on QLFT outcome are limited to discussions amongst stakeholders and anecdotal reports. This review finds a significant research gap in this area, and thorough understanding of human errors in QLFT is needed to assure the safety of this method in workplace fit testing programmes.

Qualitative fit testing is a manual process, dependent upon the taste response of a human wearer and the skills and ability of a human fit test operator. This review has highlighted that the competency of the fit tester is widely recognised as being crucial in assuring proper application of the QLFT method. Whilst there are strong indicators from commentators in this field that fit testers may lack the appropriate skills and knowledge, very little is known about errors made by fit testers in real-world settings. Further research is needed to assess the prevalence and types of errors made by fit testers and whether these can lead to false pass results. Studies are also required to better understand why fit testers make errors and whether further controls can be implemented to increase the reliability of the test outcome.

This review has revealed that the factors contributing to errors in QLFT are multi-faceted, with a number of contributory themes. The safety-critical nature of fit testing is such that a thorough understanding of each of these factors is both vital and urgent. Employers have a duty of care to ensure that employees are adequately protected from respiratory hazards at work, and workers have the right to expect this to be done properly. Fit testing is just one link in the chain of respiratory protection, but that chain is only as strong as its weakest link.

2.6 Conclusion

Errors in QLFT can be described through three factors: statistical errors inherent in the method, errors arising from the subjective nature of the method, and human errors in real-world application. This review has identified very few gaps under the theme of statistical errors, with numerous studies providing a rigorous understanding in this area. Several gaps were revealed pertaining to the theme of errors arising from the subjective nature of the method, with further research required to better understand a possible placebo effect and the variation in human taste characterisation. The most significant gaps were identified for the theme of human errors in the application of the method in real-world settings. Gaps in research in this area indicate that whilst much is known about the QLFT method under laboratory test conditions, little is known about its application in the real world. If the safety of workplace QLFT, and hence the safety of the wearer, is to be assured then all sources of error must be properly understood, including those resulting from human error.

2.7 References

1. Health and Safety Executive. HSG53 Respiratory protective equipment at work. A practical guide. (Fourth edition). HSE Books; 2013. Available at: <http://www.hse.gov.uk/pubns/priced/hsg53.pdf>
2. British Standards Institution. BS EN 149:2001+A1:2009 Respiratory protective devices. Filtering half masks to protect against particles - requirements, testing, marking. London: British Standards Institution; 2002.
3. British Standards Institution. BS EN 13274-7:2019 Respiratory protective devices - methods of test. Part 7: Determination of particle filter penetration. London: British Standards Institution; 2019.
4. Rengasamy S., Eimer B. Nanoparticle penetration through filter media and leakage through face seal interface of N95 filtering facepiece respirators. The Annals of Occupational Hygiene. 2012; Available at: DOI:10.1093/annhyg/mer122
5. British Standards Institution. BS EN 140:1999 Respiratory protective devices. Half masks and quarter masks - requirements, testing, marking. London: British Standards Institution; 1999.
6. British Standards Institution. BS EN 136:1998 Respiratory protective devices. Full face masks - requirements, testing, marking. London: British Standards Institution; 1998.
7. Occupational Safety and Health Administration. 1910.134 Appendix A - Fit testing procedures (mandatory). 2004. Available at: <https://www.osha.gov/laws-regs/regulations/standardnumber/1910/1910.134AppA>
8. Coffey CC., Lawrence RB., Campbell DL., Zhuang Z., Calvert CA., Jensen PA. Fitting Characteristics of Eighteen N95 Filtering-Facepiece Respirators. Journal of Occupational and Environmental Hygiene. April 2004; 1(4): 262–271. Available at: DOI:10.1080/15459620490433799

9. Duling MG., Lawrence RB., Slaven JE., Coffey CC. Simulated workplace protection factors for half-facepiece respiratory protective devices. *Journal of Occupational and Environmental Hygiene*. 2007; 4(6): 420–431. Available at: DOI:10.1080/15459620701346925
10. Danyluk Q., Hon C-Y., Neudorf M., Yassi A., Bryce E., Janssen B., et al. Health care workers and respiratory protection: is the user seal check a surrogate for respirator fit-testing? *Journal of Occupational and Environmental Hygiene*. 2011; 8(5): 267–270. Available at: DOI:10.1080/15459624.2011.566016
11. Health and Safety Executive. INDG479 Guidance on respiratory protective equipment (RPE) fit testing. HSE Books; 2019. Available at: <https://www.hse.gov.uk/pubns/indg479.pdf>
12. British Safety Industry Federation. Fit2Fit Companion to HSE INDG479 guidance on respiratory protective equipment (RPE) fit testing, qualitative method. June 2021. Available at: <https://www.fit2fit.org/resources>
13. British Safety Industry Federation. Fit2Fit Companion to HSE INDG479 guidance on respiratory protective equipment (RPE) fit testing, quantitative ambient particle counting method. May 2021. Available at: <https://www.fit2fit.org/resources>
14. British Safety Industry Federation. Fit2Fit Companion to HSE INDG479 guidance on respiratory protective equipment (RPE) fit testing, quantitative controlled negative pressure method. May 2021. Available at: <https://www.fit2fit.org/resources>
15. TSI Incorporated. Portacount Pro Model 8030 Respirator Fit Tester theory of operation. Application note RFT-002. 2013. Available at: www.tsi.com
16. OHD LLLP. QuantiFit2 Owners Manual. Available at: <https://ohdusa.com/wp-content/uploads/2021/01/QuantiFit2-Manual.pdf> (Accessed: 6 May 2021)
17. Crutchfield CD., Fairbank EO., Greenstein SL. Effect of test exercises and mask donning on measured respirator fit. *Applied Occupational and Environmental Hygiene*. 1999; 14(12): 827–837. Available at: DOI:10.1080/104732299302062

18. Macfarlan Smith Ltd. Bitrex. Available at: <https://www.bitrex.com/> (Accessed: 6 May 2021)
19. 3M. Qualitative fit test apparatus FT10 (sweet) and FT30 (bitter). Available at: <https://multimedia.3m.com/mws/media/252309O/3m-adflo-qualitative-fit-test-apparatus-ft-10-ft-30.pdf> (Accessed: 6 May 2021)
20. 3M. Guide to using the 3M qualitative fit test kits. Available at: <https://multimedia.3m.com/mws/media/473960O/guide-to-using-the-3m-qualitative-fit-test-kits.pdf> (Accessed: 6 May 2021)
21. British Safety Industry Federation. Fit2Fit Newsletter December 2020. Available at: www.fit2fit.org/resources (Accessed: 15 February 2021)
22. Coffey CC., Lawrence RB., Zhuang Z., Campbell DL., Jensen PA., Myers WR. Comparison of five methods for fit-testing N95 filtering-facepiece respirators. *Applied Occupational and Environmental Hygiene*. 2002; 17(10): 723–730. Available at: DOI:10.1080/10473220290107002
23. Revoir W. Correlation of results of qualitative respirator facepiece fitting tests using isoamyl acetate vapor as test agent with results of quantitative respirator facepiece fitting tests using polydisperse diethylhexyl phthalate aerosol as test agent. *Journal of the International Society for Respiratory Protection*. 1983; 1(4): 42–59.
24. Health and Safety Executive. OC 282/28 Fit testing of respiratory protective equipment facepieces. HSE Books; 2012.
25. British Standards Institution. BS ISO 16975-3:2017 Respiratory protective devices. Selection, use and maintenance. Part 3: Fit-testing procedures. London: British Standards Institution; 2017.
26. American National Standards Institute. ANSI/AIHA Z88.10-2010 Respirator Fit Testing Methods. Fairfax: American Industrial Hygiene Association; 2010.
27. Campbell DL., Coffey CC., Lenhart S. Respiratory protection as a function of respirator fitting characteristics and fit-test accuracy. *American Industrial Hygiene*

Association Journal. 2001; 62(1): 36–44. Available at:

DOI:10.1080/15298660108984607

28. Marsh JL. Evaluation of saccharin qualitative fitting test for respirators.

American Industrial Hygiene Association Journal. 1984; 45(6): 371–376. Available at:

DOI:10.1080/15298668491399965

29. Mullins H., Danisch S., Johnston A. Development of a new qualitative test for fit testing respirators. American Industrial Hygiene Association Journal. 1995; 56(11): 1068–1073. Available at: DOI:10.1080/15428119591016278

30. Fakherpour A., Jahangiri M., Seif M., Charkhand H., Abbaspour S., Floyd EL. Quantitative fit testing of filtering face-piece respirators during the COVID-19

pandemic reveals anthropometric deficits in most respirators available in Iran. Journal of Environmental Health Science and Engineering. 2021; Available at:

DOI:10.1007/s40201-021-00648-3

31. Schaller G., Nayar SK., Erotocritou M., Overton A., Stelzhammer T., Berber O. Efficacy of surgical helmet systems for protection against COVID-19: a double-blinded randomised control study. International Orthopaedics. 2021; 45(1): 39–42. Available at:

DOI:10.1007/s00264-020-04796-3

32. Fakherpour A., Jahangiri M., Yousefinejad S., Seif M., Banaee S. Assessment of aloe vera for qualitative fit testing of particulate respirators: a logistic regression approach. Industrial Health. 2020; 58(1): 46–53. Available at:

DOI:10.2486/indhealth.2109-0019

33. Wallis S. Binomial confidence intervals and contingency tests: mathematical fundamentals and the evaluation of alternative methods. Journal of Quantitative Linguistics. 2013; 20(3): 178–208. Available at: DOI:10.1080/09296174.2013.799918

34. Coffey CC., Zhuang Z., Campbell DL. Evaluation of the Bitrex™ Qualitative Fit Test Method Using N95 Filtering- Facepiece Respirators. Journal of the

International Society for Respiratory Protection. 1998; 16(1): 48–55. Available at:

<https://www.isrp.com/the-isrp-journal/complete-journal-archive-for-members-only/616-vol-16-no-1-4-1998/file>

35. Coffey CC., Lawrence RB., Zhuang Z., Duling MG., Campbell DL. Errors associated with three methods of assessing respirator fit. *Journal of Occupational and Environmental Hygiene*. 2006; 3(1): 44–52. Available at: DOI:10.1080/15459620500455398
36. Hon C-Y., Danyluk Q., Bryce E., Janssen B., Neudorf M., Yassi A., et al. Comparison of qualitative and quantitative fit-testing results for three commonly used respirators in the healthcare sector. *Journal of Occupational and Environmental Hygiene*. 2017; 14(3): 175–179. Available at: DOI:10.1080/15459624.2016.1237030
37. Janssen LL., Luinenburg MD., Mullins HE., Nelson TJ. Comparison of three commercially available fit-test methods. *American Industrial Hygiene Association Journal*. 2002; 63(6): 762–767. Available at: DOI:10.1080/15428110208984767
38. Frost S., Mogridge R., Roff M. RR1029 Review of fit test pass criteria for filtering facepieces Class 3 (FFP3) respirators. Health and Safety Executive; 2015. Available at: <https://www.hse.gov.uk/research/rrpdf/rr1029.pdf>
39. McKay RT., Davies E. Capability of respirator wearers to detect aerosolized qualitative fit test agents (sweetener and Bitrex) with known fixed leaks. *Applied Occupational and Environmental Hygiene*. 2000; 15(6): 479–484. Available at: DOI:10.1080/104732200301269
40. Mennella JA., Spector AC., Reed DR., Coldwell SE. The bad taste of medicines: overview of basic research on bitter taste. *Clinical Therapeutics*. August 2013; 35(8): 1225–1246. Available at: DOI:10.1016/j.clinthera.2013.06.007
41. Occupational Safety and Health Administration. Minnesota Mining and Manufacturing Company post hearing data submission on qualitative fit testing protocols under the lead standard before OSHA. Docket No. H-049A. 1981.
42. Mitchell BG., Wells A., McGregor A., McKenzie D. Can homemade fit testing solutions be as effective as commercial products? *Healthcare infection*. December 2012; 17(4): 111–114. Available at: DOI:10.1071/HI12019

43. Fakharpour A., Jahangiri M., Yousefinejad S., Seif M. Feasibility of replacing homemade solutions by commercial products for qualitative fit testing of particulate respirators: a mixed effect logistic regression study. *MethodsX*. 2019; 6: 1313–1322. Available at: DOI:10.1016/j.mex.2019.05.034
44. Seligman H., Zaman S., Pitcher DS., Shun-Shin MJ., Hepworth Lloyd F., Androschuk V., et al. Reusable snorkel masks adapted as particulate respirators. Shou D (ed.) *PLOS ONE*. 5 April 2021; 16(4): e0249201. Available at: DOI:10.1371/journal.pone.0249201
45. Regli A., Sommerfield A., Ungern-Sternberg BS. The role of fit testing N95/FFP2/FFP3 masks: a narrative review. *Anaesthesia*. 2021; 76(1): 91–100. Available at: DOI:10.1111/anae.15261
46. Weed J. Training is the key to a good respirator fit test program. *Journal of the International Society for Respiratory Protection*. 1996; 14(4): 8–10.
47. British Safety Industry Federation. Fit2Fit candidate handbook. 2014. pp. 1–8. Available at: www.fit2fit.org/resources (Accessed: 1 October 2020)
48. McKay RT. Respirator fit test methods - are faster protocols equivalent to OSHA? *Journal of Occupational and Environmental Hygiene*. 2018; 15(7): D53–D57. Available at: DOI:10.1080/15459624.2018.1463098
49. Clayton M., Vaughan N. Fit for purpose? The role of fit testing in respiratory protection. *The Annals of Occupational Hygiene*. 2005; Available at: DOI:10.1093/annhyg/mei046
50. Report of the stakeholders meeting on respiratory protective equipment fit testing. 22nd April 2005. Health and Safety Laboratory; Available at: <https://www.hse.gov.uk/asbestos/assets/docs/stakeholdermeeting220405.pdf>
51. McKay RT. Respiratory Protection Newsletter. Respiratory Protection Newsletter. 2016. Available at: www.fit2fit.org/resources (Accessed: 16 February 2021)
52. Robertson S., Ramsdale MP. Audit of qualitative fit testing for FFP3 respirators. *British Dental Journal*. 2021; : 1–4. Available at: DOI:10.1038/s41415-021-2716-6

53. Rollings L. FFP3 respirator face fit testing - what is it all about? *British Dental Journal*. 2020; 7(8): 30–33. Available at: DOI:10.1038/s41407-020-0404-z
54. British Safety Industry Federation. Fit2Fit scheme syllabus. Available at: <https://www.fit2fit.org/resources> (Accessed: 6 May 2021)
55. British Safety Industry Federation. Fit2Fit Technical Committee meeting minutes February 2021.

3 Human errors in qualitative fit testing: a study of real-world fit testers

3.1 Abstract

Fit Testing is an essential part of any respiratory protective device program and provides assurance that a selected make, model, and size of facepiece is suitable for an individual wearer. Although qualitative fit test methods are assumed to be simple and easy to use, it remains critical that the methodology is applied in accordance with the protocols set out in guidance to ensure that a correct ‘pass’ or ‘fail’ result is assigned. An incorrectly assigned pass result increases the risk to the wearer of exposure to hazardous respiratory substances in the workplace putting their health or life in danger. Fit testing stakeholders, manufacturers and regulatory bodies have raised concerns about the quality of fit testing conducted in the UK workplace, but there have been no published research studies in this area to date. This article presents results from a study into the errors made by real-world fit testers, using the data gathered during Fit2Fit practical assessments over a ten-year period from 2009 to 2019. When application of the qualitative fit test method was measured against the HSE INDG479 protocol, assessment records indicate that fit testers made a median of 4 errors, and a high of 29 errors. Fit2Fit accredited fit testers appear to make half as many errors as non-accredited fit testers, yet the scheme remains voluntary in the UK. The sources of errors were analysed in detail and revealed that the most common errors were made in instructing the wearer, donning and fit checking RPE, verifying the nebuliser function, and assuring the wearer’s taste detection. Results from this study suggest that errors made by real-world fit testers are common, and steps to mitigate the assignment of false pass fit test results arising from these errors are needed if wearer health is to be better protected.

3.2 Background

Fit testing is an essential part of any effective respiratory protective device (RPD) program and is a crucial link in the chain of assuring worker protection. Tight-fitting facepieces, which include disposable masks, and reusable half masks and full-face masks, must be suitably matched to a wearer's individual facial features to minimise leakage via the face seal. The Approved Codes of Practice supporting the Control of Substances Hazardous to Health (COSHH) Regulations 2002 require that all reasonable steps be taken to prevent or reduce exposure to substances hazardous to health [1]. Where RPD is selected on the basis of a risk assessment, it must provide adequate protection against the hazard and, in order to be suitable for the wearer, the initial selection of tight-fitting facepieces should include a fit test.

Guidance on fit testing of RPD facepieces is set out in the Health and Safety Executive (HSE) industry guidance document HSE INDG479 [2]. This guidance is supported through the British Safety Industry Federation (BSIF) Fit2Fit Companions which provide complementary practical guidance in a series of documents which accompany the main three fit test methods: qualitative, quantitative (ambient particle counting) and quantitative (controlled negative pressure) [3][4][5]. Quantitative fit test (QNFT) methods can be used to fit test all classes of tight-fitting facepieces and employ the use of an instrument to assess the amount of face seal leakage, producing a numerical value known as a fit factor (FF) [6]. Qualitative fit test (QLFT) methods are restricted in their use to disposable and reusable half masks, however due to their relatively low cost and ease of use, these methods are in widespread use in the UK workplace.

QLFT methods use the wearer's sense of taste (or less commonly, smell) to identify face seal leakage of a bitter- or sweet-tasting challenge agent and produce an outcome that is a simple pass or fail result. The importance of properly assigning a 'pass' or 'fail' fit test result cannot be underestimated; RPD facepieces may be the last (and sometimes only) line of defence against a respiratory hazard, and a wearer that is assigned a pass result in error (a false pass result) potentially risks increased exposure in the workplace. False pass results can occur in all fit testing methods and arise when a wearer with an unsuitable (leaking) facepiece erroneously passes a fit test. It is critical to reduce errors that may lead to false pass results to assure the health and safety of RPD wearers.

Sources of errors in the QLFT method are complex, and a literature review has identified three key themes which are crucial to an overall investigation: 1) statistical errors inherent in the QLFT method, 2) errors arising from the subjective nature of the method, and 3) human errors made in real-world application of the method [7]. Whilst the first two of these areas are largely well understood, further research is needed, especially in understanding the wide range of statistical errors reported and the problems of subjectivity associated with this method. The biggest gap identified by the literature review was the paucity of research into human errors in QLFT. Almost all research into the QLFT methodology has thus far been conducted under laboratory conditions, with those studies which have been conducted in the workplace focussing on the outcomes of workplace protection factors studies or respirator type. Although many commentators have expressed concerns and cited anecdotal evidence about human errors in real-world application of the QLFT method, there have been no formal research studies conducted in this area to date. The strongest indicator of the prevalence of human errors in QLFT comes from data provided by the BSIF RPE Fit Test Providers 'Fit2Fit' Accreditation Scheme (Fit2Fit scheme). Figures published by the BSIF in October 2021 reported that 30% of the fit testers who were assessed through the scheme failed to meet the standard of competence set against HSE INDG479 guidance [8]. This suggests that human errors in workplace QLFT may be far more common than those identified in laboratory settings and hence the workplace safety of this method must be questioned.

This study is the first to use the primary data gathered during Fit2Fit assessments as a research tool for addressing three important research questions: 1) how frequently do real-world fit testers make errors in application of the QLFT method; 2) what are the main sources of human errors in QLFT; and 3) what are the implications and recommendations, if any, for use of the QLFT method in the UK workplace? Research into human errors in QLFT, along with an understanding of where and how errors are made is crucial if false pass results are to be mitigated and the workplace safety of this fit test method assured. This in turn is essential for ensuring that wearers of tight-fitting facepieces are afforded the protection they expect in the workplace.

3.3 Method

3.3.1 Data collection

Under UK guidance, fit testing should only be conducted by a competent person [2]. A competent person as defined by the British Standard BS ISO 16975-3 is, “*person with suitable and sufficient experience and with practical and theoretical knowledge of fit-test methods*” [6]. Although not mandatory, HSE INDG479 guidance cites the Fit2Fit scheme as one way in which fit testers can demonstrate their competence in fit testing. The Fit2Fit scheme has a set of well-defined assessment criteria which closely reference HSE INDG479, providing a consistent and controlled evaluation mechanism [9]. The primary purpose of the records (assessment papers) gathered during the Fit2Fit assessment process is to enable a judgement to be made of a fit tester’s competence. Since recording is made meticulously against traceable criteria, these assessment papers also provide a detailed insight into those areas of the fit testing process which the fit tester conducts correctly, and those in which errors are made. This thereby offers a unique opportunity for secondary research into human errors in real-world QLFT.

The process of Fit2Fit assessment which results in Fit2Fit accreditation is shown in Figure 3-1. Fit testers may apply for assessment in any (or all) of the three of the fit testing methods which are described in HSE INDG479: qualitative, quantitative (ambient particle counting) and quantitative (controlled negative pressure).

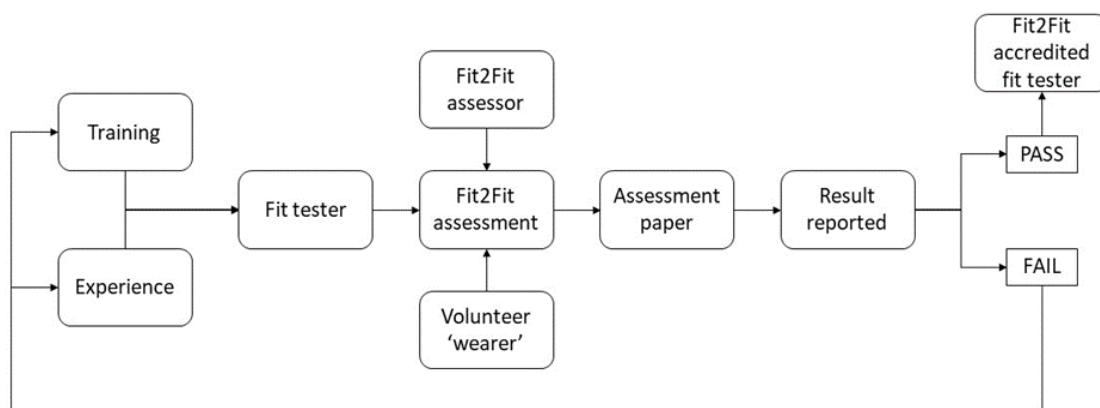


Figure 3-1 Schematic of the process leading to Fit2Fit accreditation

Notes: Fit2Fit accreditation is awarded by the British Safety Industry Federation (BSIF).

The Fit2Fit assessment process is designed to examine the knowledge, understanding and practical skills of RPD fit testers through a theory examination and practical assessment. Successful candidates are awarded the 'Fit2Fit accreditation' and are required to re-sit the practical assessment every 3-years, and the theory examination every 6 years to maintain their accreditation status. The theory examination is designed to assess a fit tester's knowledge of relevant guidance, and the practical assessment is designed to assess a fit tester's practical skill and ability to apply the method. Since this research study is concerned with investigating human errors in the application of the method, analysis was undertaken for the practical assessment only. Further studies to assess the fit tester's knowledge using the theory examination may be useful in providing information to authors of guidance and Standards, or fit test training courses.

The primary data used in this research (assessment papers) were recorded by nine different Fit2Fit assessors during Fit2Fit practical assessments conducted between 2009 and 2019 at numerous different locations around the UK. A complete dataset of 656 QLFT assessment papers, for 396 different subjects, was made available for secondary research at BSIF offices, along with supporting protocols and procedures. This dataset provided a record of the observations of over 800 hours of practical QLFT for a wide range of different filtering facepieces and reusable half mask respirators. Wearers participating in these assessments had received no previous training or experience of using RPD in the workplace and were asked to read the pre-assessment documentation provided by Fit2Fit and give consent for participation. During the practical assessment, the fit tester was required to prepare and conduct one or more fit tests for a wearer using a range of facepieces applicable to the fit test method. Each practical assessment was observed and marked by a Fit2Fit assessor who used an assessment paper to record compliance or deviation (an error). Assessment was conducted against a set of prescribed assessment criteria, which reference the HSE INDG479 protocol (Appendix B) and Fit2Fit accreditation was subsequently awarded on the basis of the assessor's judgement.

For the purposes of this study, each of the 656 available assessment papers was investigated using a methodical paper-by-paper secondary analysis process. Data was anonymised and transferred from each hard-copy paper to an electronic spreadsheet for secondary inquiry and statistical analysis. Various versions of the assessment paper were

found to have been used over the 10-year period, with additional sub-criteria and varying recording fields. Integration of all versions was necessary to allow analysis of the entire dataset, and therefore sub-criteria were merged and recording fields were dummy-coded. Of the total dataset available, 17 assessment papers were found to be either incomplete or partially illegible and were therefore removed, leaving a total of 639 assessment papers included in this study.

3.3.2 Scoring and assessment criteria

A Fit2Fit assessment paper equips the Fit2Fit assessor with a set of 49 assessment criteria and three recoding options against each: '*Fully covered*', '*Partially covered*', and '*Not covered*'. Since the purpose of this research was to investigate the prevalence and sources of human errors, these options were streamlined and dummy-coded into a binary scoring system. This scoring system used '0' to denote a '*Fully covered*' outcome (indicating compliance with the HSE INDG479 protocol), and '1' to denote a '*Partially covered*' or '*Not fully covered*' outcome (an 'error', indicating deviation from HSE INDG479 guidance).

To provide clear terms of reference, each of the 49 Fit2Fit assessment criteria were referenced by linking to the corresponding criteria in both the Fit2Fit syllabus and HSE INDG479, and these are detailed in Appendix B. The data were analysed both at individual-criterion level and in groups according to the assessment paper section headings, which facilitated analysis of specific criteria and wider trends. Data were handled in spreadsheet software (Microsoft Excel, Microsoft Office 365) and symmetry of the data was investigated using a skewness test. Appropriate descriptive statistics were used to examine the data and non-parametric statistical analysis was performed using IBM SPSS software (version 26). A Mann-Whitney U test was used to compare the error scores for sub-groups according to accreditation status of non-accredited and Fit2Fit accredited subjects (significance level = 0.05).

3.4 Results

The outcomes of this study reveal that of the 639 Fit2Fit QLFT assessment papers investigated in this study, 30.0% ($n = 192$) indicated that the fit tester failed to meet the required standard to award Fit2Fit accreditation. This figure is in line with the data reported by BSIF in a recent communication [10]. Table 3-1 displays the descriptive statistics for observations recorded at Fit2Fit QLFT assessments and provides additional data about accreditation outcome. The mean number of errors recorded in assessments for fit testers that were judged to have failed the Fit2Fit assessment was 11.2 ($n = 192$), compared to a mean of 3.2 ($n = 448$) for those that passed the Fit2Fit assessment.

Fit2Fit outcome	Recorded error scores		
	<i>M</i>	<i>Mdn</i>	Range
All	5.6	4	0 – 29
Pass ^a	3.2	2	0 – 22
Fail ^b	11.2	11	1 – 29

Table 3-1 Descriptive statistics by outcome for Fit2Fit QLFT practical assessments during the period 2009 – 2019

Notes: $N = 639$. ^a $n = 447$. ^b $n = 192$.

The frequency of errors recorded for fit testers attending Fit2Fit QLFT practical assessments for the period 2009 – 2019 is illustrated by the histogram in Figure 3-2. Possible error scores ranged from zero to 49, with a higher score indicating a greater number of errors and a higher deviation from the HSE INDG479 protocol. Error scores in the data set showed a minimum score of zero errors and a maximum score of 29 errors. The mean score ($M = 5.6$) was greater than the median ($Mdn = 4$), indicating that the distribution had a longer tail towards higher scores. This positive skew was illustrated by the shape of the frequency histogram in Figure 3-2 and confirmed by a skewness statistic that was positive ($skew = 1.34$), indicating that the data was not normally distributed. These data show that one or more errors were recorded for fit testers in 83.7% ($n = 535$) of assessments, and an interquartile range of 1 to 9, shows that between 1 and 9 errors were recorded for fit testers in half of all assessments.

The data further indicate that five or more errors were recorded in almost half of all assessments (46.3%, $n = 296$), ten or more errors in one fifth of assessments (20.7%, $n = 132$), and twenty or more errors in 2.7% ($n = 17$) of assessments.

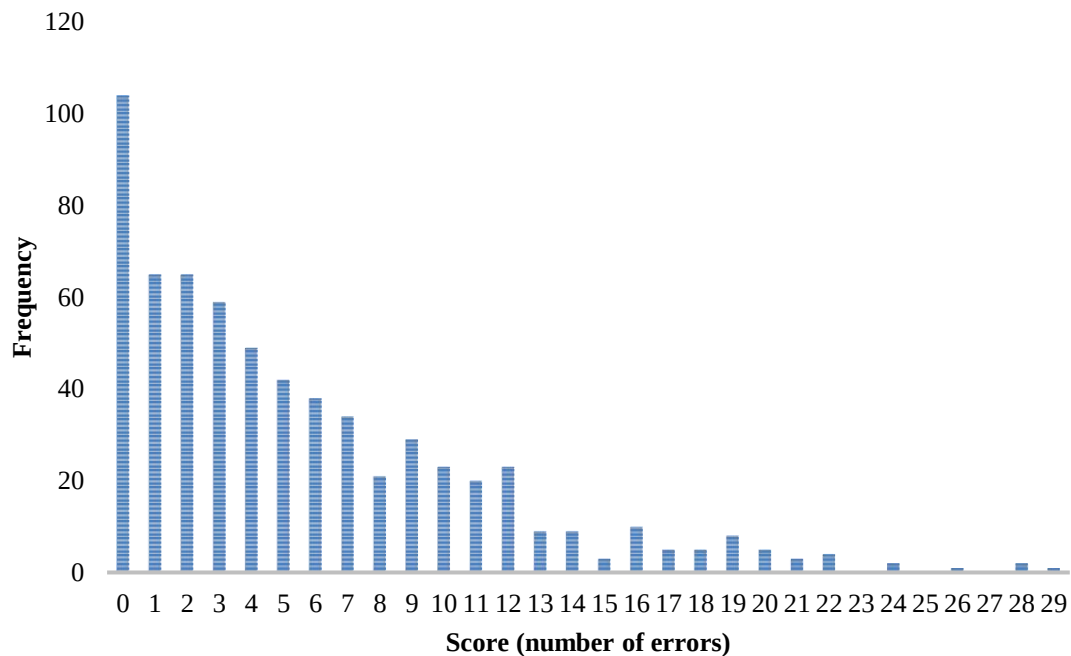


Figure 3-2 Frequency of errors recorded in Fit2Fit QLFT practical assessment papers, 2009 – 2019

Notes: $N = 639$. Maximum score = 49.

Table 3-2 provides a summary of the descriptive statistics for each year of the study period (2009 – 2019). These results indicate a trend towards a broader range of error scores over this period of study, with the five highest mean and median scores in the latest five years. When the year of assessment (time variable) was regressed on the number of errors observed, the resulting unstandardised regression coefficient was $b = 0.372$. The value of this coefficient showed that, for every one-year increase since the start of the scheme, there was an increase of 0.372 observed errors. The R^2 value for the regression model was 0.03, indicating that year of assessment accounted for only about 3% of the variance in error score.

Year	Recorded error scores			
	<i>n</i>	<i>M</i>	<i>Mdn</i>	Range
2009	20	2.8	3	0 - 10
2010	10	1.6	1	0 - 5
2011	32	5.1	3	0 - 29
2012	42	4.5	3	0 - 18
2013	54	5.0	4	0 - 19
2014	63	4.4	3	0 - 15
2015	100	6.4	6	0 - 19
2016	98	5.7	4	0 - 28
2017	105	6.5	5	0 - 21
2018	104	6.4	4	0 - 26
2019	11	7.0	5	0 - 20

Table 3-2 Descriptive statistics for Fit2Fit QLFT practical assessment papers, 2009 – 2019

Notes: *N* = 639.

Data gathered in this study also provided the opportunity to examine the descriptive statistics according to ‘accreditation status’, these results are summarised in Table 3-3. Of the 639 assessment papers investigated in this study, approximately three quarters (74.8%, *n* = 478) were conducted for non-accredited fit testers, and a quarter (25.2%, *n* = 161) were for those already Fit2Fit accredited and returning for a three-yearly re-assessment.

Accreditation status	Recorded error scores		
	<i>M</i>	<i>Mdn</i>	Range
Non-accredited ^a	6.4	5	0 - 29
Fit2Fit accredited ^b	3.0	2	0 - 22
All	5.6	4	0 - 29

Table 3-3 Descriptive statistics by accreditation status for Fit2Fit QLFT practical assessments, 2009 – 2019

Notes: *N* = 639. ^a*n* = 494, ^b*n* = 145.

The range of error scores recorded in assessments for the group of Fit2Fit accredited fit testers (0 - 22) was found to be similar to that for the group of non-accredited fit testers (0 - 29). However, the mean and median error scores for the group of Fit2Fit accredited fit testers ($M = 3.0$, $Mdn = 2$) was lower than that for the non-accredited group ($M = 6.4$, $Mdn = 5$) suggesting that Fit2Fit accredited fit testers achieve higher compliance with the HSE INDG479 protocol. A Mann-Whitney test indicated a significant difference in the distributions between the two groups (Mann–Whitney $U = 22\,311.5$, $n_1 + n_2 = 639$, $p < 0.05$ two-tailed). The result is significant at $p < 0.05$, suggesting that the group of Fit2Fit accredited fit testers make fewer errors compared to the group of non-accredited fit testers.

This study was also designed to examine the sources of errors made by fit testers in Fit2Fit assessments. Table 3-4 shows the percentage of total errors for each of the Fit2Fit sections. The highest number of errors were observed during the sections covering: *Conduct fit test* (27.4%), *Recognise wearer related factors prior to carrying out a fit test* (15.5%), and *Conduct sensitivity test* (13.4%).

Section	Fit2Fit section heading	% (n)
1	Ensure suitability of facilities	1.5 (55)
2	Preliminary actions prior to performing a qualitative fit test	3.3 (119)
3	Brief wearer on purpose of fit testing and test exercises	7.0 (252)
4	Examine the face piece for suitability prior to fit testing	11.2 (402)
5	Confirms the wearer has presented themselves as required	8.7 (311)
6	Conduct sensitivity test	13.4 (481)
7	Recognises wearer related factors prior to a fit test	15.5 (555)
8	Conduct fit test	24.7 (887)
9.1	Response if tasted (i.e., a fail)	11.2 (403)
9.2	Response if nothing tasted (i.e., a Pass)	3.2 (113)
10	Post Fit Test Procedures	0.2 (6)

Table 3-4 Frequency of errors by section recorded in Fit2Fit QLFT practical assessment papers, 2009 – 2019

Notes: $N = 3584$. % (n) = percentage (number) of all errors.

Detailed analysis of data for each of the 49 Fit2Fit criteria reveals additional information about the sources of errors made by fit testers. Table 3-5 shows the ‘top ten’ errors recorded in assessment papers investigated in this study (full results can be found in Appendix C). These criteria are ranked in order from highest to lowest prevalence with higher scores indicating a higher deviation from the HSE INDG79 protocol. The most common errors made by fit testers, each recorded in almost one third of assessment papers were: *Ensures fit check is performed* (32.2%, $n = 206$), and *Demonstrates instruction on correct method of donning an RPD including a fit check* (31.3%, $n = 200$). Approximately one quarter of all assessment papers indicated that the fit tester made errors in each of the criteria: *Verify correct function of nebuliser during the test* (25.2%, $n = 161$) and *Questions the wearer about the use of other PPE* (24.4%, $n = 156$). The least common errors recorded by assessment papers investigated in this study were: *Clean Hood* (0.5%, $n = 3$), *Dispose of Solutions* (0.3%, $n = 2$), and *Rinse & clean nebulisers* (0.2%, $n = 1$).

Fit2Fit Assessment Criteria	% (<i>n</i>)
Ensures fit check is performed	32.2 (206)
Demonstrates correct method of donning an RPD including a fit check	31.3 (200)
Verify correct function of nebuliser during the test	25.2 (161)
Questions the wearer about the use of other PPE.	24.4 (156)
Advises wearer to rinse mouth and clean face after sensitivity test	24.3 (155)
Repeat test sensitivity Test	23.2 (148)
Verify palette is clear of taste	21.0 (134)
Explains the purpose of fit testing	20.3 (130)
Explains the fit test exercises and procedure	19.1 (122)
Administer solution as per sensitivity level	18.2 (116)

Table 3-5 Top ten sources of error recorded in Fit2Fit QLFT practical assessment papers, 2009 – 2019

Notes: $N = 639$. % (n) = percentage (number) of assessment papers that recorded an error.

When Fit2Fit outcome (binary Pass or Fail) was compared with the assessor’s observation (binary compliance or deviation) against each of the 49 assessment criteria, the strongest relationship was found for *Administer top ups during fit test* ($\phi = 0.508$, $p < 0.001$). The correlation between Fit2Fit outcome and *Administer fit test solution as per sensitivity*

level was also strongly statistically significant ($\phi = 0.474$, $p < 0.001$), as was the correlation between Fit2Fit outcome and *Verify correct functioning of nebuliser* ($\phi = 0.451$, $p < 0.001$). Whilst this finding indicates that these areas are most strongly correlated to Fit2Fit outcome, this may also suggest that these areas are judged most critically by the assessor.

The secondary data used in this study were recorded by nine different Fit2Fit assessors over the ten-year period. Table 3-6 shows the recorded error scores by assessor, where identifying data have been removed to assure anonymity.

Fit2Fit Assessor	Recorded error scores				
	<i>n</i>	<i>M</i>	<i>Mdn</i>	Range	% Fail
1	36	8.1	8	0 - 19	50.0
2	236	6.2	4.5	0 - 29	26.3
3	138	8.3	7	0 - 29	44.2
4	122	3.4	2	0 - 22	23.0
5	15	5.3	5	0 - 10	26.7
6	3	7.3	7	3 - 12	66.7
7	25	1.8	1	0 - 5	4.0
8	32	2.6	2	0 - 12	21.9
9	32	2.7	1	0 - 17	25.0
All	639	5.6	4	0 - 29	30.0

Table 3-6 Descriptive statistics by Fit2Fit assessor for Fit2Fit QLFT practical assessments, 2009 – 2019

Notes: $N = 639$. % Fail = percentage of assessment papers that recorded a fail overall result.

Error scores recorded by the nine assessors show considerable variation, with mean error values ranging from 1.8 (assessor 7) to 8.3 (assessor 3). Similarly, the number of Fit2Fit failures judged by the nine assessors varied widely with a range of 4.0% (assessor 7) to 66.7% (assessor 6). The large variation for assessors 6 and 7 may be a function of a small sampling population since the number of assessments conducted by these assessors was low ($n = 3$, $n = 25$) compared to the total ($N = 639$). By contrast, the error scores for assessors 2, 3 and 4, who together have conducted 77.6% of all assessments, show more consistency ($M = 3.4 - 8.3$, % Fail = 23.0 – 44.2).

3.5 Discussion

Since the widespread requirement for fit testing under the COSHH Regulations 2002 there has been a huge uptake of fit testing in the UK workplace. Once the remit of just a handful of experts, RPD Fit Testing is now offered by in-house fit testers as well as a multitude of external providers. False passes are by far the most significant from the perspective of wearer-health and therefore the need for information about the prevalence and sources and of human errors in QLFT is of paramount importance. The variety and number of errors made in the practical application of QLFT revealed by this study is therefore of serious concern.

The results of the study into the frequency with which errors are made indicates that human errors are common and widespread, suggesting that the majority of fit testers do not apply the QLFT method strictly in accordance with the HSE INDG479 protocol. Less than one fifth of all assessment papers indicated that the fit tester applied the protocol wholly correctly (making zero errors), with over four-fifths recording errors against one or more of the Fit2Fit assessment criteria, and the worst-performing recording errors in 29 out of a possible 49 criteria. These results are all the more surprising given that the Fit2Fit scheme is voluntary and hence subjects in this study may have come from a population that deem themselves ‘good enough’ to apply for examination and value the ‘best practice’ approach supported by Fit2Fit. This indicates that human errors made in the application of QLFT in the wider UK workplace may be far more prevalent than those observed in this study.

It might be expected that the number of errors made during Fit2Fit assessments would have decreased over time with greater awareness, availability of training resources and improved practical guidance from HSE and Fit2Fit [3]. Analysis of the trend in data over the 10-year period in this study do not appear to support this expectation however, and instead show a trend towards higher deviation from the HSE protocol with increasing raw error scores and range over the period. However, due to the lower number of assessments conducted in earlier years, it is difficult to know whether this is a true effect, or a function of the lower sampling population. Another factor to consider here is that the fit testers presenting for assessment in earlier years are more likely to have been key stakeholders, including regulatory bodies, RPD manufacturers and full-time fit test providers. These

groups of fit testers could be expected to have a better knowledge of guidance, more practical experience, and better preparation, all of which may have contributed to improved performance and lower error scores. The true picture of the performance of real-world fit testers may therefore be closer to the data found for later years, where the scheme had reached those in the wider UK workplace.

The results show that the individual ability of fit testers varies considerably and those that are Fit2Fit accredited appear to make half as many errors, and hence achieve higher compliance with the HSE INDG479 protocol than their non-accredited counterparts. These results support the Fit2Fit mission statement “*to improve the respiratory health of those wearing tight fitting Respiratory Protective Equipment through being face fit tested by a Fit2Fit Accredited provider*” [11]. Although Fit2Fit accredited fit testers appear to deviate less from the HSE protocol than non-accredited fit testers, this study was not designed to examine the reasons for this apparent difference in performance. Fit2Fit accredited fit testers in this study were returning for a 3-year reaccreditation, and therefore had the benefit of attending and passing a previous accreditation plus at least 3 years’ experience. The post-assessment feedback provided by the assessor and scheme administrator provides valuable information which the more diligent candidate may use in improving their performance in future assessments, hence demonstrating fewer errors. Of the 639 assessment papers analysed in this study, 243 papers were recorded for candidates that had returned for a first, second or third re-accreditation or re-attempt. Future research using this dataset could be used to reveal trends in error scores and types of errors made during repeat assessments. Analysis of these trends alongside information about further study could be used to indicate to what extent any improvement in competence is achieved following re-training.

The second important area investigated by this study was an examination of the sources of errors which was evaluated through analysis of Fit2Fit sections and individual assessment criteria. The overlap in the data relating to Fit2Fit ‘pass’ or ‘fail’ outcomes provides evidence to suggest that the source of errors is as, or more, important than the number of errors. Analysis shows that the best-performing fit testers in the ‘fail’ group made only 2 errors, whereas the worst-performing fit testers in the ‘pass’ group made 22 errors. One note of caution here: since a pass / fail outcome is awarded on the basis of the

assessor's judgement alone, then there is the potential for assessor bias and further analysis including additional assessor data would be required to answer this question. If indeed assessors judge some errors to be more critical than others, then the scheme would be more robust and would benefit from a clearer set of guidelines for assessors, perhaps indicating which criteria are to be judged most critically.

The secondary data used in this study were recorded by nine different Fit2Fit assessors. The variation in error scores recorded by some assessors may be indicative of differences in harshness of judgement which could therefore skew the data used in this study. In addition, cognitive biases originating from their own skills, experience, perceptions, or points of view, may impact an assessor's subjective view of performance. Such biases are difficult to overcome in practice, however a systematic approach to ensure that assessors receive regular training, monitoring and joint observation could all help in setting and maintaining a consistent standard for Fit2Fit. Since assessors typically operate in different geographical locations, this may result in the data for some assessors to be skewed by industry / sector type. Further analysis of this secondary dataset could examine and compare the prevalence and sources of errors made by fit testers from within specific sector types. This information may be useful to regulators in focussing awareness and training campaigns.

It is perhaps not surprising that most errors were found for the section *Conduct fit test*, since this section is the crux of the assessment process, and accounts for 10 of the total 49 individual assessment criteria. Indeed, of the ten Fit2Fit sections, those which involve the set-up, use and application of fit test equipment accounted for more than half of all errors recorded during Fit2Fit assessments. This is further emphasised by the data which shows that almost a third of assessment papers recorded that the fit tester failed to verify the function of the nebuliser at any point during the procedure. Analysis of the individual assessment criteria reveal that errors are made across a wide variety of different sections. Surprisingly, four of the top ten most common errors relate to the instruction given by the fit tester to the wearer in donning / fit checking the mask or explaining the purpose of the fit test and exercises. At first sight, these areas may not appear to directly influence the ability of the method to determine the fit of a respirator, however the wearer's understanding of the principles of this method is crucial to ensuring an effective fit test.

Since this method is subjective and relies wholly upon the wearer's cooperation and honesty in taste detection, their comprehension of the test protocol, correctly reporting detection of taste, and function of the test exercises cannot be underestimated and errors in these areas may have critical consequences. This is further highlighted by assessor comments against some of these criteria that noted "*a complete lack of explanation of the basic tenet of taste-test QLFT*", and even "*told wearer that the aim is NOT to taste the solution!*".

The ability of the fit tester to correctly administer the taste solution, and the wearer to correctly report taste are key to the reliability of the QLFT method. The fact that four of the most common errors relate to delivery of the taste aerosol and management of the wearer's palate is therefore of concern. The QLFT method relies solely on the wearer's taste response in determining face seal leakage, and errors which affect the wearer's ability to detect the challenge agent have the potential to lead to false pass results and may have serious consequences for the wearer's health. The prevalence of errors relating to taste detection supports other researchers' suggestions for inclusion of a placebo control during sensitivity testing and further research is needed in this area [12]. This modification may improve the reliability of the test by eliminating unreliable responses that may lead to false pass results. In addition, failure to administer the taste aerosol correctly has the potential to limit the ability of the fit test to identify face seal leakage, allowing a poor fitting respirator to pass the fit test. Assuring proper generation of the taste aerosol by verification of nebuliser function is fundamental to ensuring that the respirator seal is properly challenged, and this area must be emphasised during fit tester training programmes.

Very few assessment papers recorded errors in the condition, positioning, and post-test cleaning of the test hood. The role of the test hood in containing the challenge agent is critical, and there have been concerns that the size and design some of the test hoods on the market may render them unfit for purpose. A recent report found some hoods to have a volume 140% larger than the recommended hood volume and expressed concerns over the use of these hoods in terms of test sensitivity [13]. Current Fit2Fit criteria assess the suitability of the test hood in terms of overall condition but not absolute volume, and this may be a recommendation for future revisions of the assessment criteria.

3.6 Conclusion

These results indicate that a significant proportion of Fit2Fit assessment papers recorded errors made by the fit tester in application of the QLFT method. Fit2Fit accredited fit testers appear to make fewer errors than non-accredited fit testers, but the results indicate no improvement in performance over the ten-year period of this study. The most common individual errors made during assessments investigated in this study were in ensuring proper donning and fit checking RPD, and improvements in the fit tester's instruction to the wearer are needed in this area. The indication that areas involving the set-up, use and application of fit test equipment accounted for more than half of all errors recorded during Fit2Fit assessments highlights a potentially serious problem and these areas require further investigation to ascertain the impacts of such errors. The purpose of this study was to investigate the prevalence and sources of human errors to gain a complete understanding of errors in QLFT and ultimately to improve the workplace safety of this method. Whilst many of the errors made by fit testers have the potential to undermine the test protocol, further research is needed to evaluate the impact on QLFT outcome and hence the likelihood of these errors leading to false passes in workplace fit testing programmes.

3.7 References

1. Health and Safety Executive. L5 Control of substances hazardous to health. The Control of Substances Hazardous to Health Regulations 2002 (as amended). Approved Code of Practice and guidance L5 (Fifth edition). HSE Books; 2005. Available at: <http://www.hse.gov.uk/pubns/priced/l5.pdf>
2. Health and Safety Executive. INDG479 Guidance on respiratory protective equipment (RPE) fit testing. HSE Books; 2019. Available at: <https://www.hse.gov.uk/pubns/indg479.pdf>
3. British Safety Industry Federation. Fit2Fit Companion to HSE INDG479 guidance on respiratory protective equipment (RPE) fit testing, qualitative method. June 2021. Available at: <https://www.fit2fit.org/resources>
4. British Safety Industry Federation. Fit2Fit Companion to HSE INDG479 guidance on respiratory protective equipment (RPE) fit testing, quantitative ambient particle counting method. May 2021. Available at: <https://www.fit2fit.org/resources>
5. British Safety Industry Federation. Fit2Fit Companion to HSE INDG479 guidance on respiratory protective equipment (RPE) fit testing, quantitative controlled negative pressure method. May 2021. Available at: <https://www.fit2fit.org/resources>
6. British Standards Institution. BS ISO 16975-3:2017 Respiratory protective devices. Selection, use and maintenance. Part 3: Fit-testing procedures. London: British Standards Institution; 2017.
7. Long K., Mai N., Williams M. Human errors in qualitative fit testing: a literature review [Unpublished paper]. 2019.
8. British Safety Industry Federation. Respiratory Protection Group meeting minutes October 2021.
9. British Safety Industry Federation. Fit2Fit scheme syllabus. Available at: <https://www.fit2fit.org/resources> (Accessed: 6 May 2021)

10. British Safety Industry Federation. Fit2Fit Technical Committee meeting minutes February 2021.
11. British Safety Industry Federation. Fit2Fit Mission Statement. Available at: www.fit2fit.org (Accessed: 13 May 2021)
12. McKay RT., Davies E. Capability of respirator wearers to detect aerosolized qualitative fit test agents (sweetener and Bitrex) with known fixed leaks. *Applied Occupational and Environmental Hygiene*. 2000; 15(6): 479–484. Available at: DOI:10.1080/104732200301269
13. Health and Safety Executive., Bowry A. Effect of hood size on a qualitative Fit Test. 2018. Available at: [https://www.hsl.gov.uk/media/1270950/hood size powerpoint.pdf](https://www.hsl.gov.uk/media/1270950/hood_size_powerpoint.pdf) (Accessed: 15 February 2021)

4 Human errors in qualitative fit testing: a theoretical study of the impact of human errors on fit test outcome

4.1 Abstract

It has been observed that real-world fit testers make numerous errors in their application of the qualitative fit test method when measured against the criteria used by Fit2Fit, however the impact of these errors has not been studied until now. This study assesses the theoretical impact of errors on the outcome of a qualitative fit test through evaluation against the key principles of the method. Fit2Fit criteria were filtered to reveal a set of ‘critical errors’ which have the potential to impact one or more of the key method principles and hence to produce a false pass fit test outcome. When applied to previous observations of real-world fit testers, these results suggest that a significant proportion of fit testers make one or more critical errors that could lead to a false pass result and incorrect assignment of a poor fitting respirator.

4.2 Background

Fit testing is an essential part of any respiratory protective device (RPD) programme and ensures that a specific make, model, and size of tight-fitting RPD facepiece is *suitable* in providing an adequate facial fit to an individual wearer. Tight-fitting RPD including those incorporating disposable, half and full facepieces are used by millions of workers in the UK in a wide variety of business sectors from healthcare to building and construction. Due to the low cost of equipment and perceived simplicity of use, the qualitative fit test (QLFT) method is the most commonly used method for fit testing of disposable and reusable half facepieces, often by those with little or no experience [1]. Fit testing guidance, manufacturers' instructions, national and international standards provide detailed advice on proper application of the QLFT method and caution that any deviation from the required protocol may invalidate fit test results [2][3][4]. Data published by the British Safety Industry Federation (BSIF) in October 2021 revealed that 30% of applicants to the Fit Test Providers 'Fit2Fit' Accreditation Scheme (Fit2Fit scheme) failed to meet the standard required to pass the QLFT practical assessment [5]. This highlights potential problems with application of the QLFT method when measured against the Fit2Fit criteria which derive from the guidance set out by the Health and Safety Executive (HSE) in guidance document INDG479 [6][7].

A previous study has highlighted the extent to which real-world fit testers make errors when measured against these criteria, but that study was not designed to provide information about the potential impacts of these errors upon fit test outcome [8]. Errors which affect the outcome of a fit test by producing a false pass result may allow a poor fitting facepiece to be assigned to a wearer erroneously, potentially leading to increased exposure to hazardous substances in the workplace [9]. The probability that the QLFT method will return a false pass fit test result has been extensively investigated in laboratory studies and is reported by researchers as the 'Beta error' [10]. Human errors in workplace QLFT have the potential to increase the Beta error beyond that observed in laboratory studies, and work is needed in this area to better understand their possible impact. This paper is the first to assess the theoretical impact of human errors on qualitative fit test outcome and aims to inform a better understanding of human errors in QLFT to improve the safety of this method. A theoretical analysis is crucial to understanding the underlying principles of the method, how human errors can undermine

these principles and what, if any, is the theoretical impact on fit test outcome. This research will inform areas of investigation for subsequent studies into the practical impacts of errors, providing a focus for targeted research. It is critical that human errors in the QLFT method are eliminated to assure safe application of this procedure in the UK workplace and hence to protect wearers' health.

4.3 Method

4.3.1 Selection of evaluation criteria

With a continuous programme of improvement towards ‘gold standard’ best practice, current fit testing guidance has developed far beyond the basic tenets of the original QLFT protocol which was proposed by the Minnesota Mining and Manufacturing Company, and later validated for use as a real-world method [11][12][13]. The gradual broadening of the fit tester’s role has been encouraged by HSE, Fit2Fit and others in an attempt to address some of the many issues surrounding incorrect use, maintenance, and application of RPD in the workplace, along with wearer-related factors such as facial hair [14][15][16]. As such, current guidance now places many additional expectations on fit testers which sit beyond the principles of the original validated QLFT method [6].

Whilst some human errors in QLFT may be deemed important from the standpoint of an effective RPD programme, they may not have any real impact upon the outcome of a fit test. A theoretical evaluation of the impacts of errors should therefore begin by addressing whether, and how, errors may affect the key principles of the method - those upon which the sensitivity of the method depends.

The original QLFT protocol set out three key principles which were judged to be essential to the ability of the method to detect a poor fitting facepiece:

- 1) a sensitivity check to assure that the respirator wearer is able to taste a low level of challenge agent should a leakage occur,*
- 2) procedures to assure that the aerosol is generated in a consistent and controlled manner, and*
- 3) steps to assure that the test atmosphere is contained around the respirator and is controlled [11].*

The first key principle concerns the wearer’s ability to correctly identify and detect the taste of a challenge agent - a factor critical to the sensitivity of the method. Human errors that have a theoretical impact on the first key principle are termed ‘subjectivity errors’ in this study. These types of human errors have the potential to compromise the human taste

response; an element which essential to the ability of the wearer to detect face seal leakage.

The second and third key principles concern the proper use of QLFT equipment to assure the correct generation and control of a suitable aerosol concentration used to challenge a facepiece. These key principles require the appropriate use and application of equipment including handheld nebulisers, a test hood and two solutions. Errors made against these two principles (in this study termed ‘equipment errors’) have the potential to limit the in-hood challenge aerosol concentration, a factor critical to the ability of the method to properly identify poor fitting respirators.

In theoretical terms, any human error made in workplace fit testing which reduces the efficacy of one or more of these three key principles has the potential to decrease the Test Sensitivity (ability of the method to correctly identify poor fitting facepieces) and increase the Beta error (defined as ‘1 – Test Sensitivity’) [17].

4.3.2 Theoretical assessment of impact

A theoretical assessment of real-world errors was conducted by evaluating the potential impact of an error on any, or all, of the three key method principles. Using the assessment criteria set out by Fit2Fit, a set of real-world errors was theorised, and the potential impacts were analysed and evaluated against the key method principles. The data used in this chapter comprise a series of ‘judgements’ made by the author which were based on extensive experience, referenced to current guidance, and subjected to peer-review by Fit2Fit and academic colleagues. Equitable judgement was applied to every sample of the data, allowing equal confidence to be ascribed to the dataset as a whole.

Fit2Fit assessment criteria provide the most exhaustive available list of a fit tester’s tasks and can be clearly and traceably referenced to HSE guidance and QLFT protocols [8]. Taking each Fit2Fit criterion descriptor in turn, a ‘human error’ was theorised as the null response to the descriptor through reference to the guidance provided in manufacturers’ user instructions, British Standard BS ISO 16975-3, HSE guidance document INDG479, Fit2Fit Companion documents and personal practical experience [3][4][2][7].

One or more ‘worst-case’ error scenarios were identified for each of the theorised human errors and the outcome of each was hypothesised. The impact of each error scenarios was then judged against the key method principles set out in section 4.3.1. Figure 4-1 shows an example of this judgement process for Fit2Fit criterion 8d *Administer fit test solution as per sensitivity level*. Full results for all Fit2Fit criteria are shown in Appendix D.

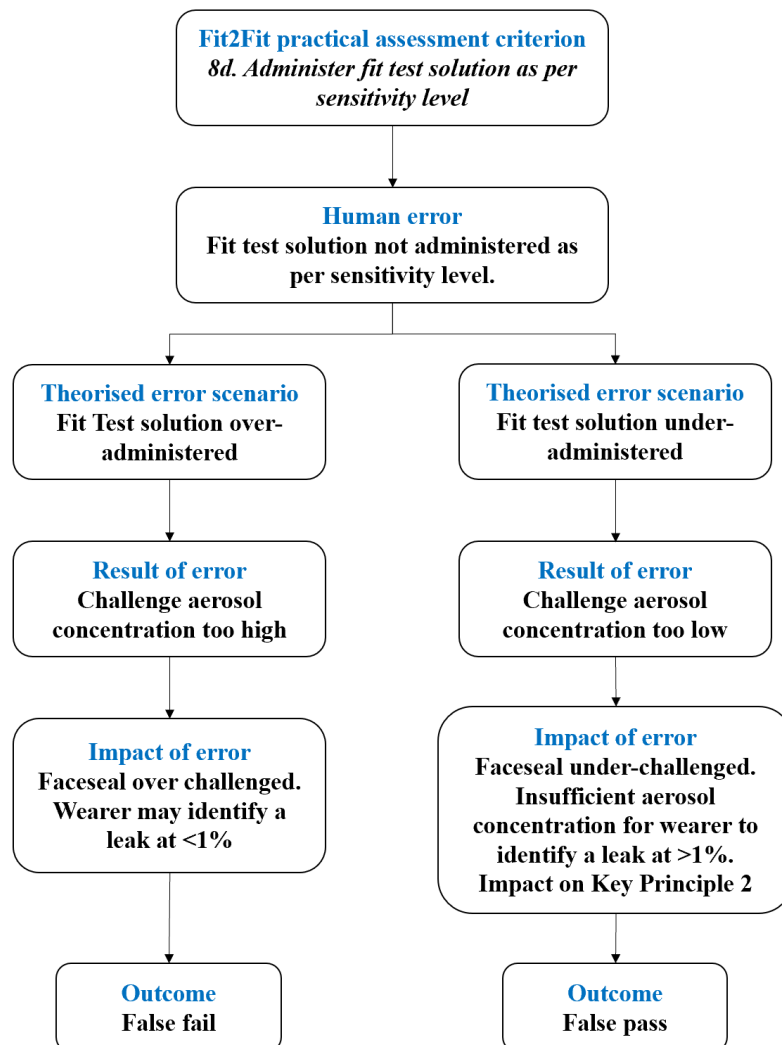


Figure 4-1 Process for theorising and judging the impact of human errors in QLFT for Fit2Fit criterion 8d

The theorised set of error scenarios aimed to include possible human errors in all parts of the QLFT procedure and encompass errors which derive both from a lack of knowledge / understanding or accidental misapplication. The resulting list of error scenarios was peer-reviewed by Fit2Fit and academic colleagues to ensure that as many scenarios were included as possible.

User instructions were accessed through authenticated manufacturer's websites to ensure the most up to date version was available for study [3]. The British Standard was accessed via the British Standards Institution rather than through the International Organization for Standards (ISO) website so as to include any National Foreword pertaining to local or national guidance [4]. Fit testing guidance documents and companion guides were accessed through the HSE and Fit2Fit websites [2][7].

Many of the Fit2Fit assessment criteria were found to relate to a single specific task, e.g., criterion 6a: *Position hood correctly*, whereby an error could be clearly theorised as the single human error scenario, *Hood positioned incorrectly*. Other assessment criteria were found to be more complex and to encompass a number of different error scenarios, e.g., criterion 8d: *Administer solution as per sensitivity test*, where two possible error scenarios included *Solution under-administered* and *Solution over-administered*, each having a different potential impact on in-hood challenge aerosol concentration. Logical reasoning, practical experience and peer-review were used to assess whether each of the identified error scenarios could have a theoretical impact on one or more of the key principles. The term 'impact' is defined for the purposes of this study as: a factor which has the potential to contravene one of the key principles of the QLFT method such that a false pass fit test result may be assigned.

Outcomes of the theoretical assessment were used to compile a list of error scenarios against each of the Fit2Fit criteria and key method principles. Where duplicate scenarios were found against more than one Fit2Fit criteria, these were removed to leave a unique list of errors, which were termed 'critical errors'. For the purposes of this study a critical error is defined as: an error which has the potential to adversely affect one or more of the key method principles. Results of this evaluation were applied to the findings of a previous study of real-world fit testers to provide an indication of the theoretical impact of the errors previously observed [8].

4.4 Results

The theoretical assessment process identified a total of sixty-six error scenarios against the set of Fit2Fit criteria. Full results of this process are detailed in Appendix D which provides references to sources, reasoning, key principle affected (or none) and impact. The proportion of error scenarios found against each key principle (or none) is shown in Figure 4-2.

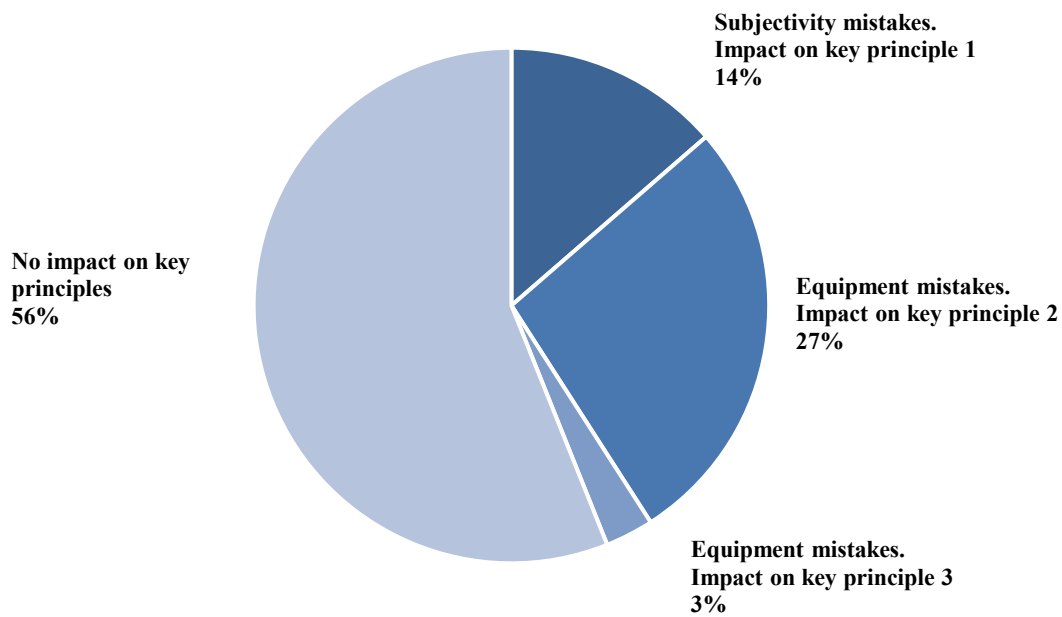


Figure 4-2 Distribution of error scenarios for QLFT key principles 1), 2), 3) or none

Notes: $N = 66$. Key principles: 1) a sensitivity check to assure that the respirator wearer is able to taste a low level of challenge agent should a leakage occur, 2) procedures to assure that the aerosol is generated in a consistent and controlled manner, 3) steps to assure that the test atmosphere is contained around the respirator and is controlled.

These data indicate that over half (56.1%, $n = 37$) of all error scenarios were judged to have no theoretical impact upon any of the three key principles. This indicates the extent to which the Fit2Fit scheme has developed beyond the original protocol to include peripheral tasks which are assigned to the fit tester's role. Just under half (43.9%, $n = 29$) of the error scenarios were judged to have a theoretical impact on one or more of the three key principles. Of these, the data suggest that 13.6% ($n = 9$) were 'subjectivity errors', having the potential to impact key principle 1). A further 27.3% ($n = 18$) were judged to have the potential to impact key principle 2), and 3.0% ($n = 2$) have the potential to impact

key principle 3). These data therefore indicate that just under a third (30.3%, $n = 20$) of all error scenarios identified were ‘equipment errors’ (those which have the potential to impact key principles 2 and 3).

Several Fit2Fit criteria were identified as being susceptible to numerous different critical errors, suggesting that these criteria are highly vulnerable to error. Table 4-1 provides a summary of these criteria, all of which relate to key principle 2).

Fit2Fit ref.	Fit2Fit practical assessment criteria	Critical error	Key principle
2b	Fills the nebulisers with the correct solutions and checks volume	Fit test nebuliser filled with sensitivity solution	2
		Fit test nebuliser under-filled	2
2c	Checking nebuliser functioning	Fit test nebuliser clogged	2
		Fit test nebuliser bottom bung not removed	2
		Fit test nebuliser Venturi mis-seated	2
8g	Monitors wearer and equipment through test to ensure correct protocol HSE INDG479	Fit test nebuliser clogged	2
		Fit test nebuliser bottom bung not removed	2
		Fit test nebuliser Venturi mis-seated	2
8i	Administer top-ups correctly throughout test.	Fit test solution under-administered	2
		Fit test nebuliser not upright	2
8j	Verify correct function of nebuliser during the test	Fit test nebuliser clogged	2
		Fit test nebuliser bottom bung not removed	2
		Fit test nebuliser Venturi mis-seated	2
		Fit test nebuliser not actuated with full squeezes	2

Table 4-1 Fit2Fit criteria with two or more theorised critical error scenarios

Notes: Fit2Fit ref. = criteria used during a Fit2Fit practical assessment. Critical error is an error which has the potential to adversely affect one or more of the key method principles. Key principle 2: procedures to assure that the aerosol is generated in a consistent and controlled manner.

The method of theorising error scenarios described in Figure 4-1 produced several identical error scenarios for different Fit2Fit criteria (see Appendix D). For example, *Fit test nebuliser clogged*, was identified as a possible error scenario against four separate Fit2Fit criteria (2c, 8g, 8j and 10c). In practical terms however, *Fit test nebuliser clogged*

represents a single error type, and therefore once the process of theorising error all scenarios was complete, duplicates were removed. This process identified a total of eighteen unique ‘critical errors’: seven subjectivity errors (affecting key principle 1) and eleven equipment errors (affecting key principles 2 or 3). A summary of the critical errors identified by key method principle is shown in Table 4-2.

Critical error	Key principle affected
Subjectivity errors	
Room poorly ventilated	1
Wearer has smoked, eaten, or drunk before test	1
Wearer not advised to rinse mouth / clean face	1
Clear palate not verified	1
Wearer not instructed to breathe through mouth	1
Sensitivity test not repeated	1
Hood not cleaned	1
Equipment errors	
Fit test nebuliser filled with sensitivity solution	2
Fit test nebuliser under-filled	2
Fit test nebuliser clogged	2
Fit test nebuliser bottom bung not removed	2
Fit test nebuliser Venturi mis-seated	2
Fit test solution under-administered	2
Fit test timing incorrect	2
Fit test nebuliser not upright	2
Fit test nebuliser not actuated with full squeezes	2
Hood damaged / unsuitable	3
Hood positioned incorrectly	3

Table 4-2 Critical errors identified against the key principles of the QLFT method

Notes: Appendix D provides full results. Critical error is an error which has the potential to adversely affect one or more of the key method principles. Key principles: 1) a sensitivity check to assure that the respirator wearer is able to taste a low level of challenge agent should a leakage occur, 2) procedures to assure that the aerosol is generated in a consistent and controlled manner, 3) steps to assure that the test atmosphere is contained around the respirator and is controlled.

A deeper examination of the subjectivity errors suggests two distinct categories: errors that may affect the wearer's taste prior to the sensitivity test (two scenarios), and errors that may affect the wearer's taste prior to or during the fit test (five scenarios). A similar examination of equipment errors reveals that these relate either to set up of the equipment (five scenarios), or to application of the equipment during the fit test (six scenarios).

When applied to the outcomes of the real-world previous study, results suggest that 39% ($n = 1400$) of the errors made by fit testers during Fit2Fit QLFT practical assessments during the period 2009 - 2019 could be considered to be 'critical errors' [8]. These data are shown in Table 4-3, and further details with references to the results of the previous study can be found in Appendix E.

Key principle	% (n) of observed errors
1	19.9 (715)
2	17.7 (636)
3	1.4 (49)
Total	39 (1,400)

Table 4-3 Critical errors applied to total errors made by real-world fit testers during Fit2Fit QLFT practical assessments [8]

Notes: $N = 3584$. % (n) = percentage (number) of all errors observed during Fit2Fit QLFT practical assessments, 2009 – 2019 [8]. Critical error is an error which has the potential to adversely affect one or more of the key method principles. Key principles: 1) a sensitivity check to assure that the respirator wearer is able to taste a low level of challenge agent should a leakage occur, 2) procedures to assure that the aerosol is generated in a consistent and controlled manner, 3) steps to assure that the test atmosphere is contained around the respirator and is controlled.

When critical errors were analysed at individual fit test level, the results suggest that 71.5% ($n = 457$) of fit tests observed during Fit2Fit QLFT practical assessments 2009 – 2019 included one or more critical errors [8]. These results are shown in Table 4-4 and further details of this analysis are provided in Appendix E.

Key principle	% (<i>n</i>) of fit tests observed
1	55.6 (355)
2	49.5 (316)
3	6.9 (44)
All	71.5 (457)

Table 4-4 Critical errors applied to fit tests observed during Fit2Fit QLFT practical assessments

Notes: $N = 639$. % (n) = percentage (number) of all fit tests observed during Fit2Fit practical assessments, 2009 – 2019. Critical error is an error which has the potential to adversely affect one or more of the key method principles. Key principles: 1) a sensitivity check to assure that the respirator wearer is able to taste a low level of challenge agent should a leakage occur, 2) procedures to assure that the aerosol is generated in a consistent and controlled manner, 3) steps to assure that the test atmosphere is contained around the respirator and is controlled.

The data suggest that almost one half (49.5%, $n = 316$) of all fit tests observed included critical errors made against key principle 2) regarding proper use of the nebulisers and solutions in correct generation of the test aerosol. The data further indicates that over one half (55.6%, $n = 355$) of all fit tests observed in this study were found to include critical errors against key principle 1) affecting the wearer's taste detection.

4.5 Discussion

A previous study has highlighted the extent to which fit testers make errors in applying the QLFT protocol under Fit2Fit assessment conditions, but the potential impact of these errors has not been studied until now [8]. False pass fit test results should be of real concern to fit testers, regulatory bodies, and equipment manufacturers. More crucially, these unsafe outcomes should be of concern to RPD wearers whose health or life may depend upon the protection afforded by a tight-fitting facepiece. The results of this study reveal a list of critical errors that have the theoretical potential to produce a false pass result in QLFT. When applied to the previous study, these data suggest that critical errors may have been made in over 70% of fit tests observed during Fit2Fit QLFT practical assessments for the period 2009 - 2019. This provides a worrying insight into the application of the QLFT method by real-world fit testers and poses some questions to the Fit2Fit scheme about the use of assessment criteria in the identification or poor practice and the prevention of false pass results.

Factors which affect the wearer's ability to properly detect taste (key principle 1) must be treated with the utmost importance if a reliable fit test outcome is to be assured since detection of the challenge agent is used as a proxy for face seal leakage. The impacts of these 'subjectivity errors' are likely to be broad and wearer-specific due to the well-reported and widely varied human taste-response mechanism [18]. Issues with human taste detection were discussed in a previous literature review which revealed gaps in research into errors in QLFT including the highly subjective nature of the method [19]. That review highlighted concerns about the possibility of masking or mischaracterisation of taste and concluded that further research is crucial to a better understanding of taste factors in QLFT. Although beyond the scope of this study, further research and collaboration between fit testing researchers and medical experts would be welcome in addressing this complex issue. It is clear that both the fit tester and the wearer have a role to play in following the guidance provided by HSE and manufacturers, and also in understanding the importance of taste detection in the QLFT method. Any measures which aim to mitigate the potential impacts of errors involving the wearer's taste response must therefore address both practical measures and training needs.

Perhaps of greatest concern are the results which reveal the extent to which ‘equipment errors’ appear throughout the entirety of the test procedure including set up, sensitivity testing, fit testing and cleaning. This indicates that the use and application of QLFT equipment is highly susceptible to human error. The error scenarios *Fit test nebuliser clogged*, *Fit test nebuliser bottom bung not removed* and *Fit test nebuliser Venturi mis-seated* were the most prevalent and appeared across three different Fit2Fit criteria. This suggests that these areas are highly impactable by human errors, i.e., one error kind could manifest itself in several mechanisms that lead to false pass fit test outcomes. Although QLFT equipment manufacturers provide detailed instruction on proper use of QLFT equipment, limited attention is paid to potential problems and troubleshooting in the guides available. The presence of clear warnings about the potential effects of misuse would perhaps focus the fit tester’s attention and reduce human errors. The ISO 16975-3 Standard provides an equipment specification based upon design and performance characteristics but sets no requirement for supporting materials such as user instructions. Currently there is no requirement in the UK to register QLFT kits to ensure they contain a proper set of equipment and instructions, and neither is there a requirement to register individual components as being ‘fit for purpose’. A formal specification and registration process for both QLFT equipment and supporting materials would close a gap in these resources which currently appear insufficient in preventing or reducing human errors

Five of the Fit2Fit criteria were found to be susceptible to multiple critical errors, and all of these were associated with incorrect use of the fit test nebuliser. This suggests that proper completion of these five criteria should be regarded as vital in awarding a pass during Fit2Fit practical assessments. This also indicates that use of the fit test nebuliser is highly vulnerable to human error and therefore is a skill that must be learned and applied meticulously. This further suggests that these criteria are open to compound errors which could have an even greater impact on fit test outcome than errors made individually.

One interesting observation from this investigation is the feature that over half of the Fit2Fit criteria were judged not to have the potential to lead to a false negative outcome. There are two indications that arise from this observation: firstly, the increasing breadth of the Fit2Fit syllabus, and secondly the possibility of false positive (false fail) outcomes.

It has been discussed previously that ongoing evolution of the Fit2Fit syllabus has given rise to an increase in the number of tasks a fit tester must complete beyond that of simply ‘conducting a fit test’. The previous study found four different versions of the Fit2Fit practical assessment paper over the course of the ten-year sample. A re-examination of these versions shows an increase in the number of assessment criteria of over 20% during the period which is indicative of the creep towards increasing the complexity of fit testers’ tasks. Although it may be viewed as helpful from a health and safety programme standpoint, broadening the fit tester’s role may be counterproductive overall. The burden on fit testers should not be increased such that their primary function is lost; distraction, time pressures and added tasks may all increase the likelihood of errors being made. The second potentially more concerning outcome from this part of the data is the theoretical possibility of false fail (false positive) outcomes. Of the sixty-six error scenarios identified, eleven were judged to affect the fit testing procedure so as to be a potential source of a false fail result, these are indicated by an asterisk (N*) in Annex 4.1. False fail results occur in QLFT when a wearer with a well-fitting facepiece (fit factor ≥ 100 for half mask RPD by quantitative fit testing) indicates detection of the challenge agent during the fit test and hence fails the fit test falsely. The consequences of false fail outcomes are largely considered to be wasteful of time, effort, and resources, but nonetheless benign. This general perception may be misleading however, and the ramifications of false fails may in fact be more serious than lost time or earnings. Each time a false failure occurs, the wearer must undertake a new fit test, and each time a new fit test is conducted there is another opportunity for a false pass to occur. The compound effect of a series of failed and repeated tests was highlighted by Coffey et al (2006) who concluded that *‘retesting may cause a poorly fitting respirator to be erroneously assigned to a wearer because most programs do not disqualify a person from wearing a respirator after failing a fit test the first time.’* [10].

Theoretical studies such as this are inherently limited in their nature, and caution should be exercised when considering the practical implications. This study is limited in scope to a theoretical impact of errors on fit testing outcome; the range and severity of errors made by fit testers in the real-world is likely to be far more varied. In addition, some errors may be habitual (for example as a result of poor learning), whilst others may be intermittent (for example due to distraction or time pressure). When observing fit testers

in real-world studies, the ‘Hawthorne effect’ may cause subjects to alter their behaviour due to their awareness of being observed. This effect may result in an improvement in observed action or behaviour, especially during Fit2Fit assessments. It is important therefore to recognise that this study has theorised the impact of ‘worst-case scenarios’ which may over-estimate those observed in the real-world. Nonetheless, this study has revealed those areas which have the theoretical potential to lead to a false pass result, and these areas should be viewed as critical when applying the method in the workplace.

Further research to investigate the practical impacts of the critical errors theorised in this study is important to gain an understanding of the impact of human errors on QLFT outcome in real-world settings. Laboratory studies which track in-hood aerosol concentration would be useful in understanding to what extent aerosol generation and containment may be affected by human errors. In addition, qualitative and quantitative fit test studies using volunteer wearers would provide an indication of the impact of human errors on their ability to detect face seal leakage. Although a significant undertaking, studies which assess the real-world impact of human errors on the level of protection provided in the workplace would ultimately bring the investigation full circle. Such studies could be conducted through a case-study approach, where an in-house programme of QLFT is observed against the criteria set by Fit2Fit, and then facial fit is immediately ‘checked’ using a QNFT method to assess the impact of any errors observed.

4.6 Conclusion

This study has proposed a list of critical errors which have the theoretical potential to impact one or more of the three key principles of the QLFT method and lead to a false pass result. Comparison with a previous real-world study suggest that over 39% of the errors observed amongst real-world fit testers during Fit2Fit QLFT practical assessments may have been critical errors, and these were recorded in over 70% of the fit tests observed. This theoretical assessment revealed that errors relating to the use and application of QLFT equipment were the most numerous and widespread throughout the fit testing procedure. Whilst this research identifies the critical errors which have the theoretical potential to adversely affect fit testing outcome, empirical studies are needed to evaluate whether these can indeed translate to practical impacts which affect the outcome of a fit test.

4.7 References

1. Murray A. UK best practices on RPD fit testing (Fit2Fit programme). National Fit Testing Programmes. 18th May 2021. International Society for Respiratory Protection; Available at: <https://www.isrp.com/minutes>
2. Health and Safety Executive. INDG479 Guidance on respiratory protective equipment (RPE) fit testing. HSE Books; 2019. Available at: <https://www.hse.gov.uk/pubns/indg479.pdf>
3. 3M. Qualitative fit test apparatus FT30 user instructions. Available at: <https://multimedia.3m.com/mws/media/1464265O/3m-qualitative-fit-test-apparatus-ft-30-user-instruction.pdf> (Accessed: 16 October 2021)
4. British Standards Institution. BS ISO 16975-3:2017 Respiratory protective devices. Selection, use and maintenance. Part 3: Fit-testing procedures. London: British Standards Institution; 2017.
5. British Safety Industry Federation. Fit2Fit Newsletter December 2020. Available at: www.fit2fit.org/resources (Accessed: 15 February 2021)
6. British Safety Industry Federation. Fit2Fit scheme syllabus. Available at: <https://www.fit2fit.org/resources> (Accessed: 6 May 2021)
7. British Safety Industry Federation. Fit2Fit Companion to HSE INDG479 guidance on respiratory protective equipment (RPE) fit testing, qualitative method. June 2021. Available at: <https://www.fit2fit.org/resources>
8. Long K., Mai N., Williams M. Human errors in qualitative fit testing: a study of real-world fit testers [Unpublished paper]. 2020.
9. Campbell DL., Coffey CC., Lenhart S. Respiratory protection as a function of respirator fitting characteristics and fit-test accuracy. American Industrial Hygiene Association Journal. 2001; 62(1): 36–44. Available at: DOI:10.1080/15298660108984607
10. Coffey CC., Lawrence RB., Zhuang Z., Duling MG., Campbell DL. Errors associated with three methods of assessing respirator fit. Journal of Occupational and

Environmental Hygiene. 2006; 3(1): 44–52. Available at:
DOI:10.1080/15459620500455398

11. Occupational Safety and Health Administration. Minnesota Mining and Manufacturing Company post hearing data submission on qualitative fit testing protocols under the lead standard before OSHA. Docket No. H-049A. 1981.
12. Marsh JL. Evaluation of saccharin qualitative fitting test for respirators. American Industrial Hygiene Association Journal. 1984; 45(6): 371–376. Available at:
DOI:10.1080/15298668491399965
13. Mullins H., Danisch S., Johnston A. Development of a new qualitative test for fit testing respirators. American Industrial Hygiene Association Journal. 1995; 56(11): 1068–1073. Available at: DOI:10.1080/15428119591016278
14. Report of the stakeholders meeting on respiratory protective equipment fit testing. 22nd April 2005. Health and Safety Laboratory; Available at:
<https://www.hse.gov.uk/asbestos/assets/docs/stakeholdermeeting220405.pdf>
15. Bell N., Vaughan N., Hopkinson J., Hill H. RR798 Factors influencing the implementation of RPE programmes in the workplace. Health and Safety Executive; 2010. Available at: <http://www.hse.gov.uk/research/rrpdf/rr798.pdf>
16. Health and Safety Executive. RR1052 The effect of wearer stubble on the protection given by Filtering Facepieces Class 3 (FFP3) and Half Masks. 2015. Available at: <https://www.hse.gov.uk/research/rrhtm/rr1052.htm>
17. American National Standards Institute. ANSI/AIHA Z88.10-2010 Respirator Fit Testing Methods. Fairfax: American Industrial Hygiene Association; 2010.
18. Mennella JA., Spector AC., Reed DR., Coldwell SE. The bad taste of medicines: overview of basic research on bitter taste. Clinical Therapeutics. August 2013; 35(8): 1225–1246. Available at: DOI:10.1016/j.clinthera.2013.06.007
19. Long K., Mai N., Williams M. Human errors in qualitative fit testing: a literature review [Unpublished paper]. 2019.

20. Nebuliser check list. QLTk1 - qualitative fit test kit supplement. Annan: Alpha Solway Ltd; 2020.

5 Human errors in qualitative fit testing: a practical study of the impact of human errors on fit test outcome

5.1 Abstract

This study assesses the practical impact of errors made in the use and application of qualitative fit test equipment on fit test outcome. Practical impact is studied through the effect of errors on in-hood aerosol concentration and shortened qualitative fit tests in the presence of a known leak. Results indicate that many of the errors involving incorrect use and application fit test nebulisers, solutions and hood have the potential to have a practical impact on fit test outcome, possibly leading to false pass results. This study indicates that the error scenarios *Fit test nebuliser filled with sensitivity solution*, *Fit test nebuliser clogged* and *Fit test nebuliser not upright and under-filled* have the highest impact on outcome and must be eliminated if this fit test method is to be more reliable. Findings from an earlier study are combined to provide an estimate of the workplace Beta error of the QLFT method of 0.43. The value of the widely used ‘reveal test’ is questioned in the light of this study which appears to show that, rather than mitigating errors, it could contribute to false pass outcomes.

5.2 Background

The rapid increase in the use of class 3 filtering facepiece (FFP3) respirators by healthcare workers during the recent SARS-CoV-2 pandemic caused a surge in demand for fit testers and fit test equipment that far outstripped supply in every part of the nation. Despite national and global supply issues, the Health and Safety Executive (HSE) maintained the requirement for fit testing of all tight-fitting respiratory protective device (RPD) facepieces including filtering facepieces, reusable half masks and full-face masks [1]. Pressures on resources raised concerns about a lack of competency and poor practice within the industry, with anecdotal evidence from the fit testing community about the emergence of ‘pop-up’ fit testers with little or no experience or expertise. If a tight-fitting facepiece is to afford the correct protection to the wearer in the workplace, it is vital that it is demonstrated to be capable of providing an adequate facial seal to the wearer. Given the importance of properly fitted RPD, it would therefore be reasonable for wearers to expect that a fit test should be carried out correctly, and that the outcome of the fit test can be trusted.

HSE guidance document INDG479 provides detailed guidance for fit testers, including the test protocol and pass / fail criteria [2]. Unlike quantitative fit tests (QNFT) which provide a numerical fit factor result, qualitative fit tests (QLFT) offer only a simple pass or fail outcome. An incorrect QLFT outcome can therefore fall into two categories: false negatives (false passes), and false positives (false fails). The most serious of these errors are false passes which occur when a wearer with a poor fitting facepiece passes a fit test. The consequences of false pass results are potentially health critical as they may allow an unsuitable facepiece to be used in the workplace. Eradication or reduction of errors which may lead to false passes in QLFT is therefore of the utmost importance.

A study which used assessment papers collected by the British Safety Industry Federation (BSIF) Fit Test Providers ‘Fit2Fit’ Accreditation Scheme (Fit2Fit scheme), provided the first observation-based record of the human errors in QLFT, and indicated widespread poor application of the method [3]. A further study into the theoretical impacts of human errors, identified eighteen different ‘critical errors’ that were judged to have the potential to impact one or more of the key method principles and hence potentially lead to a false pass [4]. When combined with results from the earlier observational study, the findings

suggested that over a third of the errors made by fit testers during QLFT Fit2Fit practical assessments for the period 2009 – 2019 may have been critical errors, potentially leading to false pass outcomes [4].

Whilst a theoretical study has indicated the potential for critical errors to impact upon fit test outcome, the practical impacts of these errors have not been studied to date. Empirical studies are therefore required to assess whether critical errors can lead to false pass results in practice and lead to an increase in the Beta error beyond that observed in laboratory studies. The theoretical study postulated that human errors could be considered as two groups:

1. Subjectivity errors (impact the wearer's ability to detect the challenge agent)
2. Equipment errors (impact the proper use or application of QLFT equipment)

The purpose of this study is to examine the practical impact of human errors which relate to the second of these two groups: equipment errors. Equipment errors are of particular interest when considering false passes in qualitative fit testing. Indications from earlier studies suggests that these types of human errors are most numerous and widespread throughout each part of the fit test procedure and are directly influenced by the skills of the fit tester, thus making them highly vulnerable to human error. Errors resulting from poor set-up or use of QLFT equipment theoretically can impact the generation and containment of a consistent and controlled concentration of challenge aerosol – factors that are critical in the ability of the QLFT method to detect a leak [5]. Since the potential therefore exists for these errors to affect fit test outcome, a proper understanding of their practical impact is crucial in ensuring workplace safety of this method.

This study investigates the practical impacts of equipment errors through the staging and testing of a 'control' and various 'error scenarios'. Aerosol mass concentration measurements are used to assess the impact of error scenarios on in-hood challenge aerosol concentration, and shortened fit tests are used to assess the impact on fit test outcome for wearers using facepieces with a known leak. Results from these two studies are synthesised to inform an overall discussion of the practical impact of equipment errors on QLFT outcome and an estimate of the workplace Beta error for the QLFT method is calculated.

5.3 Method

5.3.1 Set up of control and error scenarios

The practical impacts of the equipment errors theorised by an earlier study were assessed through the managed set up and testing of a series of ‘error scenarios’ against a ‘control scenario’ [4]. The control scenario was set up according to conditions detailed in the HSE INDG479 protocol [2]. The conditions used for each of the error scenarios were standardised to minimise variability and allow comparison with the control. All tests were conducted by a single Fit2Fit accredited fit tester in order to minimise operator-to-operator variation. Table 5-1 provides a list of the equipment errors identified by an earlier study along with the conditions used for the error and control scenarios. A diagram of a BS ISO 16975-3 nebuliser is shown in Appendix F to provide common nomenclature.

Equipment error [4]	Control scenario conditions	Error scenario conditions
Fit test nebuliser filled with sensitivity solution	Fit test nebuliser filled with fit test solution	Fit test nebuliser filled with sensitivity solution
Fit test nebuliser under-filled	Fill volume 5ml	Fill volume 2.5ml
Fit test nebuliser clogged	No clogging	Partial blockage of Venturi
Fit test nebuliser bottom bung not removed	Bottom bung removed	Bottom bung not removed
Fit test nebuliser Venturi mis-seated	Venturi fully seated	Venturi seated 3mm high
Fit test solution under-administered	5 squeezes every 30 seconds	2, then 3 squeezes every 30 seconds
Fit test timing incorrect	5 squeezes every 30 seconds	5 squeezes every 45 seconds
Fit test nebuliser not upright	Nebuliser held upright	Nebuliser held at 45°
Fit test nebuliser not actuated with full squeezes	Squeezes applied using whole hand	Squeezes applied using only thumb and forefinger
Hood damaged / unsuitable	Hood as per specification in BS ISO 16975-3	Hood volume 30% larger than BS ISO 16975-3 specification
Hood positioned incorrectly	Application hole positioned centre	Application hole positioned by right ear

Table 5-1 Conditions for control and error scenarios for assessment of practical impact of equipment errors in QLFT

Selection of the conditions for set up of the various error scenarios was made in two ways. Firstly, where binary choices of control / error scenarios were available, these were used, e.g., *Fit test nebuliser bottom bung not removed*, was the direct opposite of the control scenario which required removal of the bottom bung in line with HSE INDG479. Secondly, where binary choices were not available, selection of error scenario conditions was made with reference to manufacturer's instructions, HSE guidance and observations recorded by Fit2Fit assessors. The error scenario *Fit test nebuliser under-filled* was set up using 2.5ml of fit test solution instead of the control volume of 5ml (measured using a pipette). A 2.5ml volume was selected since numerous manufacturers' instructions refer to use of a single 'ampoule' of commercially available fit test solution, supplied in 2.5ml quantities. The error scenario *Fit test nebuliser not upright* was set up using a clinometer to position the nebuliser at 45° to the vertical, instead of the control position of upright. This scenario was tested using both the control volume of 5ml and the reduced 'error' volume of 2.5ml. This additional test, *Fit test nebuliser not upright and under-filled*, was included to determine the compound effect of errors made in both of these areas due to concerns highlighted by HSE guidance during the SARS-CoV-2 pandemic [6].



Figure 5-1 Set up of control and error scenarios for nebuliser function tests

Notes: (Right) control scenario: bottom bung removed, (Centre) error scenario: bottom bung not removed, (Left top): control scenario: Venturi correctly seated, (Left bottom): error scenario Venturi seated 3mm high.

One manufacturer's supplement recently provided warning information about the need to properly seat the Venturi inside the nebuliser [7]. The error scenario *Fit test nebuliser Venturi mis-seated* was therefore set up by assembling the Venturi at 3mm above the nebuliser base, facing towards the nebuliser outlet, rather than the control scenario which was fully seated (Figure 5-1). HSE Covid guidance for fit testers requires the nebulisers to be fully cleaned by submersion in disinfecting solution after every fit test, and therefore it was not possible to create 'clogging' of the nebuliser through crystallisation which could be used repeatedly. The error scenario *Fit test nebuliser clogged* was instead simulated by introducing a partial blockage using a water-resistant silicone sealant.

The control scenario was conducted in accordance with the timings and applications of the HSE INDG479 protocol which requires 5 squeezes every 30 seconds for the lowest sensitivity banding. Conditions for the error scenario *Fit test timing incorrect* increased this time interval to 45 seconds, and the error scenario *Fit test solution under-administered* decreased the number of squeezes by a half to an alternating 2 / 3 squeezes (Appendix J). *Fit test nebuliser not actuated with full squeezes* was studied using partial squeezes of the nebuliser using just the thumb and forefinger, instead of the full hand. A study by the Health and Safety Laboratory suggested that some test hoods were larger than that specified in BS ISO 16975-3 and raised concerns that this may lead to unreliable fit test results [8]. In addition, some test hood models are supplied without a collar and require 'shaping' by the fit tester prior to use. For this study, a range of test hoods was acquired and the 'shaped' and 'unshaped' (where applicable) volume of each was measured using external dimensions. The largest was found to have an unshaped volume that was 130% larger than the standard hood (a figure in line with that of previous research) and was used for the error scenario *Hood damaged / unsuitable*. Finally, the error scenario *Hood positioned incorrectly* was set up with the collar turned against the test hood such that the hole through which the nebuliser should be applied sat near the ear position, rather than centrally.

Alpha Solway 2530V FFP3 filtering facepiece respirators were used for all studies, and a new facepiece was checked and used for each test. QLFT equipment comprised a new 3M FT30 Bitrex (denatonium benzoate) qualitative fit test kit which was used for all control and error scenarios [9]. The kit included sensitivity solution (13.5 mg Bitrex in

100 ml of a 5 % sodium chloride solution by weight in distilled water), and fit test solution (337.5 mg of Bitrex in 200 ml of a 5 % sodium chloride solution by weight in distilled water) [10]. The frequency and depth of actuation of the nebuliser was standardised between tests, and application was always directed towards the centre of the facepiece to avoid spraying directly in the path of any sample lines. A digital stopwatch was used to monitor test timings and all equipment was stored in accordance with manufacturer's instructions.

5.3.2 Aerosol mass concentration studies

A series of aerosol mass concentration tests were conducted to provide data about the impact of each of the error scenarios on the in-hood challenge aerosol concentration versus the control scenario. These tests were conducted using a mannequin connected to a pump working at 10 litres per minute. The pump was used instead of a human wearer to increase the reproducibility of the filtering effects of inhalation. The mannequin was fitted with a T-shirt to prevent slipping of the test hood, and the temperature and humidity of the test environment were controlled through air conditioning. A new Alpha Solway 2530V respirator was fitted to the mannequin, and a sample line was fixed adjacent outside of the right lateral nose bridge to simulate the common leak position for half mask respirators [11][12]. The 3M qualitative fit test hood was fitted over the mannequin to contain the aerosol, taking care not to disturb the sample line or respirator.

Aerosol mass concentration measurements were conducted using a DustTrak DRX model 8534 manufactured by TSI Incorporated, Shoreview, MN [13]. The DustTrak DRX is an optical aerosol monitor capable of estimating size-segregated mass fraction across channels corresponding to size PM₁, PM_{2.5}, PM₄ and PM₁₀ over a concentration range 0.001–150 mg/m³ [14][15]. PM_{2.5} data was selected for use in this study since it is most representative of the geometric mean particle size of the Bitrex fit test and sensitivity aerosol, determined by Mullins et al. to be 2.38µm, σ 1.45 (standardised using 5% NaCl in distilled water) [16]. Instrument-to-instrument variation was eliminated by use of the same DustTrak for each test, and a zero-calibration and background measurement was

conducted prior to each testing session. The DustTrak was factory-calibrated to ISO 12103-1, A1 test dust (formerly known as ‘Arizona test dust’) and an additional size correction factor was calculated to improve the relative accuracy against the fit test aerosol [13]. The DustTrak was programmed to take a sample measurement every one second for a period of seven minutes (the normal duration of a qualitative fit test) and report a mean aerosol mass concentration result.

A series of seven-minute simulated qualitative fit tests were conducted for the mannequin configuration with either the control scenario or one of the error scenarios from Table 5-1 introduced. Each of the scenarios were set up and tested twenty times in randomised order over a period of two months. All apparatus was cleaned, dried and re-set between each test and equipment was inspected for condition after each use. Figure 5-2 provides a flowchart of the procedure used for the mass aerosol concentration studies.

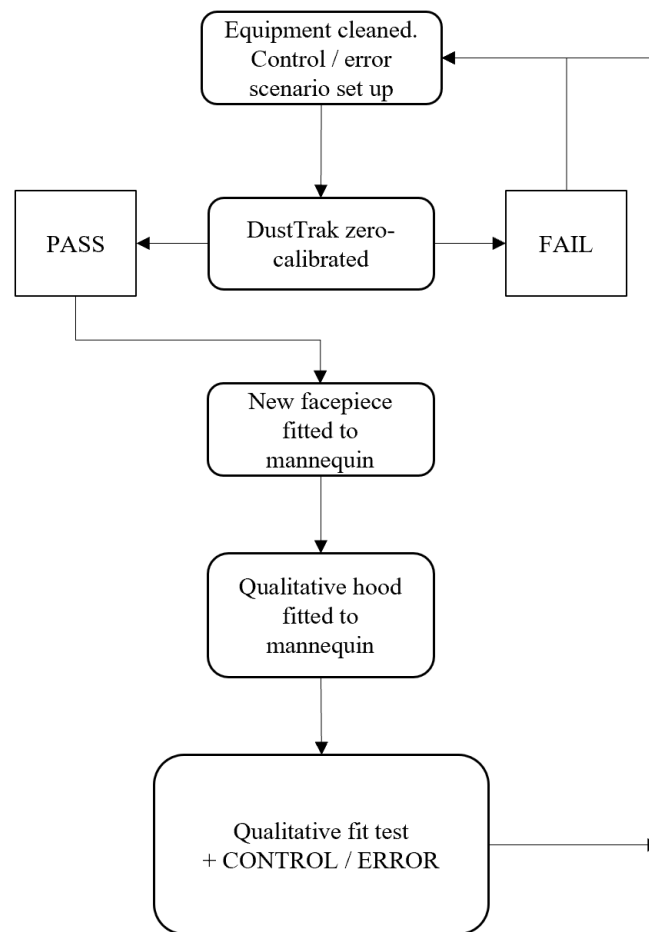


Figure 5-2 Procedure for aerosol mass concentration studies

Data output from the TSI DustTrak was handled using TSI TrakPro software version 4.7.2 and exported into a spreadsheet created in Microsoft Excel (Microsoft Office 365) [17]. Symmetry of the data was investigated using a skewness test and appropriate descriptive statistics were used to examine the data using IBM SPSS software (version 26). A two-tailed t-test was used to compare results from the two studies using a significant level of 0.05.

5.3.3 Shortened fit test studies

A series of shortened qualitative fit tests was conducted to evaluate the impact of the error scenarios on fit test outcome for wearers using Alpha Solway 2530V FFP3 respirators that were adapted to include a known leak of a size 1 - 2%.

Respirators were fitted with two pluggable ports: a sample port; and an adjustable leak port, these are shown in Figure 5-3.

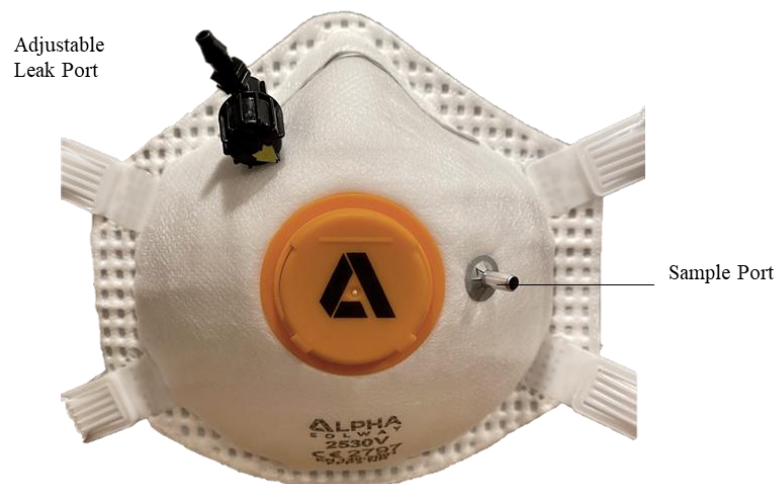


Figure 5-3 Location of the sample port and adjustable leak port

The sample port was created using an unmodified fit test sample probe (model 8025-N95, TSI Incorporated, Shoreview, MN) and positioned between the nose / mouth area as close as possible to the wearer's breathing zone in accordance with the method detailed in INDG479 [2]. The adjustable leak port was created using a fit test sample probe (model 8025-N95, TSI Incorporated, Shoreview, MN) modified with an in-line adjustable flow valve. This design was subjected to multiple tests so that a leak could be obtained when

fitted of 1 – 2% (equivalent to a quantitative fit factor of 50 – 100). The leak port was positioned at the bridge of the nose on the opposite side to the sample port to prevent streaming of sample air.

Quantitative fit factors (QNFF) were measured using a TSI Portacount Pro+ Model 8038 Respirator Fit Tester manufactured by TSI Incorporated, Shoreview, MN [18]. The same Portacount was used for each test to eliminate instrument-to-instrument variation. The Portacount was factory calibrated in accordance with manufacturer's instructions and a 'Daily Check' was conducted prior to each testing session to ensure that the instrument was functioning correctly. The ambient particle concentration was maintained at a stable level above 3,000 particles / cm³ throughout each test as recommended by INDG479.

Wearers were recruited from male and female staff, associates, family members and friends. No attempt was made to select wearers according to anthropometric data representative of the UK respirator-wearer population, instead recruitment was made on the basis of available volunteers. Wearers were provided with a written procedure and asked to give informed consent for inclusion in this study (Appendix G). In order to reduce the effects of wearer-generated particles on quantitative fit factor measurement and to ensure a clean palate prior to qualitative fit tests, wearers were instructed to refrain from eating, drinking (except plain water), smoking or vaping for 60 minutes prior to the test.

Each wearer undertook a sensitivity test to determine their taste threshold to Bitrex, plus a water-placebo to check their response. Wearers that incorrectly characterised the taste of the placebo were instructed to wait a further thirty minutes and the process was repeated. Wearers that were unable to detect Bitrex during the sensitivity test were eliminated from the study. If a correct sensitivity response to Bitrex was given, wearers were instructed to wipe their face, cleanse their palate by drinking water and wait at least 15 minutes (or longer until their palate had cleared) before proceeding to the next stage. Wearers were provided with a mirror and instructed to don the respirator and check for fit in accordance with manufacturer's instructions. A five-minute comfort assessment period was used to allow the wearer time to adjust the respirator for comfort and fit, after which time no adjustment was permitted.

The Portacount sample tube was connected to the sample port and supported using a lanyard to prevent disturbance of fit on the wearer's face. Using the Portacount 'Real-Time' mode, the leak port was adjusted to create a leak of 1 - 2%. A QNFF was then measured as the average of two 80-second normal breathing exercises to quantify leak size. Causing as little disturbance as possible, and without removing or adjusting the respirator, the Portacount sample tube was disconnected and plugged, leaving the leak port open.

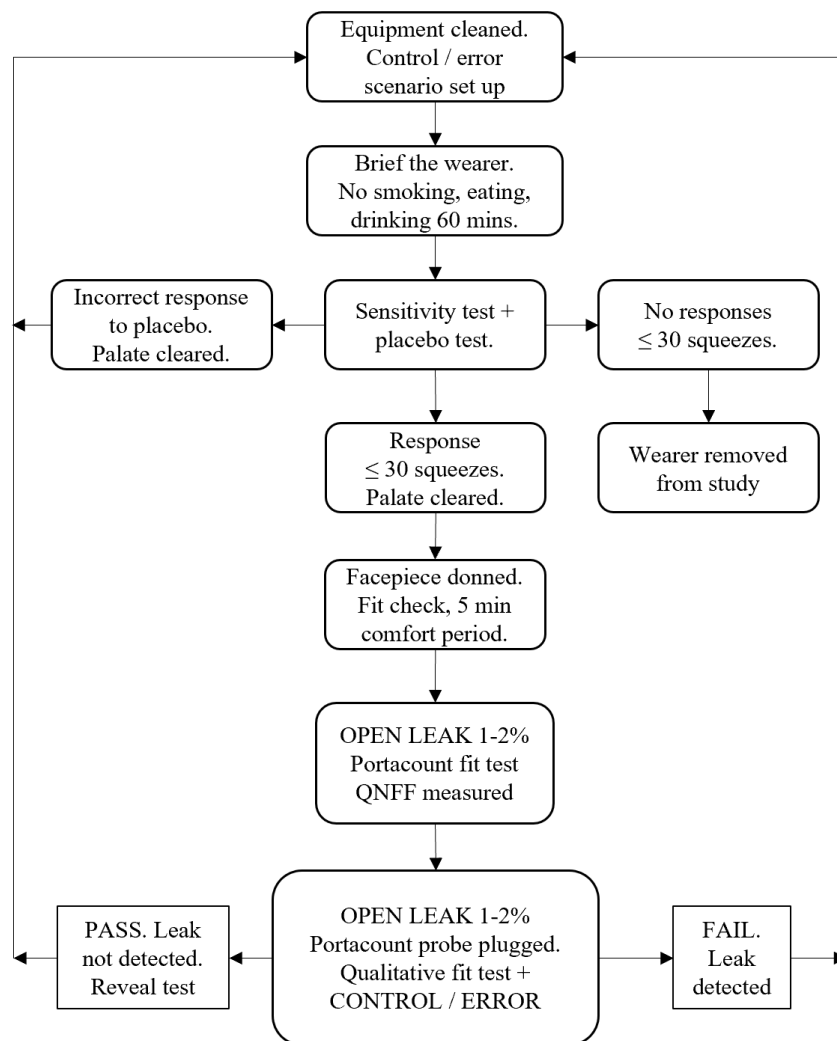


Figure 5-4 Process for shortened fit tests

Wearers were then placed inside the QLFT hood, and a two-minute shortened qualitative fit test, incorporating one of the ‘error scenarios’ from Table 5-1 or the ‘control scenario’ was conducted. Wearers were instructed to breathe normally through their mouth with no head movements or exercises, and to indicate any detection of taste immediately by raising their hand. If taste was detected, wearers were instructed to remove the respirator, cleanse their palate by drinking water, wipe their face and wait until their palate had cleared before proceeding to a new test. In line with guidance issued by the BSIF Fit2Fit Scheme, if no taste was detected, wearers were instructed to carry out a reveal test, by reaching inside the hood, breaking the seal between the mask and their face and breathe in through their mouth. Results of the reveal test were recorded alongside the outcome of the test. The procedure for the shortened fit tests study is shown in Figure 5-4.

Each error scenario, plus the control scenario, was set up and tested in randomised order over a period of three months for eleven different wearers. Data was handled using TSI FitPro+ software (version 3.3.0) and exported into a spreadsheet created in Microsoft Excel (Microsoft Office 365) [19]. Symmetry of the data was investigated using a skewness test and statistical analysis was performed using IBM SPSS software (version 26). Data from the aerosol mass concentration and shortened fit test studies were compared using a Spearman correlation test to investigate the degree of association between the data sets.

5.4 Results

5.4.1 Aerosol mass concentration studies

A total of twelve error scenarios plus one control scenario were evaluated using mass concentration measurements to provide an indication of in-hood aerosol concentration. Full descriptive statistics for the aerosol mass concentration studies are reported in Appendix H. Analysis of the individual error scenarios compared with the control scenario reveals the differences in the mean values, standard deviation, and the statistical significance ($\alpha = 0.05$), these are displayed in Table 5-2. The mean mass aerosol concentration for the control scenario was found to be 11.17mg/m³ ($SD = 2.43$) and this was similar to the median value observed ($Mdn = 10.75$ mg/m³). The similarity of the mean and median values for the control scenario indicates an approximately normal distribution which was confirmed by a skewness test ($skew = 1.39$). Descriptive statistics for each of the error scenarios also show little difference between the mean and median values, and the skewness values reported in Appendix H suggest an approximately normal distribution for all data.

Error scenario	Mean mass concentration ^a <i>Error scenario</i>	Mean mass concentration ^b <i>Control scenario</i>	Difference between means		
	$M_1 (SD_1)$	$M_2 (SD_2)$	$M_2 - M_1$	$t(19)$	p
Fit test nebuliser filled with sensitivity solution	11.02 (1.57)	11.17 (2.43)	0.15	0.24	0.810
Fit test nebuliser under-filled	10.94 (1.69)	11.17 (2.43)	0.23	0.34	0.736
Fit test nebuliser clogged	2.12 (0.54)	11.17 (2.43)	9.05	16.28	<0.001
Fit test nebuliser bottom bung not removed	4.07 (0.90)	11.17 (2.43)	7.10	12.87	<0.001
Fit test nebuliser Venturi mis-seated	4.60 (0.97)	11.17 (2.43)	6.57	11.97	<0.001
Fit test solution under-administered	7.56 (1.28)	11.17 (2.43)	3.61	5.80	<0.001
Fit test timing incorrect	7.10 (0.89)	11.17 (2.43)	4.07	7.12	<0.001
Fit test nebuliser not upright	10.01 (1.30)	11.17 (2.43)	1.16	2.02	0.058
Fit test nebuliser not upright and under-filled	3.16 (0.92)	11.17 (2.43)	8.01	13.25	<0.001
Fit test nebuliser not actuated with full squeezes	3.95 (1.03)	11.17 (2.43)	7.22	12.46	<0.001
Hood damaged / unsuitable	3.34 (0.88)	11.17 (2.43)	7.83	12.79	<0.001
Hood positioned incorrectly	5.30 (0.75)	11.17 (2.43)	5.87	10.11	<0.001

Table 5-2 Results of aerosol mass concentration tests for control and error scenarios

Notes: Mass concentration measured in mg/m³. ^a $n = 20$. ^b $n = 20$. $\alpha = 0.05$.

The error scenario *Fit test nebuliser clogged* was found to have the greatest impact on aerosol mass concentration when compared to the control scenario. The observed difference between means was 9.05 ($t(19) = 16.28, p < 0.001$), indicating an in-hood aerosol mass concentration for the error scenario that was 81% lower than the control scenario. Indeed, the mean values for nine of the twelve error scenarios were statistically significantly different from the control scenario, with only *Fit test nebuliser filled with sensitivity solution*, *Fit test nebuliser under-filled* and *Fit test nebuliser not upright* found not to be significantly different. The smallest difference in aerosol mass concentration against the control was found for *Fit test nebuliser filled with sensitivity solution*, where the observed difference in means was 0.15 ($t(19) = 0.24, p = 0.81$). This result is unsurprising since the aerosol produced by the sensitivity and fit test solutions is standardised in commercially available solutions with sodium chloride. Of greater interest will be the effect of using sensitivity solution on the ability of wearers to detect a known known leak during shortened fit test studies which is discussed later in this paper.

The results for error scenarios *Fit test nebuliser under-filled* and *Fit test nebuliser not upright* are of further interest. When these scenarios were tested separately, they were found to produce a limited reduction in aerosol concentration from the control (2.07% and 10.39% respectively). However, when these scenarios were tested concurrently, *Fit test nebuliser not upright and under-filled*, a decrease of 71.1% was found against the control, with an observed difference between means of 8.01 ($t(19) = 13.25, p < 0.001$).

5.4.2 Shortened fit test studies

Of the twelve volunteer wearers that were recruited in total, one was unable to detect Bitrex during sensitivity testing and was removed from the study. All eleven remaining volunteer wearers elicited a taste response to Bitrex during sensitivity testing in the lowest threshold bracket (1-10 applications of the sensitivity nebuliser), and all elicited the correct (null) response to the water placebo. Quantitative fit factors (QNFF) for the control scenario in the presence of the known leak ranged from 54 to 81 with a mean of 70.4 ($SD = 9.69$), equivalent to a leak of 1.42%. All eleven wearers reported a taste of Bitrex (Fail result) during QLFT under the control scenario indicating that the known leak

was correctly detected. All twelve error scenarios plus one control scenario were tested using the eleven volunteer wearers in a total of 143 shortened fit test studies.

A summary of the mean QNFF for error scenarios along with the number of wearers reporting taste during QLFT is shown in Table 5-3 and full results of this study are shown in Appendix I.

Error scenario	QNFF <i>Error scenario^a</i>		QNFF <i>Control scenario^b</i>		Difference between means		
	<i>M₁ (SD₁)</i>	QLFT	<i>M₂ (SD₂)</i>	QLFT	<i>M₂ - M₁</i>	<i>t(19)</i>	<i>p</i>
Fit test nebuliser filled with sensitivity solution	70.3 (14.14)	0	70.4 (9.69)	11	0.1	0.03	0.980
Fit test nebuliser under-filled	74.4 (14.88)	11	70.4 (9.69)	11	-4.0	-0.66	0.526
Fit test nebuliser clogged	69.9 (7.27)	0	70.4 (9.69)	11	0.5	0.14	0.896
Fit test nebuliser bottom bung not removed	68.8 (12.69)	6	70.4 (9.69)	11	1.6	0.28	0.782
Fit test nebuliser Venturi mis-seated	75.4 (9.95)	3	70.4 (9.69)	11	-5.0	-1.13	0.286
Fit test solution under-administered	70.8 (11.34)	6	70.4 (9.69)	11	-0.4	-0.09	0.933
Fit test timing incorrect	73.7 (8.57)	9	70.4 (9.69)	11	-3.3	-0.90	0.391
Fit test nebuliser not upright	70.6 (13.35)	11	70.4 (9.69)	11	-0.2	-0.05	0.961
Fit test nebuliser not upright and under-filled	73.7 (8.80)	2	70.4 (9.69)	11	-3.3	-1.26	0.236
Fit test nebuliser not actuated with full squeezes	74.0 (8.80)	5	70.4 (9.69)	11	-3.6	-0.80	0.444
Hood damaged / unsuitable	72.4 (12.79)	8	70.4 (9.69)	11	-2.0	-0.33	0.749
Hood positioned incorrectly	75.5 (11.69)	7	70.4 (9.69)	11	-5.1	-1.09	0.303

Table 5-3 Summary of results for quantitative and qualitative fit tests for control and error scenarios

Notes: QNFF *M* = mean quantitative fit factor measured using a TSI Portacount model 8038 respirator fit tester. QLFT = number of qualitative fit tests that resulted in a FAIL outcome (the wearer was able to detect the challenge agent via the open leak).^a*n* = 11. ^b*n* = 11. α = 0.05.

Mean QNFF values for the twelve error scenarios ranged from 68.8 (*SD* = 12.69) to 75.5 (*SD* = 11.69), equivalent to a leak of between 1.45% and 1.32% which should have been

detectable by QLFT. Comparison of the mean QNFF values for the error scenarios against the control scenario showed no significant difference across the study with p values ranging from 0.286 – 0.980 ($\alpha = 0.05$). The similarity of the mean QNFF values for the control and error scenarios therefore allows the data for the number of wearers reporting taste detection to be compared side by side.

The number of wearers reporting detection of taste during QLFT ranged from zero (none of the wearers were able to correctly detect the open leak) to 11 (all wearers were able to correctly detect the open leak). A mean value of 5.7 for all error scenarios indicates that the leaking facepiece was identified by just over half (51.5%) of wearers in this study. This in turn means that just under a half (48.5%) of wearers were not able to detect the leaking facepiece when an error was introduced. The greatest impact on taste detection was found for the error scenarios *Fit test nebuliser filled with sensitivity solution* and *Fit test nebuliser clogged*. Data for these scenarios indicates that none of the eleven wearers reported taste detection during QLFT, despite the presence of a known leak of sizes 1.42% and 1.43% respectively. The result for *Fit test nebuliser clogged* is in line with the results reported by the mass concentration study which found this error to have the greatest impact on in-hood aerosol concentration.

The error scenarios *Fit test nebuliser under-filled* and *Fit test nebuliser not upright* appeared to have no effect on QLFT outcome in this study, with all wearers correctly indicating detection of taste through the open leak. However, when these errors were compounded in the scenario *Fit test nebuliser not upright and under-filled*, nine of the eleven wearers failed to detect the open leak suggesting that these errors become pivotal if made simultaneously. This result affirms the data from the mass concentration study which also found that these two errors were significant if compounded. Some effect on taste detection was noted for the error scenario *Fit test timing incorrect* with nine of the eleven wearers correctly detecting the open leak, but a much larger effect was noted for *Fit test solution under-administered* where only six of the eleven wearers correctly reported taste. Of further note are the results for the error scenario *Fit test nebuliser Venturi mis-seated* in which nine of the eleven wearers did not report detection of taste through the open leak during QLFT, a result which agrees with the mass aerosol concentration measurements.

In order to gain an overall view of the data, the results of the two studies were compared using the Spearman correlation coefficient. For each error scenario, the mean mass concentration reported by the aerosol concentration study was compared with the number of wearers not reporting taste during the shortened fit test study. When data for all error scenarios were included, the association between the results of the two studies was found not to be statistically significant ($r_s = 0.40774$, p (2-tailed) = 0.18827). However, when the contradictory data for *Fit test nebuliser filled with sensitivity solution* were removed, the correlation was found to be statistically significant ($r_s = 0.78083$, p (2-tailed) = 0.00456) indicating good agreement between the results of the two studies.

The validated QLFT protocol is designed to identify leaks at $> 1\%$ (equivalent to a QNFF < 100) and therefore this study can be used to assess whether leaks at or above this value can be detected when a known error is introduced during shortened qualitative fit tests. Results of the aerosol mass concentration studies (Table 5-2) and shortened fit test studies (Table 5-3) were combined to provide a holistic view of impact. The data suggest that all error scenarios, except the non-compounded errors of *Fit test nebuliser under-filled* and *Fit test nebuliser not upright*, were found to have an impact on the reliability of the fit test outcome. The data also indicate a list of errors that are most likely to lead to an unreliable fit test outcome, these are shown in Table 5-4.

Error scenario	% Change in aerosol mass concentration ^a	Number of wearers correctly reporting detection of Bitrex ^b
Fit test nebuliser filled with sensitivity solution	1.36*	0
Fit test nebuliser clogged	81.02	0
Fit test nebuliser Venturi mis-seated	58.80	3
Fit test nebuliser not upright and under-filled	71.70	2
Fit test nebuliser not actuated with full squeezes	64.68	5

Table 5-4 Error scenarios with the highest impact on the ability of wearers to detect a known leak in QLFT

Notes: * sodium chloride used in commercial solutions to maintain a stable geometric mean particle size between sensitivity and fit test solutions. ^a $n = 20$. ^b $n = 11$.

These studies suggest that the error scenarios with the greatest impact on both the aerosol mass concentration and / or ability of wearers to detect a known leak outcome were *Fit test nebuliser filled with sensitivity solution*, *Fit test nebuliser clogged* and *Fit test nebuliser not upright and under-filled*.

When data from this study were applied to the observations of real-world fit testers from an earlier study an estimate of the workplace impact can be calculated. The results suggest that 51.2% ($n = 327$) of the records include errors against one or more of the critical error areas relating to the use of QLFT equipment. The shortened fit test study has indicated that, on average, these equipment errors may result in 48.5% of wearers being unable to detect the presence of a leak. When combined, these data therefore suggest that a 24.8% of fit tests observed in the earlier study may have resulted in a false pass result, indicating a workplace Beta value (probability of a false pass) for these equipment errors of 0.25. Although useful, this method of estimation is fairly crude and takes no account of the impact and probability of each type of error. A better estimate can be achieved by multiplying the probability of a particular error being made by the proportion of false passes observed for that error and summing these over the eleven individual error types. This calculation indicates a more reliable workplace Beta value (probability of a false pass) for equipment errors of 0.43, a figure significantly higher than that observed in laboratory studies. Since errors are rarely made in isolation however, this method of calculation may result in both an over-estimation (since errors are likely to be made in combination), and an under-estimation (since the impacts of combined errors may be worse than the sum of their individual parts).

5.5 Discussion

The results of this empirical study into human errors affirm the observations of an earlier theoretical study which suggested that some human errors in QLFT (termed ‘critical errors’) have the potential to adversely affect the outcome of the test. QLFT procedures are low-tech methods which employ rudimentary equipment in the form of a set of nebulisers, a test hood and two solutions. This basic appearance leads many fit testers to assume that little or no operator experience is required, and these methods are often advertised as ‘simple’, with one commentator suggesting QLFT methods are “*convenient and easy to perform*” [20]. This outward appearance belies the truth that QLFT methods require extensive input from the fit tester and rely upon manual operation, often with little or no training. Observation of real-world fit testers highlights a serious lack of knowledge of how nebulisers work, the function of components and, importantly how they might fail.

Fit testers have been forced to become better acquainted with their fit testing equipment since the start of the SARS-CoV-2 pandemic. HSE guidance has recently advised fit testers to dismantle nebulisers after each fit test, disinfect the bulb section and submerge the plastic components in disinfecting solution [6]. This guidance of course also requires fit testers to re-assemble the components correctly - a skill which is currently not observed or assessed during Fit2Fit assessments and is not discussed in many examples of manufacturer’s guidance. One possible outcome of this process is the potential for fit testers to re-assemble the critical Venturi poorly by failing to seat the component securely. A recent communication from one of the major QLFT equipment manufacturer cautioned fit testers to check that the Venturi is properly seated inside the nebuliser to ensure proper aerosolisation of the fit test solution [7]. The effect of a poorly seated Venturi is apparent in the compounded error *Fit test nebuliser not upright and under-filled* where the action of holding the nebuliser at a 45° angle causes the fit test solution to run to one side of the nebuliser such that the Venturi is no longer submerged. It appears that the standard fill volume of 5ml mitigates the effects of holding the nebuliser at an angle by ensuring that the Venturi remains submerged in the fit test solution and hence is able to maintain proper aerosol generation.

The results for the error scenarios *Fit test solution under-administered* and *Fit test timing incorrect* are of interest when considering in-hood aerosol concentration. Use of the control scenario conducted as per the HSE INDG479 protocol results in 75 full squeezes of the nebuliser (applications) over the total test time of 7 minutes (Appendix J). The error scenario *Fit test solution under-administered* used half the control number of applications resulting in 42 full squeezes over 7 minutes, and *Fit test timing incorrect* used an increased timing between applications resulting in 55 full squeezes over 7 minutes (Appendix J). Whilst both scenarios appear to produce a dramatic reduction in the mean aerosol concentration, the impact of increasing the test timings on fit test outcome appears to be less severe, possibly as a result of the higher number of applications over the total test time. A further reason may involve the manner in which the challenge agent is administered, whereby in-hood concentration is maintained by applying the challenge agent in batches rather than a continuous series of individual applications. The standard number of applications for a wearer with the lowest sensitivity banding would be batches of five squeezes administered every 30 seconds. In the scenario *Fit test timing incorrect*, the size of the batch (five squeezes) was maintained, but with a longer interval between batches (45 seconds). An open leak of $\geq 1\%$ may therefore have been more detectable when the batch size was maintained, possibly due to a threshold of taste detection.

One question which these studies were not designed to answer is whether errors which are carried through both the sensitivity test and fit testing procedure are accounted for by an increase in sensitivity banding. An example of this would be the mistake (perhaps through poor training) of leaving the bottom bung in the nebulisers used for both the sensitivity test and fit test. The reduction in fit test aerosol concentration observed during this study might be expected to be mirrored by a reduction in sensitivity aerosol concentration if the bottom bung was also left in place. This may lead to a sensitivity test threshold banding which is higher than would be observed if the bung had been removed. This would have the subsequent effect of increasing the number of applications used during the fit test which may not affect the fit test outcome. Further research would be needed to answer this question and assess the impact of carry-through errors.

A supplementary observation of the shortened fit test studies was the ability of wearers to detect taste during the 'reveal test' which is part of the recommendation in Fit2Fit

guidance and manufacturer's user instructions when no taste is detected during the fit test. Despite not being able to detect a taste during the fit test, almost all wearers reported detection of taste at the reveal test for error scenarios including *Fit test nebuliser filled with sensitivity solution* and *Fit test nebuliser not actuated with full squeezes* (denoted with * in Appendix I). This potentially provides false assurance to the fit tester and the wearer that the test had been applied correctly and the outcome was accurate, when results of this study indicate otherwise. Although not the main purpose of this study, the observations made during shortened fit tests provide an interesting insight into the value of the 'reveal test' in offering validation of fit testing outcomes. Fit2Fit Companion guidance advises fit testers that "*The "reveal" is an important element of fit testing with the qualitative method and its utilisation will help mitigate against false passes.*" The observations made during the shortened fit test studies appear to question this guidance and merit further discussion. BS ISO 16975-3 recommends inclusion of the 'reveal test' as part of the standard QLFT procedure and provides instructions on the fit tester's actions following this test,

"If the fit-test aerosol taste is not detected after breaking the RPD seal, the fit test is null and void and the reason why the person did not taste the fit-test aerosol shall be identified. If the fit-test aerosol taste is detected after breaking the seal, the fit test is valid and the RI fits the person." [10].

Whilst the results of this study agree with the first of these actions in recommending the fit test to be declared "*null and void*" if taste is not identified, they do not agree with the second. Observations made during the shortened fit test studies suggest that, contrary to guidance, wearers can detect a taste during a reveal test despite a serious error in the application of the test. An explanation for this may be found in understanding the relative concentration of sensitivity and fit test solutions, and size of leak involved. The diluted solution used during sensitivity testing ensures that wearers are able detect a challenge agent equivalent to a 1% or greater leak of the more concentrated fit test solution used during the fit test. The action of breaking the mask seal during a reveal test would be likely to result in a transient 'leak' of far greater than 1% of the more concentrated fit test solution. This means that an error would only be picked up by the wearer at the reveal test if the in-hood aerosol concentration fell below the wearer's taste threshold so as to be

undetectable. It could therefore be very dangerous to assume that detection of taste implies a valid fit test, indeed use of the reveal test as an indicator of correct application is questionable and may contribute to reporting of false pass results.

This study has several limitations which should be discussed. Use of a continuous pump during the aerosol mass concentration studies may have overestimated the impact of an error on in-hood aerosol concentration during workplace fit testing. During normal use of a particulate-filtering respirator, the action of inhalation has the effect of reducing the particulate concentration within the immediate environment. In this study, the continuous pump was used to control inhalation filtration so that the mass aerosol concentration could be studied separately as an effect of the error scenarios introduced. In normal use however, the action of exhalation would also occur, and would have the effect of both mixing the atmosphere within the immediate environment and causing a slight increase in particulate concentration (as some particles are released from the filtering medium and valves). Future studies may consider the use of a simulated human breathing machine but will need to account for exhaled particles and mixing when evaluating aerosol mass concentration.

A further limitation of this study is the number of shortened fit tests that were conducted for each of the twelve error scenarios which was low ($n = 11$) thereby limiting the significance of the outcome. A larger population size would increase the statistical significance and give higher confidence in the results. Although the order of the tests was randomised and conducted over an extended period, there is the possibility that ‘taste fatigue’ may have had an effect on wearers with the more tests they conducted. The issue of taste fatigue is discussed by a number of researchers and is thought to stem from an evolutionary factor of toxin-avoidance [21]. The use of a water placebo during the shortened fit test studies was conducted in line with previous research in an attempt to assure the wearer’s taste response [22]. Although other researchers have reported a moderate placebo effect, with wearers incorrectly characterising the taste of plain water, this study found no such effect amongst the eleven wearers [23][24]. This could be a result of the low sample population and larger studies would be required to answer this question with higher statistical confidence.

Furthermore, this study was limited in scope to assessment of the impact of equipment errors and did not attempt to assess the impact of subjectivity factors. An earlier study suggested that critical errors in QLFT could be separated into two types: subjectivity errors (those with the potential to impact the wearer's ability to taste the challenge agent) and equipment errors (those that have the potential to impact the proper use and application of QLFT equipment) [4]. Numerous researchers, including those involved in the original validation of the method, have described how the QLFT method is highly dependent on the wearer's taste response, which has been reported to be both variable and unreliable [25][22][26]. The subjective nature of the QLFT method places critical emphasis on the management and response of the wearer to taste – factors which are largely outside the control of the fit tester. Practical errors in this category may include the fit tester not asking whether the wearer has a clear palate before the sensitivity and fit test, and not instructing the wearer to breathe through the mouth. Whilst a fit tester can attempt to manage these issues by asking the appropriate questions, their ability to control or affect the response is limited since they rely entirely upon the wearer's cooperation and honesty. Therefore, although these errors may indeed be critical to the outcome of a fit test and contribute to an increased workplace Beta error, they are not controllable by the fit tester alone and instead must involve the full cooperation and understanding of the wearer. A proper understanding of the influence of subjectivity errors on QLFT outcome is required through further study whose scope may include both examination of practical application as well as an understanding of the psychology and attitudes to taste response.

5.6 Conclusion

The results of this study indicate that errors involving the use and application of QLFT equipment have the potential to have a practical impact on both the in-hood aerosol concentration and the wearer's ability to detect a known leak. An estimated workplace Beta value of 0.43 for equipment errors indicates that the impacts of human operator errors are far greater in the real world than have been suggested by laboratory studies. The biggest impacts were found for *Fit test nebuliser filled with sensitivity solution*, *Fit test nebuliser clogged* and *Fit test nebuliser not upright and under-filled* and these errors must be eliminated if QLFT outcomes are to be more reliable.

5.7 References

1. Murray A. UK best practices on RPD fit testing (Fit2Fit programme). National Fit Testing Programmes. 18th May 2021. International Society for Respiratory Protection; Available at: <https://www.isrp.com/minutes>
2. Health and Safety Executive. INDG479 Guidance on respiratory protective equipment (RPE) fit testing. HSE Books; 2019. Available at: <https://www.hse.gov.uk/pubns/indg479.pdf>
3. Long K., Mai N., Williams M. Human errors in qualitative fit testing: a study of real-world fit testers [Unpublished paper]. 2020.
4. Long K., Mai N., Williams M. Human errors in qualitative fit testing: theoretical impact of errors on fit test outcome [Unpublished paper]. 2021.
5. Occupational Safety and Health Administration (OSHA). Minnesota Mining and Manufacturing Company post hearing data submission on qualitative fit testing protocols under the lead standard before OSHA Docket No. H- 049A 1981.
6. Health and Safety Executive. Fit testing face masks to avoid transmission during the coronavirus (COVID-19) pandemic. Available at: <https://www.hse.gov.uk/coronavirus/ppe-face-masks/face-mask-ppe-rpe.htm#minimisetherisk> (Accessed: 12 November 2020)
7. Nebuliser check list. QLTK1 - qualitative fit test kit supplement. Annan: Alpha Solway Ltd; 2020.
8. Health and Safety Executive., Bowry A. Effect of hood size on a qualitative Fit Test. 2018. Available at: <https://www.hsl.gov.uk/media/1270950/hood size powerpoint.pdf> (Accessed: 15 February 2021)
9. 3M. Qualitative fit test apparatus FT30 user instructions. Available at: <https://multimedia.3m.com/mws/media/1464265O/3m-qualitative-fit-test-apparatus-ft-30-user-instruction.pdf> (Accessed: 16 October 2021)

10. British Standards Institution. BS ISO 16975-3:2017 Respiratory protective devices. Selection, use and maintenance. Part 3: Fit-testing procedures. London: British Standards Institution; 2017.
11. Oestenstad RK., Bartolucci AA. Factors affecting the location and shape of face seal leak sites on half-mask respirators. *Journal of Occupational and Environmental Hygiene*. 2010; 7(6): 332–341. Available at: DOI:10.1080/15459621003729909
12. Oestenstad RK., Kenneth Dillion H., Perkins LL. Distribution of facesal leak sites on a half-mask respirator and their association with facial dimensions. *American Industrial Hygiene Association Journal*. 1990; 51(5): 285–290. Available at: DOI:10.1080/15298669091369664
13. TSI Incorporated. DustTrak DRX Aerosol Monitor operation and service manual. Available at: https://tsi.com/getmedia/3699890e-4adf-452f-9029-f3725612d5d1/8533-8534-DustTrak_DRX-6001898-Manual-US?ext=.pdf (Accessed: 15 September 2021)
14. TSI Incorporated. DustTrak DRX Aerosol Monitor theory of operation. Available at: https://tsi.com/getmedia/1efa3785-c52d-491e-9cd1-10f30a01c997/EXPMN-002_DustTrak_DRX_Theory_of_Operation?ext=.pdf (Accessed: 15 September 2021)
15. Wang X., Chancellor G., Evenstad J., Farnsworth JE., Hase A., Olson GM., et al. A novel optical instrument for estimating size segregated aerosol mass concentration in real time. *Aerosol Science and Technology*. 2009; 43(9): 939–950. Available at: DOI:10.1080/02786820903045141
16. Mullins H., Danisch S., Johnston A. Development of a new qualitative test for fit testing respirators. *American Industrial Hygiene Association Journal*. 1995; 56(11): 1068–1073. Available at: DOI:10.1080/15428119591016278
17. TSI Incorporated. TrakPro data analysis software. Version 4.7.2. Available at: <https://tsi.com/support/tsi-software-and-firmware/> (Accessed: 18 October 2020)

18. TSI Incorporated. Portacount Pro 8030 and Portacount Pro+ 8038 Respirator Fit Testers operation and service manual. P/N 6001868, revision P, March 2015. Available at: <https://tsi.com/products/respirator-fit-testers/portacount®-respirator-fit-tester-8030/> (Accessed: 4 June 2020)
19. TSI Incorporated. FitPro+ software for TSI Portacount Pro 8030 and Portacount Pro+ 8038 Respirator Fit Testers. Version 3.3.0. Available at: <https://tsi.com/support/tsi-software-and-firmware/> (Accessed: 18 October 2021)
20. Colton CE. Respirator fit testing: choosing the best method. *Metal Finishing*. 2001; 99(9): 53–55. Available at: DOI:10.1016/S0026-0576(01)81436-7
21. Mennella JA., Spector AC., Reed DR., Coldwell SE. The bad taste of medicines: overview of basic research on bitter taste. *Clinical Therapeutics*. August 2013; 35(8): 1225–1246. Available at: DOI:10.1016/j.clinthera.2013.06.007
22. McKay RT., Davies E. Capability of respirator wearers to detect aerosolized qualitative fit test agents (sweetener and Bitrex) with known fixed leaks. *Applied Occupational and Environmental Hygiene*. 2000; 15(6): 479–484. Available at: DOI:10.1080/104732200301269
23. Fakharpour A., Jahangiri M., Yousefinejad S., Seif M. Feasibility of replacing homemade solutions by commercial products for qualitative fit testing of particulate respirators: a mixed effect logistic regression study. *MethodsX*. 2019; 6: 1313–1322. Available at: DOI:10.1016/j.mex.2019.05.034
24. Mitchell BG., Wells A., McGregor A., McKenzie D. Can homemade fit testing solutions be as effective as commercial products? *Healthcare infection*. December 2012; 17(4): 111–114. Available at: DOI:10.1071/HI12019
25. Marsh JL. Evaluation of saccharin qualitative fitting test for respirators. *American Industrial Hygiene Association Journal*. 1984; 45(6): 371–376. Available at: DOI:10.1080/15298668491399965

26. Hon C-Y., Danyluk Q., Bryce E., Janssen B., Neudorf M., Yassi A., et al. Comparison of qualitative and quantitative fit-testing results for three commonly used respirators in the healthcare sector. *Journal of Occupational and Environmental Hygiene*. 2017; 14(3): 175–179. Available at: DOI:10.1080/15459624.2016.1237030

6 Human errors in qualitative fit testing: an investigation of current controls

6.1 Abstract

Mistakes, oversights, and omissions are frequently viewed as inevitable by-products of any human operation. In some processes, these human errors are nothing more than wasteful or inconvenient, whereas in others there may literally be life or death consequences. Earlier studies have suggested that a significant proportion of real-world fit testers make errors in areas which may be critical to the ability of the QLFT method to properly identify a leak (critical errors). This poses the very real possibility that some QLFT fit testers make errors which can lead to a false pass fit test result and hence assign an ill-fitting, leaky respiratory protective device to a wearer. This paper aims to assess the reasons why some fit testers make errors by investigating the gaps in current engineering and administrative controls. It is hoped that the gaps revealed will inform stakeholders in implementing further controls aimed at reducing number of critical errors that fit testers make in applying the QLFT method, and ultimately assist in protecting the wearer's health in the UK workplace.

6.2 Background

Recent news reports highlighting a spate of incorrect laboratory SARS-CoV-2 Polymerase Chain Reaction (PCR) test results has given wide publicity to the phrase ‘false negative’ results [1][2]. More commonly termed ‘false passes’, these types of erroneous results are particularly dangerous since a false assurance of safety is offered, potentially putting health at risk. In the UK workplace, a Health and Safety Executive (HSE) prescribed fit testing protocol is used to provide assurance to a wearer that a given make, model, and size of tight-fitting respiratory protective device (RPD) provides a suitable fit for their face [3]. Since tight-fitting RPD are frequently used as a last resort in reducing a wearer’s exposure to an airborne hazard, a suitable facial fit is vital in protecting the wearer’s health [4][5]. A false pass result in this context gives the wearer false assurance that the selected RPD provides a suitable facial fit for them when in truth, it does not.

The qualitative fit test (QLFT) method is by far the most common RPD fit testing procedure employed in the UK workplace today [6]. This method uses the wearer’s taste response to a strong-tasting challenge agent (usually BitrexTM, denatonium benzoate) to detect face seal leakage and hence identify a poor fitting respirator [7]. The widespread use of this fit testing method across different industry sectors can be attributed both to the low entry cost of equipment and the widely held view that QLFT is a ‘simple’ procedure [8]. This perception of simplicity is dangerous however since it disguises the more complex principles that sit behind the method and underestimates the care that must be taken in application if assignment of false pass results is to be avoided. Indeed, human errors resulting in misapplication of the validated method were found by an earlier study to be the least well-understood of the complexity of factors that can lead to errors in workplace QLFT and required further investigation [9].

This series of papers has aimed to provide a better understanding of human errors in QLFT with the ultimate aim of assuring better workplace safety for RPD wearers. Earlier studies postulated a list of ‘critical errors’, and the findings indicate that these may affect one or more of the fundamental key principles of the QLFT method which impact the ability of the method to identify a poor fitting respirator [10]. When referenced against the observations recorded of real-world fit testers, studies suggest that the prevalence of

human errors is widespread, indicating a problem which remains largely unchecked and poorly controlled in the workplace [11].

In any safety-critical process, if risks cannot be mitigated by elimination or substitution of the hazard, proper controls are necessary to safeguard against the worst outcomes of human error. Effective controls which mitigate the most serious of fit tester errors (those that have the potential to lead to a false pass result), are therefore crucial in assuring a safe fit test outcome for the wearer. Controls typically may be administrative in nature (information, instruction, warnings etc), or may involve hard-engineering solutions (fail-safe devices, adaptations to equipment / procedures etc). Engineering controls are usually prioritised above administrative controls in risk mitigation strategies such as the Hierarchy of Controls since these methods tackle risks at source rather than requiring a change to action or behaviour [12].

Any attempt at mitigation of human errors in QLFT must therefore start with an understanding of the current control strategies in order to identify and address possible gaps. Indications from earlier studies points to a lack of awareness and understanding amongst some fit testers when addressing outcome-critical actions, suggesting the presence of gaps in the current administrative controls [11]. Concerns raised by other researchers point to potentially serious gaps in engineering controls aimed at preventing or reducing the possibility of human errors occurring – controls which might be expected given the critical nature of a false pass outcome [13].

This paper aims to assess the current administrative and engineering controls available to real-world fit testers to evaluate the extent to which human errors are mitigated in workplace QLFT. Results of these studies are synthesised to reveal gaps in the controls provided which then offer indications about current failings and sources for improvement. The extent of current controls has not been studied to date and this research is crucial if results of workplace qualitative fit tests are to be assured. The findings of this study will be useful to stakeholders to assist in mitigation of human errors in QLFT through improvement, modification, or supplementation of the current controls.

6.3 Method

Two studies were used to identify the presence or absence (gaps) of controls which could mitigate one or more of the ‘critical errors’ which were proposed by an earlier study as potentially leading to a false pass outcome [10]:

1. An assessment of current administrative controls;
2. An assessment of current engineering controls.

A full list of the critical errors used for this study is shown in Appendix K, along with an ID key to facilitate referencing within this document. No attempt was made in this study to assess the efficacy of any controls identified, but simply to investigate the presence or absence of mitigating information (administrative control) or engineering solutions (engineering control). Similarly, this investigation was limited to the list of critical errors which was theorised by an earlier study. Whilst every effort was made previously to ensure this list was exhaustive, it is recognised that gaps may be present which therefore are not investigated in this study.

6.3.1 Assessment of current administrative controls

Information provided by the British Safety Industry Federation (BSIF) Fit2Fit scheme suggests that there are four main administrative resources available to qualitative fit testers which provide information instruction and guidance in the QLFT procedure:

- 1) Information provided by QLFT equipment user instructions;
- 2) Guidance provided in HSE INDG479;
- 3) Guidance provided in the Fit2Fit QLFT Companion; and
- 4) Instruction provided by fit tester training courses [14]

These administrative resources were assessed in turn to identify the presence or absence (gap) of controls useful in mitigating critical errors in QLFT.

QLFT user instructions were obtained from manufacturers’ websites and were referenced alongside hard-copy instructions available in the QLFT equipment kits [15][16][17].

Manufacturers of QLFT equipment in the UK marketplace were identified through a survey of known fit test equipment providers, RPD manufacturers and other RPD suppliers (stockists, wholesalers etc). Further information was requested from the Respiratory Protection Product Group (a special interest working group of the BSIF which represents UK RPD manufacturers and suppliers) to ensure that the list of UK QLFT equipment providers was as complete as possible. The survey sought to include both original manufacturers as well as other equipment manufacturers (OEMs) that provide QLFT equipment in a re-branded form. OEM providers were included in this survey to allow assessment of the controls provided in OEM user instructions which may be re-written as part of the re-branding process.

A new set of QLFT equipment (a 'kit') was obtained directly from each original manufacturer / OEM for the purposes of these studies and a batch number was allocated and recorded to allow identification of the source, time, and date. Although some QLFT kit manufacturers offer both 'bitter' and 'sweet' versions, this study was limited to 'bitter' versions to allow comparison across all brands. Each kit and set of user instructions was dummy-coded to allow reporting of results with removal of manufacturer-identifying data for purposes of potential commercial sensitivity.

Assessment of guidance resources was limited to analysis of the Health and Safety Executive (HSE) document INDG479 and the BSIF Fit2Fit Companion guide to qualitative fit testing which were accessed through the originators' websites [3][18]. Analysis of training course content was limited in this study to courses whose content had been judged to meet the requirements of the BSIF Fit2Fit Approved Training Course scheme [19]. A complete evaluation of the content of all fit tester training courses from every provider in the UK was not possible due to the large and increasing number of different course providers, and restricted access to private company training materials. Instead, the Fit2Fit Approved Training Course provides a syllabus and reference form which is available in the public domain and allows analysis of a 'gold standard' course [20]. The scheme is designed to provide a signpost to fit testers who want to access best practice training, with the aim of ensuring that the content of such training courses, '*enable satisfactory RPE face fit testing competence.*' [19]. Requirements of these

training courses are set out in the scheme's syllabus which provided the basis for assessment of training course content in this study.

All administrative resources were handled in electronic format to facilitate cataloguing and search functions. Each document was interrogated for information pertaining to the mitigation of any of the critical errors listed in Appendix K. A list of search words and key terms was compiled to standardise the search criteria for each document, and this is detailed in Appendix L. The inclusion of some American-English words, e.g., 'nebulizer', was found to be useful since several manufacturers produce generic instructions for their US and European markets. For all documents, a simple scoring system was used to record whether mitigating information was present or absent (a gap); information was only recorded as 'present' where specific information relating to mitigation of the critical error was provided. As an example, the general instruction, 'ensure proper functioning of the nebuliser' was considered insufficient to record actions relating to nebuliser components as 'present'. A total score was recorded for each administrative resource to allow comparison and assessment of gaps. The results of these studies were compiled in a table (Microsoft Excel, Microsoft Office 365) for cross-comparison and analysis of gaps.

6.3.2 Assessment of current engineering controls

The main equipment resources available to a qualitative fit tester are the components included in the QLFT kit which is used to conduct the sensitivity test and fit test procedures. As observed during Fit2Fit practical assessments, many fit testers will choose to supplement their kit with additional resources such as wipes, water, a stopwatch etc [11]. Since use of additional resources is unrestricted and varied, this investigation is restricted to the resources included as standard in the UK commercially available QLFT kits identified for assessment in this study.

An investigation was conducted for each QLFT kit in turn to identify the presence or absence of controls intended to mitigate any of the critical errors detailed in Appendix K. Each kit was dummy coded with an identification number (ID) which was recorded alongside the QLFT equipment manufacturer details in spreadsheet format. The ID number was used to hard label all major components, sub-components and removable

parts within each kit. This method of identification was used to allow side-by-side comparison of components and sub-components whilst ensuring that their origin could be traced to the equipment provider if required.

HSE guidance document INDG479 requires QLFT equipment used in the UK to meet the specification detailed in British Standard BS ISO 16975-3 [21]. Using this specification as a reference, each component was inspected to identify any adaptations which could assist mitigation of any of the critical errors detailed in Appendix K. A binary scoring system was used to record presence or absence (a gap) of engineering controls, and a total 'score' was calculated to allow comparison and assessment of gaps. The results of these studies were compiled in a table (Microsoft Excel, Microsoft Office 365) for analysis. Following completion of both studies, results of the assessments of administrative and engineering controls were synthesised to allow an overall assessment of current controls and gaps.

6.4 Results

Seven different QLFT kit providers were identified in the UK marketplace; six original manufacturers and one re-branded OEM provider. Table 6-1 provides a list of these providers along with the version numbers of user instructions supplied.

QLFT equipment manufacturer	Manufacturer type	Item code	User instruction version
3M	Original	FT-30	34-8704-6582-9
Scott	Original	2017070	595116-01 Rev B 02/20
Alpha Solway	Original	QLTK1	Revision 14 (June 2020)
Moldex	Original	010301	EN Rev11-20
JSP	Original	BPT050-000-000	RFGZ-005-900_19-08
Draeger	Original	4055625	1004 Rev A
ARCO	OEM	1E1100	Revision 15 (June 2020)

Table 6-1 UK QLFT equipment (kit) providers

All kits were found to contain two nebulisers, a test hood, sensitivity solution and fit test solution. Some kits contained solutions in 50ml or 55ml bottles, whereas others provided individual 2.5ml single-use ampoules. One set of user instructions was identified within each kit; some as hard-copy documents, and others supplied on a USB flash-drive and / or Quick Response (QR) code.

6.4.1 Current administrative controls

The administrative resources investigated in this study comprised a total of ten documents: three guidance/training documents (HSE INDG479, Fit2Fit QLFT companion, Approved Training syllabus), plus seven QLFT user instructions (Table 6-1). Assessment of these administrative resources reveals a wide variation in the controls available to fit testers, with no single resource offering a complete set of controls in mitigating critical errors.

Table 6-2 displays the results of the assessment of the ten administrative resources against the list of eighteen critical errors, and full details are provided in Appendix M. Possible

‘present’ scores ranged from zero to eighteen, with a higher score indicating control of a higher number of critical errors. ‘Present’ scores in the data set showed a minimum score of four and a maximum score of 14. These data show that the HSE INDG479 guidance provided controls for the fewest number of critical errors (score = 4), meaning that gaps were identified for 14 of the possible 18 critical errors. The highest number of administrative controls were found to be provided by the Fit2Fit companion guidance and user instructions supplied with QLFT kits 3 and 7 (score = 14). These data further indicate that all seven QLFT user instructions provided controls for nine or more critical errors ($M = 10.6$), all scoring higher than the Approved training syllabus (score = 8).

Identification of administrative control	Guidance/Training			QLFT user instructions						
	HSE INDG479	Fit2Fit Companion	Approved Training	Kit 1	Kit 2	Kit 3	Kit 4	Kit 5	Kit 6	Kit 7
Present	4	14	8	13	13	14	9	12	10	14
Absent	14	4	10	5	5	4	9	6	8	4

Table 6-2 Administrative controls identified in QLFT guidance, training, and equipment user instructions

Notes: $N = 18$. Present = critical errors for which control was identified. Absent = critical errors for which no control was identified.

This study also provided the opportunity to examine how well each of the eighteen critical errors were mitigated by the ten different administrative resources investigated, these data are displayed in Table 6-3. Possible ‘present’ scores ranged from zero to ten, with a higher score indicating mitigating information in a higher number of administrative resources. ‘Present’ scores in the data set showed a minimum score of 2 and a maximum score of 10. The highest score was found for the critical error *Wearer has smoked, eaten, or drunk before test* (ID key 2), for which mitigating information was found to be present in all ten administrative resources. Critical errors 13 and 14 relating to the dosage and timing of challenge aerosol application were also found to be high scoring, with controls identified in 9 out of 10 resources. The lowest score was found for the critical error *Room poorly ventilated*, for which mitigating controls were identified in only two of the ten resources, followed by *Fit test nebuliser Venturi mis-seated*, identified in only three of the ten

resources. Indeed, no information was identified in half or more of the administrative resources for eight of the eighteen critical errors, suggesting numerous gaps.

ID Key	Critical error	Identification of administrative control	
		Present	Absent
1	Room poorly ventilated	2	8
2	Wearer has smoked, eaten or drunk before test	10	0
3	Wearer not advised to rinse mouth / clean face	5	5
4	Clear palate not verified (after sensitivity test)	7	3
5	Wearer not instructed to breathe through mouth	9	1
6	Sensitivity test not repeated (after failed fit test)	8	2
7	Hood not cleaned (after fit test)	6	4
8	Fit test nebuliser filled with sensitivity solution	6	4
9	Fit test nebuliser under-filled	5	5
10	Fit test nebuliser clogged	8	2
11	Fit test nebuliser bottom bung not removed	5	5
12	Fit test nebuliser Venturi mis-seated	3	7
13	Fit test solution under-administered	9	1
14	Fit test timing incorrect	9	1
15	Fit test nebuliser not upright	5	5
16	Fit test nebuliser not actuated with full squeezes	4	6
17	Hood damaged / unsuitable	4	6
18	Hood positioned incorrectly	6	4

Table 6-3 Analysis of the control of critical errors in QLFT through administrative resources

Notes: $N = 10$. Present = control identified in administrative resource. Absent = control not identified in administrative resource.

Of interest are the numerous gaps in controls found in HSE INDG479 guidance and Fit2Fit Approved Training content. Directions provided in these documents suggests that they are meant to be read in conjunction with QLFT user instructions. However, analysis of the administrative controls referenced in HSE INDG479 guidance and Fit2Fit Approved Training content versus the various QLFT manufacturers' user instructions, reveals gaps that remain between these documents. The frequency of these gaps is displayed in Figure 6-1.

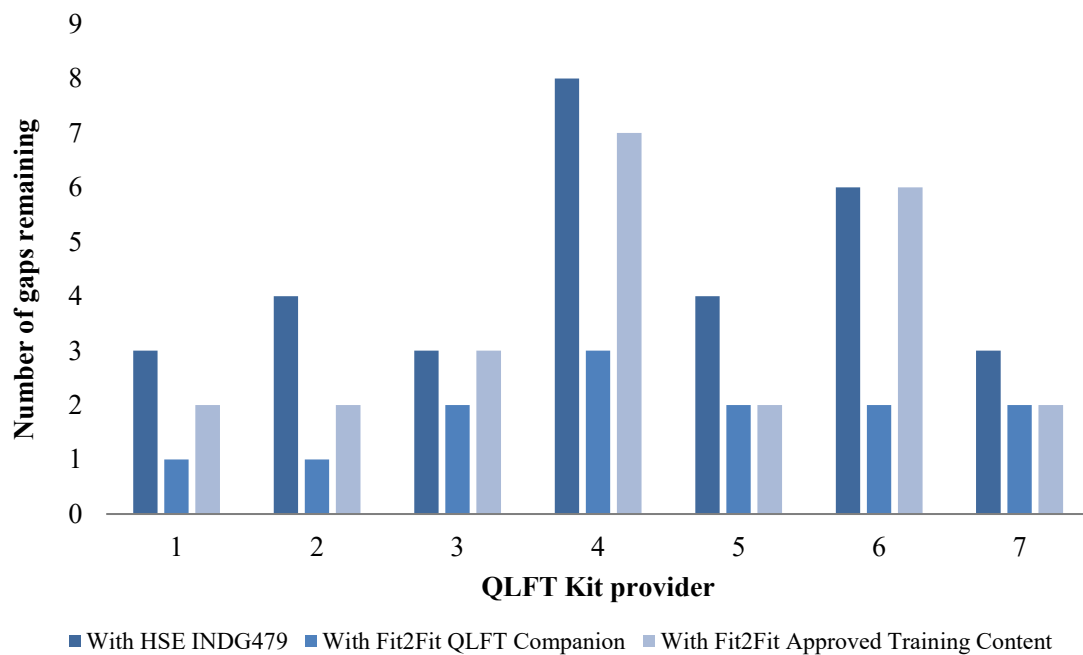


Figure 6-1 Gaps in administrative controls between QLFT equipment manufacturers' user instructions and INDG479, Fit2Fit QLFT Companion and Fit2Fit Approved Training content

These results indicate that whilst some user instructions fill the majority of the gaps in controls identified in guidance and training documents, others fail to do so, and significant gaps remain. The user instructions supplied with QLFT kit 4 offer the poorest administrative controls when analysed in conjunction with HSE INDG479, with eight of the eighteen critical errors remaining uncontrolled. These uncontrolled errors include the need to remove the bottom bung of the nebuliser, ensure the Venturi is properly seated, hold the nebuliser upright and ensure full squeezes - all shown by earlier studies to be fundamental factors in ensuring proper aerosol generation.

6.4.2 Current engineering controls

Assessment of current engineering resources reveals very few mitigating controls and numerous gaps suggesting huge scope for improvement in this area. These results further indicate inconsistency across different QLFT kit manufacturers and failure to provide engineering controls for the majority of critical errors.

Table 6-4 displays the results of the assessment of equipment resources investigated in this study and full details are provided in Appendix N. Seven different QLFT kits were examined against eighteen critical errors. Possible ‘present’ scores ranged from zero to eighteen, with a higher score indicating control of a higher number of critical errors. ‘Present’ scores in the data set showed a minimum score of two and a maximum score of 4 ($M = 2.7$) indicating widespread failure to provide engineering controls. The best-performing QLFT kit (Kit 1) was found to provide engineering controls for only four of the eighteen critical errors, while the worst-performing (Kits 2, 4 and 6) for just two. For the worst-performing kits, this study therefore identified gaps in controls against sixteen of the eighteen critical errors, indicating poor mitigation of human errors.

Identification of control for critical error	QLFT equipment						
	Kit 1	Kit 2	Kit 3	Kit 4	Kit 5	Kit 6	Kit 7
Present	4	2	3	2	3	2	3
Absent	14	16	15	16	15	16	15

Table 6-4 Engineering controls in QLFT kits

Notes: $N = 18$. Present = critical errors for which control was identified. Absent = critical errors for which no control was identified.

Table 6-5 provides a focussed analysis of the controls available for individual critical errors along with observations from the investigation. When these data were analysed according to error type, the results indicate that engineering controls in QLFT are focused on mitigation of just four of the eighteen errors: *Fit test nebuliser filled with sensitivity solution* (ID key 8, seven kits), *Fit test nebuliser under-filled* (ID key 9, one kit), *Fit test nebuliser clogged* (ID key 10, seven kits), and *Hood positioned incorrectly* (ID key 18, four kits).

No engineering controls were identified in any of the resources for fourteen of the eighteen critical errors. These errors include *Fit test nebuliser not upright**, *Fit test nebuliser bottom bung not removed*, and *Fit test nebuliser Venturi mis-seated*, all of which were found by a practical study to have a significant impact on in-hood aerosol concentration and fit test outcome (*when tested in conjunction with Fit test nebuliser under-filled) [22]. No engineering controls were identified for critical errors ID key 1-7

which relate to subjectivity factors, affecting the wearer's ability to detect the taste of the challenge agent. The gaps in engineering controls for these areas suggest a potential for uncontrolled false passes to occur as a result of errors made in managing the wearer's taste detection.

ID Key	Critical error	<i>n</i>	Observations
1	Room poorly ventilated	0	No controls identified
2	Wearer has smoked, eaten or drunk before test	0	No controls identified
3	Wearer not advised to rinse mouth / clean face	0	No controls identified
4	Clear palate not verified (after sensitivity test)	0	No controls identified
5	Wearer not instructed to breathe through mouth	0	No controls identified
6	Sensitivity test not repeated (after failed fit test)	0	No controls identified
7	Hood not cleaned (after fit test)	0	No controls identified
8	Fit test nebuliser filled with sensitivity solution	7	Inconsistent colour-coding, labelling
9	Fit test nebuliser under-filled	1	Graduated fill lines
10	Fit test nebuliser clogged	7	Pin and spare Venturi supplied
11	Fit test nebuliser bottom bung not removed	0	No controls identified
12	Fit test nebuliser Venturi mis-seated	0	No controls identified
13	Fit test solution under-administered	0	No controls identified
14	Fit test timing incorrect	0	No controls identified
15	Fit test nebuliser not upright	0	No controls identified
16	Fit test nebuliser not actuated with full squeezes	0	No controls identified
17	Hood damaged / unsuitable	0	No controls identified
18	Hood positioned incorrectly	4	Collar with 'shoulder' seats

Table 6-5 Control of critical errors through engineering resources provided in QLFT kits

Notes: $N = 7$, n = number of kits in which control was identified.

All kits were found to include engineering controls aimed at mitigating errors in the incorrect use of the sensitivity solution (ID key 8) through labelling, colouring, or numbering. However, observations indicate a highly inconsistent approach between the various kits, with a different approach used for each different kit provider. Figure 6-2 highlights the labelling / colouring systems for the nebulisers found in QLFT kits 1 and 5, which reveal the opposite use of colouring systems between the sensitivity and fit test nebulisers supplied in the two kits.



Figure 6-2 Labelling / colouring of nebulisers identified as a control in mitigating the critical error *Fit test nebuliser filled with sensitivity solution*

Notes: Sensitivity nebuliser (red labelling) and Fit test nebuliser (blue labelling) QLFT kit 1 (left). Sensitivity nebuliser (blue labelling) and fit test nebuliser (red labelling) QLFT kit 5 (right).

All QLFT kits were found to include two sets of pins and new Venturi. These were identified as controls to assist in mitigating a clogged nebuliser (critical error ID 10), although comparison with the earlier assessment of administrative controls reveals that only three of the kits provided instructions for use of these items in their user instructions. Only one QLFT kit (Kit 1) was found to provide engineering controls to mitigate an error in under-filling the fit test nebuliser (critical error ID 9) by providing marked graduated fill-lines. Figure 6-3 shows the presence of graduated fill lines provided in nebulisers supplied in kit 1, compared with those supplied in kit 3 which feature no fill lines.



Figure 6-3 Engineering controls identified as assisting in proper nebuliser filling

Notes: Nebuliser saucers with graduated fill lines supplied in QLFT kit 1 (left), nebuliser saucers with no fill lines supplied in QLFT kit 3 (right).

When the results from the studies of administrative and engineering controls were synthesised, the critical errors for which the fewest controls were identified across all resources were revealed. Figure 6-4 provides an overview of this synthesis, displaying the number of gaps identified across the seventeen combined administrative and engineering resources. These results indicate significant gaps in the controls available to fit testers through the ten administrative resources (guidance, training, and user instructions) and seven engineering resources (QLFT equipment kits).

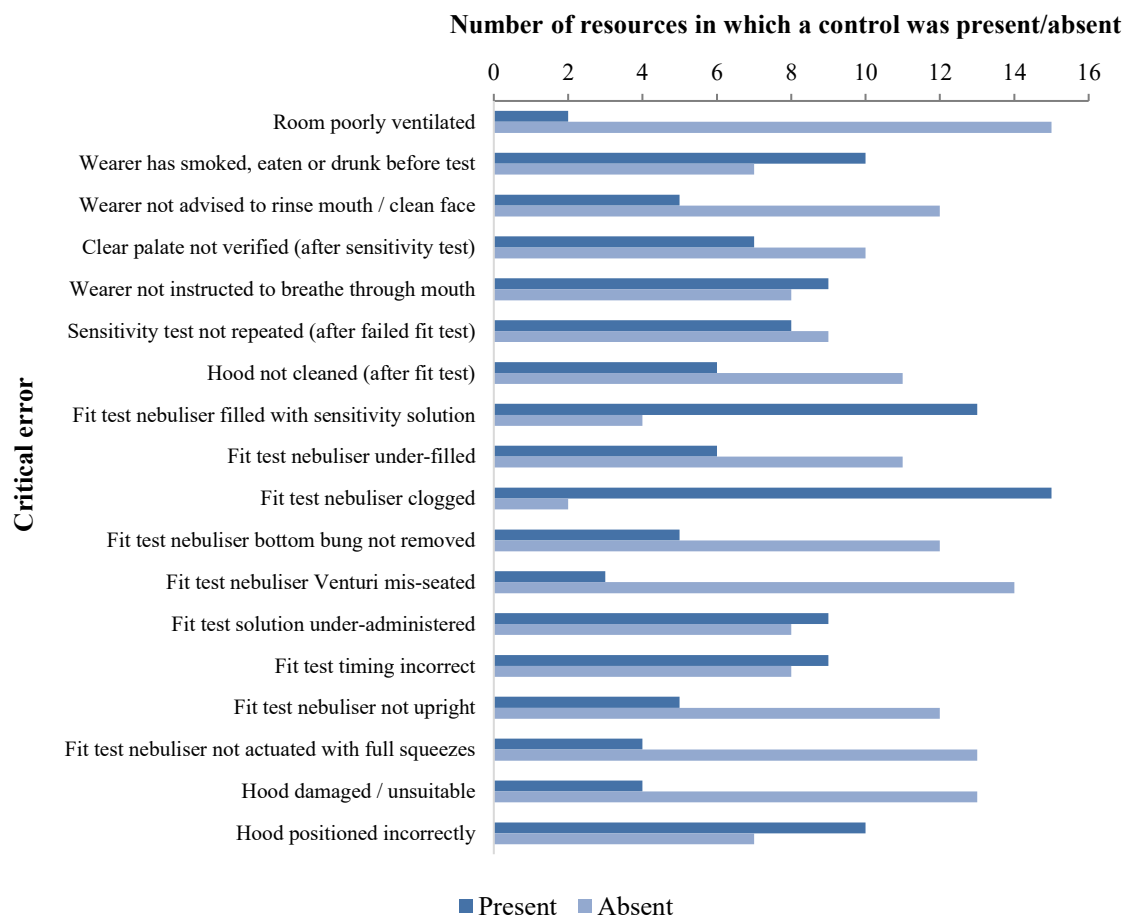


Figure 6-4 Presence / absence of controls for critical errors in QLFT

Notes: $N = 17$. Present = number of administrative or engineering resources in which control was identified. Absent = number of administrative or engineering resources in which no control was identified. Administrative resources ($n = 10$) relate to information provided in HSE INDG479 guidance, Fit2Fit QLFT companion guidance, Approved training course content, and user instructions from QLFT providers in the UK. Engineering resources ($n = 7$) relate to the components and sub-components provided in QLFT equipment kits in the UK market.

The highest number of gaps were found for *Room poorly ventilated* for which controls were found only in two of the seventeen resources examined. Numerous gaps were also found relating to critical errors *Fit test nebuliser Venturi mis-seated*, *Fit test nebuliser not actuated with full squeezes*, *Hood damaged / unsuitable*, *Fit test nebuliser bottom bung not removed*, and *Fit test nebuliser not upright** for which controls were identified in five or fewer resources. These errors all were found by an earlier study to have a significant effect on the in-hood aerosol concentration and fit test outcome (*when tested in combination with Fit test nebuliser under-filled), and the gaps identified are therefore important and must be addressed [22].

6.5 Discussion

The results of this study indicate that much more can be done to assist fit testers in mitigating human errors by addressing the gaps identified in both administrative and engineering controls in QLFT. Improvements in controls against the gaps identified would serve to offer better resources to real-world fit testers and hence reduce the potential for false pass fit test results in workplace QLFT.

Administrative controls provided in guidance, training course content and manufacturer's user instructions play a critical role in the development of a competent fit tester by building essential skills and knowledge. To be effective, these controls must include information about how and why errors occur, in order to aid understanding and drive a safety-conscious work attitude. It is perhaps surprising then, that the HSE sets no requirement for formally-certified training for fit testers, stating only "*Fit testing should be conducted by a competent person*" [3]. Improving the information provided to fit testers through the administrative resources available is central to building confidence in a properly administered information process.

In terms of guidance, HSE document INDG479 is the principal reference guide which provides advice to both fit testers and employers on fit testing in the UK workplace. The gaps in administrative controls found in HSE guidance are surprising and are indicative of a creep towards generalisation of HSE information which seeks to place the requirement for providing greater detail onto manufacturers and industry bodies such as BSIF through the Fit2Fit scheme. A worrying outcome of the results provided by this study is therefore the indication that manufacturers' user instructions are currently failing to address the numerous gaps in HSE guidance, meaning that many critical areas of the QLFT process remain uncontrolled. It is essential that the gaps which occur between HSE guidance and manufacturers' user instructions are addressed with more specific instruction provided to fit testers on the correct, effective, and safe use of equipment if false pass fit test results are to be avoided. A further source of concern is a blurring of the overlap between guidance, fit tester training course content and manufacturer's user instructions whose boundaries are currently poorly defined with confusion about the role they each play in providing information to a fit tester.

Whilst the controls provided in administrative resources seek to mitigate errors by influencing a fit tester's behaviour, engineering controls aim to mitigate errors at source. In terms of risk management, engineering controls are seen to be more effective (and hence more favourable) when compared to administrative controls because they remove human influence [4]. The results of this study suggest that engineering controls are currently under-used in QLFT equipment and numerous gaps exist where controls could be applied. QLFT equipment is rudimentary in nature, and engineering controls which eliminate human errors in the use of this equipment should provide a huge benefit in reducing false passes and assuring the reliability of the QLFT method. Although efforts have been made to provide controls for four elements of current QLFT equipment (such as providing fill lines on some nebulisers), the majority of critical errors remain entirely uncontrolled and rely solely on the ability, knowledge and skills of the fit tester to ensure mitigation of errors.

The finding that a different labelling / colour-coding system is used for the sensitivity and fit test nebulisers and solutions in each different QLFT kit is both surprising and concerning. Many fit testers will choose to use QLFT kits from different manufacturers (especially during periods of short supply such as a pandemic) or may supplement their kits with top-up solutions from a different manufacturer. The current labelling system is therefore potentially both confusing and dangerous. Earlier studies have suggested that incorrect substitution of the fit test solution for the sensitivity solution has the potential to lead to a false pass result, and clear labelling / warning is therefore essential if this is to be avoided [22]. In medical fields such as anaesthesiology and paediatrics, harm or death caused by dosing errors has resulted in calls for improvement in labelling, standardised colour coding and even pictogram systems [23][24]. Standardised use of colours and labels in QLFT kits requires cooperation between different manufacturers but is a small change that could have an important impact on mitigating dosing errors in QLFT.

At first appearance, it may not be surprising that no engineering controls were identified against the critical errors relating to subjectivity factors since these are less tangible than those relating to the use of equipment and focus instead on management of the wearer's taste palate. However, it would be erroneous to conclude that controls are impossible in

mitigating errors relating to these factors when considering the extent to which systems such as gas detection and filter breakthrough technology have evolved to negate the need for a human sensory system. Current engineering controls are failing to engage with, and address, subjectivity issues in QLFT, despite the wide understanding that the reliability of human taste detection is a known problem in this method.

The results of this study reveal the gaps in current administrative and engineering controls and provide the first suggestion of some of the reasons why fit testers may make errors in the workplace. The complete failure to provide controls for the majority of the most important elements of the QLFT procedure means that this method is extremely vulnerable to human error. QLFT kit manufacturers and providers can, and must, do much more to provide engineering controls in this procedure if human errors are to be reduced or eliminated. The outcome of this study suggests that limited information is available to fit testers about key areas of the fit testing process indicating that fit testers may not be equipped with the knowledge required to mitigate false pass results in this method. Evidence from earlier studies suggests that some fit testers may not refer to guidance or user instructions and may not ever attend a training course [11]. Administrative controls such as these require active engagement on the part of the fit tester and hence may be less successful as a result. On the other hand, almost all fit testers use QLFT equipment provided by one of the UK equipment manufacturers and therefore engineering controls which eliminate human errors at source are likely to be the most successful in mitigating real-world problems. Redesign of fit testing equipment e.g., nebulisers, could provide engineering controls that preclude apparent operation if the equipment is not properly assembled and used. Such controls should provide a visible, audible or other easily identifiable means for the fit tester to determine proper functioning. Innovative engineering controls could also include the use of technology such as an automated dispenser to assure proper and controlled dosing, and a pressure sensor to monitor proper actuation of the nebuliser.

6.6 Conclusion

The findings of this study suggest that numerous gaps are present in the current administrative and engineering controls aimed at mitigating critical human errors in QLFT. The identification of very few engineering controls in QLFT equipment kits, suggests that manufacturers and providers are currently failing to address human errors, and have an important role to play in ensuring that engineering controls are used to best advantage in mitigating errors. In addition to QLFT kit providers, other guidance and training stakeholders can all provide better administrative controls by improving the information resources provided to fit testers in mitigating critical errors.

6.7 References

1. Mayers C., Baker K. Impact of false-positives and false-negatives in the UK's COVID-19 RT-PCR testing programme. 2020. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/895843/S0519_Impact_of_false_positives_and_negatives.pdf (Accessed: 28 October 2021)
2. Torjesen I. Covid-19: PCR testing is suspended at private laboratory after high rate of false negatives. BMJ. 2021. p. n2535. Available at: DOI:10.1136/bmj.n2535
3. Health and Safety Executive. INDG479 Guidance on respiratory protective equipment (RPE) fit testing. HSE Books; 2019. Available at: <https://www.hse.gov.uk/pubns/indg479.pdf>
4. Health and Safety Executive. L5 Control of substances hazardous to health. The Control of Substances Hazardous to Health Regulations 2002 (as amended). Approved Code of Practice and guidance L5 (Fifth edition). HSE Books; 2005. Available at: <http://www.hse.gov.uk/pubns/priced/l5.pdf>
5. Health and Safety Executive. HSG53 Respiratory protective equipment at work. A practical guide. (Fourth edition). HSE Books; 2013. Available at: <http://www.hse.gov.uk/pubns/priced/hsg53.pdf>
6. Murray A. UK best practices on RPD fit testing (Fit2Fit programme). National Fit Testing Programmes. 18th May 2021. International Society for Respiratory Protection; Available at: <https://www.isrp.com/minutes>
7. Macfarlan Smith Ltd. Bitrex. Available at: <https://www.bitrex.com/> (Accessed: 6 May 2021)
8. Colton CE. Respirator fit testing: choosing the best method. Metal Finishing. 2001; 99(9): 53–55. Available at: DOI:10.1016/S0026-0576(01)81436-7
9. Long K., Mai N., Williams M. Human errors in qualitative fit testing: a literature review [Unpublished paper]. 2019.

10. Long K., Mai N., Williams M. Human errors in qualitative fit testing: theoretical impact of errors on fit test outcome [Unpublished paper]. 2021.
11. Long K., Mai N., Williams M. Human errors in qualitative fit testing: a study of real-world fit testers [Unpublished paper]. 2020.
12. Health and Safety Executive. Hierarchy of risk controls. Available at: <https://www.hse.gov.uk/construction/lwit/assets/downloads/hierarchy-risk-controls.pdf>
13. Frost S., Mogridge R., Roff M. RR1029 Review of fit test pass criteria for filtering facepieces Class 3 (FFP3) respirators. Health and Safety Executive; 2015. Available at: <https://www.hse.gov.uk/research/rrpdf/rr1029.pdf>
14. British Safety Industry Federation. Fit2Fit candidate handbook. 2014. pp. 1–8. Available at: www.fit2fit.org/resources (Accessed: 1 October 2020)
15. 3M. Qualitative fit test apparatus FT30 user instructions. Available at: <https://multimedia.3m.com/mws/media/1464265O/3m-qualitative-fit-test-apparatus-ft-30-user-instruction.pdf> (Accessed: 16 October 2021)
16. KG MA& C. Qualitative fit test instructions. Available at: https://www.moldex-europe.com/fileadmin/user_upload/files/user-instructions/Moldex-Fit-Test-EN-Rev11-20.pdf (Accessed: 2 October 2021)
17. JSP Ltd. JSP Qualitative Fit Test Apparatus. Available at: http://docs.jsp.co.uk/UI-RFGZ-005-900_QualitativeFitTest.pdf (Accessed: 20 October 2021)
18. British Safety Industry Federation. Fit2Fit Companion to HSE INDG479 guidance on respiratory protective equipment (RPE) fit testing, qualitative method. June 2021. Available at: <https://www.fit2fit.org/resources>
19. British Safety Industry Federation. Fit2Fit Approved Training Courses. Available at: <https://www.fit2fit.org/fit2fit-approved-training-providers/>
20. British Safety Industry Federation. Fit2Fit Training Approval Scheme. Syllabus reference form. Available at: www.fit2fit.org/resources

21. British Standards Institution. BS ISO 16975-3:2017 Respiratory protective devices. Selection, use and maintenance. Part 3: Fit-testing procedures. London: British Standards Institution; 2017.
22. Long K., Mai N., Williams M. Human errors in qualitative fit testing: practical impacts of human errors on fit test outcome [Unpublished paper]. 2021.
23. Orser BA., Oxorn DC. An anaesthetic drug error: minimizing the risk. *Canadian Journal of Anaesthesia*. 1994; 41(2): 120–124. Available at: DOI:10.1007/BF03009804
24. Yin HS., Mendelsohn AL., Fierman A., van Schaick L., Bazan IS., Dreyer BP. Use of a pictographic diagram to decrease parent dosing errors with infant acetaminophen: a health literacy perspective. *Academic Pediatrics*. 2011; 11(1): 50–57. Available at: DOI:10.1016/j.acap.2010.12.007

7 Conclusions and summary

The primary aim of this thesis was to better understand human errors in qualitative fit testing (QLFT) and relate the findings to the impact on fit test outcome, with the core research question “How do human errors affect the outcome of a qualitative fit test and hence the potential for false pass results?” This chapter discusses the findings of the work in a real-world context, reports the contribution this thesis makes to existing knowledge and offers proposals for improvements to current resources. Finally, the chapter will present suggestions for future research extending from this thesis.

7.1 Summary of research findings

The overall findings of this thesis indicate that workplace qualitative fit testing is highly vulnerable to human errors which may result in the assignment of false passes. Although human errors appear to be the least well understood of the complexity of errors in the QLFT method, this study has provided evidence that they appear to be widespread amongst real-world fit testers.

Theoretical assessment of the impacts of human errors suggests that some are critical to the ability of the QLFT method to properly identify a leaking facepiece, and practical studies provide further evidence to support this claim. A workplace Beta value of 0.43 was estimated when the measured impacts of errors in the use of QLFT equipment were applied to the observations of real-world fit testers. This value exceeds the Beta value reported by laboratory-based studies (0.04 – 0.12), suggesting that the QLFT method becomes a less reliable procedure for detecting poor fitting facepieces when applied outside of laboratory settings. The thesis finds that the most critical of human errors are currently poorly controlled through administrative and engineering resources, and urgent action is required to provide better controls to mitigate these errors and assure the workplace safety of this fit test method.

The findings of this thesis were gathered through five separate, but complementary studies which sought to answer the five research objectives identified. The objective of the first study was to evaluate the current understanding of sources of error in QLFT

through a review of the literature, and thus to identify research gaps. This study revealed three key themes in the literature relating to errors in QLFT: statistical errors inherent in the method, errors arising from the subjective nature of the QLFT method, and human errors in real-world application of the QLFT method. The review found that statistical errors were well described and have been extensively studied in laboratory settings. The Beta error value (probability of the method returning a false pass result) has been widely reported by numerous researchers, however the influence of facepiece type and reference QNFT method contribute to the large range of published values, many of which fall outside of the American National Standards Institute (ANSI) criteria for new fit test methods [1][2][3]. Errors relating to the subjective nature of the qualitative fit test method have been broadly discussed amongst researchers in this field, and reliance of the method on the human taste response mechanism is widely acknowledged [4][5]. A key finding of this review is the significant gap in the literature pertaining to human errors in QLFT. The potential impacts of human errors in QLFT have been commented upon by several researchers, but this review found no studies which published research in this area [6]. Discussed only in passing, these human errors and their impacts were found to be poorly described in the literature with references largely limited to anecdotal reports and discussion amongst stakeholders [7][8]. The study highlighted the importance of filling this research gap to better understand how human errors affect the outcome of a qualitative fit test, and whether this can lead to an increase in false passes beyond those reported as the Beta error in laboratory studies.

The second study sought to further the overall research aim by investigating the prevalence and sources of human errors amongst real-world QLFT operators. This study used the secondary records of observations of fit testers during Fit2Fit QLFT practical assessments to provide an understanding of the extent and types of human errors. A major finding from this study was the indication that over 83% of the assessments reported the presence of one or more human errors, with the worst-performing recording errors in 29 out of the 49 areas observed. Although Fit2Fit accredited fit testers appear to make fewer errors than non-accredited fit testers, the results indicate no improvement in mean observed error rate over the ten-year period studied. The most common errors were found to relate to proper donning and fit checking of RPD prior to the commencement of the fit testing process. Of the ten Fit2Fit sections, those which involve the set-up, use and

application of fit test equipment accounted for more than half of all errors recorded during Fit2Fit assessments. This was further emphasised by the finding that in almost one third of cases, the fit tester failed to verify the function of the nebuliser at any point during the procedure. This study highlighted that further investigation was required to ascertain whether the errors observed amongst real-world fit testers during Fit2Fit assessments have the potential to undermine the test protocol and lead to false passes in workplace fit testing.

The third study expanded on the observations gathered from the previous study and aimed to understand and assess the theoretical impacts of human errors upon QLFT outcome. This study proposed a list of ‘critical errors’ which were judged to have the theoretical potential to impact one or more of the three key principles of the QLFT method and lead to a false pass result. Analysis conducted in conjunction with the previous study indicated that over 39% of the errors observed amongst real-world fit testers during Fit2Fit QLFT practical assessments may have been critical errors, and these were recorded in over 70% of the fit tests observed. The study also suggested that the potential for human errors relating to the use and application of QLFT equipment are the most numerous and widespread and are therefore highly vulnerable to human error. Whilst this research identified those errors which have the theoretical potential to adversely affect fit testing outcome, the study highlighted the need for empirical research to evaluate whether these can translate to practical impacts which affect the outcome of a fit test.

The purpose of the fourth study was to extend the previous theoretical study by evaluating the practical impacts of human operator errors upon QLFT outcome and hence assess the potential for false pass results. Practical impacts were assessed through evaluation of errors both upon the in-hood aerosol concentration, and the ability of wearers to detect the presence of a known leak. The results of this study suggest that critical errors relating to the use and application of QLFT equipment have a measurable practical impact upon in-hood aerosol concentration and the wearer’s ability to detect a known leak. The biggest impacts were found for the errors *Fit test nebuliser filled with sensitivity solution*, *Fit test nebuliser clogged* and the compound error *Fit test nebuliser not upright and under-filled*. The key finding of this study is the indication that these errors are likely to affect the ability of the fit test to identify a leaking facepiece and therefore potentially lead to

assignment of false pass results in real-world fit testing. When these findings were applied to observations of real-world fit testers, the data suggest that a workplace Beta value for human errors in the use of QLFT equipment may be estimated at 0.43, a value far exceeding that reported in laboratory settings.

The fifth and final study of this thesis aimed to assess the current control measures in QLFT and identify gaps to inform future improvements in design and methodology aimed at mitigating the effects of human errors. A survey of current administrative and engineering resources highlighted the presence of numerous gaps in controls aimed at mitigating errors in QLFT. A key finding of this study was the almost total absence of engineering controls available to fit testers using commercially-sourced QLFT equipment kits. This finding suggests that QLFT kit manufacturers are currently failing to address the potential for human operator errors, and therefore the possibility of false passes in this method. In addition, the study found that current administrative controls fail to properly address the range of human operator errors in QLFT, and improvements to information resources provided to fit testers is crucial in mitigating human errors.

7.2 Contribution to knowledge

The separate and collective findings of the five studies included in this thesis contribute to the existing knowledge in several important ways which can be considered at the theoretical and empirical levels.

From a theoretical perspective, this research contributes to the knowledge of errors in the QLFT method by extending the current understanding of laboratory-based studies into real-world application. By means of an extensive literature review, this thesis examines the multi-faceted body of current research and identifies a major research gap in the knowledge pertaining to human operator errors. Importantly, this thesis expands current knowledge by addressing errors that are introduced as a result of the human operator - a factor that has not previously been considered. By relating the impacts of human errors back to the fundamental principles of the original methodology, the findings show consistency with the key tenets that underpin the method. Contravention of these fundamental principles has been demonstrated by this thesis to impact upon outcome and hence the reliability of the method. Additionally, use of the Fit2Fit criteria for the impact studies considered in this thesis provided a solid theoretical basis for examination of the effects of human errors, with clear references to Health and Safety Executive (HSE) INDG479 guidance [9]. By theorising a list of possible error scenarios against these criteria, this thesis postulates a set of critical errors which have the potential to undermine the fundamental principles of the method. This therefore provides a theoretical foundation for other researchers to evaluate the impacts of fit tester errors in a systematic manner. In applying the measured impacts of errors to the observations of real-world fit testers, this thesis offers the first estimate of a workplace Beta value for human errors in use of QLFT equipment. This offers a significant contribution to knowledge which thus far has been limited to a measure of the Beta value of QLFT conducted under laboratory conditions.

At the empirical level, this thesis is the first to conceptualise and implement secondary analysis of data acquired during Fit2Fit QLFT practical assessments as a tool for assessing human errors. A significant contribution to knowledge is made in this area, with the findings providing data about the prevalence and sources of errors made by real-world fit testers. The real-world impacts and implications of this research are an underlying principle throughout this thesis, with the knowledge that the qualitative method is the

most common method used for fit testing of tight-fitting disposable and reusable half facepieces. The findings presented in this thesis provide evidence to further support the view that the role of the fit test operator is crucial in ensuring the efficacy and reliability of the QLFT method. In turn, this also enhances the need for proper instruction, guidance, and support resources for fit testers in workplace settings. Through examination of current control measures, this thesis identifies the gaps present in administrative and engineering resources, offering an insight into the possible reasons for human errors.

7.3 Recommendations

This thesis has demonstrated that human errors are common in real-world QLFT and the most critical of these errors has the potential to severely impact upon the ability of the method to identify a leaking facepiece.

The high prevalence of human errors in QLFT is indicative of a failure of current resources to educate, control and mitigate errors in this method. This thesis has identified numerous gaps in current administrative and engineering controls which must be filled if false passes are to be reduced and eliminated in workplace fit testing. In this section, two new strategies are proposed for mitigation of human errors in QLFT:

1. Recommendations for new administrative controls
2. Recommendations for improved engineering controls

The proposed new administrative controls are designed to fill gaps in current administrative resources (HSE / Fit2Fit guidance, user instructions and training content), and the recommendations for improved engineering controls to fill the gaps identified in current QLFT equipment. For both strategies, the overall aim is to prevent, reduce or eliminate human errors, and hence false passes, in qualitative fit testing.

7.3.1 Recommendations for new administrative controls

A summary of the gaps identified in current administrative controls is shown in Table 7-1. To address these gaps, a series of new administrative controls are recommended. The aim of these new controls is to prevent, reduce or eliminate the most critical human errors in QLFT and hence mitigate false passes.

Summary of Gaps identified in current administrative controls	Proposed new administrative controls
QLFT user instructions Inconsistent and incomplete control of critical errors across various manufacturers	<ul style="list-style-type: none"> • Provide clear information about critical errors and the risks associated. • Provide clear images to aid component identification. • Add a procedure and checklist for pre-use kit inspection. • Include information about how to check for function and rectify problems. • Provide a 'Daily Check List' for fit testers
HSE INDG479 and Fit2Fit QLFT Companion Limited guidance provided about control of critical errors	<ul style="list-style-type: none"> • Provide additional guidance in HSE INDG479 relating to mitigation of critical errors. • Provide additional guidance in Fit2Fit QLFT companion relating to mitigation of critical errors.
Fit tester training course content Lack of detail and incomplete control of critical errors	<ul style="list-style-type: none"> • Provide improved information in training course content relating to mitigation of critical errors. • Add a 'Daily Check List' practical session to training courses. • Add a fault-finding practical session to training courses.

Table 7-1 Gaps identified in current administrative controls and proposed new controls

To aid dissemination of the proposed new administrative controls, a new Industry Guidance note has been developed. This document is directed towards workplace qualitative fit testers with the aim of filling the gaps in the fragmented and inconsistent advice provided by current administrative resources. The key message provided by this guidance note is the importance of correctly using and applying qualitative fit test equipment to avoid human errors in workplace fit testing.

The new guidance provides information through a multi-level approach:

1. Awareness of how human errors can occur in QLFT
2. Education about the risks of these errors
3. Instruction about how to avoid errors.

The Industry Guidance note and a summary of the conclusions of this thesis have been discussed with policymakers at the Health and Safety Executive and the British Safety Industry Federation (BSIF), the governing board of the Fit2Fit scheme and key stakeholders in industry. Since the Fit2Fit scheme is built upon competence in fit testing, the new guidance would sit well as a supplement to one of the scheme's informative newsletters which are disseminated regularly to the fit testing community. Designed to be manufacturer non-specific, the new Industry Guidance note could further be incorporated as a supplement to existing guidance provided by HSE, Fit2Fit and QLFT kit manufacturer's user instructions.

Large, labelled diagrams, bullet points and clear instructions have been included to ensure that this document is a user-friendly reference for workplace fit testers. Component identification is aided through clearly labelled images and a pre-use checklist is provided to aid proper set-up and inspection of QLFT equipment. Further informative text has been provided to support each set of guidance along with references to existing documents. Attention is drawn to the requirement to assure the wearer's taste response through guidance which highlights key stages in the fit testing process. A summary 'Daily Check List' provides the fit tester with an aide-memoire of the key points along with an auditable record of daily equipment inspection; information that is currently not provided by any QLFT equipment manufacturer in the UK.

7.3.1.1 Proposed Industry Guidance note

Reducing human errors in the use of qualitative fit test equipment

Qualitative fit testing is an important health and safety procedure which must be conducted correctly to ensure reliable results. Fit testing is a process which assesses whether a specific make, model, and size of tight-fitting facepiece is suitable for an individual wearer. The UK qualitative fit test protocol is described in Health and Safety Executive guidance INDG479 and manufacturer's user instructions.

This Industry Guidance note provides supplementary information to fit testers about the safe use and application of qualitative fit test equipment with the aim of reducing human errors. A checklist of key points is provided at the end of this guide which can be used prior to and during each fit testing session.

Qualitative fit test equipment

Qualitative fit test equipment used in the UK should meet the specification in BS ISO 16975-3. A list of major parts included in a qualitative fit test kit and their purpose is described below.

Sensitivity solution:	A dilute solution of Bitrex or Saccharin used during the Sensitivity Test. The contents of this solution are described in BS ISO 16975-3.
Fit test solution:	A more concentrated solution of Bitrex or Saccharin used during the Fit Test. The contents of this solution are described in BS ISO 16975-3.
Sensitivity nebuliser:	Used during the Sensitivity Test stage to produce a fine mist of sensitivity solution. Proper function of this nebuliser is important to ensure that the correct amount of sensitivity aerosol is produced.
Fit test nebuliser:	Used during the Fit Test stage to produce a fine mist of fit test solution. Proper function of this nebuliser is critical to ensure that the correct amount of fit test aerosol is produced.
Hood / collar:	A head-worn enclosure with a clear front panel used during both the Sensitivity Test and Fit Test to contain the sensitivity and fit test aerosol.
Spares:	Qualitative kits are supplied with two spare Venturi and small-gauge wire which may be used for clearing blockages in the nebuliser. See the Troubleshooting section for further information.

Daily Inspection Checks

Fit test equipment needs to be inspected prior to use to ensure that components are clean, present and in good working order. Dirty, missing, or damaged parts may affect proper function and lead to invalid fit test results.

Table A provides details of the checks that should be carried out prior to each use, or if the wearer is unable to detect taste at the sensitivity test or reveal test.

Table A Inspection Checks for qualitative fit test equipment

Major part	Sub-component	What to check
Sensitivity solution	Ampoules or bottles	<ul style="list-style-type: none"> • Within manufacturer's use-by date • Within manufacturer's open time (bottles) OR • Unopened (ampoules) • Free from dirt or debris
Fit test solution	Ampoules or bottles	<ul style="list-style-type: none"> • Within manufacturer's use-by date • Within manufacturer's open time (bottles) OR • Unopened (ampoules) • Free from dirt or debris
Nebulisers*	Bulb	<ul style="list-style-type: none"> • Undamaged, clean and dry • Connected securely to the nebuliser lower body
	Lower body	<ul style="list-style-type: none"> • Undamaged, clean and dry • Connected securely to the bulb • Nozzle free from blockage** • O-ring present (see below) • Venturi present (see below)
	O-ring	<ul style="list-style-type: none"> • Undamaged, clean and dry • Fully seated in nebuliser lower body
	Venturi	<ul style="list-style-type: none"> • Undamaged, clean and dry. • Present and seated fully down onto the Nozzle • Aperture free from blockage**
	Upper body	<ul style="list-style-type: none"> • Undamaged, clean and dry • Fits securely onto the nebuliser lower body
	Top and bottom bungs	<ul style="list-style-type: none"> • Both bungs must be able to be removed when nebuliser is in use
Hood / collar	Hood body	<ul style="list-style-type: none"> • Undamaged, clean, and dry • Able to maintain shape
	Collar (if separate)	<ul style="list-style-type: none"> • Undamaged, clean and dry • Securely attached to hood body • Aligned with front of hood body

Notes: *Location of nebuliser parts is shown in Figures 1 and 2 below. **See Checking and Clearing nebuliser blockages section.

Inspection and set up of nebulisers

Nebulisers are an essential part of the qualitative fit test kit and special attention should be paid to ensure their correct function. The location of the main nebuliser components and sub-components is shown in Figures 1 and 2.



Figure 1 Nebuliser main components



Figure 2 Nebuliser sub-components

Filling sensitivity and fit test nebulisers

It is essential that the correct solutions are used for the sensitivity and fit test stages of the test. Failure to apply the correct solution will invalidate the fit test. Nebulisers must be adequately filled with 5ml of solution to ensure proper aerosol generation.

Always:

- Use the sensitivity solution with the sensitivity nebuliser.
- Use the fit test solution with the fit test nebuliser.
- Ensure that the matching tasting solution is used for both stages of the test
- Ensure that both nebulisers are filled with 5ml of solution.

For kits containing bottles:

Using a clean, dry pipette or other measuring device, transfer 5ml of sensitivity solution to the sensitivity nebuliser. Use a second clean, dry pipette to transfer 5ml of fit test solution to the fit test nebuliser.

For kits containing ampoules:

Fill the sensitivity nebuliser with the contents of TWO 2.5ml ampoules of sensitivity solution. Fill the fit test nebuliser with the contents of TWO 2.5ml ampoules of fit test solution.

Checking the function of nebulisers prior to use

To ensure proper aerosol generation, both nebulisers must be checked prior to each use, periodically during the fit test, and if the wearer is unable to detect taste either at the sensitivity test or the reveal test.

Function test for sensitivity and fit test nebulisers:

1. Ensure that components are inspected and assembled correctly (see Table 1).
2. Ensure that each nebuliser is filled with 5ml of the appropriate solution.
3. Ensure that the bottom and top bungs are removed.
4. Hold the nebuliser in an upright position against a dark background.
5. Squeeze the bulb firmly several times, fully depressing and releasing.
6. A cloud of white aerosol should be visible against the dark background.

If no cloud of aerosol is visible, refer to the section below.

Checking and Clearing Blockages

Blockages may arise from crystallisation of the taste solution and must be cleared to ensure proper aerosol generation. To verify the nebuliser is free from blockages:

- Ensure that the nebuliser is fully cleaned and dried after each use.
- Remove the Venturi from the Nozzle.
- Use the small-gauge wire to rectify blockages. The wire should be used on both the Nozzle (Figure 3) and the Aperture (Figure 4).
- Reassemble the nebuliser and repeat the function test.

If the nebuliser is still not producing aerosol, do not use. Seek replacement parts from the manufacturer.

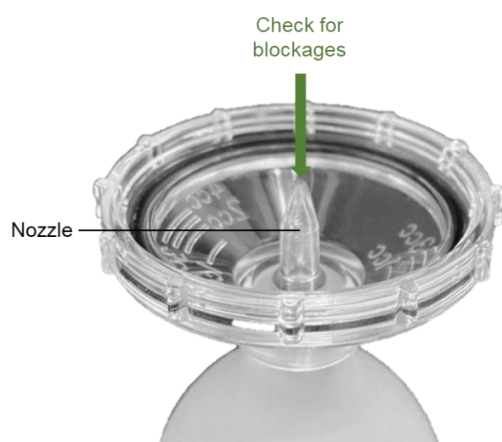


Figure 3 Checking the Nozzle for blockages

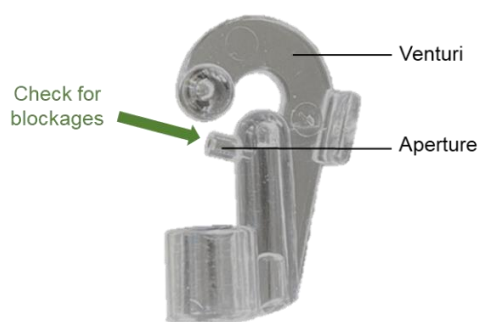


Figure 4 Checking the Aperture for blockages

Managing the wearer's taste response

It is essential for the wearer's taste palate to be properly managed to ensure that a leaking facepiece can be correctly identified. Failure to manage the wearer's taste palate may affect taste detection and invalidate the fit test.

To ensure proper management of the wearer's taste palate, always:

- **Ensure that the room used for fit testing is well ventilated.**

This prevents a build-up of aerosol which may affect the wearer's taste.

- **Ensure that the wearer's palate is clear before the sensitivity test.**

Wearers must avoid smoking, vaping, eating, and drinking for 30 minutes prior to the sensitivity test to ensure that they have a clear palate.

- **Advise the wearer to rinse mouth / clean face after the sensitivity test**

This ensures helps to clear the wearer's palate and removes any residue.

- **Verify that the wearer has a clear palate before commencing the fit test**

This helps to ensure a clear palate before the fit test.

- **Instruct and remind the wearer to breathe through their mouth**

The qualitative fit test method relies on the wearer's ability to detect a leakage using their taste response.

- **Ensure that the sensitivity test is repeated after a failed fit test**

The sensitivity test must always be repeated after a failed fit test since the wearer's taste response must be re-established.

Operation and use of fit test equipment

Fit test equipment must be set up and used in accordance with manufacturer's instructions. The qualitative fit test protocol is described in HSE INDG479.

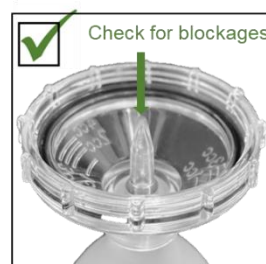
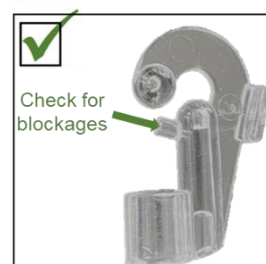
During use, nebulisers must be **held upright with both bungs removed** to ensure proper aerosol generation. The hood must be positioned with the **clear visor facing forwards**. Throughout the fit test, nebulisers must be used in accordance with the INDG479 protocol using the **required number of squeezes every 30 seconds** for the duration of the test.

Cleaning and disinfection of fit test equipment

Fit test equipment must be cleaned and disinfected in accordance with manufacturer's instructions and HSE guidance. Proper cleaning of equipment is essential to remove any residue of taste solution, ensure components are free from blockages, and ensure proper function for the next fit test. See the Checking and Clearing Blockages section for information on how to check and remove nebuliser blockages.

Daily Check List

- ☐ Room well ventilated
- ☐ Wearer not smoked, eaten or drunk 30 mins before test
- ☐ All fit test equipment checked upon set up
- ☐ Nebuliser Aperture and Nozzle checked for blockages
- ☐ Venturi fully seated on the Nozzle
- ☐ Nebulisers filled with 5ml of the correct solution
- ☐ Nebulisers held upright when in use
- ☐ Both bungs removed from nebulisers when in use
- ☐ Nebuliser bulb squeezed fully for every application
- ☐ Nebuliser function checked before and during each test
- ☐ Hood positioned correctly on the wearer
- ☐ Solutions and timings applied as per INDG479 protocol
- ☐ Wearer advised to clean palate after the sensitivity test
- ☐ Wearer's clear palate verified before the fit test
- ☐ Wearer instructed to breathe through mouth
- ☐ Sensitivity test repeated after a failed fit test
- ☐ Equipment cleaned and disinfected after each use



7.3.2 Recommendations for new engineering controls

This thesis also proposes a series of improvements to the engineering controls in QLFT equipment to enable mitigation of human errors at source. Current QLFT kits provide very few controls aimed at mitigating human errors in QLFT, and improvements in this area are urgently needed. Table 7-2 provides a summary of the gaps identified and recommendations for improvements in commercially available QLFT kits. These recommendations are directed at QLFT kit manufacturers but will also be of interest to other stakeholders and regulators.

A range of new controls are proposed which encompass both the use of existing resources and the possibilities for innovation in this field. Simple additional resources such as a reusable 5ml measuring pipette would provide a relatively inexpensive control to aid fit testers in properly filling the nebulisers. Minor re-design of QLFT nebulisers to include an aid such as a bubble-level would provide a visual reminder to the wearer to keep the nebuliser upright. Designing a positive ‘seat’ or permanent fixing for the Venturi would eliminate human errors in this area and aid proper aerosol generation. Fixing the joint between the top and bottom bungs would eliminate the possibility of leaving the bottom bung in place and it is somewhat surprising that this has not already been enacted.

There appear to be numerous possibilities for innovation in the design of QLFT equipment whilst maintaining the fundamental principles of the method. Perhaps the most effective may be a switch from a human-operated manual nebuliser to an automated nebuliser, and this field is ripe for innovation. Automated nebulisers are frequently used in medical applications for drug delivery systems where they are used to deliver accurate doses of drugs to treat a wide variety of respiratory illnesses [10]. Full or semi-automation of aerosol dispensing would be likely to have a high impact on mitigating critical errors relating to nebuliser function and use, since the need for manual application by the fit tester would be removed. Interestingly such a method has been attempted previously with the QFit™ Respirator Fit Tester (TSI Inc, Shoreview, MN) which provided both fully automated and semi-automated options [11]. The product was short-lived however, as the market failed to see the benefits of the automated process at almost double the price of the manual handheld version. Whilst full automation of the QLFT procedure almost entirely removes the need for human operation, such a system does come at an expense,

thus limiting the wide availability which is the current draw of the QLFT method. Success of such a system would require better efforts at education and adoption by the main QLFT manufacturers.

Summary of gaps identified in current engineering controls	Recommended new engineering controls
Inconsistent labelling and differentiation of sensitivity / fit test solution.	<ul style="list-style-type: none"> Standardise labelling / colours across different manufacturers.
Limited controls available to aid adequate filling of nebuliser.	<ul style="list-style-type: none"> Provide 5ml single-use ampoules. <i>or</i> Provide nebulisers with graduated fill markings. <i>or</i> Include 5ml pipette with QLFT kit.
No controls available to ensure removal of bottom bung.	<ul style="list-style-type: none"> Provide fixed connection between top and bottom bungs.
No controls available to ensure proper seating of Venturi.	<ul style="list-style-type: none"> Re-design nebuliser to include a firm seat for the Venturi. <i>or</i> Re-design nebuliser to eliminate the removable Venturi.
No controls available to aid proper administering of fit test solution.	<ul style="list-style-type: none"> Design a fully- or semi-automated QLFT process.
No controls available to assist in accurate fit test timings.	<ul style="list-style-type: none"> Design a fully- or semi-automated QLFT process. Include digital stopwatch with QLFT kit.
No controls available to ensure fit test nebuliser is held upright during use.	<ul style="list-style-type: none"> Fit mini bubble-level to each nebuliser.
No controls available to ensure that fit test nebuliser is actuated with full squeezes.	<ul style="list-style-type: none"> Fully or semi-automated QLFT process. <i>and</i> Device to check proper aerosol generation.
No controls to ensure that hood is undamaged / suitable.	<ul style="list-style-type: none"> Provide single-use disposable hoods.
Limited controls to assist in correct hood positioning.	<ul style="list-style-type: none"> All manufacturers add a collar with shoulders to hood to indicate correct positioning.

Table 7-2 Gaps identified in current engineering controls and proposed new controls

When considering real-world implementation of the recommendations proposed in this thesis, several factors must be taken into account. Additional administrative controls (such as the recommendations for improved HSE guidance) may face resistance given the growing reluctance of regulators to take on new legislative responsibility. The key to countering this resistance will be through discussion of the potential health and safety risks of inaction in this important area. Further market-sensitive opposition to changes in QLFT user instructions and equipment design is to be expected, however the threat of litigation (especially from those based in North American markets) may provide the impetus required to provide better controls. Perhaps the biggest change may be driven through fit testers themselves. The UK fit testing community, especially the 446 Fit2Fit accredited fit testers (figure from October 2021), has strong links and communication channels [12]. Changes to guidance and manufacturer's recommendations are often swiftly disseminated through social media and the strong sense of workplace safety is a common theme. An improvement in knowledge and understanding from within this group through better education could result in significant reduction in fit tester errors and hence improve the safety of workplace fit testing.

7.4 Research impact and dissemination

This thesis makes important contributions to the theoretical and empirical understanding of human errors in qualitative fit testing, which will have wide impacts in the academic world and workplace fit testing programmes. The prevalence and risks of human errors in QLFT should no longer be ignored and the important issues raised in this thesis must be addressed. In order to disseminate the findings of this thesis, specific efforts have been made to engage with both researchers and policymakers in this field. Research work into human errors made by real-world fit testers has been presented to audiences of both stakeholders and academics at a variety of European Section meetings of the International Society of Respiratory Protection (ISRP) [13].

Further presentations and panellist contributions were made in a series of webinars and conferences held online during the Covid-19 pandemic, including to a wider audience of researchers and policymakers from outside the UK [14]. Key findings have been discussed with policymakers within the National Health Service (NHS) and the Department of Health and Social Care (DHSC) through presentations to meetings, and discussions are ongoing as to how best to implement change through policy in the healthcare sector. In terms of effecting change at an individual fit tester level, key findings have been discussed with the Fit2Fit scheme and efforts have been made to influence development of new policy and procedures through the Fit2Fit technical committee of experts. A summary of findings and recommendations for a new Industry Guidance note have been discussed with the British Safety Industry Federation and policymakers from the UK Health and Safety Executive.

7.5 Limitations

Whilst each paper in this thesis has provided its own discussion of limitations, this section discusses these considerations from an overall perspective.

7.5.1 Real-world data

A view of human errors in qualitative fit testing was gained by examining the prevalence and types of errors observed amongst real-world fit testers during Fit2Fit practical assessments. Whilst this study provided a robust method for analysis of data against the Fit2Fit criteria, it had several limitations which must be considered in the wider context.

Firstly, these data were gathered under exam conditions where the fit tester's actions were observed and marked by an assessor. Under these conditions, anxiety due to exam pressure may have had the effect of increasing the number of errors exhibited by fit testers when compared to their normal, unobserved, workplace fit testing activity. Conversely, knowledge that they are being observed and assessed may cause the fit tester to be on 'best behaviour', thus exhibiting fewer errors than they would do in their normal activity. The resulting data from this study is therefore likely to both underestimate and overestimate workplace errors which will largely depend upon how well an individual fit tester performs under exam conditions. Secondly, since the Fit2Fit scheme is voluntary, fit testers attending the Fit2Fit practical assessment had presented themselves for examination. These fit testers may have come from a population that deemed themselves 'good enough' to pass; a population which therefore may exhibit fewer errors when compared to the entire population of UK fit testers. Data more representative of the population of workplace fit testers in the UK would require consideration of this factor whilst maintaining a robust methodology.

Several limitations should also be considered when examining the data presented for the theoretical and practical impacts of human errors on fit test outcome. The first of these is that real-world human errors are unlikely to fall neatly into 'impact' or 'no impact' categories, but rather will sit on a sliding scale of severity. Further studies will be needed to assess where real-world errors may sit on this scale to evaluate overall impact. In addition, real-world errors may be compounded by the effects of other errors leading to

impacts that are greater than the sum of the individual errors. This effect was demonstrated in the compound error *fit test nebuliser not upright and underfilled* whose impact was observed to be far greater than the sum the two individual critical errors. This study is therefore limited to assessment of individual errors, and further studies would be needed to assess and evaluate any compound effect.

7.5.2 Choice of the TSI Portacount quantitative reference method

Use of the TSI Portacount instrument (TSI Incorporated, Shoreview, MN) as the QNFT reference device during practical shortened qualitative fit tests (section 5.4.2) has a number of limitations which should be addressed.

The Portacount was selected for use in this study due to its ease of use, capability of fit testing filtering facepiece respirators, and highly portable nature [15]. Based on condensation particle counting (CPC) technology, the Portacount makes use of environmental (ambient) particles which act as the facepiece challenge agent. During fit testing, the Portacount calculates and reports a fit factor (FF) for each of test exercise as the ratio of particle concentration outside and inside a facepiece [16]. This is given as the equation:

$$FF = \frac{C_{out}}{C_{in}}$$

Equation 7-1 Fit factor calculation

Notes: FF = Fit Factor, where:

C_{out} = challenge aerosol concentration outside the facepiece.

C_{in} = challenge aerosol concentration inside the facepiece

The overall quantitative (QNFF) result is calculated as the harmonic mean of the individual fit factors for each test exercise. The overall QNFF for the HSE INDG479 protocol based on seven test exercises is shown below [16]:

$$QNFF = \frac{7}{\frac{1}{FF_1} + \frac{1}{FF_2} + \frac{1}{FF_3} + \frac{1}{FF_4} + \frac{1}{FF_5} + \frac{1}{FF_6} + \frac{1}{FF_7}}$$

Equation 7-2 Overall quantitative fit factor calculation for the HSE INDG479 protocol based on seven test exercises

Notes: QNFF = Quantitative fit factor, where:

FF_1 = Fit factor for test exercise 1.

FF_2 = Fit factor for test exercise 2, etc.

As with any CPC instrument, the Portacount is unable to differentiate between particles deriving from various sources, and instead reports the C_{in} measurement as the total of all particles detected in the air sampled from inside a facepiece. This measurement includes particles derived from true face seal leakage, as well as extraneous particles arising from filter penetration and the wearer's exhaled breath. Numerous studies have discussed this issue, and it has been estimated that particles in the wearer's breath can range from <0.1 to 4 particles / cm³ [17][18]. Although particle-classifiers such as TSI's N95 technology allow quantitative fit testing of facepieces with filters of less than 99% efficiency, all aerosol-based quantitative methods suffer from under-reporting of QNFF values due to counting of extraneous particles.

During the shortened fit test study, several measures were taken to reduce the impact of extraneous particles on the resulting QNFF measurements:

- Tests were conducted using high-filtering efficiency FFP3 respirators to minimise particle penetration through the filtering material.
- The ambient particle concentration was maintained above 3,000 particles / cm³ to minimise the effects of wearer-generated particles.
- Wearers were asked to refrain from smoking, eating, drinking, and vaping for 60 minutes prior to commencement of the fit test.

Despite these measures, it is likely that some in-mask particles will have derived from sources other than face seal leakage, and that these will have been detected by the Portacount. It must therefore be assumed that the C_{in} concentration measurement will be

biased high and hence the reported QNFF value will be biased low. This means that the true face seal leakage will be higher than reported, and hence the true QNFF values will be closer to the required fit factor value. Any comparative fit tests conducted by the QLFT method should therefore have been less likely to result in a failure. The impact of the errors reported in this study may therefore have been an underestimate of true impact, adding further weight to their importance in determining the effects of human errors.

7.6 Suggestions for future research

By extending the understanding of QLFT to include the impacts of human errors, this thesis opens several avenues for new research in this field. The first direction of research may be to complement the outcomes of this thesis through additional real-world observational studies of QLFT fit testers to further examine human errors. A deeper understanding of the drivers for human errors could be gained through studies which examine the behaviour of fit testers, including their learning and motivational patterns. Research into awareness and understanding of human errors amongst fit testers would provide additional information about the risks in workplace fit testing and support the call for further guidance in this area. Such studies may also seek to better understand the effects of increased time and work pressure upon a fit tester's role, and how these additional demands impact the fit tester's behaviour. The unprecedented increase in the use and fit testing of FFP3 respirators during the Covid-19 pandemic potentially may provide a new and large source of data, especially in healthcare settings. At a governance level, data gathered by healthcare providers may provide information on the hierarchy of management as well as the supervision and training requirements of an in-house fit testing programme in these settings. At a workplace protection level, gaining permission from the Department of Health to access to data regarding workplace Covid-19 infection prevention would allow researchers to estimate the efficacy of workplace fit testing programmes.

A second avenue for new research would be to address the area of subjectivity in QLFT. The highly subjective nature of the QLFT method is well-documented and a number of researchers have reported false response rates to a water placebo [19][20]. The use of a water placebo as a pre-qualifying stage to the sensitivity test was included as part of the original validation of the protocol and has further been recommended by a number of researchers in this field but has yet to gain any traction into guidance or best practice advice [5]. Anecdotal evidence from fit testers within the healthcare sector during the SARS-CoV-2 pandemic has indicated a low-confidence in the wearer's reporting of taste response, with attempts to 'cheat' the outcome of the test [21]. Research into the effectiveness of including a water placebo as part of the QLFT protocol would provide greater confidence in the method and help to mitigate errors for which engineering controls are more difficult, for example assuring the wearer's clear palate. This additional

step is likely to have minimal time and financial impact but would serve to make the QLFT method more reliable and less subjective. An additional avenue for research into subjectivity in QLFT would be to investigate the psychological factors associated with taste detection. This research would provide a foundation for better understanding the complexity of factors that affect the human taste response in QLFT. Research is needed to understand the wide range of human sensitivity levels, taste fatigue, voluntary and involuntary responses, and pressures of time or employment.

The third area for new research stemming from this thesis is innovation in QLFT design. The equipment used for QLFT is rudimentary in design, requires extensive manual input and lacks proper engineering controls – all of which leave this method very vulnerable to human errors. As a result, new research into innovative aerosol delivery and containment methods would be welcome and would serve to control errors at source, making this method more robust for workplace use. In addition, innovative methods of solving the earlier-discussed problems of subjectivity are urgently needed. This could perhaps be achieved through the use of new challenge agents which are detectable by non-human methods such as colour-change indicators or simple electronic sensors. Research would be needed to verify the efficacy of any new equipment, challenge agents or detectors proposed, and this may require a full re-validation against the criteria set out by the American National Standards Institute (ANSI) in ANSI Z88.10 [22].

7.7 References

1. Coffey CC., Lawrence RB., Zhuang Z., Campbell DL., Jensen PA., Myers WR. Comparison of five methods for fit-testing N95 filtering-facepiece respirators. *Applied Occupational and Environmental Hygiene*. 2002; 17(10): 723–730. Available at: DOI:10.1080/10473220290107002
2. Frost S., Mogridge R., Roff M. RR1029 Review of fit test pass criteria for filtering facepieces Class 3 (FFP3) respirators. Health and Safety Executive; 2015. Available at: <https://www.hse.gov.uk/research/rrpdf/rr1029.pdf>
3. Janssen LL., Luinenburg MD., Mullins HE., Nelson TJ. Comparison of three commercially available fit-test methods. *American Industrial Hygiene Association Journal*. 2002; 63(6): 762–767. Available at: DOI:10.1080/15428110208984767
4. Marsh JL. Evaluation of saccharin qualitative fitting test for respirators. *American Industrial Hygiene Association Journal*. 1984; 45(6): 371–376. Available at: DOI:10.1080/15298668491399965
5. McKay RT., Davies E. Capability of respirator wearers to detect aerosolized qualitative fit test agents (sweetener and Bitrex) with known fixed leaks. *Applied Occupational and Environmental Hygiene*. 2000; 15(6): 479–484. Available at: DOI:10.1080/104732200301269
6. Campbell DL., Coffey CC., Lenhart S. Respiratory protection as a function of respirator fitting characteristics and fit-test accuracy. *American Industrial Hygiene Association Journal*. 2001; 62(1): 36–44. Available at: DOI:10.1080/15298660108984607
7. Clayton M., Vaughan N. Fit for purpose? The role of fit testing in respiratory protection. *The Annals of Occupational Hygiene*. 2005; Available at: DOI:10.1093/annhyg/mei046
8. Report of the stakeholders meeting on respiratory protective equipment fit testing. 22nd April 2005. Health and Safety Laboratory; Available at: <https://www.hse.gov.uk/asbestos/assets/docs/stakeholdermeeting220405.pdf>

9. Health and Safety Executive. INDG479 Guidance on respiratory protective equipment (RPE) fit testing. HSE Books; 2019. Available at: <https://www.hse.gov.uk/pubns/indg479.pdf>
10. Carvalho TC., McConville JT. The function and performance of aqueous aerosol devices for inhalation therapy. *Journal of Pharmacy and Pharmacology*. 2016; Available at: DOI:10.1111/jphp.12541
11. TSI Incorporated. QFit Respirator Fit Tester. Available at: <https://tsi.com/discontinued-products/qfit-automatic-respirator-fit-tester-deluxe-kit-805001/> (Accessed: 15 November 2021)
12. British Safety Industry Federation. Respiratory Protection Group meeting minutes October 2021.
13. Long K. False passes in qualitative fit testing. Spring Meeting. International Society for Respiratory Protection. 2019. Available at: <https://www.isrp.com/europe-members/stoop-april-2019/1104-false-passes-in-qualitative-fit-testing>
14. Long K. One size doesn't fit all. 2020 Webinar series. International Society for Respiratory Protection. Available at: <https://www.isrp.com/2020-webinar-recordings>
15. TSI Incorporated. Portacount Pro 8030 and Portacount Pro+ 8038 Respirator Fit Testers operation and service manual. P/N 6001868, revision P, March 2015. Available at: <https://tsi.com/products/respirator-fit-testers/portacount®-respirator-fit-tester-8030/> (Accessed: 4 June 2020)
16. British Standards Institution. BS ISO 16975-3:2017 Respiratory protective devices. Selection, use and maintenance. Part 3: Fit-testing procedures. London: British Standards Institution; 2017.
17. Fairchild CI., Stampfer JF. Particle concentration in exhaled breath. *American Industrial Hygiene Association Journal*. 1987; 48(11): 948–949. Available at: DOI:10.1080/15298668791385868

18. Duling MG., Lawrence RB., Slaven JE., Coffey CC. Simulated workplace protection factors for half-facepiece respiratory protective devices. *Journal of Occupational and Environmental Hygiene*. 2007; 4(6): 420–431. Available at: DOI:10.1080/15459620701346925
19. Mitchell BG., Wells A., McGregor A., McKenzie D. Can homemade fit testing solutions be as effective as commercial products? *Healthcare infection*. December 2012; 17(4): 111–114. Available at: DOI:10.1071/HI12019
20. Fakherpour A., Jahangiri M., Yousefinejad S., Seif M. Feasibility of replacing homemade solutions by commercial products for qualitative fit testing of particulate respirators: a mixed effect logistic regression study. *MethodsX*. 2019; 6: 1313–1322. Available at: DOI:10.1016/j.mex.2019.05.034
21. Regli A., Sommerfield A., Ungern-Sternberg BS. The role of fit testing N95/FFP2/FFP3 masks: a narrative review. *Anaesthesia*. 2021; 76(1): 91–100. Available at: DOI:10.1111/anae.15261
22. American National Standards Institute. ANSI/AIHA Z88.10-2010 Respirator Fit Testing Methods. Fairfax: American Industrial Hygiene Association; 2010.

8 Appendices

8.1 Appendices A-N

Appendix A Keyword combinations for themes

Keywords	Combination
	Theme 1
qualitative fit test* respirator	beta
Bitrex fit test* respirator	false
saccharin fit test* respirator	error
respirator face seal leakage	sensitiv*
	Theme 2
qualitative fit test* respirator	detect*
Bitrex fit test* respirator	subjectiv*
saccharin fit test* respirator	placebo
respirator face seal leakage	threshold
	taste
	wearer
	Theme 3
qualitative fit test* respirator	operat*
Bitrex fit test* respirator	Human error
saccharin fit test* respirator	application
respirator face seal leakage	work
	healthcare
	industry

Table A-1 Keyword combinations for literature survey themes

Notes: Themes: 1) Statistical errors inherent in the QLFT protocol, 2) errors arising from the subjective nature of the QLFT method, 3) human errors in the real-world application of the QLFT method.

Appendix B Fit2Fit Criteria referencing

Fit2Fit ref.	Fit2Fit assessment criteria	HSE INDG479	Fit2Fit QLFT
[1]		[2]	[3]
Section 1: Ensure suitability of facilities			
1a	Ensure room and facilities provided are suitable for the test, why?	H30	F7
Section 2: Preliminary actions prior to performing a qualitative fit test			
2a	Preparing the hood. Checks hood for condition and suitability	H66	F13
2b	Fills the nebulizers with the correct solutions and checks volume	H30	F14
2c	Checking nebuliser functioning	H70	F14
Section 3: Brief wearer on purpose of fit testing and test exercises			
3a	Explains the purpose of fit testing	H42	F14
3b	Explains the fit test exercises and procedure	H71	F15
Section 4: Examine the facepiece for suitability prior to fit testing			
4a(f)	FFP: Visually verify condition of straps, facepiece and valves	H37	F11
4b(f)	FFP: Identify make, model, size and classification	H37	F11
4a(h)	Half Mask: Ensure particle filters are correctly fitted	H37	F11
4b(h)	Half Mask: Visually verify condition of straps, facepiece and valves	H37	F11
4c	Half Mask: Identify make, model, size and classification	H37	F11
Section 5: Confirm the wearer has presented themselves as required			
5a	Confirms wearer is clean shaven	H45	F8
5b	Confirms wearer has not smoked, eaten, drunk for the required period	H43	F12
5c	Questions the wearer about the use of other PPE.	H44	F11
5d	Takes appropriate action to the response of the wearer to above	H44	F11

Section 6: Conduct sensitivity test			
6a	Position hood correctly, sensitivity test	H67	F13
6b	Instruct wearer to breathe through mouth and indicate taste	H68	F14
6c	Verify correct function of nebuliser before sensitivity test	H67	F14
6d	Administers solution and count squeezes	H68	F13
6e	Takes appropriate corrective actions (where applicable)	H68	F13
6f	Advises wearer to rinse mouth and clean face after sensitivity test	H69	F13
6g	Explains the fit test exercises and procedure	H71	F15
Section 7: Recognise wearer related factors prior to a fit test			
7a	Demonstrates correct method of donning an RPD including a fit check	H48	F10
7b	Observes the wearer don the facepiece and examines the fit	H49	F10
7c	Ensures fit check is performed	H11	F10
7d	Identifies a poorly fitting or fitted respiratory protective device	H48	F10
7e	Apply corrective actions for a poorly fitted RPD	H48	F10
Section 8: Conduct fit test			
8a	Verify palette is clear of taste	H69	F12
8b	Position hood correctly, fit test	H67	F13
8c	Instruct wearer to breathe through their mouth.	H68	F3
8d	Administer solution as per sensitivity level	H68	F13
8e	Note time/start stopwatch	H30	F19
8f	Maintains control of the wearer during the test	H71	F15
8g	Monitors wearer and equipment to ensure protocol as HSE INDG479	H71	F15
8h	Ensures exercises are completed to standard set out in HSE INDG479	H76	F19
8i	Administer top-ups correctly throughout test.	H30	F19
8j	Verify correct function of nebuliser during the test	H70	F14
Section 9.1: Response if tasted (i.e., a fail)			
9.1a	Look for obvious leaks following failed fit test	H30	F16
9.1b	Record fit test result, details of the RPD and advise the wearer	H26	F16
9.1c	Advise to rinse mouth and clean face	H43	F16
9.1d	Re-Inspect RPD following failed fit test	H37	F16
9.1e	Repeat test sensitivity Test	H69	F16
9.1f	Repeat test process in full, including all exercises	H69	F16

Section 9.2: Response if nothing tasted (i.e., a Pass)			
9.2a	Record fit test result, details of the RPD, and advise the wearer	H26	F18
9.2b	Advise wearer of result, implications and limitations	H73	F17
Section 10: Post Fit Test procedures			
10a	Clean Hood	H67	F16
10b	Dispose of solutions	H67	F16
10c	Rinse and clean nebulisers	H67	F16

Table B-1 Fit2Fit assessment criteria referenced against HSE INDG479 and Fit2Fit qualitative companion

Notes: ‘Fit2Fit ref’ is the criteria used in British Safety Industry Federation Fit2Fit scheme practical qualitative assessment papers [1]. ‘Fit2Fit QLFT’ is the Fit2Fit Qualitative companion [2]. HSE INDG479 [3].

Appendix C Sources of error during Fit2Fit QLFT practical assessments, 2009 – 2019

Fit2Fit ref. [1]	Fit2Fit Assessment Criteria	% (n)
Section 1: Ensure suitability of facilities		
1a	Ensure room and facilities provided are suitable for the test, and why?	8.6 (55)
Section 2: Preliminary actions prior to performing a qualitative fit test		
2a	Preparing the hood. Checks hood for condition and suitability	2.8 (18)
2b	Fills the nebulizers with the correct solutions and checks volume	2.5 (16)
2c	Checking nebuliser functioning	13.3 (85)
Section 3: Brief wearer on purpose of fit testing and test exercises		
3a	Explains the purpose of fit testing	20.3 (130)
3b	Explains the fit test exercises and procedure	19.1 (122)
Section 4: Examine the facepiece for suitability prior to fit testing		
4a(f)	FFP: Visually verify condition of straps, facepiece and valves	8.6 (55)
4b(f)	FFP: Identify make, model, size and classification	8.1 (52)
4a(h)	Half Mask: Ensure particle filters are correctly fitted	18.0 (115)
4b(h)	Half Mask: Visually verify condition of straps, facepiece and valves	17.5 (112)
4c	Half Mask: Identify make, model, size and classification	10.6 (68)
Section 5: Confirm the wearer has presented themselves as required		
5a	Confirms wearer is clean shaven	6.1 (39)
5b	Confirms wearer has not smoked, eaten, drunk for the required period	12.1 (77)
5c	Questions the wearer about the use of other PPE.	24.4 (156)
5d	Takes appropriate action to the response of the wearer to above	6.1 (39)
Section 6: Conduct sensitivity test		
6a	Position hood correctly, sensitivity test	5.9 (38)
6b	Instruct wearer to breathe through mouth and indicate taste	8.3 (53)
6c	Verify correct function of nebuliser before sensitivity test	12.7 (81)
6d	Administers solution and count squeezes	7.2 (46)
6e	Takes appropriate corrective actions (where applicable)	3.8 (24)
6f	Advises wearer to rinse mouth and clean face after sensitivity test	24.3 (155)
6g	Explains the fit test exercises and procedure	13.1 (84)

Section 7: Recognise wearer related factors prior to a fit test		
7a	Demonstrates correct method of donning an RPD including a fit check	31.3 (200)
7b	Observes the wearer don the facepiece and examines the fit	5.9 (38)
7c	Ensures fit check is performed	32.2 (206)
7d	Identifies a poorly fitting or fitted respiratory protective device	9.5 (61)
7e	Apply corrective actions for a poorly fitted RPD	7.8 (50)
Section 8: Conduct fit test		
8a	Verify palette is clear of taste	21.0 (134)
8b	Position hood correctly, fit test	4.9 (31)
8c	Instruct wearer to breathe through their mouth.	10.6 (68)
8d	Administer solution as per sensitivity level	18.2 (116)
8e	Note time/start stopwatch	12.8 (82)
8f	Maintains control of the wearer during the test	9.2 (59)
8g	Monitors wearer and equipment to ensure protocol as HSE INDG479	10.0 (64)
8h	Ensures exercises are completed to standard set out in HSE INDG479	9.5 (61)
8i	Administer top-ups correctly throughout test.	17.4 (111)
8j	Verify correct function of nebuliser during the test	25.2 (161)
Section 9.1: Response if tasted (i.e., a fail)		
9.1a	Look for obvious leaks following failed fit test	11.6 (74)
9.1b	Record fit test result, details of the RPD and advise the wearer	11.1 (71)
9.1c	Advise to rinse mouth and clean face	11.7 (75)
9.1d	Re-Inspect RPD following failed fit test	4.1 (26)
9.1e	Repeat test sensitivity Test	23.2 (148)
9.1f	Repeat test process in full, including all exercises	1.4 (9)
Section 9.2: Response if nothing tasted (i.e., a Pass)		
9.2a	Record fit test result, details of the RPD, and advise the wearer	12.2 (78)
9.2b	Advise wearer of result, implications, and limitations	5.5 (35)
Section 10: Post Fit Test procedures		
10a	Clean Hood	0.5 (3)
10b	Dispose of solutions	0.3 (2)
10c	Rinse and clean nebulisers	0.2 (1)

Table C-1 Sources of error during Fit2Fit QLFT practical assessments, 2009 – 2019

Notes: $N = 639$. % (n) = percentage (number) of assessment papers that recorded an error.

Appendix D Theoretical impacts of errors for Fit2Fit QLFT practical assessment criteria

Fit2Fit ref. [1]	Fit2Fit practical assessment criteria [1]	Error scenario	Reasoning	Key principle affected	Impact
Ensure suitability of facilities					
1a	Ensure room and facilities provided are suitable for the test, and why?	Room poorly ventilated (B,F)	Challenge agent present in the test environment, wearer's palate is not clear of taste before start of test.	1	Y
Preliminary actions prior to performing a qualitative fit test					
2a	Preparing the hood. Checks hood for condition and suitability	Hood damaged / unsuitable (H,B,F,M)	Hood is too large / leaky, challenge aerosol therefore is not properly contained.	3	Y
2b	Fills the nebulisers with the correct solutions and checks volume	Fit test nebuliser filled with sensitivity solution (H,B,F,M)	Sensitivity solution used for the fit test, facepiece is not properly challenged.	2	Y
		Sensitivity nebuliser filled with fit test solution	Wearer experiences a very strong taste immediately; test is void and must be repeated.		N
		Fit test nebuliser under-filled (H,B,F,M)	Fit test nebuliser fill volume too low, nebuliser function is affected resulting in low challenge aerosol concentration.	2	Y
2c	Checking nebuliser functioning	Sensitivity nebuliser clogged	Sensitivity nebuliser function is affected, sensitivity threshold is reported high. Challenge aerosol concentration too high at fit test.		N*
		Sensitivity nebuliser bottom bung not removed	Sensitivity nebuliser function is affected, sensitivity threshold is reported high. Challenge aerosol concentration too high at fit test.		N*
		Sensitivity nebuliser Venturi mis-seated	Sensitivity nebuliser function is affected, sensitivity threshold is reported high. Challenge aerosol concentration too high at fit test.		N*
		Fit test nebuliser clogged (H,B,F)	Fit test nebuliser function is affected, challenge aerosol concentration is too low.	2	Y
		Fit test nebuliser bottom bung not removed (F,M)	Fit test nebuliser function is affected, challenge aerosol concentration is too low.	2	Y
		Fit test nebuliser Venturi mis-seated (O)	Fit test nebuliser function is affected, challenge aerosol concentration is too low.	2	Y

	Brief wearer on purpose of fit testing and test exercises				
3a	Explains the purpose of fit testing	Purpose of test not explained to wearer	Wearer does not understand the importance of fit testing.		N
3b	Explains the fit test exercises and procedure	Exercises / procedure not explained to wearer	Wearer does not understand the test procedure and need to conduct exercises correctly.		N
	Examine the facepiece for suitability prior to fit testing				
4a(f)	FFP: Visually verify condition of straps, facepiece and valves	FFP not visually checked	Faulty or poorly maintained respirator. Leakage through components leading to a fit test fail.		N
4b(f)	FFP: Identify make, model, size and classification	FFP misidentified	Incorrect facepiece details recorded		N
4a(h)	Half Mask: Ensure particle filters are correctly fitted	Incorrect filters fitted to half mask	Gas / vapour filters used instead of particulate filters. Leakage through filters leading to a fit test fail.		N
4b(h)	Half Mask: Visually verify condition of straps, facepiece and valves	Half mask not visually checked	Faulty or poorly maintained respirator. Leakage through components leading to a fit test fail.		N
4c	Half Mask: Identify make, model, size and classification	Half mask misidentified	Incorrect facepiece details recorded		N
	Confirm the wearer has presented themselves as required				
5a	Confirms wearer is clean shaven	Wearer is not clean shaven	RPD does not seal adequately to wearer's face, face seal leakage leading to a fit test fail.		N
5b	Confirms wearer has not smoked, eaten, drunk for the required period	Wearer has smoked, eaten or drunk before test (H,B,F,M)	Wearer's palate is not clear of taste before the start of the test.	1	Y
5c	Questions the wearer about the use of other PPE.	Wearer not questioned about use of other PPE	Other PPE may interfere with fit leading to a fit test fail.		N
5d	Takes appropriate action to the response of the wearer to above	Appropriate action not taken in response	Appropriate action not taken in response		N
	Conduct sensitivity test				
6a	Position hood correctly, sensitivity test	Hood positioned incorrectly	Sensitivity aerosol not properly contained; sensitivity threshold is reported high. Challenge aerosol concentration too high at fit test.		N*
6b	Instruct wearer to breathe through mouth and indicate taste	Wearer not instructed to breathe through mouth	Sensitivity threshold is reported high. Challenge aerosol concentration too high at fit test.		N*
		Wearer not instructed to indicate detection of taste	Sensitivity threshold is reported high. Challenge aerosol concentration too high at fit test.		N*

6c	Verify correct function of nebuliser before sensitivity test	Sensitivity nebuliser blocked / poorly functioning	Sensitivity nebuliser function is affected, sensitivity threshold is reported high, challenge aerosol concentration too high at fit test.		N*
6d	Administers solution and count squeezes	Sensitivity level reported low, fit test solution under-administered (H,B,F,M)	Sensitivity threshold is reported low. Challenge aerosol concentration too low.	2	Y
		Sensitivity solution administered / counted high	Sensitivity threshold is reported high. Challenge aerosol concentration too high at fit test.		N*
		Sensitivity nebuliser not actuated with full squeezes	Sensitivity threshold is reported high. Challenge aerosol concentration too high at fit test.		N*
6e	Takes appropriate corrective actions (where applicable)	Appropriate corrective action not taken	Appropriate corrective action not taken		N
6f	Advises wearer to rinse mouth and clean face after sensitivity test	Wearer not advised to rinse mouth / clean face (H,B,F,M)	Wearer's palate is not clear of taste before the start of the test.	1	Y
6g	Explains the fit test exercises and procedure	Exercises / procedure not explained to wearer	Wearer does not understand the test procedure and need to conduct exercises correctly.		N
Recognises wearer related factors prior to carrying out a fit test					
7a	Demonstrates correct method of donning an RPD including a fit check	RPD donning not properly demonstrated	Facepiece donned incorrectly, face seal leakage leading to fit test fail.		N
7b	Observes the wearer don the facepiece and examines the fit	Donning not observed	Facepiece donned incorrectly, face seal leakage leading to fit test fail.		N
7c	Ensures fit check is performed	Fit Check not performed	Facepiece donned incorrectly, face seal leakage leading to fit test fail.		N
7d	Identifies a poorly fitting or fitted respiratory protective device	Poorly fitting facepiece not identified	Facepiece donned incorrectly, face seal leakage leading to fit test fail.		N
7e	Apply corrective actions for a poorly fitted RPD	Corrective actions not taken	Facepiece donned incorrectly, face seal leakage leading to fit test fail.		N
Conduct fit test					
8a	Verify palette is clear of taste	Clear palate not verified (H,B,F,M)	Wearer's palate is not clear of taste before the start of the test.	1	Y
8b	Position hood correctly, fit test	Hood positioned incorrectly (H,B,F,M)	Hood leaks, fit test aerosol is not properly contained, challenge aerosol concentration too low.	3	Y

8c	Instruct wearer to breathe through their mouth.	Wearer not instructed to breathe through mouth (H,B,F,M)	Wearer unable to taste, facepiece leakage is not detected.	1	Y
8d	Administer solution as per sensitivity level	Fit test solution under-administered (H,B,F,M)	Challenge aerosol concentration is too low.	2	Y
		Fit test solution over-administered	Challenge aerosol concentration is too high.		N*
8e	Note time/start stopwatch	Fit test timing incorrect (H,B,F,M)	Test time / timing of top up puffs too long, challenge aerosol concentration is too low.	2	Y
8f	Maintains control of the wearer during the test	Wearer not controlled during fit test	Uncontrolled talking / movement. Increased facepiece leakage leading to a fit test fail.		N
8g	Monitors wearer and equipment through test to ensure correct protocol HSE INDG479	Fit test nebuliser clogged (H,B,F,M)	Fit test nebuliser function is affected, challenge aerosol concentration is too low.	2	Y
		Fit test nebuliser bottom bung not removed (F,M)	Fit test nebuliser function is affected, challenge aerosol concentration is too low.	2	Y
		Fit test nebuliser Venturi mis-seated (O)	Fit test nebuliser function is affected, challenge aerosol concentration is too low.	2	Y
8h	Ensures exercises are performed & completed to standard set out in HSE INDG479	Exercises / procedure not conducted correctly	Wearer does not understand the test procedure and need to conduct exercises correctly.		N
8i	Administer top-ups correctly throughout test.	Fit test solution under-administered (H,B,F,M)	Challenge aerosol concentration is too low.	2	Y
		Number of top up squeezes too high	Challenge aerosol concentration is too high.		N*
		Fit test nebuliser not upright (H,B,F,M)	Challenge aerosol concentration is too low.	2	Y
8j	Verify correct function of nebuliser during the test	Fit test nebuliser clogged (H,B,F,M)	Fit test nebuliser function is affected, challenge aerosol concentration is too low.	2	Y
		Fit test nebuliser bottom bung not removed (F,M)	Fit test nebuliser function is affected, challenge aerosol concentration is too low.	2	Y
		Fit test nebuliser Venturi mis-seated (O)	Fit test nebuliser function is affected, challenge aerosol concentration is too low.	2	Y
		Fit test nebuliser not actuated with full squeezes (H,B,F)	Fit test nebuliser function is affected, challenge aerosol concentration is too low.	2	Y

	Response if tasted (i.e., a fail)				
9.1a	Look for obvious leaks following a failed fit test	Leaks not visually checked following failed fit test	Facepiece donned incorrectly, face seal leakage leading to fit test fail.		N
9.1b	Record the fit test result, details of the RPD and advise the wearer	Poor / limited recording of data	Incorrect / missing data		N
9.1c	Advise to rinse mouth and clean face	Wearer not advised to rinse mouth / clean face (H,B,F,M)	Wearer's palate is not clear of taste before the start of a new test.	1	Y
9.1d	Re-Inspect RPD following a failed fit test	RPD not re-inspected following failed fit test	Faulty or poorly maintained respirator. Leakage through components leading to a fit test fail.		N
9.1e	Repeat test sensitivity Test	Sensitivity test not repeated (H,B,F,M)	Wearer's palate is not clear of taste before the start of the test.	1	Y
9.1f	Repeat test process in full, including all exercises	Fit test not repeated in full	Fit test not conducted as INDG479		N
	Response if nothing tasted (i.e., a Pass)				
9.2a	Record the result, including details of the RPD, wearer and other PPE	Poor / limited recording of data	Incorrect / missing data		N
9.2b	Advise the wearer of the result, implications and limitations	Wearer not advised of result, implications and limitations	Use of incorrect facepiece / need for new fit test.		N
	Post Fit Test procedures				
10a	Clean Hood	Hood not cleaned (H,B,F,M)	Wearer's palate is not clear of taste before the start of the test.	1	Y
10b	Dispose of Solutions	Solutions not discarded	Solutions contaminated.		N
10c	Rinse & Clean nebulisers.	Fit test nebuliser clogged (H,B,F,M)	Fit test nebuliser function is affected, challenge aerosol concentration is too low.	2	Y

Table D-1 Theoretical impacts of errors for Fit2Fit QLFT practical assessment criteria

Notes: Letters shown in parentheses indicate source of reference, H = HSE INDG479 [3], B = BS ISO 16975-3 [4], F = Fit2Fit Qualitative Companion [2], M = 3M FT30 user instruction [5], O = Other information supplied by manufacturer [6]. Key principles: 1) a sensitivity check to assure that the respirator wearer is able to taste a low level of saccharin should a leakage occur; 2) procedures to assure that the aerosol is generated in a consistent and controlled manner; 3) steps to assure that the test atmosphere is contained around the respirator and is controlled [7]. Impact: Y = theoretical impact on fit test outcome that may lead to a false pass, N = no theoretical potential for a false pass, N* = no theoretical potential for a false pass but may lead to false fail.

Appendix E Theoretical impacts of errors applied to results of real-world study

Fit2Fit ref [1]	Fit2Fit assessment criteria [1]	% (n)	Critical Error
Section 1: Ensure suitability of facilities			
1a	Ensure room and facilities provided are suitable for the test, and why?	8.6 (55)	1
Section 2: Preliminary actions prior to performing a qualitative fit test			
2a	Preparing the hood. Checks hood for condition and suitability	2.8 (18)	3
2b	Fills the nebulizers with the correct solutions and checks volume	2.5 (16)	2
2c	Checking nebuliser functioning	13.3 (85)	2
Section 3: Brief wearer on purpose of fit testing and test exercises			
3a	Explains the purpose of fit testing	20.3 (130)	
3b	Explains the fit test exercises and procedure	19.1 (122)	
Section 4: Examine the facepiece for suitability prior to fit testing			
4a(f)	FFP: Visually verify condition of straps, facepiece and valves	8.6 (55)	
4b(f)	FFP: Identify make, model, size and classification	8.1 (52)	
4a(h)	Half Mask: Ensure particle filters are correctly fitted	18.0 (115)	
4b(h)	Half Mask: Visually verify condition of straps, facepiece and valves	17.5 (112)	
4c	Half Mask: Identify make, model, size and classification	10.6 (68)	
Section 5: Confirm the wearer has presented themselves as required			
5a	Confirms wearer is clean shaven	6.1 (39)	
5b	Confirms wearer has not smoked, eaten, drunk for the required period	12.1 (77)	1
5c	Questions the wearer about the use of other PPE.	24.4 (156)	
5d	Takes appropriate action to the response of the wearer to above	6.1 (39)	
Section 6: Conduct sensitivity test			
6a	Position hood correctly, sensitivity test	5.9 (38)	
6b	Instruct wearer to breathe through mouth and indicate taste	8.3 (53)	
6c	Verify correct function of nebuliser before sensitivity test	12.7 (81)	
6d	Administers solution and count squeezes	7.2 (46)	2
6e	Takes appropriate corrective actions (where applicable)	3.8 (24)	
6f	Advises wearer to rinse mouth and clean face after sensitivity test	24.3 (155)	1
6g	Explains the fit test exercises and procedure	13.1 (84)	

Section 7: Recognise wearer related factors prior to a fit test			
7a	Demonstrates correct method of donning an RPD including a fit check	31.3 (200)	
7b	Observes the wearer don the facepiece and examines the fit	5.9 (38)	
7c	Ensures fit check is performed	32.2 (206)	
7d	Identifies a poorly fitting or fitted respiratory protective device	9.5 (61)	
7e	Apply corrective actions for a poorly fitted RPD	7.8 (50)	
Section 8: Conduct fit test			
8a	Verify palette is clear of taste	21.0 (134)	1
8b	Position hood correctly, fit test	4.9 (31)	3
8c	Instruct wearer to breathe through their mouth.	10.6 (68)	1
8d	Administer solution as per sensitivity level	18.2 (116)	2
8e	Note time/start stopwatch	12.8 (82)	2
8f	Maintains control of the wearer during the test	9.2 (59)	
8g	Monitors wearer and equipment to ensure protocol as HSE INDG479	10.0 (64)	2
8h	Ensures exercises are completed to standard set out in HSE INDG479	9.5 (61)	
8i	Administer top-ups correctly throughout test.	17.4 (111)	2
8j	Verify correct function of nebuliser during the test	25.2 (161)	2
Section 9.1: Response if tasted (i.e., a fail)			
9.1a	Look for obvious leaks following failed fit test	11.6 (74)	
9.1b	Record fit test result, details of the RPD and advise the wearer	11.1 (71)	
9.1c	Advise to rinse mouth and clean face	11.7 (75)	1
9.1d	Re-Inspect RPD following failed fit test	4.1 (26)	
9.1e	Repeat test sensitivity Test	23.2 (148)	1
9.1f	Repeat test process in full, including all exercises	1.4 (9)	
Section 9.2: Response if nothing tasted (i.e., a Pass)			
9.2a	Record fit test result, details of the RPD, and advise the wearer	12.2 (78)	
9.2b	Advise wearer of result, implications and limitations	5.5 (35)	
Section 10: Post Fit Test procedures			
10a	Clean Hood	0.5 (3)	1
10b	Dispose of solutions	0.3 (2)	
10c	Rinse and clean nebulisers	0.2 (1)	2

Table E-1 Theoretical impacts of errors applied to results of real-world study

Notes: $N = 639$. % (n) = percentage (number) of fit testers that made an error against each criterion. Key principles: 1) a sensitivity check to assure that the respirator wearer is able to taste a low level of saccharin should a leakage occur; 2) procedures to assure that the aerosol is generated in a consistent and controlled manner; 3) steps to assure that the test atmosphere is contained around the respirator and is controlled [7].

Appendix F Anatomy of BS ISO 16975-3 qualitative fit test nebuliser

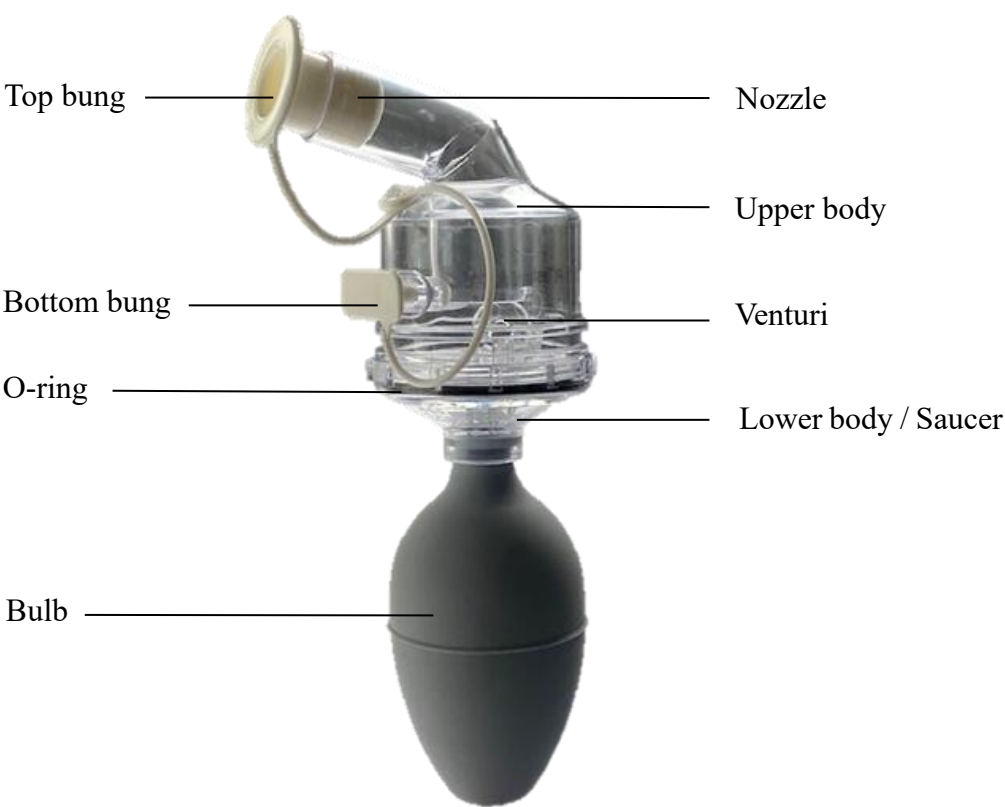


Figure F-1 Anatomy of BS ISO 16975-3 qualitative fit test nebuliser

Appendix G Informed consent form

Title of the project:	Human errors in qualitative fit testing
Name of the researcher:	Karen Long
Researcher's contact details:	karen.long@cranfield.ac.uk
Participant number:	(assigned randomly)
Date:	

Part 1: Participation and withdraw

1. I confirm that I have been informed about this research project and I agree to take part.
2. I understand that all personal information I provide will be treated with confidence and my name will not be used in any report, publication, or presentation.
3. I have been provided with a participant number as shown above. The researcher(s) will record data against my participant number instead of recording my name. The file linking my name to my participant number will be accessible only to the main researchers and will be securely destroyed after two years.
4. I understand that I can withdraw from this project at any stage by informing a member of the research team, for whom contact details have been provided. I also understand that I can withdraw my data for a period of up to two years from today, as after this time it will not be possible to identify my individual data from the aggregated results as explained in point 3.

Part 2: Data security

- I understand that the data I provide will be used by Cranfield University for the purpose of research. The data will be stored on the University's network that can only be accessed by authorised users, in line with UK Data Protection Act 2018. Data will also be stored off network on a local computer, secured via password protection.

Part 3: Use of data

- Anonymised fit test data will be securely stored on the Cranfield University network and a local computer as csv files. These files will be securely deleted as soon as they are no longer required.
- Analytical software will be used to aggregate the results of the research and every reasonable step will be taken to anonymise the data.

I understand that the aggregated data will be published in support of the research findings.

☐

I confirm that I have read and understand the information provided on this form and give my consent to taking part in this research.

Participant's signature:		Date:	
Participant's name:			
Researcher's signature:		Date:	

One copy of this form must be given to the participant and one copy held by the researcher.

Outline of the study

This study will investigate whether the results of qualitative fit tests are affected by faults in the set up and use fit testing equipment.

Fit testing is a requirement for all workers that need to wear disposable, half or full-face masks. This study will use only the disposable (FFP3) type of mask which is typically worn by workers in construction, industry, and healthcare. The fit tests conducted in this study will be operated by a trained, competent fit tester who is accredited to the British Safety Industry Federation (BSIF) Fit2Fit scheme. Further details about this accreditation scheme can be found here: www.fit2fit.org

The procedure used in this study is a shortened version of the protocol which is provided by the Health and Safety Executive (HSE). Further details about the HSE protocol can be found here: <https://www.hse.gov.uk/pubns/indg479.pdf>

Precautions will be taken to prevent the transmission of the Covid-19 virus using the protocol set out by HSE which can be found here: <https://www.hse.gov.uk/coronavirus/ppe-face-masks/face-mask-ppe-rpe.htm#minimisetherisk>

What you can expect

- You will be asked to read this document which provides an outline of the study, and then sign the consent form to indicate your understanding and willingness to participate if you are happy to proceed.
- The fit tester will provide a verbal description of the purpose, scope and procedure of the study. You will have the opportunity to ask any questions you may have.
- You will undergo a sensitivity screening test to check your level of taste response to a bitter-tasting spray containing a very diluted solution of Bitrex. You will be asked to wipe your face and have a drink of water to remove any residual taste.
- You will be provided with a new, clean disposable FFP3 mask and given instruction on how to correctly don and check the mask for fit.
- The fit tester will measure the amount of leakage through the mask in a two-minute period. You may remain seated or standing, but you will be asked not to talk or adjust your mask during this period.
- You will be asked to put on a large hood which will rest comfortably on your shoulders.
- The fit tester will conduct a series of tests using a spray, and you should indicate immediately if you are able to detect the taste of the spray. The fit tester will ask you to breathe through your mouth at all times, with your tongue extended.
- After each test, you will be asked to remove and discard the FFP3 mask and have a drink of water to ensure that you have no residual taste.
- You must inform the fit tester immediately if you feel unwell during any of the tests.
- You may withdraw from this study at any time.

Appendix H Results of aerosol mass concentration studies

Test	Scenario, mass aerosol concentration, PM2.5, mg/m ³												
	Control	1	2	3	4	5	6	7	8	9	10	11	12
1	11.80	10.70	11.00	2.78	2.95	7.02	9.17	5.87	12.40	4.82	4.53	4.43	6.38
2	8.45	11.80	11.50	2.28	5.64	4.66	10.00	6.24	9.14	3.50	4.27	2.75	5.24
3	12.20	13.60	9.42	2.55	2.36	3.27	9.48	6.12	9.45	4.04	2.96	4.37	6.44
4	9.11	13.90	13.60	3.20	4.24	3.32	6.64	6.13	9.75	2.27	2.47	2.70	5.82
5	10.30	10.60	8.13	3.19	3.33	3.58	6.51	6.89	11.20	3.54	2.20	3.56	6.25
6	11.20	11.60	10.80	2.84	3.51	5.02	5.90	7.23	8.92	2.50	4.63	1.82	4.51
7	11.60	9.67	8.21	2.05	4.56	4.09	5.83	6.27	10.30	2.34	3.81	2.97	5.70
8	11.40	10.70	9.44	1.54	4.03	4.07	5.92	7.17	13.20	3.96	2.51	2.79	4.24
9	9.78	12.10	11.70	1.46	3.59	4.77	6.49	8.07	10.30	3.39	2.98	2.82	4.49
10	8.75	11.50	9.96	1.57	3.64	5.70	6.92	7.17	10.10	3.44	5.79	2.32	4.99
11	16.20	11.60	12.60	1.84	5.61	5.16	8.16	9.02	8.96	3.97	5.76	3.28	5.25
12	17.60	12.90	10.50	2.22	4.26	5.04	6.85	5.72	10.90	2.21	3.47	2.36	4.81
13	13.90	10.00	10.20	2.01	5.01	5.01	8.21	7.28	9.92	2.22	3.79	2.37	5.08
14	12.40	9.45	11.70	1.94	3.45	5.47	7.25	8.03	8.53	3.09	3.98	3.55	4.68
15	9.14	10.80	9.78	1.51	2.62	4.55	8.35	6.89	8.66	3.90	4.12	3.51	4.16
16	9.45	8.39	13.60	1.96	4.55	3.55	6.09	7.04	9.96	3.39	3.43	3.61	4.75
17	9.75	12.10	9.38	1.95	4.82	4.69	8.51	7.19	10.10	4.28	3.64	5.38	6.61
18	11.20	7.69	12.90	1.51	4.45	3.01	8.50	8.38	11.00	1.00	4.45	4.01	5.36
19	8.92	11.20	10.90	1.96	4.68	5.43	8.62	8.06	7.75	2.60	5.25	4.14	5.79
20	10.30	10.10	13.50	2.04	4.04	4.65	7.74	7.19	9.69	2.77	4.88	3.97	5.41
<i>M</i>	11.17	11.02	10.94	2.12	4.07	4.60	7.56	7.10	10.01	3.16	3.95	3.34	5.30
<i>SD</i>	2.43	1.57	1.69	0.54	0.90	0.97	1.28	0.89	1.30	0.92	1.03	0.88	0.75
<i>Mdn</i>	10.75	11.00	10.85	1.99	4.14	4.68	7.50	7.17	9.94	3.39	3.90	3.40	5.25
<i>Skew</i>	1.39	-0.18	0.14	0.78	-0.08	0.38	0.24	0.36	0.77	-0.41	0.12	0.42	0.28

Table H-1 Results of aerosol mass concentration studies

Notes: $N = 20$. Error scenarios: 1) Fit test nebuliser filled with sensitivity solution, 2) Fit test nebuliser under-filled, 3) Fit test nebuliser clogged, 4) Fit test nebuliser bottom bung not removed, 5) Fit test nebuliser Venturi mis-seated, 6) Fit test solution under-administered, 7) Fit test timing incorrect, 8) Fit test nebuliser not upright, 9) Fit test nebuliser not upright and under-filled, 10) Fit test nebuliser not actuated with full squeezes, 11) Hood damaged / unsuitable, 12) Hood positioned incorrectly.

Appendix I Results of shortened fit test studies

Test	Control		1		2		3		4		5		6		7		8		9		10		11		12	
	QNFF	QLFT	QNFF	QLFT	QNFF	QLFT	QNFF	QLFT	QNFF	QLFT	QNFF	QLFT	QNFF	QLFT	QNFF	QLFT	QNFF	QLFT	QNFF	QLFT	QNFF	QLFT	QNFF	QLFT	QNFF	QLFT
1	56	F	62	P*	66	F	68	P	71	P*	84	P	70	P*	75	F	74	F	77	P	69	F	69	F	89	F
2	61	F	66	P*	90	F	77	P	77	F	86	P	73	F	81	F	69	F	65	F	84	P*	86	F	82	P*
3	66	F	64	P*	77	F	61	P	53	F	69	F	66	F	63	F	87	F	72	P	76	P*	84	F	65	P*
4	76	F	88	P*	72	F	80	P	88	F	73	P	57	F	84	P*	80	F	75	P	66	P*	74	P*	76	P
5	72	F	50	P*	68	F	78	P	59	F	70	P*	80	P*	78	F	53	F	89	P	69	P*	67	P*	63	F
6	79	F	92	P*	86	F	73	P	72	P*	86	F	61	P*	62	F	65	F	79	P	88	F	79	F	57	F
7	72	F	59	P*	62	F	60	P	58	F	70	P*	70	F	68	F	90	F	78	P*	83	F	60	F	86	P*
8	81	F	71	P*	49	F	70	P	63	P*	68	P*	52	F	87	P*	81	F	77	P	61	P*	56	F	90	F
9	78	F	90	P*	61	F	75	P	78	P*	60	P	77	F	65	F	53	F	72	F	69	P*	75	F	67	F
10	54	F	58	P*	92	F	66	P	86	F	71	F	86	P*	71	F	71	F	54	P	81	F	93	P*	70	F
11	79	F	73	P*	95	F	61	P	52	P*	92	P*	87	P*	77	F	54	F	73	P*	68	F	53	F	86	F
<i>M</i>	70.4		70.3		74.4		69.9		68.8		75.4		70.8		73.7		70.6		73.7		74.0		72.4		75.5	
<i>SD</i>	9.69		14.14		14.88		7.27		12.69		9.95		11.34		8.57		13.35		8.80		8.80		12.79		11.69	
<i>Mdn</i>	72		66		72		70		71		71		70		75		71		75		69		74		76	
<i>Skew</i>	-0.69		0.48		-0.07		-0.12		0.14		0.37		-0.13		0.04		-0.11		-0.81		0.27		-0.03		-0.18	

Table I-1 Results of shortened fit test studies

Notes: QNFF is the fit factor measured using a TSI Portacount model 8038 respirator fit tester. QLFT result is the outcome of a qualitative fit test. * denotes that wearer identified taste at the Reveal Test. Error scenarios: 1) Fit test nebuliser filled with sensitivity solution, 2) Fit test nebuliser under-filled, 3) Fit test nebuliser clogged, 4) Fit test nebuliser bottom bung not removed, 5) Fit test nebuliser Venturi mis-seated, 6) Fit test solution under-administered, 7) Fit test timing incorrect, 8) Fit test nebuliser not upright, 9) Fit test nebuliser not upright and under-filled, 10) Fit test nebuliser not actuated with full squeezes, 11) Hood damaged / unsuitable, 12) Hood positioned incorrectly

Appendix J Number of applications and timings for control scenario and error scenarios Fit test solution under-administered and Fit test timing incorrect

Time	Control Scenario [3]	Error scenario	
		Fit test solution under-administered	Fit test timing incorrect
00:00	10	10	10
00:15			
00:30	5	2	
00:45			5
01:00	5	3	
01:15			
01:30	5	2	5
01:45			
02:00	5	3	
02:15			5
02:30	5	2	
02:45			
03:00	5	3	5
03:15			
03:30	5	2	
03:45			5
04:00	5	3	
04:15			
04:30	5	2	5
04:45			
05:00	5	3	
05:15			5
05:30	5	2	
05:45			
06:00	5	3	5
06:15			
06:30	5	2	
06:45			5
07:00	Reveal test	Reveal test	Reveal test
Total applications	75	42	55

Table J-1 Number of applications and timings for control scenario and error scenarios Fit test solution under-administered and Fit test timing incorrect

Appendix K Critical errors identified against key principles of the validated QLFT method

Critical error ID	Critical error made by fit tester which may lead to a false pass result	Key principle affected
1	Room poorly ventilated	1
2	Wearer has smoked, eaten or drunk before test	1
3	Wearer not advised to rinse mouth / clean face	1
4	Clear palate not verified (after sensitivity test)	1
5	Wearer not instructed to breathe through mouth	1
6	Sensitivity test not repeated (after failed fit test)	1
7	Hood not cleaned (after fit test)	1
8	Fit test nebuliser filled with sensitivity solution	2
9	Fit test nebuliser under-filled	2
10	Fit test nebuliser clogged	2
11	Fit test nebuliser bottom bung not removed	2
12	Fit test nebuliser Venturi mis-seated	2
13	Fit test solution under-administered	2
14	Fit test timing incorrect	2
15	Fit test nebuliser not upright	2
16	Fit test nebuliser not actuated with full squeezes	2
17	Hood damaged / unsuitable	3
18	Hood positioned incorrectly	3

Table K-1 Critical errors identified against key principles of the validated QLFT method

Notes: Critical error is an error which has the potential to adversely affect one or more of the key method principles. Key principles: 1) a sensitivity check to assure that the respirator wearer is able to taste a low level of challenge agent should a leakage occur, 2) procedures to assure that the aerosol is generated in a consistent and controlled manner, 3) steps to assure that the test atmosphere is contained around the respirator and is controlled [7]. Fit test nebuliser under-filled and Fit test nebuliser not upright were assessed through a previous study and results indicate that these two errors only have an impact on outcome if made concurrently. Critical error ID = means of identification for analysis and discussion within this paper.

Appendix L Search terms and key phrases for critical errors identified against key principles of the validated QLFT method

Critical error	Additional search words / terms
Room poorly ventilated	environment, ambient, atmosphere, ventilation, window, fresh, door, open, air
Wearer has smoked, eaten or drunk before test	clear palate, refrain, ingest, inhale, vape, smoking, food, drink, taste
Wearer not advised to rinse mouth / clean face	wipe, lips, drink, water, clear, palate, taste
Clear palate not verified	wipe, lips, drink, water, clear, palate, taste
Wearer not instructed to breathe through mouth	tongue, extend, inhale, open, taste
Sensitivity test not repeated	threshold, screening, entire, procedure, taste
Hood not cleaned	wipe, inside, disinfect
Fit test nebuliser filled with sensitivity solution	nebulizer, threshold, screening
Fit test nebuliser under-filled	nebulizer, 2.5ml, ampoule, solution
Fit test nebuliser clogged	nebulizer, block
Fit test nebuliser bottom bung not removed	nebulizer, plug
Fit test nebuliser Venturi mis-seated	nebulizer, atomiser
Fit test solution under-administered	nebulizer, nebuliser, top-up, squeezes, actuate, actuation
Fit test timing incorrect	stopwatch, time, seconds, minutes
Fit test nebuliser not upright	nebulizer, angle, hold, held, vertical
Fit test nebuliser not actuated with full squeezes	nebulizer, hand, actuate, administer, squeeze
Hood damaged / unsuitable	shape, size, form
Hood positioned incorrectly	collar, front

Table L-1 Search terms and key phrases for critical errors identified against key principles of the validated QLFT method

Appendix M Current administrative and engineering controls available to qualitative fit testers

ID Key	Critical mistake	Administrative resources										Engineering resources						
		Guidance/Training			QLFT user instructions							QLFT equipment kit						
		INDG479 [3]	Fit2Fit Companion [2]	Training Content [8]	Kit 1	Kit 2	Kit 3	Kit 4	Kit 5	Kit 6	Kit 7	Kit 1	Kit 2	Kit 3	Kit 4	Kit 5	Kit 6	Kit 7
1	Room poorly ventilated	X	✓	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2	Wearer has smoked, eaten or drunk before test	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	X	X	X	X	X	X	X
3	Wearer not advised to rinse mouth / clean face	X	X	X	✓	X	✓	✓	✓	X	✓	X	X	X	X	X	X	X
4	Clear palate not verified (after sensitivity test)	✓	✓	✓	X	✓	✓	X	X	✓	✓	X	X	X	X	X	X	X
5	Wearer not instructed to breathe through mouth	X	✓	✓	✓	✓	✓	✓	✓	✓	✓	X	X	X	X	X	X	X
6	Sensitivity test not repeated (after failed fit test)	X	✓	X	✓	✓	✓	✓	✓	✓	✓	X	X	X	X	X	X	X
7	Hood not cleaned (after fit test)	X	✓	✓	✓	✓	X	✓	X	✓	X	X	X	X	X	X	X	X
8	Fit test nebuliser filled with sensitivity solution	X	✓	X	✓	✓	✓	X	✓	X	✓	✓	✓	✓	✓	✓	✓	✓
9	Fit test nebuliser under-filled	X	✓	X	✓	X	✓	X	✓	X	✓	✓	X	X	X	X	X	X
10	Fit test nebuliser clogged	✓	✓	X	✓	✓	✓	✓	✓	X	✓	✓	✓	✓	✓	✓	✓	✓
11	Fit test nebuliser bottom bung not removed	X	✓	X	X	✓	✓	X	X	✓	✓	X	X	X	X	X	X	X
12	Fit test nebuliser Venturi mis-seated	X	X	X	X	✓	✓	X	X	X	✓	X	X	X	X	X	X	X
13	Fit test solution under-administered	X	✓	✓	✓	✓	✓	✓	✓	✓	✓	X	X	X	X	X	X	X
14	Fit test timing incorrect	X	✓	✓	✓	✓	✓	✓	✓	✓	✓	X	X	X	X	X	X	X
15	Fit test nebuliser not upright	X	✓	X	✓	X	✓	X	✓	X	✓	X	X	X	X	X	X	X
16	Fit test nebuliser not actuated with full squeezes	X	X	X	✓	✓	X	X	✓	✓	X	X	X	X	X	X	X	X
17	Hood damaged / unsuitable	✓	✓	✓	X	X	X	✓	X	X	X	X	X	X	X	X	X	X
18	Hood positioned incorrectly	X	X	X	✓	✓	✓	X	✓	✓	✓	✓	X	✓	X	✓	X	✓

Table M-1 Current administrative and engineering controls available to qualitative fit testers

Notes: Key: ✓ = control identified, X = no control identified

Appendix N Analysis of engineering controls in QLFT equipment

ID Key	Critical mistake	QLFT equipment number							Summary
		1	2	3	4	5	6	7	
8	Fit test nebuliser filled with sensitivity solution	Red labelling for sensitivity. Black labelling for Fit Test	Red labelling for sensitivity. Blue labelling for Fit Test	Red 'A' labelling for sensitivity. Blue 'B' labelling for Fit Test	Red labelling for sensitivity. Blue labelling for Fit Test	Blue 'A' labelling for sensitivity. Red 'B' labelling for Fit Test	Red labelling for sensitivity. Blue labelling for Fit Test	'A' labelling for sensitivity. 'B' labelling for Fit Test.	Inconsistent controls to ensure clear labelling and differentiation of sensitivity / fit test solution
9	Fit test nebuliser under-filled	Graduated markings on nebuliser	-	-	-	-	-	-	Limited controls to aid adequate filling of nebuliser
10	Fit test nebuliser clogged	Pin and new Venturi supplied	Pin and new Venturi supplied	Pin and new Venturi supplied	Pin and new Venturi supplied	Pin and new Venturi supplied	Pin and new Venturi supplied	Pin and new Venturi supplied	Pin supplied in all kits to removed clogging
11	Fit test nebuliser bottom bung not removed	-	-	-	-	-	-	-	No controls to ensure removal of bottom bung
12	Fit test nebuliser Venturi mis-seated	-	-	-	-	-	-	-	No controls to ensure proper seating of Venturi
13	Fit test solution under-administered	-	-	-	-	-	-	-	No controls to aid proper administering of fit test solution
14	Fit test timing incorrect	-	-	-	-	-	-	-	No controls to assist in accurate fit test timings
15	Fit test nebuliser not upright	-	-	-	-	-	-	-	No controls to ensure nebuliser is held upright during use
16	Fit test nebuliser not actuated with full squeezes	-	-	-	-	-	-	-	No controls to ensure fit test nebuliser is actuated with full squeezes
17	Hood damaged / unsuitable	-	-	-	-	-	-	-	No controls to ensure hood is undamaged / suitable
18	Hood positioned incorrectly	Collar with shoulders	-	Collar with shoulders	-	Collar with shoulders/straps	-	Collar with shoulders.	Limited controls to aid correct hood positioning

Table N-1 Analysis of engineering controls in QLFT equipment

8.2 References

1. British Safety Industry Federation. Fit2Fit Mission Statement. Available at: www.fit2fit.org (Accessed: 13 May 2021)
2. British Safety Industry Federation. Fit2Fit Companion to HSE INDG479 guidance on respiratory protective equipment (RPE) fit testing, qualitative method. June 2021. Available at: <https://www.fit2fit.org/resources>
3. Health and Safety Executive. INDG479 Guidance on respiratory protective equipment (RPE) fit testing. HSE Books; 2019. Available at: <https://www.hse.gov.uk/pubns/indg479.pdf>
4. British Standards Institution. BS ISO 16975-3:2017 Respiratory protective devices. Selection, use and maintenance. Part 3: Fit-testing procedures. London: British Standards Institution; 2017.
5. 3M. Qualitative fit test apparatus FT30 user instructions. Available at: <https://multimedia.3m.com/mws/media/1464265O/3m-qualitative-fit-test-apparatus-ft-30-user-instruction.pdf> (Accessed: 16 October 2021)
6. Nebuliser check list. QLTK1 - qualitative fit test kit supplement. Annan: Alpha Solway Ltd; 2020.
7. Occupational Safety and Health Administration (OSHA). Minnesota Mining and Manufacturing Company post hearing data submission on qualitative fit testing protocols under the lead standard before OSHA Docket No. H- 049A 1981.
8. British Safety Industry Federation. Fit2Fit Training Approval Scheme. Syllabus reference form. Available at: www.fit2fit.org/resources