ADULT MENTAL HEALTHCARE STAFF EXPERIENCES OF PSYCHOLOGY-LED SUPERVISION GROUPS FOR MANAGING RISK WITHIN ASSERTIVE OUTREACH AND INPATIENT REHABILITATION TEAMS.

By

FAREHA BEGUM

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF CLINICAL PSYCHOLOGY

Centre for Applied Psychology

School of Psychology

The University of Birmingham

June 2022

Thesis overview

This thesis is submitted in fulfilment of the requirements for the degree of Doctorate in Clinical Psychology at the University of Birmingham. The thesis consists of three chapters.

Chapter one reflects the findings of a systematic literature review exploring the reliability of the subscales for the Professional Quality of Life (ProQOL) Scale.

Chapter two reports findings from an empirical study which aimed to explore the experiences of adult mental healthcare staff working in Assertive Outreach Teams (AOTs) and Inpatient Rehabilitation when using psychology-led supervision groups to manage risk.

Chapter three is a public dissemination document which provides an accessible lay summary of both the systematic literature review and the empirical study.

Acknowledgements

I would like to sincerely thank my supervisor Dr Andy Fox for the unswerving support and encouragement throughout. Although it was a tough process, with lots of bumps and challenges along the way, your confidence in me did not falter and motivated me to keep going.

To my sisters Saira and Ayesha and my dearest friend Sharmin, you have been my rock throughout this journey and without you, it would not have been possible to complete. For all the times you saw me despair, you helped me get back up and find the courage to continue. You reminded me of the light at the end of the tunnel and how far I had come, always giving me strength and hope to endure the (seemingly never ending!) journey.

Finally, thank you to my wonderful participants who were very honest and forthcoming in sharing their experiences; without you, it would not have been possible to complete this insightful piece of research!

Contents page

Analysis of the Compassion fatigue Subscale

Analysis of the Compassion Satisfaction Subscale

CHAPTER 2 - Empirical Paper

Adult mental healthcare staff experiences of psychology-led supervision groups within Assertive Outreach Teams (AOTs) and inpatient rehabilitation teams in managing risk.

CHAPTER ONE

A systematic literature review exploring the reliability of the subscales of the Professional Quality of Life (ProQOL) Scale

Abstract

Background

Professional quality of life (PQoL) is a term that can be used to refer to both positive and negative emotions an individual experiences in his/her job as a helper and consists of three dimensions including compassion fatigue, burnout and compassion satisfaction. Evidence suggests health professionals working in critical care services, emergency departments, and in mental health and direct patient care, such as nurses and frontline health workers, are at increased risk of developing burnout and compassion fatigue. PQol is commonly assessed using the Professional Quality of Life (PRoQOL) Scale. However, little is known about the reliability of the scale.

Aim

To explore the reported reliability (Cronbach's Alpha) for each of the three subscales of the ProQOL.

Search databases

A systematic search of the literature was carried out from 1976-May 2021 using three databases; 146 studies were retrieved, 37 of which were included in this review.

Findings

The burnout subscale possessed the lowest reliability (0.48-0.94) of the three subscales. The compassion fatigue subscale had better reliability (0.64-0.91), with the compassion satisfaction subscale demonstrating the strongest reliability (0.77-0.96).

Conclusions

The ProQoL subscales showed adequate reliability, however, the burnout subscale demonstrated more variability than the other subscales. Most of the studies included in this

review did not explore reliability of the ProQoL scale as a whole (as recommended by the authors of the scale). Recommendations can only be made, therefore, on the use of the subscales separately, and not as an overall score. The reported variability of the burnout subscale and lack of conceptual clarity of some of the concepts underpinning the tool (such as compassion fatigue) mean that further research to better understand these is recommended.

Introduction:

Quality of life (QoL) can be defined as individual reflections on life achievements and standing, related to cultural views and beliefs and different value systems which may be linked to expectations, standards and goals one may have (World Health Organization. Division of Mental & Prevention of Substance, 1997). Definitions of QoL tend to be multifaceted and can include both negative and positive aspects of life. Importantly, QoL is subjective and will vary between individuals (Kim, 2014).

The general assumption is that an increased sense of QoL will be reflected in a lack of physiological symptoms, such as sleep difficulties, increased psychological well-being and the activities pursued by an individual. For instance, a better QoL would be indicated in lower levels of psychological distress which typically includes anxiety, depression and rational thinking, and may impact everyday activities including social skills, participation in leisure activities, employment and satisfaction with differing roles (Daundasekara et al., 2020). Fulfilment

Similarly, professional quality of life (PQoL) is a term that can be used to refer to both pleasant and unpleasant emotions experienced by professionals who have a caring role (Kim et al., 2015). Stamm (2010) defined PQoL as "the quality one feels in relation to their work as a helper" and can be understood as the extent to which individuals view their occupational roles and experiences as also fulfilling their personal needs (Roney & Acri, 2018). PQoL consists of three dimensions, including compassion fatigue, burnout and compassion satisfaction (Ruiz-Fernández et al., 2021).

Compassion fatigue is most commonly described as "a state of exhaustion and dysfunction biologically, psychologically, and socially as a result of prolonged exposure to

compassion stress and all it invokes" (Figley, 2013). It can result from "the change in empathetic ability of the caregiver in reaction to the prolonged and overwhelming stress of caregiving" (Lynch & Lobo, 2012). Certain professional groups who witness traumatic incidences more frequently, such as emergency workers and healthcare staff, are more likely to develop compassion fatigue (Cocker & Joss, 2016), with nurses deemed to be at higher risk (Adriaenssens et al., 2015).

There are various indicators of compassion fatigue impacting an individual themselves as well as those they care caring for. These include feelings of exhaustion, anger and irritability and generally a reduced sense of satisfaction with their occupational role and surroundings. There is a negative impact on decision making skills which affects the quality of care delivered to patients. Specifically, there is a significant negative impact when communicating key information to other colleagues, families and patients themselves (Nolte et al., 2017). Other indicators of compassion fatigue may be manifested in negative coping behaviours including the misuse of drugs and alcohol (Mathieu, 2007).

The majority of current research on compassion fatigue has included medical health care providers, social workers, and counselors (Andersen et al., 2015). There has also been significant research conducted on compassion fatigue in nursing, with evidence to suggest higher levels of compassion fatigue is linked with deteriorations in health and quality of life, with many nurses reporting a range of negative experiences including difficulties related to the digestive system, anxiety, low mood and depression, with a further negative impact on sleep (Fu et al., 2018).

More recent research on compassion fatigue (Norrman Harling et al., 2020) found psychologists reported feeling negatively affected by patients in both their personal and working lives, specifically when feeling and expressing compassion. Furthermore, tolerance levels were found to be reduced, parallel to increased compassion fatigue, with these psychologists experiencing higher levels of annoyance quicker and an impaired ability to understand distress and suffering experienced by patients.

As compassion and empathy are key aspects of the therapeutic relationship (Ardito & Rabellino, 2011) it is essential for healthcare professionals, including psychologists, to gain a better understanding of compassion fatigue. This will enable the delivery of better quality of care to patients as well as a more supportive working environment for staff. Some researchers (Ledoux, 2015) have suggested the term compassion fatigue is problematic; the notion that fatigue may be inextricably linked to compassion infers negative effects of compassionate feelings and behaviours i.e. exhaustion and with a limited capacity that is more quickly depleted (Fernando III & Consedine, 2014). This may be due to the concept of compassion being poorly defined with tools used to quantify compassion fatigue lacking construct validity (Ledoux, 2015). Indeed, there is no universal definition of compassion fatigue, with its relationship with compassion ambiguous (Horkan, 2014).

Compassion fatigue has also been used interchangeably with burnout, however, some (Rauvola et al., 2019; Sorenson et al., 2016) have rasised concerns over this, with a view that they are two distinct conditions and the latter not as a direct result of witnessing traumatic events (Adams et al., 2008). Furthermore, burnout is due to extended periods of physical and emotional fatigue, leading to a disinterest in work and relationships (Maslach, 2003). It can stem from work pressures, manifesting in difficulties achieving work goals, feeling frustrated, feeling incapable and generally negative attitudes towards work (Davis et al., 2013). As a result, many professionals experiencing burnout may change their work or career paths whereas compassion fatigue is thought to be highly treatable once professionals recognise it and take steps to address it (Slatten et al., 2011).

A third dimension of PQoL is compassion satisfaction and has been found to be positively related to psychological wellbeing which can be described as the amalgamation of positive emotions and ability to function effectively in aspects of daily living, reflected in overall good health, pleasure and prosperity (Elliott & Gramling, 1990). Evidence suggests increased job satisfaction amongst mental health professionals and when assisting patients with their difficulties, is associated with a better sense of psychological wellbeing (Amjad et al., 2020).

Compassion satisfaction is associated with feelings including happiness, fulfilment, inspiration, drive and optimism and can counteract the effects of compassion fatigue due to a

sense of satisfaction from caring for others (Hinderer et al., 2014). Compassion from professionals is further emphasised due to the impact it has on patients perceptions of their treatment experiences (Sinclair, Beamer, et al., 2017).

In recent years, research has emphasised the importance of understanding difficult and potentially traumatic experiences that staff may be exposed to within the workplace and how this may be associated with individual characteristics and styles (Stamm, 2010).

Different scales have been developed to assess PQoL. For instance, the Compassion Fatigue Self Test (CFST) (Figley, 1995) was one of the first measures to explore compassion fatigue as well as job burnout and consists of 40 items. This was further developed by Figley & Stamm (1996) who created a 66 item construct with additional items including more positively orientated questions, focusing specifically on compassion satisfaction. The CFST has been revised multiple times and was ultimately renamed as the Professional Quality of Life Scale (ProQOL) which is further discussed below. For instance, Gentry et al. (2002) proposed the Compassion Fatigue Scale Revised (CFS-R) which has 22 items measuring compassion fatigue and 8 measuring burnout. However, the psychometric properties of this measure have been questioned, resulting in the development of the Compassion Fatigue Short Scale (CF-SC), comprising of 8 items exploring compassion fatigue and 5 exploring burnout (Adams et al., 2006). The Secondary Traumatic Stress Scale (STSS) (Bride et al., 2004) was also developed to explore compassion fatigue, with secondary traumatic stress deemed to be an underlying element of compassion fatigue. Additionally, although the Impact of Event Scale (IES) (Horowitz et al., 1979) and the Impact of Event Scale revised (IES-R) (Weiss & Marmar, 2004) were designed to measure trauma experienced directly, they have been used in studies exploring compassion fatigue in those who provide services to others.

With regards to measuring burnout specifically, Maslach's Burnout Inventory (MBI) (Maslach et al., 1997), comprising of 22 items and exploring emotional exhaustion (EE), depersonalisation (DP) and personal achievement (PA) is considered the "gold standard." However, some criticisms of this tool highlight the lack of a cut-off score for burnout which can be linked to negative outcomes (West et al., 2012). Additionally, although the version of the MBI specifically for assessing burnout in healthcare professionals, the MBI Human Services Survey for Medical Personnel (MBI-HSS-MP) is the most widely used to assess

burnout in healthcare professionals, it's psychometric properties have not yet been examined (Lheureux et al., 2017).

The development of the ProQOL scale primarily aimed to address difficulties separating burnout and secondary traumatic stress as well as psychometric problems noted in the CFST, with a shorter questionnaire of 30 items also reducing respondent burden. Items that were retained are theoretically more salient and better represent the subscale construct. Secondly, a change in name focusing on positive and negative aspects made it easier to support positive changes to maintain the positive effects of providing care as well as ameliorate or prevent negative impacts of providing care. The most recent version of the ProQOL scale is ProQOL-5 (Stamm, 2009), with revisions to previous versions focusing on strengthening psychometric properties and changes to wording to reduce ambiguity and/or better reflect individual subscale constructs. Additionally, some items have been re-worded based on more recent literature exploring the subscales. For instance, item 10 on the ProQOL R-IV "I feel depressed as a result of my work as a helper" was re-worded in the ProQOL-5 to read "I feel depressed because of the traumatic experiences of a person I help." Similarly, item 29 has been reworded from "sensitive person" to "caring person" in the ProQOL-5 scale.

The ProQOL scale comprises of compassion satisfaction, burnout and compassion fatigue and is the most widely used tool of exploring the positive and negative impact on healthcare professionals (Stamm, 2010). The three dimensions have 10 items each, ranked on a 5-point Likert scale, ranging from 1 (Never) to 5 (Very Often), reflecting how the respondent felt about that item over the last 30 days.

Compassion satisfaction is measured by the sum of scores on items 3, 6, 12, 16, 18, 20, 22, 24, 27 and 30. Burnout is measured by the sum scores on items 1, 4, 8, 10, 15, 17, 19, 21, 26 and 29, whereas compassion fatigue is measured by the sum of scores on items 2, 5, 7, 9, 11, 13, 14, 23, 25 and 28. For each subscale, a sum score of 22 or less indicates low compassion satisfaction, burnout or compassion fatigue; a sum score of 23-41 indicates average levels of compassions satsifaction, burnout or compassion fatigue whilst a sum score of >42 for any of the subscales would indicate high levels of compassion satisfaction, burnout or compassion fatigue. As the three subscales are distinct, it is not possible to provide a composite score for the ProQOL scale and the complex relationship between the dimensions means respondents

can simultaneously score high for both the compassion dimensions. Additionally, items 1, 4, 15, 17 and 29 need to be reverse-scored.

Often, the ProQOL scale is used as a screening tool for people in professions that involve caring or helping others such as nursing, where there are links to job performance, including clinical performance and competence (Vagharseyyedin et al., 2011). The three dimensions also determine the key characteristics of PQoL.

Although compassion fatigue and burnout are present in a variety of different practitioner groups and specialities (Cavanagh et al., 2020), there is evidence to suggest increased risk for frontline workers and nurses working in specific settings including mental health, emergency departments and critical care services (Dubale et al., 2019; O'Connor et al., 2018). Healthcare professionals are particularly susceptible to burnout due to intense pressures to remain sensitive and compassionate, even when having to make rapid decisions related to high-risk (Burton et al., 2017). The risk of developing mental health conditions such as depression and anxiety is also increased (Drury et al., 2014) and posttraumatic stress disorder (PTSD) (Figley, 2002), which are associated with increased sickness absences, employment turnover and overall reduced effectiveness.

In addition to the direct impact on staff, healthcare professionals' quality of life also appears to impact the quality and safety of patient care they deliver. Hall et al. (2016) suggest an association between reduced well-being and burnout, resulting in reduced patient safety. If health professionals are exhausted during work hours, this is likely to impact their ability to deliver effective care, which can be dangerous for patients and lead to suboptimal outcomes for services (Heeb & Haberey-Knuessi, 2014). If they are also experiencing higher levels of compassion fatigue, desensitisation may be experienced with increased annoyance towards patients, increasing the potential for key information to be misunderstand and miscommunicated, resulting in inadequate care provided and inappropriate professional conduct (Başoğul et al., 2021).

The literature above indicates the importance of PQoL, consisting of compassion fatigue, burnout and compassion satisfaction, the impact it has on professionals, both in their personal and professional lives, as well as those they are caring for. PQoL is often measured

using the ProQOL scale. However, to date, there is minimal research conducted on the psychometric properties of the ProQOL scale, specifically exploring reliability which can include test-retest reliability, inter-rater reliability or internal consistency. Test re-test refers to administering the same test twice or more to the same sample and examining the scores obtained at each time point to compare reliability. When reliability is assumed to be good, scores across different timepoints should be similar. However, it is important to note that test re-test reliability may be impacted when using a measure or scale across different settings, populations or cultures due to different expectations, norms or cultural practices. Therefore, even if some samples have similar scores across different timepoints, the overall reliability of the scale may be reduced if other samples in different settings score lower across timepoints. Additionally, due to constant change in human experiences, there is a chance that scores may be different over time, especially if exploring constructs related to attitudes or internal experiences which again may impact reliability of a scale when assessing using test re-test. Inter-rater reliability refers to 2 (or more) raters agreeing on the same item or concept, without discussing it; the higher the percentage of agreement, the better the perceived reliability. Subjectivity is likely to be present for each rater, therefore it is important to clarify objective criteria for how constructs or items should be rated and if there are numerous raters, they should receive the same information about specific criteria to ensure more objective rating and ultimately better reliability. Essentially, this method of assessing reliability was not suitable for this study as the ProQOL scale is a self-report questionnaire. Finally, internal consistency, measured using Cronbach's Alpha (Cronbach, 1951), provides a measure of the internal reliability of an outcome measure or scale and the degree to which the same core concept is collectively measured by the items (Portney & Watkins, 2009). Cronbach's Alpha is a numerical value between 0 and 1 with a value of 1 representing perfect consistency and reliability. Acceptable alpha coefficient values range between 0.7-0.95 (Bland & Altman, 1997; Nunnally, 1994). A general rule of thumb, however, indicates alpha values <0.5 = unacceptable; 0.5-0.6 = poor; 0.6-0.7 = questionable; 0.7-0.8 = acceptable; 0.8-0.9 = good and >0.9 = excellent (Mills, 2003).

Cronbach's Alpha has been described as "one of the most important and pervasive statistics in research involving test construction and use" (Cortina, 1993) (p. 98) with research studies commonly referencing Alpha values (Schmitt, 1996). It statistically explores all

possible item combinations to accurately report reliability of a scale and is more favourable compared to other estimates of reliability such as test re-test, as it is easier to use (Cohen et al., 1996) and requires administering only once. More importantly, Cronbach's Alpha is rooted in the 'tau equivalent model' which requires all items in a scale to measure the same core concept. When this is met, there is increased confidence in estimates of reliability (Tavakol & Dennick, 2011) compared to test-retest or interrater reliability and is also the most extensively used objective measure of reliability. This was deemed important when exploring reliability of the ProQOL as there are numerous different definitions and indicators for the subscales which are psychologically contentious across different settings, populations, languages, and cultures. Therefore, the aim of this literature review is to explore reliability of the ProQOL further, specifically the reported Cronbach's Alpha for each of the three subscales of the ProQOL. Additionally, due to the limited research conducted in this area, this review was interested in any study that reported a Cronbach's Alpha, including any discipline of healthcare staff.

Method:

Identifying primary studies

Search of Electronic Databases

A systematic search of the literature was initially carried out in May 2021 using the PsycINFO, CINAHL and EMBASE databases. The aim of the search was to obtain a comprehensive overview of the literature into reported alpha coefficients for the ProQOL scale. The search terms that were used to identify the relevant studies are outlined in table 1 below.

Table 2 – search criteria

Database	**Free Text Search Terms**	**Method of Search**	**Limits**
PsycINFO	("professional quality of life" OR ProQOL).af	Free search terms All search terms combined with *OR*	Peer reviewed articles 1967-May 2021
	(Reliab* OR Cronbach* af		
	("professional quality of life" OR ProQOL).af		

	"Internal consistency"/ Reliability"/ or "internal consistency"/		
CINAHL	"Internal consistency"/ or "coefficient alpha"/"Internal consistency"/ (Reliab* OR Cronbach*).af "Interrater reliability"/ ("professional quality of life" OR ProQOL).afReliability/ Reliability/ "Cronbach alpha coefficient"/ or "internal consistency"/ "Internal consistency"/ or "coefficient alpha"/ (Reliab* OR Cronbach*).af (Reliab* OR Cronbach*).af ("professional quality of life" OR ProQOL).af	Free search terms All search terms combined with *OR*	Peer reviewed articles 1967-May 2021
EMBASE	Reliability/ "Cronbach alpha coefficient"/ or "internal consistency"/ (Reliab* OR Cronbach*).af ("professional quality of life" OR ProQOL).af	Free search terms All search terms combined with *OR*	Peer reviewed articles 1967-May 2021

Inclusion criteria

Five different versions of the ProQOL scale were identified during this analysis (ProQOL-5; ProQOL; ProQOL-21; ProQOL Revised & ProQOL-R-IV). The most recent version (ProQOL-5) has also been translated for use in twenty-eight different languages, indicating its application across different countries, with different populations. All studies that used any version of the ProQOL scale and reported an alpha coefficient for any one of the subscales or the total alpha coefficient for the ProQOL scale were included in the review.

The results of the systematic search are presented in Figure 1. The search yielded a total of 146 articles. Once duplicated studies were removed, there were 120 remaining studies. A further 6 studies were removed as these were not in English and translated versions in English could not be identified. The remaining 114 articles were then screened using the main body of the text rather than an initial screening of the title and abstract, as many of these studies were found to report the alpha coefficients in the main text but not in the title or abstract.

After screening the full texts of the 114 identified studies, 40 of these had reported alpha coefficients for the ProQOL scale. 3 studies reported only the total alpha coefficient for the ProQOL scale, without reporting alpha values for any of the subscales, and was deemed too small a number to conduct a meta-analysis on. The remaining 37 studies, all which reported alpha coefficients for one, two or all three of the subscales were included in this meta-analysis of each subscale. (See Appendix 1 for summaries of the studies included).

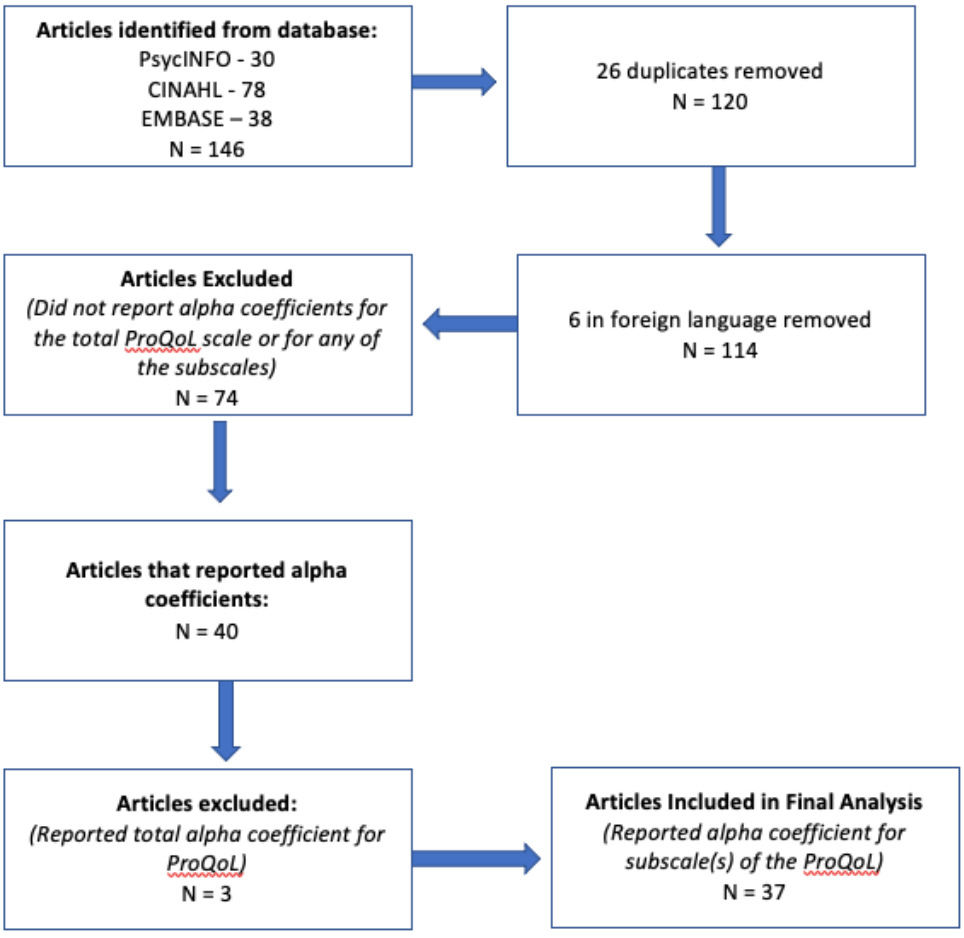


Figure 2 – Results of the systematic search and the application of the inclusion criteria

Data extraction

All data were extracted by the author. A second data extractor was not used to cross validate the data extraction methods as this was not possible due to time constraints.

Risk of Bias Assessment

A set of quality criteria were developed to assess any risk of bias within this literature. The quality criteria were adapted from existing risk of bias frameworks, including The Cochrane Collaboration Risk of Bias Tool (Higgins et al., 2011) and the Risk of Bias Assessment Tool for Nonrandomised Studies (Kim et al., 2013) and adapted from a meta-analytic review of the Toronto Alexithymia Scale-20 by Spry (2019). The current framework assesses risk of bias in six domains: selection bias, performance bias, detection bias, statistical bias, reporting bias and generalisation. The risk of bias score was calculated as the sum of two points for each area of low risk of bias, one point for each area of unclear bias and zero points for each area of high risk of bias. Therefore, the risk of bias score ranged from 0-12 points. In addition to this, every study that was designed to assess psychometric properties (and therefore likely to produce more valid psychometric characteristics) was awarded an additional 10 points. Accordingly, the quality score ranged from 0-22 points and was expressed as a percentage of the total possible score. The quality score is shown in Figure 2 along with the individual ratings for each of the areas of the risk of bias. The risk of bias in the six domains and the criteria for Low, Unclear or High risk is described in Appendix 2 and the application of these criteria are reported in Figure 2.

Selection Bias

Overall, the majority of the studies included were rated as low risk for selection bias, with only five studies (Chang & Taormina, 2011; Choi & Lee, 2017; Erkorkmaz et al., 2018; Kim et al., 2017; Sanford et al., 2018) being rated as unclear risk. All the low risk studies clearly described the characteristics of the population and how participants were recruited.

Performance Bias

Performance bias was the risk of bias criteria that had the biggest mix of risk level within the studies. Majority of studies (twenty-two) were rated as low risk, with a third of studies (thirteen) rated as unclear risk. These studies did not report confidentiality or anonymity, and it was unclear how much information was provided to participants, prior to participation. The remaining two studies were rated as high risk of performance bias (Ang et al., 2018; Hemsworth et al., 2018) as they had paid participants a reward as an incentive to participate.

Detection Bias

The majority of studies included (twenty-six) appeared to be low risk for detection bias, with the scale being administered in its original format. Although they used different versions of the English language ProQOL scale, they were all standardised scales. Eight studies were classed as unclear risk, as these were translated versions of the scale, where it was not stated if they had changed any of the wording of the items. The remaining three studies (Alhalal et al., 2020; Chang & Taormina, 2011; Choi & Lee, 2017) were rated as high risk as they specifically stated that they had made changes to the wording of some of the items. Alhalal et al. (2020) reported this was done as some phrases were not understandable outside of a Western context. This study also did not report the alpha coefficient for the compassion fatigue subscale. Furthermore, the study by Erkorkmaz et al. (2018) had an alpha coefficient for one subscale <0.6, impacting the reliability of the scale significantly.

Statistical Bias

All studies were rated as low risk as they had used appropriate methods of analysis and reported a Cronbach's Alpha value.

Reporting Bias

Majority of studies were rated as low risk, reporting data on the total sample size and analysis of the three subscales. Only three studies demonstrated unclear risk. For instance, Alhalal et al. (2020) did not provide an alpha coefficient for the compassion fatigue subscale whereas Begic et al. (2019) did not report an alpha coefficient for the compassion satisfaction

subscale. Furthermore, Duarte & Pinto-Gouveia (2017) did not report any data on the compassion satisfaction subscale.

Generalisability

Almost all of the studies were rated as low risk, having large sample sizes that were representative of the target population. Only the studies by Chang & Taormina (2011) and Begic et al. (2019) were rated as unclear risk. Although the former study had a good sample size (N=102), the idiosyncratic sample included only male Chinese soldiers following an earthquake, which would be hard to generalise to other populations. Similarly, the study by Begic et al. (2019) included 27 home visitors who serve pregnant women or new mothers but also had a much smaller sample size (N=27).

Summary

The biggest variation in risk of bias, across the studies, was apparent for the performance and detection bias criteria. Generally, however, most of the studies appeared to demonstrate low risk of bias across the different bias criteria. Eight studies were rated as low risk for all risk of bias criteria, with only one of these (Heritage et al., 2018) possessing an overall risk quality of 100% due to the study aim exploring psychometric properties of the ProQOL scale, which enhanced the rating of the study, compared to studies that did not explore psychometric properties. For instance, although the studies by both Fu et al. (2018) and Duarte (2017) only had one risk of bias criteria as unclear risk, the overall study quality score for the latter study was given a better score as it was directly exploring the psychometric properties of the ProQOL scale, whereas the study by Fu et al. (2018) did not.

Figure 3 - Ratings of risk of bias. *Red indicates high risk of bias, amber marks an unclear risk of bias and green is a low risk of bias*

Study Name	Selection Bias	Performance Bias	Detection Bias	Statistical Bias	Reporting Bias	Generalisability Bias	Quality Index
Duarte 2017	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	95%
Heritage 2018	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	100%
Samson 2016	Low risk	Unclear risk	Unclear risk	Low risk	Low risk	Low risk	91%
Ravi 2016	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	55%
Galiana 2017 - Spanish	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	95%
Galiana 2017 - Brazilian	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	95%
Galiana 2019 - Spanish	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	55%
Galiana 2019 - Argentinian	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	55%
Salloum 2018	Low risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	50%
Storm 2021	Low risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	50%
Yildirim 2021	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	55%
Alhalal 2020	Low risk	Low risk	High risk	Low risk	Unclear risk	Low risk	41%
Salimi 2020	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	50%
Lemieux-Cumberlege 2019	Low risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	50%
Kagan 2019	Low risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	50%
Xu 2019	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	55%
Begic 2019	Low risk	Low risk	Low risk	Low risk	Unclear risk	Unclear risk	45%
das Neves Borges 2019	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	50%
Ang 2018	Low risk	High risk	Low risk	Low risk	Low risk	Low risk	45%
Sanford 2018	Unclear risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	45%
Itzick 2018	Low risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	50%
Fu 2018	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	50%
Hemsworth 2018 - Australian	Low risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	95%
Hemsworth 2019 - Canadian	Low risk	High risk	Low risk	Low risk	Low risk	Low risk	91%
Hemsworth 2020 - Canadian Palliative Nurses	Low risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	95%
Choi 2017	Unclear risk	Low risk	High risk	Low risk	Low risk	Low risk	41%
Duarte & Pinto-Gouveia 2017	Low risk	Low risk	Low risk	Low risk	Unclear risk	Low risk	50%
Kim 2017	Unclear risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	45%
Jang 2016	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	50%
Lee 2016	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	55%
Hunsaker 2015	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	55%
Leners 2014	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	55%
Michalec 2013	Low risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	50%
Elkonin 2011	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	55%
Chang 2011	Unclear risk	Low risk	High risk	Low risk	Low risk	Unclear risk	36%
Goshen General Hospital (Yoder) 2010	Low risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	50%
Erkorkmaz 2018	Unclear risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	45%

Analysis

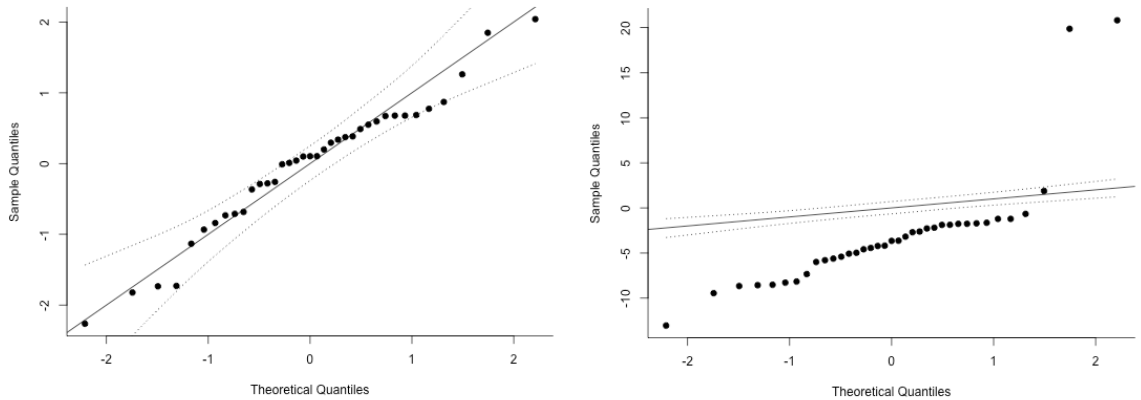
As most studies using the ProQOL scales did not report a total alpha coefficient for the scale, rather three separate alpha coefficients for the three subscales, this analysis aimed to explore reliability of the three subscales for each version of the ProQOL scale separately. This was deemed appropriate as the ProQOL manual emphasises including and exploring the three different subscales due to a complex relationship between the three subscales and no composite score for the three subscales. Additionally, the analysis aimed to analyse each subscale of the different language versions.

The data has been analysed in accordance with the centre for applied psychology meta-analysis strategy (see Appendix 5).

Analysis of the Burnout Subscale

Selection of the meta-analytic model for the burnout subscales

The distribution of primary study effects is shown in Figure 1.1 for both the fixed effects model (FEM) and the random effects model (REM). The REM between studies variance (tau2) was calculated using the DerSimonian-Laird estimator.



A B

Figure 1.1: QQ plot of the distribution of alpha coefficients within the primary studies for the burnout subscales. Chart A plots the fit of the random effects model and chart B plots the fit of the fixed effects model

As can be seen from Figure 1.1 there is clear evidence of non-linearity in the distribution of alpha coefficients for the burnout subscales, within the fixed effects model, whereas the random effects model shows a good fit to the distributional assumptions of the REM.

Therefore, the use of the random effects model using the DerSimonian-Laird estimator of between-subjects variance is an appropriate method for the synthesis of these data.

The omnibus test for the burnout subscales

The alpha coefficients described in the primary studies are reported in Table 1.1. There were 33 studies reporting 37 effects, with a total of 13,435 participants.

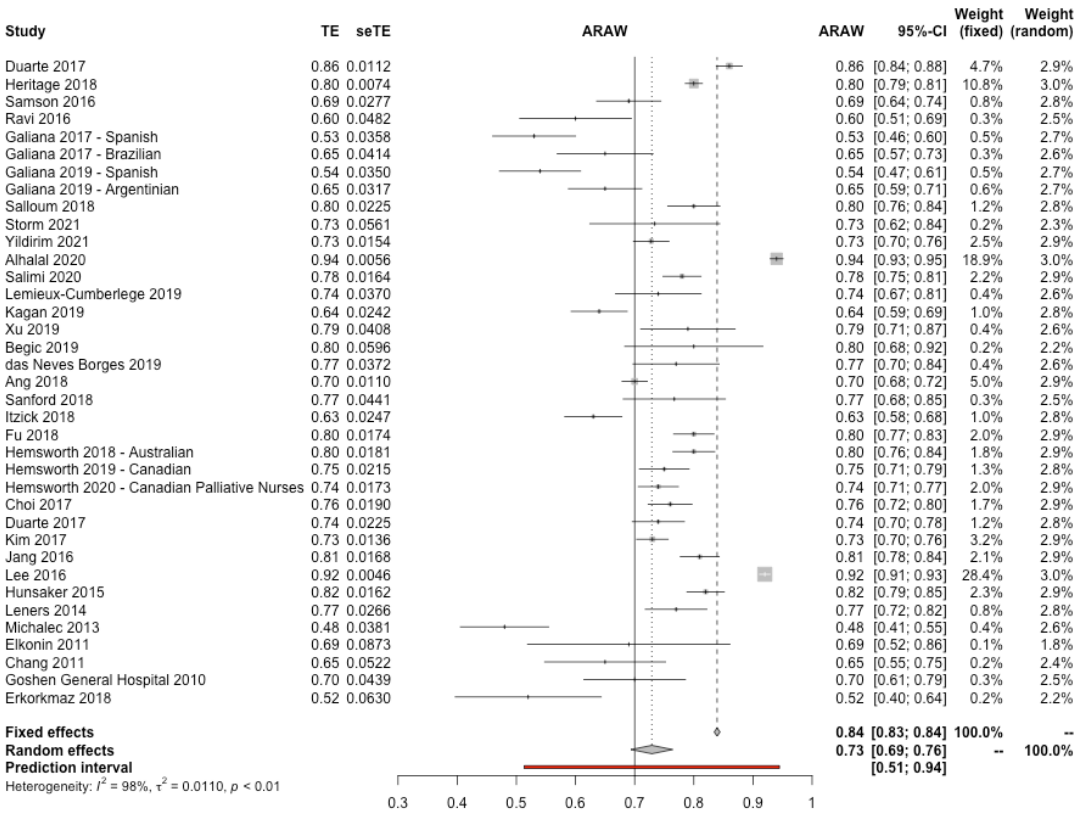
Table 1.1 – Alpha coefficients for the Burnout subscales

Study Name	Alpha Coefficient	N	ProQOL Version	Language Version
Duarte 2017	0.86	390	ProQOL-5	Portuguese
Heritage 2018	0.8	1615	ProQOL-21	English
Samson 2016	0.69	377	ProQOL	Hebrew
Ravi 2016	0.6	155	ProQOL-5	English
Galiana 2017 - Spanish	0.53	385	ProQOL	Spanish
Galiana 2017 - Brazilian	0.65	161	ProQOL	Brazilian
Galiana 2019 - Spanish	0.54	385	ProQOL	Spanish
Galiana 2019 - Argentinian	0.65	273	ProQOL	Argentinian
Salloum 2018	0.8	177	ProQOL-5	English
Storm 2021	0.734	52	ProQOL	English
Yildirim 2021	0.728	697	ProQOL R-IV	English
Alhalal 2020	0.94	255	ProQOL-5	Arabic
Salimi 2020	0.78	400	ProQOL	Persian
Lemieux-Cumberlege 2019	0.74	112	ProQOL-5	English
Kagan 2019	0.64	494	ProQOL	English
Xu 2019	0.79	61	ProQOL	English
Begic 2019	0.8	27	ProQOL	English
das Neves Borges 2019	0.77	87	ProQOL-5	Portuguese
Ang 2018	0.7	1667	ProQOL	English
Sanford 2018	0.767	64	ProQOL	English
Itzick 2018	0.63	501	ProQOL Revised	English
Fu 2018	0.8	294	ProQOL-5	Chinese
Hemsworth 2018 - Australian	0.8	273	ProQOL-5	Australian
Hemsworth 2019 - Canadian	0.75	303	ProQOL-5	Canadian
Hemsworth 2020 - Canadian Palliative Nurses	0.74	503	ProQOL-5	Canadian
Choi 2017	0.76	358	ProQOL Revised	Korean
Duarte 2017	0.74	298	ProQOL-5	English
Kim 2017	0.73	875	ProQOL-5	English

Jang 2016	0.81	285	ProQOL-5	Korean
Lee 2016	0.92	680	ProQOL	English
Hunsaker 2015	0.82	278	ProQOL-5	English
Leners 2014	0.77	168	ProQOL	English
Michalec 2013	0.48	416	ProQOL-5	English
Elkonin 2011	0.69	30	ProQOL R-IV	English
Chang 2011	0.65	102	ProQOL	Chinese
Goshen General Hospital 2010	0.7	106	ProQOL R-IV	English
Erkorkmaz 2018	0.52	131	ProQOL	English

A random effects models was calculated using the generic inverse variance method. The random effects model suggested a weighted average alpha coefficient of α=0.7290 (z = 40.38, p < 0.01) and a 95% confidence interval of between 0.69 to 0.76.

Figure 1.2: Forest plot of alpha coefficients for the burnout subscales



A high level of heterogeneity in the primary studies was observed (Higgin's I^2 = 97.8%, tau^2 = 0.0110, p = < 0.01) suggesting that the estimates of alpha coefficients in the primary studies may be biased by the presence of uncontrolled or confounding factors. Therefore, the

focus of the subsequent analyses will be upon the identification of the sources of heterogeneity between the estimates of alpha coefficients between the primary studies.

The impact of influential primary studies for the burnout subscales

The impact of studies having a disproportionate influence was assessed using a "leave-one-out" analysis, in which the random effects model was calculated with each of the primary studies removed in turn and change in weighted average effect size (i.e., influence) and the change in heterogeneity (i.e., discrepancy) was recorded. The result of this "leave-one-out" analysis is presented on the Baujat plot (Baujat, Pignon, & Hill, 2002) in Figure 1.3.

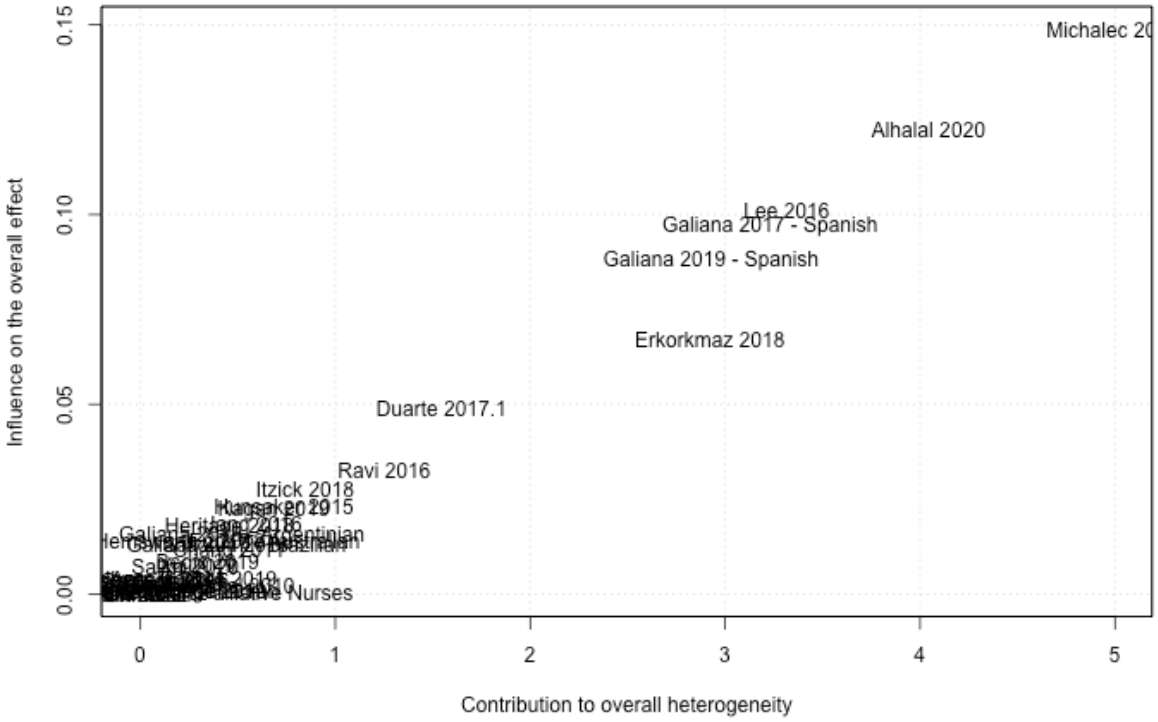


Figure 1.3: Baujat diagnostic plot of sources of heterogeneity for the burnout subscales. The vertical axis reports the influence of the study on the overall effect and the horizontal axis reports the discrepancy of the study with the rest of the literature. The shaded area in the top right quadrant of the graph is the area associated with influential and discrepant studies

The Baujot plot in Figure 1.3 above suggests the studies by Alhalal (2020) and Michalec (2013) are the most influential on the overall synthesis and the most discrepant from the rest of the literature. The random effects model was recalculated having removed the two studies showing disproportionate influence. The corrected random effects model reported a synthesis of Alpha = 0.7297 (95% CI 0.70 to 0.76). The corrected random effects model

evidenced a less than 0.0098% decrease relative to the uncorrected estimate and did not change any of the conclusions from this meta-analysis. Accordingly, this meta-analysis can be considered robust to the effect of influential and discrepant studies.

The effect of risk of bias in the primary studies for the burnout subscales

In order to assess the impact of study level risk of bias upon heterogeneity, a series of subgroup analysis were conducted on the alpha coefficients for the risk of bias ratings of "low risk" and "any risk" (i.e., unclear risk and high risk of bias combined) for each of the six types of methodological bias.

Table 1.2 – Effect of risk of bias (Burnout Subscale)

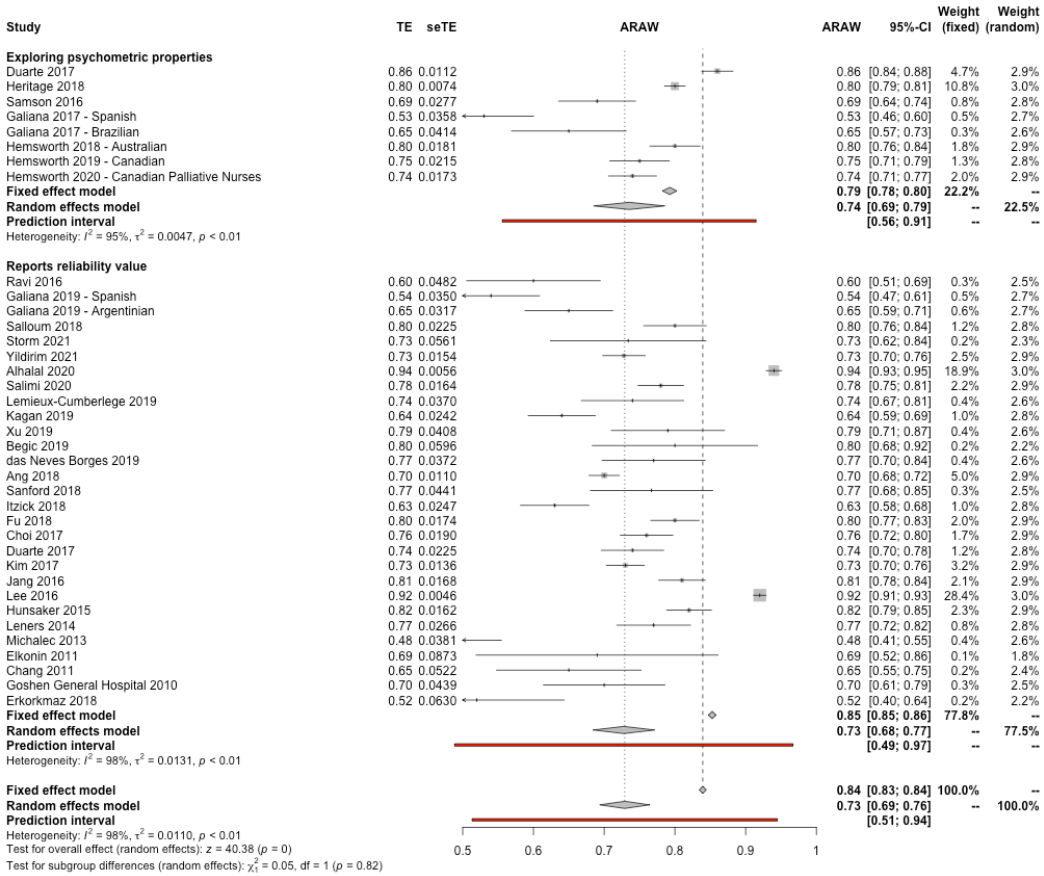
	Low Risk			Any Risk				
	Alpha	**95% CI**	**Studies**	**Alpha**	**95% CI**	**Studies**	**X^2**	**P**
Selection bias	0.7343	0.6965; 0.7721	32	0.7082	0.6546; 0.7617	5	0.61	0.4349
Performance bias	0.7517	0.7113; 0.7921	22	0.7014	0.6668; 0.7359	15	3.44	0.0636
Detection bias	7.177	0.6730; 0.7624	26	0.7539	0.6886; 0.8193	11	0.81	0.3691
Statistical bias	0.7290	0.6936; 0.7644	37				--	--
Reporting bias	0.7203	0.6841; 0.7565	34	0.8290	0.6715; 0.9865	3	1.74	0.1874
Generalisability bias	0.7293	0.6931; 0.7655	35	0.7222	0.5753; 0.8691	2	0.01	0.9267

As can be seen from Table 1.2 above, none of the risk of bias criteria evidenced significant differences between the weighted average alpha coefficients for the low and high risk of bias studies for the burnout subscale. Additionally, all studies were rated as low risk for statistical bias.

Differences based on the primary aims of the included studies for the burnout subscales

The primary studies were divided into (a) those studies that were specifically designed to assess the psychometric properties of the ProQOL scale and (b) those studies that reported psychometric properties but were designed to address a different question. There was no significant difference in the average alpha coefficients for the two types of study design (X^2 = 0.05, p=0.82) (see Figure 1.4 below).

Figure 1.4 – differences based on the primary aims of the studies for the burnout subscales

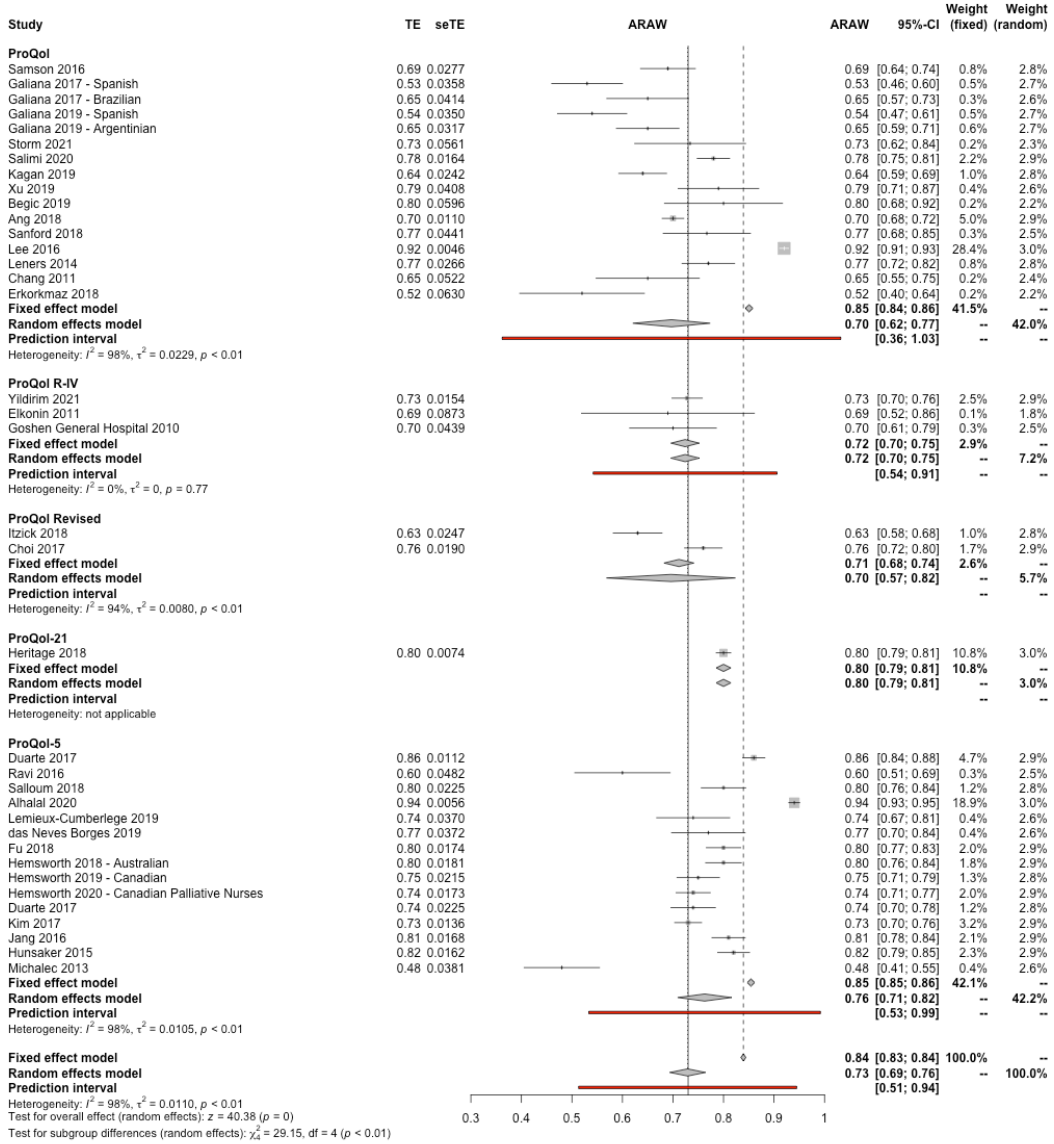


Differences between different versions of the ProQOL burnout subscales

The difference in alpha coefficients was calculated for the different between different versions of the ProQOL burnout subscales (see Figure 1.5).

A statistically significant difference was observed between the internal reliability of the different versions of the ProQOL burnout subscales (X^2 = 29.15, p <0.01). The only version that showed a meaningful difference from the overall mean was the ProQOL-21. It should be noted that the estimate of internal consistency of this version was obtained from a single study, and it is likely that the average internal consistency score of this version will change with the publication of future studies regarding the internal reliability of this measure.

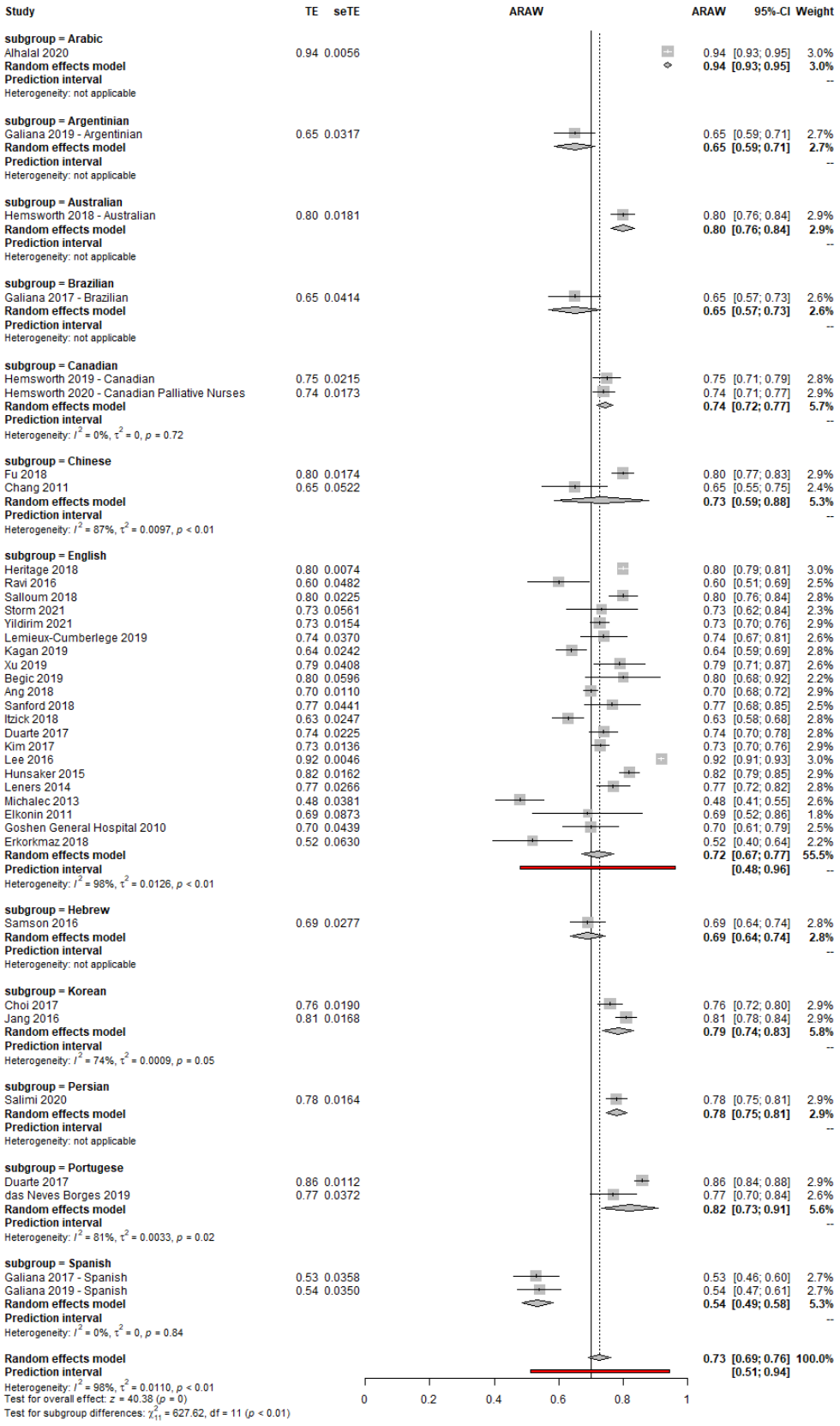
Figure 1.5: differences between different versions of the ProQOL burnout subscales



Differences between different language versions of the ProQOL burnout subscales

The difference in alpha coefficients was calculated for the different language versions of the ProQOL burnout subscales (see Figure 1.6).

Figure 1.6 – Different language versions of the ProQOL burnout subscales



A statistically significant difference was observed between the internal consistency of the different language versions of the ProQOL burnout subscales (X^2= 627.62, p <0.01).

Majority of the different language versions showed a significant difference from the overall mean including the Argentinian version with a reported alpha coefficient of 0.65 (95% CI 0.59 to 0.71); the Australian version with an alpha coefficient of 0.80 (95% CI 0.76 to 0.84); the Korean version with an alpha coefficient of 0.79 (95% CI 0.74 to 0.83); the Persian version with an alpha value of 0.78 (95% CI 0.75 to 0.81); the Portuguese version with an alpha value of 0.82 (95% CI 0.73 to 0.91), the Spanish version with an alpha value of 0.54 (95% CI 0.49 to 0.58) and the Arabic version had an alpha value of 0.94 (95% CI 0.93 to 0.95). It should be noted that the average alpha coefficient of the Argentinian and the Spanish versions of the ProQOL scale fell below the minimum accepted value of 0.7.

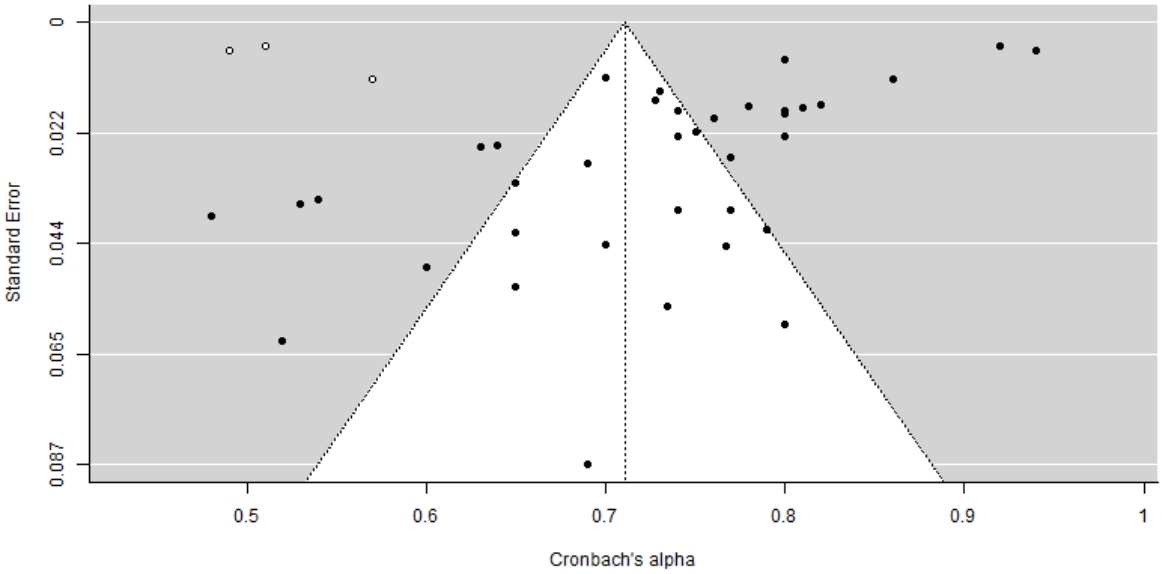
It should also be noted the internal consistency of the above language versions were obtained from no more than two studies. Therefore, it is likely that the average internal consistency score will change with the publication of future studies.

The impact of publication and small study biases for the burnout subscale

Publication bias is caused by the tendency for statistically significant results to be published and the reticence to publish papers with non-significant results. Small study bias is the tendency for studies with smaller sample sizes to show greater variability in their measurement of internal consistency. These biases can be identified in a funnel plot, which plots the magnitude of the study's reported alpha coefficients (i.e., the importance of the study in the synthesis) against an estimate of the studies deviation from the meta-analytic average (i.e., the discrepancy of the study within the literature).

If there is an absence of publication bias, studies with smaller sample sizes will show greater variability will scatter more widely at the bottom of the plot compared to studies with larger samples at the top which will lie closer to the overall meta-analytic effect, creating a symmetrical funnel shape. If there is an absence of studies in the area of the plot associated with small sample sizes and non-significant results, then it is likely there is some publication bias leading to an overestimation of the true effect. The funnel plot of study level alpha coefficients is presented in Figure 1.7.

Figure 1.7: funnel plot of the alpha coefficients for the burnout subscales. The 95% confidence interval of the expected distribution of alpha coefficients is shown as an inverted "funnel". Points filled in white represent imputed studies using the trim and fill procedure (Duval & Tweedle, 2000).



As can be seen from Figure 1.7, there is clear evidence of publication bias in the distribution of alpha coefficients for the burnout subscales. The effect of publication bias was simulated using a trim and fill procedure (Duval & Tweedle, 2000). The trim and fill procedure builds on the assumption that publication bias would lead to an asymmetrical funnel plot. The trim and fill procedure uses an iteratively removes the most extreme small studies from the side of the funnel plot associated with positive effects, re-computing the effect size at each iteration until the funnel plot is symmetric about the (corrected) effect size. While this trimming yields the adjusted effect size, it also reduces the variance of the effects, resulting in biased and narrow confidence interval. Therefore, the original studies are returned into the analysis, and the procedure imputes a mirror image for each on the side of the funnel plot associated with negative effects.

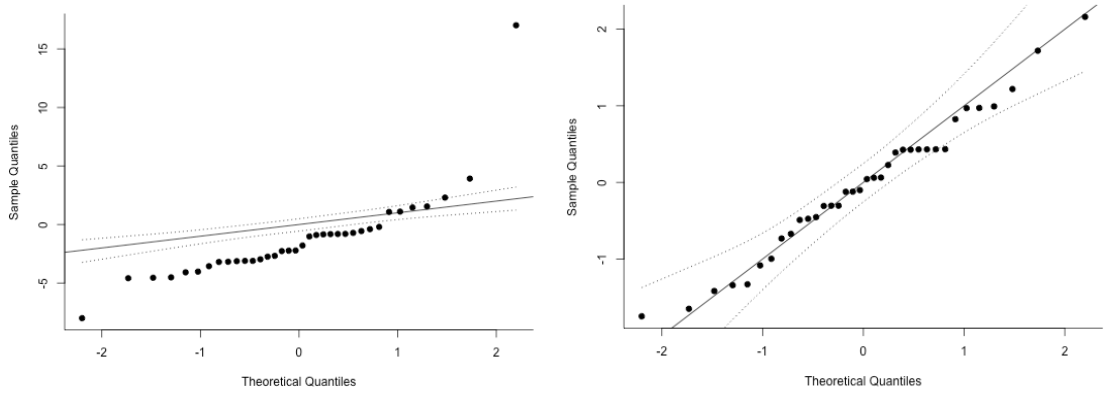
The trim and fill procedure yielded a corrected random effects model of alpha = 0.7290 (95% CI 0.6936, 0.7644). The imputed studies are shown as empty circles in Figure 1.8, and the corrected estimate inclusive of the imputed studies is alpha=0.7106 (95% CI 0.6519, 0.7694). The corrected estimate represents a -2.5203 % decrease relative to the original omnibus analysis. Nevertheless, the estimate of the average alpha coefficient correcting for

publication bias still produced an average alpha in excess of 0.7 and would not alter the interpretation of the omnibus test.

Analysis of the Compassion Fatigue Subscale

Selection of the meta-analytic model for the compassion fatigue subscales

The distribution of primary study effects is shown in Figure 2.1 for both the fixed effects model (FEM) and the random effects model (REM). The REM between studies variance (tau^2) was calculated using the DerSimonian-Laird estimator.



A B

Figure 2.1: QQ plot of the distribution of alpha coefficients within the primary studies for the compassion fatigue subscales. Chart A plots the fit of the fixed effects model and chart B plots the fit of the random effects model

As can be seen from Figure 2.1, there is clear evidence of non-normality in the distribution of alpha coefficients, for the compassion fatigue subscales, within the fixed effects model, whereas the Random effects model shows a good fit to the distributional assumptions of the REM. Therefore, the use of the random effects model using the DerSimonian-Laird estimator of between-subjects variance is an appropriate method for the synthesis of these data.

The omnibus test for the compassion fatigue subscales

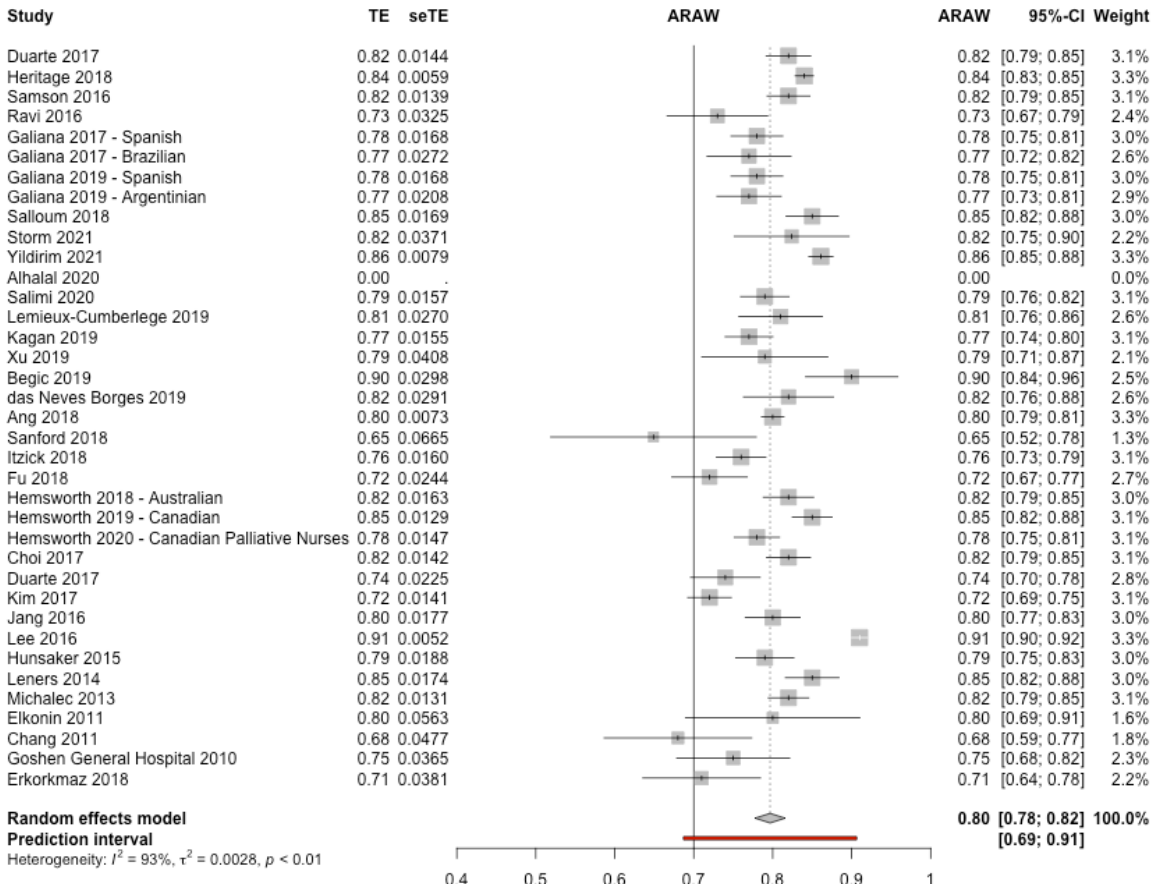
The alpha coefficients described in the primary studies are reported in Table 2.1. There were 32 studies reporting 37 effects and a total of 13,186 participants.

Table 2.1 – Effect of risk of bias (Compassion Fatigue Subscale)

Study Name	Alpha Coefficient	N	ProQOL Version	Language Version
Duarte 2017	0.82	390	ProQOL-5	Portuguese
Heritage 2018	0.84	1615	ProQOL-21	English
Samson 2016	0.82	377	ProQOL	Hebrew
Ravi 2016	0.73	155	ProQOL-5	English
Galiana 2017 - Spanish	0.78	385	ProQOL	Spanish
Galiana 2017 - Brazilian	0.77	161	ProQOL	Brazilian
Galiana 2019 - Spanish	0.78	385	ProQOL	Spanish
Galiana 2019 - Argentinian	0.77	273	ProQOL	Argentinian
Salloum 2018	0.85	177	ProQOL-5	English
Storm 2021	0.824	52	ProQOL	English
Yildirim 2021	0.861	697	ProQOL R-IV	English
Alhalal 2020		255	ProQOL-5	Arabic
Salimi 2020	0.79	400	ProQOL	Persian
Lemieux-Cumberlege 2019	0.81	112	ProQOL-5	English
Kagan 2019	0.77	494	ProQOL	English
Xu 2019	0.79	61	ProQOL	English
Begic 2019	0.9	27	ProQOL	English
das Neves Borges 2019	0.82	87	ProQOL-5	Portuguese
Ang 2018	0.8	1667	ProQOL	English
Sanford 2018	0.649	64	ProQOL	English
Itzick 2018	0.76	501	ProQOL Revised	English
Fu 2018	0.72	294	ProQOL-5	Chinese
Hemsworth 2018 - Australian	0.82	273	ProQOL-5	Australian
Hemsworth 2019 - Canadian	0.85	303	ProQOL-5	Canadian
Hemsworth 2020 - Canadian Palliative Nurses	0.78	503	ProQOL-5	Canadian
Choi 2017	0.82	358	ProQOL Revised	Korean
Duarte 2017	0.74	298	ProQOL-5	English
Kim 2017	0.72	875	ProQOL-5	English
Jang 2016	0.8	285	ProQOL-5	Korean
Lee 2016	0.91	680	ProQOL	English
Hunsaker 2015	0.79	278	ProQOL-5	English
Leners 2014	0.85	168	ProQOL	English
Michalec 2013	0.82	416	ProQOL-5	English
Elkonin 2011	0.8	30	ProQOL R-IV	English
Chang 2011	0.68	102	ProQOL	Chinese
Goshen General Hospital 2010	0.75	106	ProQOL R-IV	English
Erkorkmaz 2018	0.71	131	ProQOL	English

A random effects models was calculated using the generic inverse variance method. The random effects model suggested a weighted average alpha coefficient for the compassion fatigue subscales of α=0.7967 ($z = 82.22$, $p < 0.01$) and a 95% confidence interval of between 0.78 to 0.82.

Figure 2.2: Forest plot of alpha coefficients for the compassion fatigue subscales



A high level of heterogeneity in the primary studies was observed (Higgin's I^2 = 93%; tau^2 = 0.0028, $p < 0.01$), suggesting that the estimates of alpha coefficients in the primary studies may be biased by the presence of uncontrolled or confounding factors. Therefore, the focus of the subsequent analyses will be upon the identification of the sources of heterogeneity between the estimates of alpha coefficients between the primary studies.

The impact of influential primary studies for the compassion fatigue subscales

The impact of studies having a disproportionate influence was assessed using a "leave-one-out" analysis, in which the random effects model was calculated with each of the primary studies removed in turn and change in weighted average effect size (i.e., influence) and the

change in heterogeneity (i.e., discrepancy) was recorded. The result of this "leave-one-out" analysis is presented on the Baujat plot (Baujat, Pignon, & Hill, 2002) in Figure 2.3.

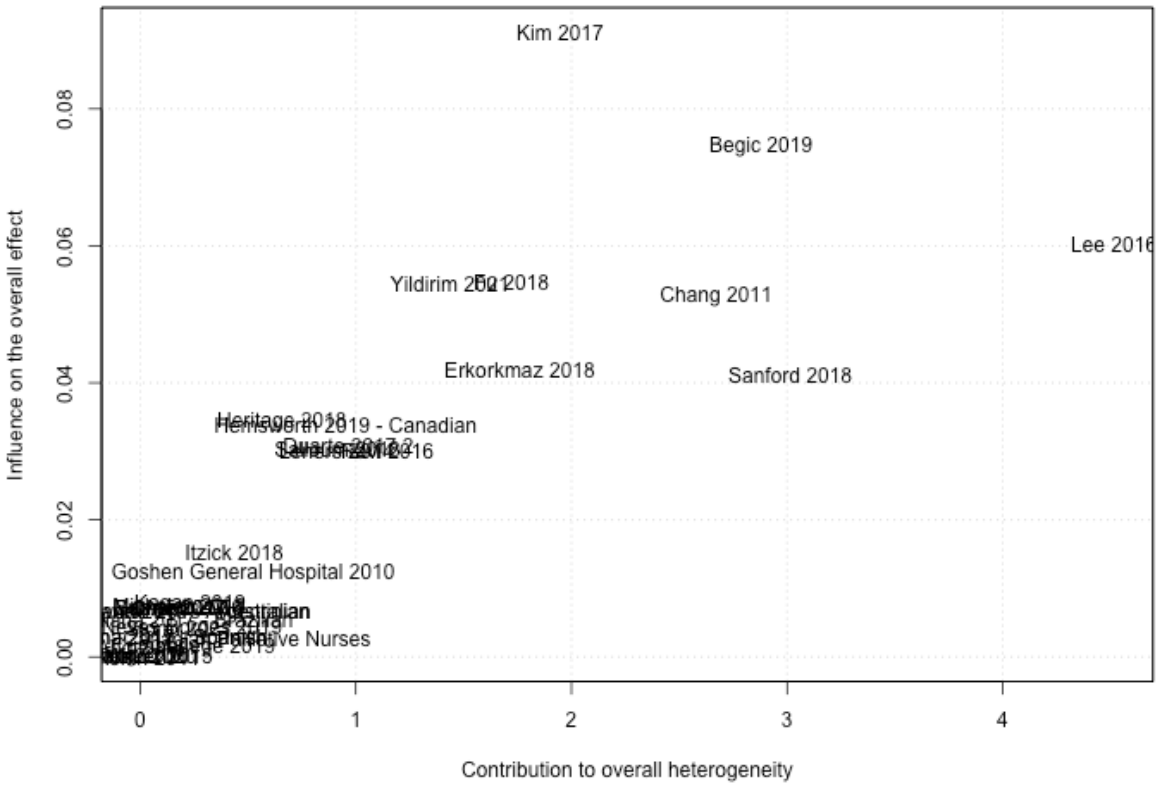


Figure 2.3: Baujat diagnostic plot of sources of heterogeneity. The vertical axis reports the influence of the study on the overall effect and the horizontal axis reports the discrepancy of the study with the rest of the literature.

The Baujat plot in Figure 2.3, suggests that the studies by Lee (2016), Begic (2019) and Kim (2017) are the most influential on the overall synthesis and the most discrepant from the rest of the literature. The random effects model was recalculated having removed the three studies showing disproportionate influence. The corrected random effects model reported a synthesis of Alpha = 0.7959 (95% CI 0.79 to 0.87). The corrected random effects model evidences a less than 0.02% decrease relative to the uncorrected estimate and did not change any of the conclusions from this meta-analysis.

The effect of risk of bias in the primary studies for the compassion fatigue subscales

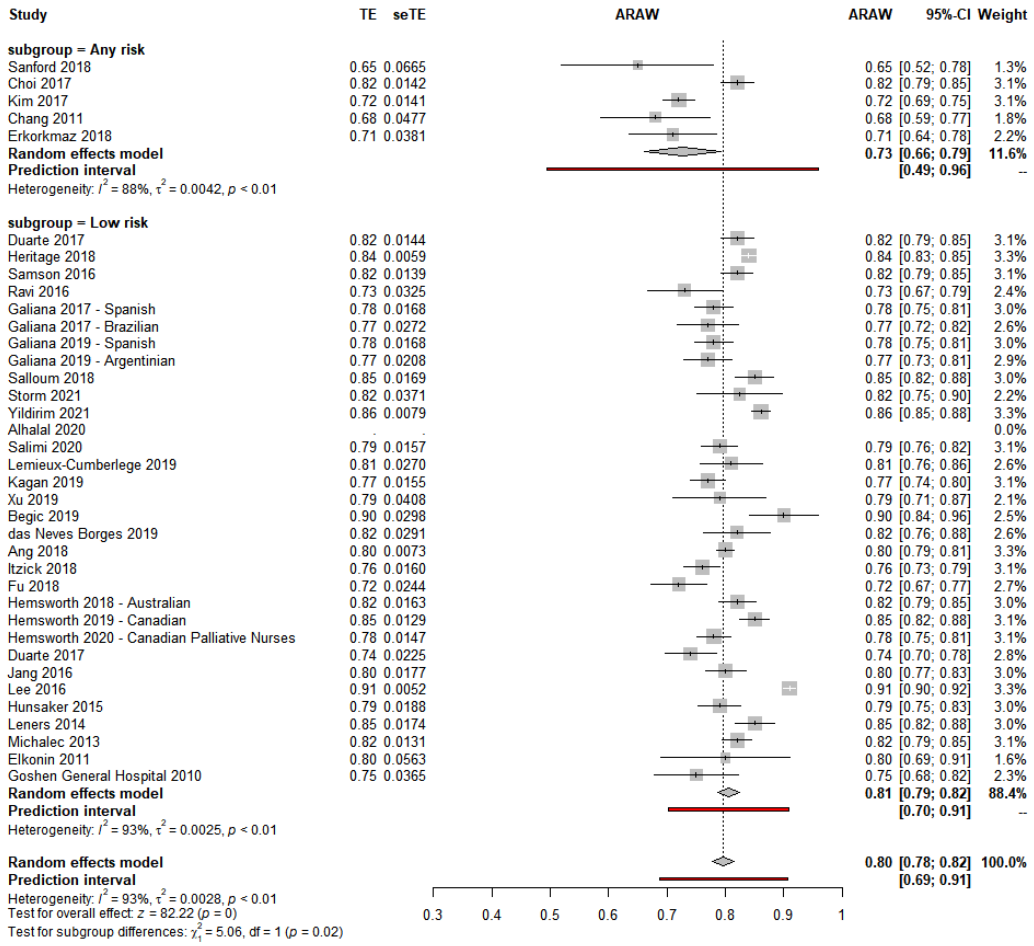
In order to assess the impact of study level risk of bias upon heterogeneity, a series of subgroup analysis were conducted on the alpha coefficients for the risk of bias ratings of "low risk" and "any risk" (i.e., unclear risk and high risk of bias combined) for each of the six types of methodological bias.

Table 2.2 – effect of risk of bias (Compassion Fatigue Subscale)

	Low Risk EFFECT	95% CI	k	Any Risk EFFECT	95% CI	k	X^2	P
Selection bias	0.8055	0.7864; 0.8246	31	0.7272	0.6616; 0.7927	5	5.06	0.0245
Performance bias	0.8020	0.7766; 0.8274	21	0.7916	0.7703; 0.8130	15	0.37	0.5406
Detection bias	0.8003	0.7769; 0.8236	26	0.7918	0.7713; 0.8123	10	0.29	0.5922
Statistical bias	0.7967	0.7777; 0.8157	36				--	--
Reporting bias	0.7957	0.7763; 0.8150	34	0.8188	0.6620; 0.9756	2	0.08	0.7741
Generalisability bias	0.7962	0.7770; 0.8155		0.7932	0.5776; 1.0000	2	0.00	0.9777

As indicated in Table 2.2. above, only selection bias evidenced statistically significant differences estimates of internal reliability for the compassion fatigue subscales, with lower levels of bias being associated with higher average alpha coefficients (see Figure 2.4).

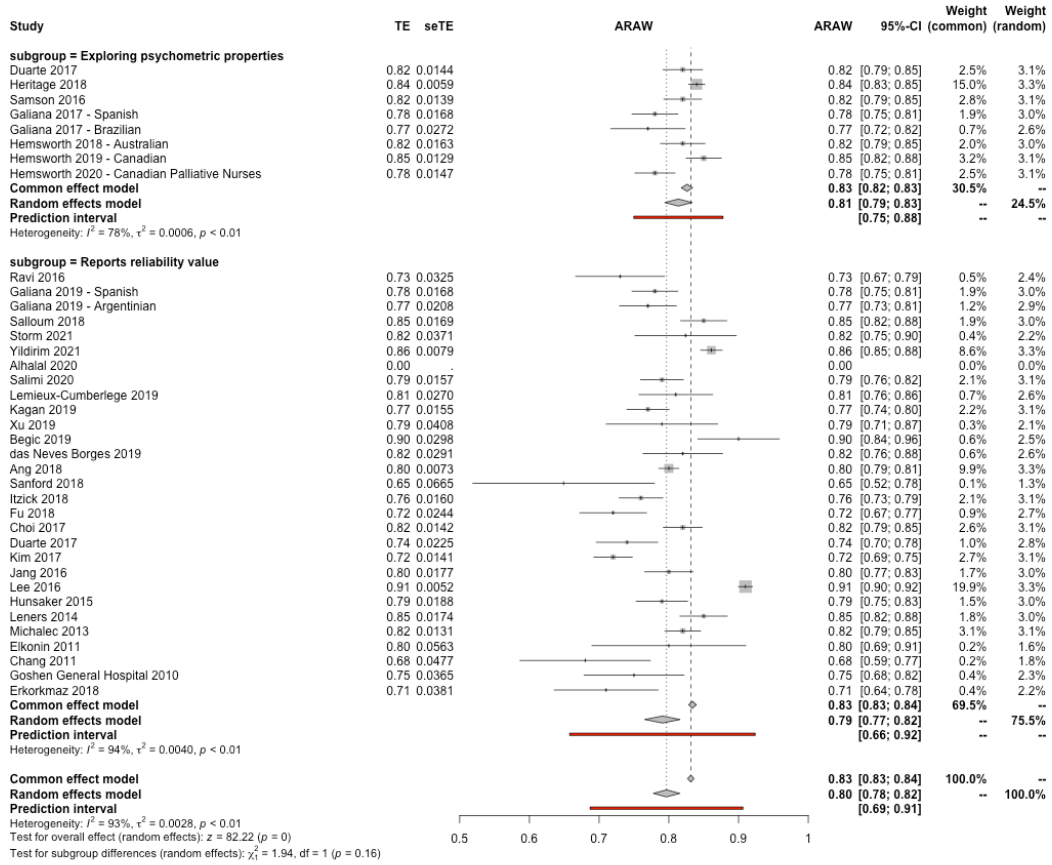
Figure 2.4: Subgroup plot of selection bias for the compassion fatigue subscales



Differences based on the primary aims of the included studies for the compassion fatigue subscales

The primary studies were divided into (a) those studies that were specifically designed to assess the psychometric properties of the ProQOL scale and (b) those studies that reported psychometric properties but were designed to address a different question. There was no significant difference in the average alpha coefficients for the two types of study design (X^2 = 1.94, p=0.16) (see Figure 2.5).

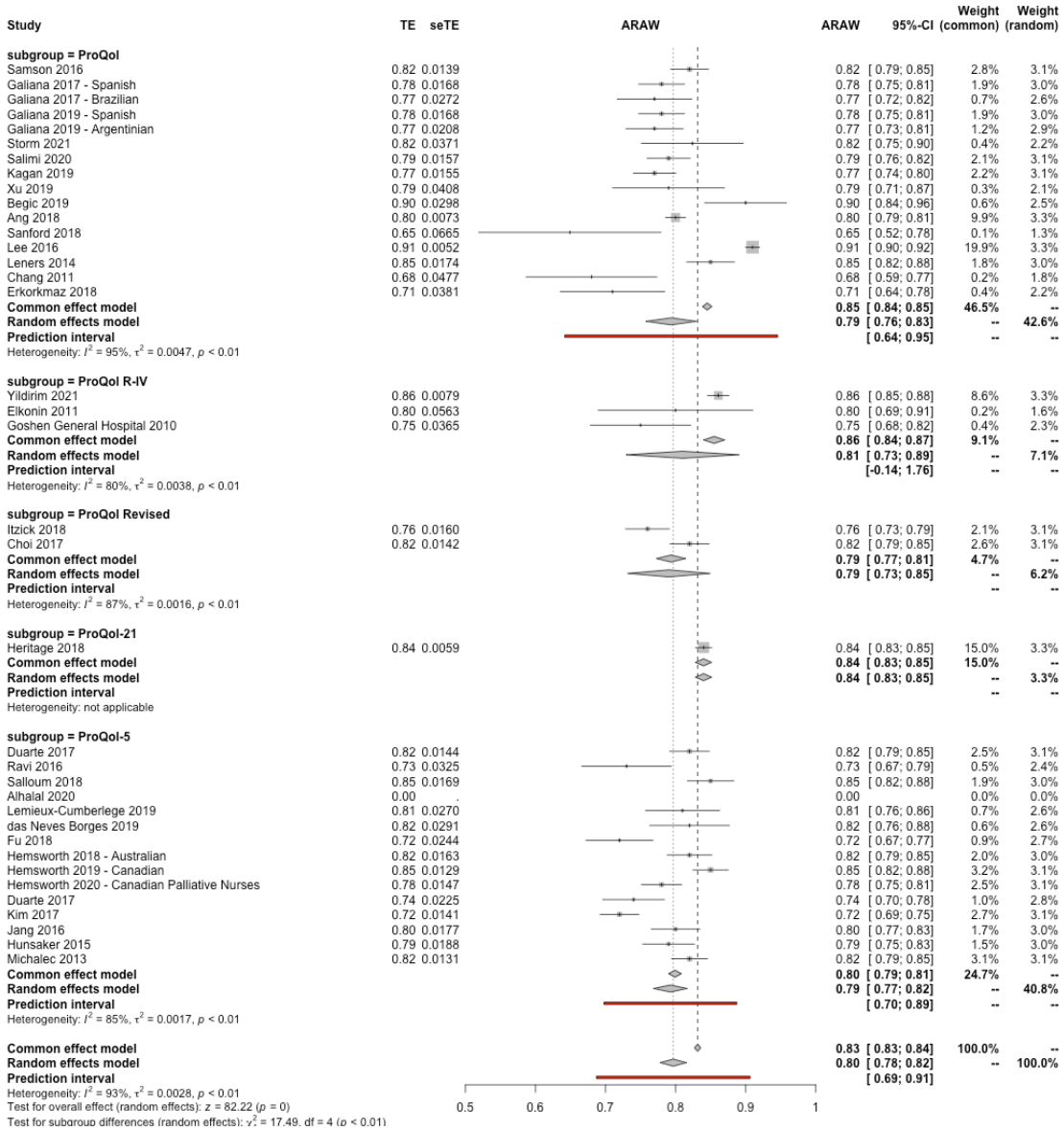
Figure 2.5 – Differences based on the primary aims of the studies



Differences between different versions of the ProQOL compassion fatigue subscales

The difference in alpha coefficients was calculated for the different versions of the ProQOL compassion fatigue subscales (see Figure 2.6).

Figure 2.6 – Differences based on different versions of the ProQOL compassion fatigue subscales

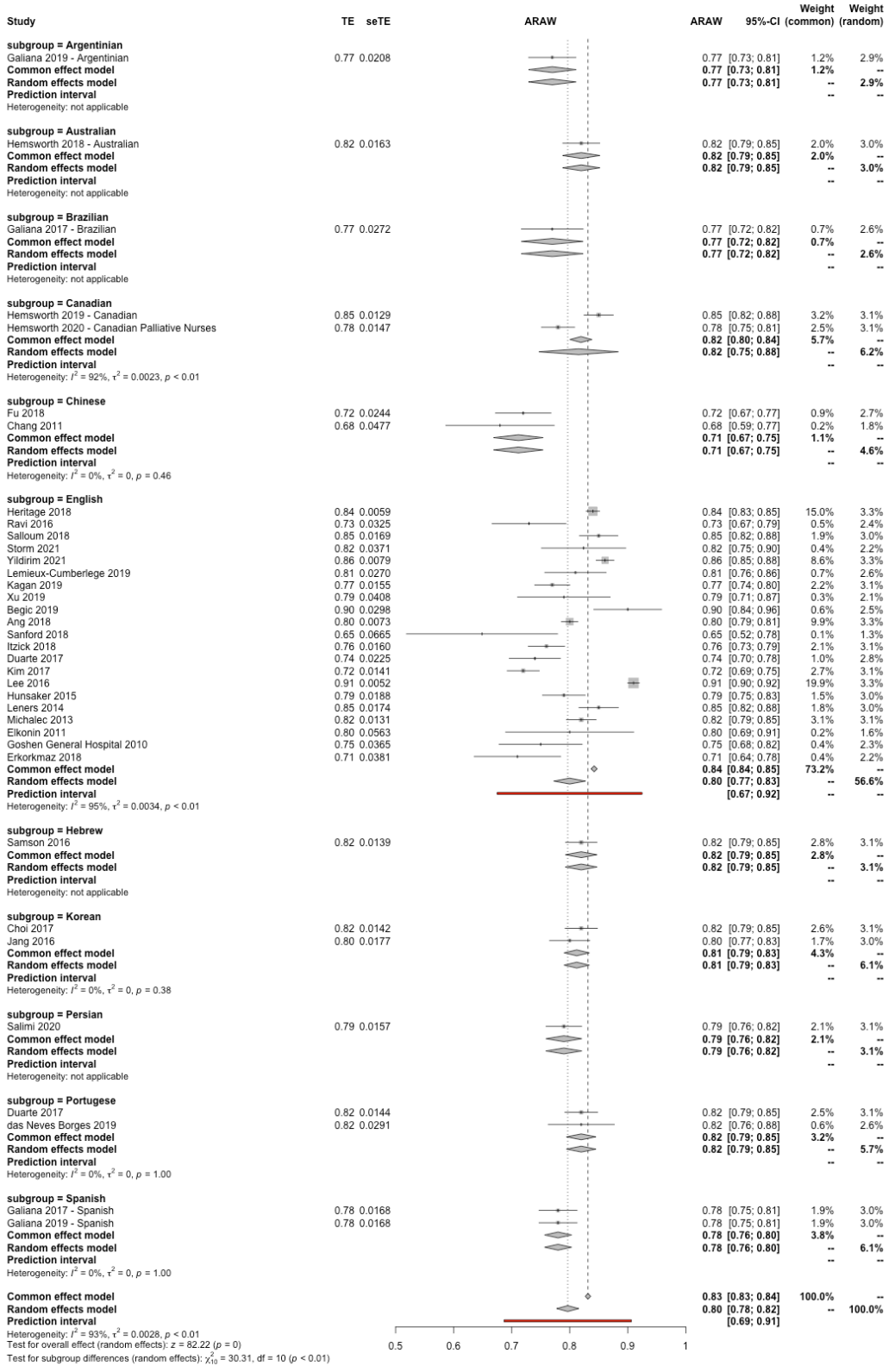


A statistically significant difference was observed between the internal consistency of the different versions of the ProQOL compassion fatigue subscales ($X^2 = 17.49$, $p < 0.01$). The only version that showed a significant difference from the overall mean was the ProQOL-21. It should be noted that the estimate of internal consistency of this version was obtained from a single study and it is likely that the average internal consistency score will vary with the publication of future studies.

Differences between different language versions of the ProQOL compassion fatigue subscales

The difference in alpha coefficients was calculated for the different language versions of the ProQOL compassion fatigues subscales (see Figure 2.7).

Figure 4.7: Subgroup plot of the difference in different language versions of the ProQOL compassions fatigue subscales

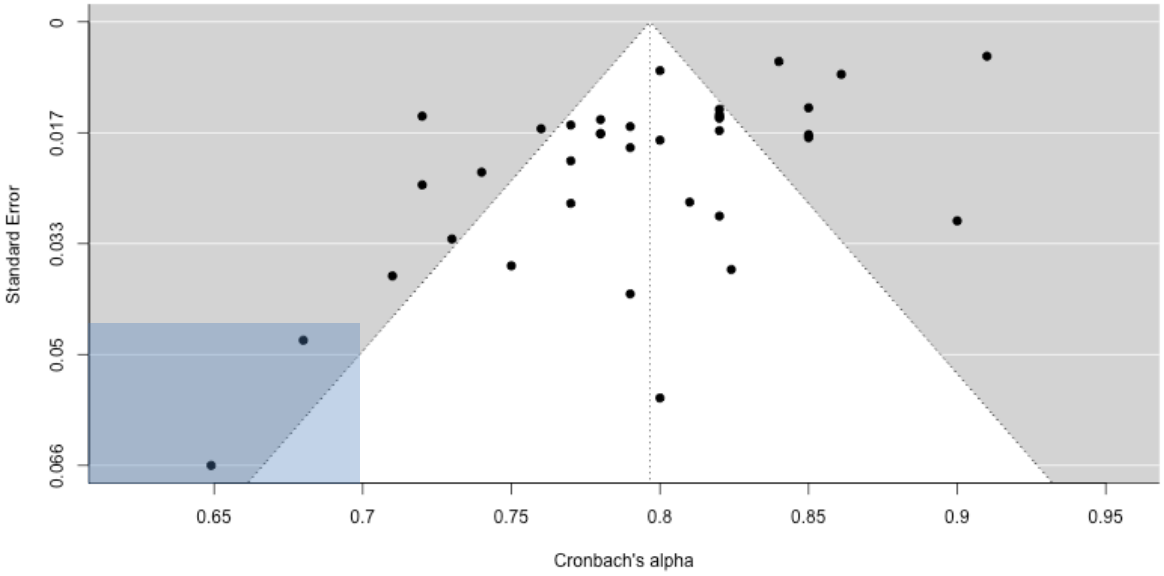


A statistically significant difference was observed between the internal consistency of the different language versions of the ProQOL compassion fatigue subscales ($X^2 = 30.31$, $p < 0.01$). The only language version that showed a significant difference from the overall mean was the Chinese version of the ProQOL scale which had a lower average alpha coefficient of alpha = 0.71 (95% CI 0.67 to 0.75). It should be noted that the estimate of internal consistency of this language version was obtained from only two studies, which both produced lower alpha coefficients than were reported for other language versions. Nevertheless, it is likely that the average internal consistency score will change with the publication of future studies.

The impact of publication and small study biases for the compassion fatigue subscales

Publication bias is caused by the tendency for statistically significant results to be published and the reticence to publish papers with non-significant results. Small study bias is the tendency for studies with smaller sample sizes to show greater variability in their measurement of internal consistency. These biases can be identified in a funnel plot, which plots the magnitude of the study's reported alpha coefficients (i.e., the importance of the study in the synthesis) against an estimate of the studies deviation from the meta-analytic average (i.e., the discrepancy of the study within the literature). If there is an absence of publication bias, the alpha coefficients reported from the studies with smaller sample sizes which show greater variability and will scatter more widely at the bottom of the plot compared to studies from larger samples which will lie closer to the overall meta-analytic effect, creating a symmetrical funnel shape. If there is an absence of studies in the area of the plot associated with small sample sizes and non-significant results, then it is likely there is some publication bias leading to an overestimation of the true effect. The funnel plot is presented in Figure 2.8.

Figure 2.8: funnel plot of the alpha coefficients for the compassion fatigue subscales. The 95% confidence interval of the expected distribution of alpha coefficients is shown as an inverted "funnel". Points filled in white represent imputed studies using the trim and fill procedure (Duval & Tweedle, 2000).

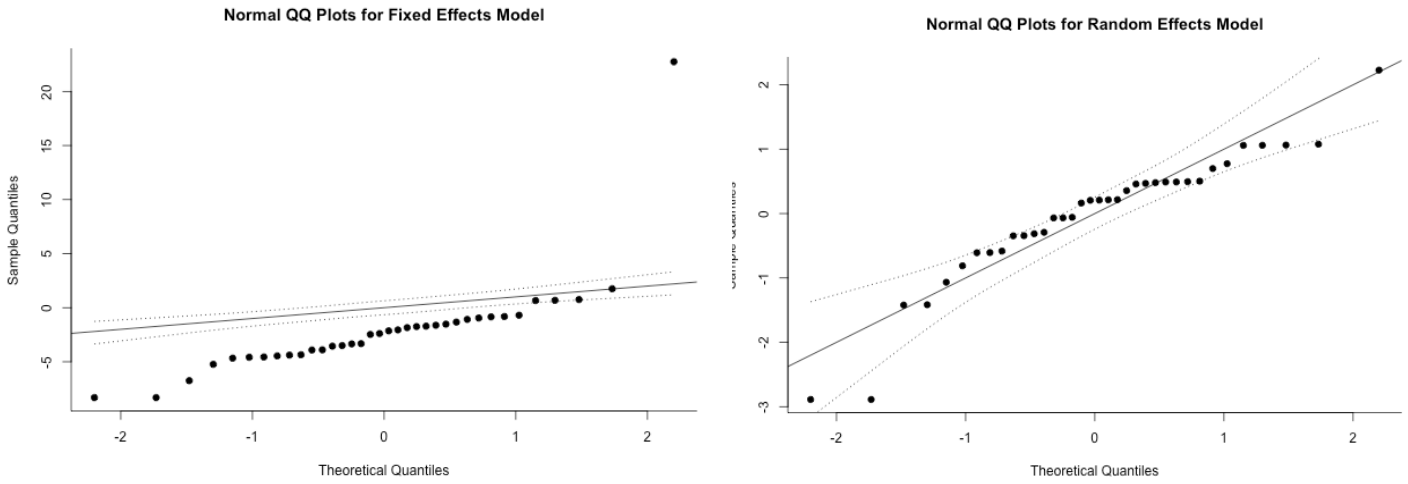


As can be seen from Figure 2.8, there is no clear evidence of publication bias in the distribution of alpha coefficients (as the small studies have tended to report lower alpha coefficients) for the compassion fatigue subscales, however, the previous noted heterogeneity in the reporting of this effect is clearly evident. Therefore, no simulation of and adjustment for publication bias and small study effects was undertaken.

Analysis of the Compassion Satisfaction Subscale

Selection of the meta-analytic model for the compassion satisfaction subscale

The distribution of primary study effects is shown in Figure 3.1 for both the fixed effects model (FEM) and the random effects model (REM). The REM between studies variance (tau2) was calculated using the DerSimonian-Laird estimator.



A B

Figure 3.1: QQ plot of the distribution of alpha coefficients within the primary studies for the compassion satisfaction subscales. Chart A plots the fit of the fixed effects model and chart B plots the fit of the random effects model

As can be seen from Figure 3.1 there is clear evidence of non-linearity in the distribution of alpha coefficients for the compassion satisfaction subscales, within the fixed effects model, whereas the random effects model shows a good fit to the distributional assumptions of the REM.

Therefore, the use of the random effects model using the DerSimonian-Laird estimator of between-subjects variance is an appropriate method for the synthesis of these data.

The omnibus test for the compassion satisfaction subscales

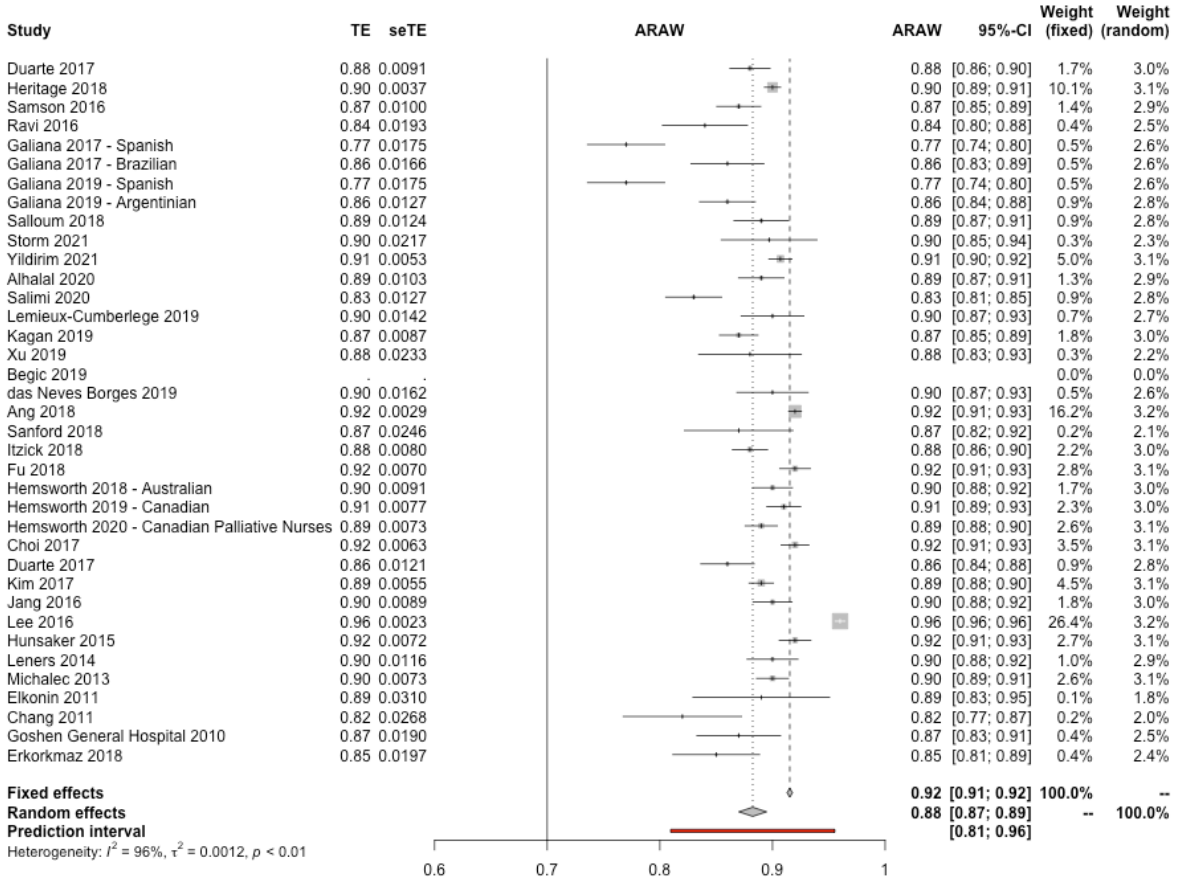
The alpha coefficients described in the primary studies are reported in Table 3.1. There were 33 studies reporting 37 effects, with a total of 13,416 participants.

Table 3.1: alpha coefficients for the compassion satisfaction subscale

Study Name	Alpha Coefficient	N	ProQOL Version	Language Version
Duarte 2017	0.88	390	ProQOL-5	Portuguese
Heritage 2018	0.9	1615	ProQOL-21	English
Samson 2016	0.87	377	ProQOL	Hebrew
Ravi 2016	0.84	155	ProQOL-5	English
Galiana 2017 - Spanish	0.77	385	ProQOL	Spanish
Galiana 2017 - Brazilian	0.86	161	ProQOL	Brazilian
Galiana 2019 - Spanish	0.77	385	ProQOL	Spanish
Galiana 2019 - Argentinian	0.86	273	ProQOL	Argentinian
Salloum 2018	0.89	177	ProQOL-5	English
Storm 2021	0.897	52	ProQOL	English
Yildirim 2021	0.907	697	ProQOL R-IV	English
Alhalal 2020	0.89	255	ProQOL-5	Arabic
Salimi 2020	0.83	400	ProQOL	Persian
Lemieux-Cumberlege 2019	0.9	112	ProQOL-5	English
Kagan 2019	0.87	494	ProQOL	English
Xu 2019	0.88	61	ProQOL	English
das Neves Borges 2019	0.9	87	ProQOL-5	Portuguese
Ang 2018	0.92	1667	ProQOL	English
Sanford 2018	0.87	64	ProQOL	English
Itzick 2018	0.88	501	ProQOL Revised	English
Fu 2018	0.92	294	ProQOL-5	Chinese
Hemsworth 2018 - Australian	0.9	273	ProQOL-5	Australian
Hemsworth 2019 - Canadian	0.91	303	ProQOL-5	Canadian
Hemsworth 2020 - Canadian Palliative Nurses	0.89	503	ProQOL-5	Canadian
Choi 2017	0.92	358	ProQOL Revised	Korean
Duarte 2017	0.86	298	ProQOL-5	English
Kim 2017	0.89	875	ProQOL-5	English
Jang 2016	0.9	285	ProQOL-5	Korean
Lee 2016	0.96	680	ProQOL	English
Hunsaker 2015	0.92	278	ProQOL-5	English
Leners 2014	0.9	168	ProQOL	English
Michalec 2013	0.9	416	ProQOL-5	English
Elkonin 2011	0.89	30	ProQOL RIV	English
Chang 2011	0.82	102	ProQOL	Chinese
Goshen General Hospital 2010	0.87	106	ProQOL R-IV	English
Erkorkmaz 2018	0.85	131	ProQOL	English

A random effects models was calculated using the generic inverse variance method. The random effects model suggested a weighted average alpha coefficient for the compassion satisfaction subscales of α=0.8824 ($z = 140.05$ $p < 0.01$) and a 95% confidence interval of between 0.8701 to 0.8948 (see Figure 3.2).

Figure 3.2: Forest plot of alpha coefficients for the compassion satisfaction subscales

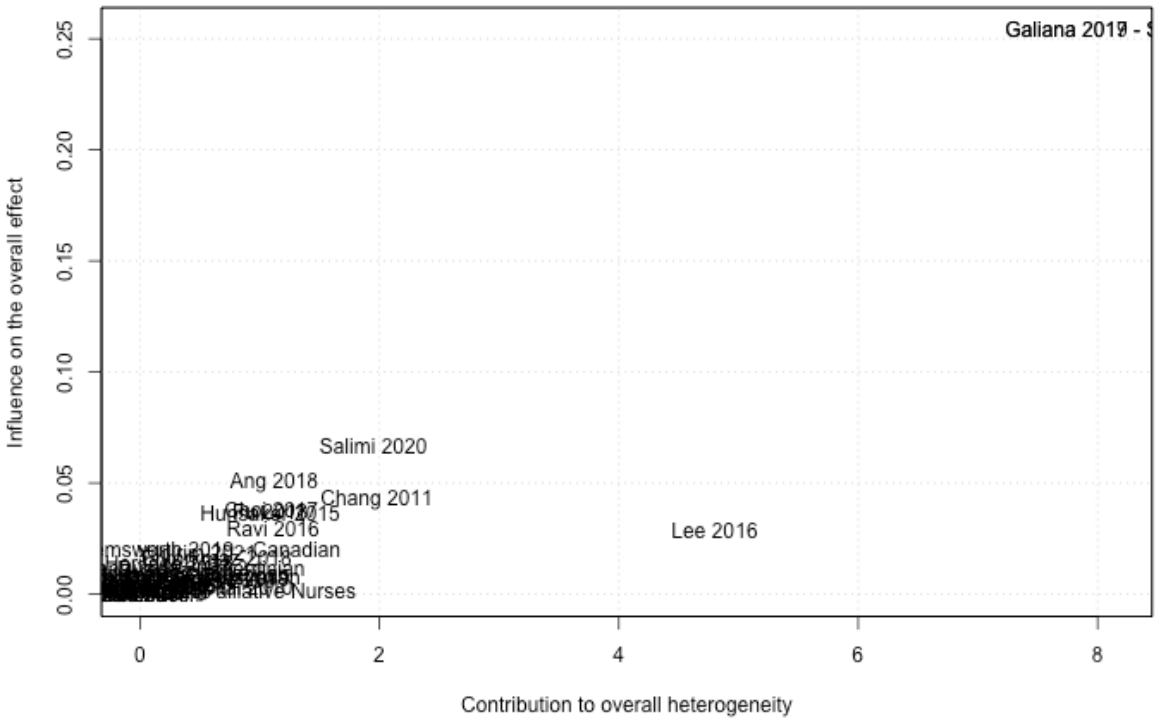


A high level of heterogeneity in the primary studies was observed (Higgin's I^2 = 96%, tau^2 = 0.0012, p = < 0.01) suggesting that the estimates of alpha coefficients in the primary studies may be biased by the presence of uncontrolled or confounding factors. Therefore, the focus of the subsequent analyses will be upon the identification of the sources of heterogeneity between the estimates of alpha coefficients between the primary studies.

The impact of influential primary studies for the compassion satisfaction subscales

The impact of disproportionately influence studies was assessed using a "leave-one-out" analysis, in which the random effects model was calculated with each of the primary studies removed in turn and change in weighted average effect size (i.e., influence) and the change in heterogeneity (i.e., discrepancy) was recorded. The result of this "leave-one-out" analysis is presented on the Baujat plot (Baujat, Pignon, & Hill, 2002) in Figure 3.3.

Figure 3.3: Baujat plot indicating impact of influential studies for the compassion satisfaction subscales



The Baujot plot in Figure 3.3 above suggests the studies by Galiana (2019) and Lee (2016) are the most influential on the overall synthesis and the most discrepant from the rest of the literature.

The random effects model was recalculated having removed the two studies showing disproportionate influence. The corrected random effects model reported a synthesis of Alpha = 0.88 (95% CI 0.877 to 0.894). The corrected random effects model evidenced a less than 0.005% decrease relative to the uncorrected estimate and did not change any of the conclusions from this meta-analysis. Accordingly, this meta-analysis can be considered robust to the effect of influential and discrepant studies.

The effect of risk of bias in the primary studies for the compassion satisfaction subscales

In order to assess the impact of study level risk of bias upon heterogeneity, a series of subgroup analysis were conducted on the alpha coefficients for the risk of bias ratings of "low risk" and "any risk" (i.e., unclear risk and high risk of bias combined) for each of the six types of methodological bias.

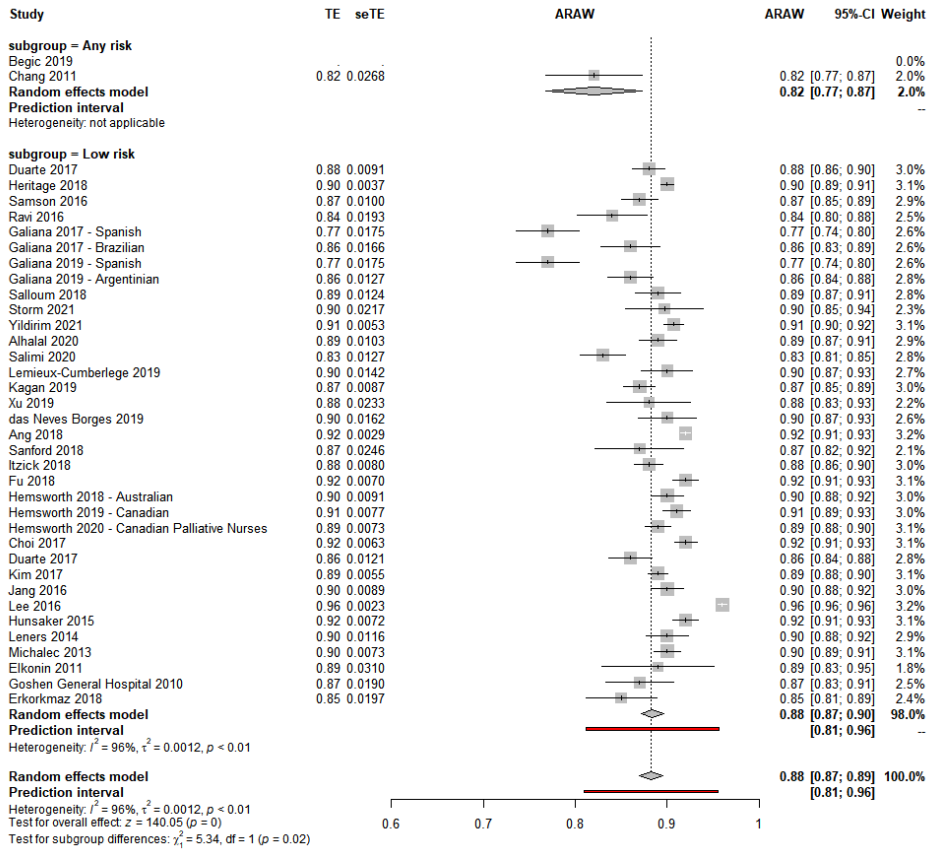
Table 3.2 – effect of risk of bias (compassion satisfaction subscale)

	Low Risk			Any Risk				
	Alpha	**95% CI**	**Studies**	**Alpha**	**95% CI**	**Studies**	**X^2**	**P**
Selection bias	0.8834	0.8698; 0.8969	31	0.8780	0.8504; 0.9056	5	0.12	0.7307
Performance bias	0.8773	0.8578; 0.8968	21	0.8896	0.8784; 0.9008	15	1.16	0.2823
Detection bias	0.8870	0.8725; 0.9015	25	0.8723	0.8499; 0.8946	11	1.18	0.2781
Statistical bias	0.8824	0.8701; 0.8948	36				--	--
Reporting bias	0.8829	0.8702; 0.8956	34	0.8757	0.8463; 0.9051	2	0.19	0.6598
Generalisability bias	0.8837	0.8713; 0.8961	35	0.8200	0.7674; 0.8726	1	5.34	0.0208

As can be seen from Table 3.2 above, only generalisability bias evidenced statistically significant differences between estimates of internal reliability for the compassion satisfaction subscales, with lower levels of bias associated with higher average alpha coefficients (see Figure 3.4).

It should be noted that the presence of generalisability bias did not result in any substantive alterations in the conclusions for this outcome.

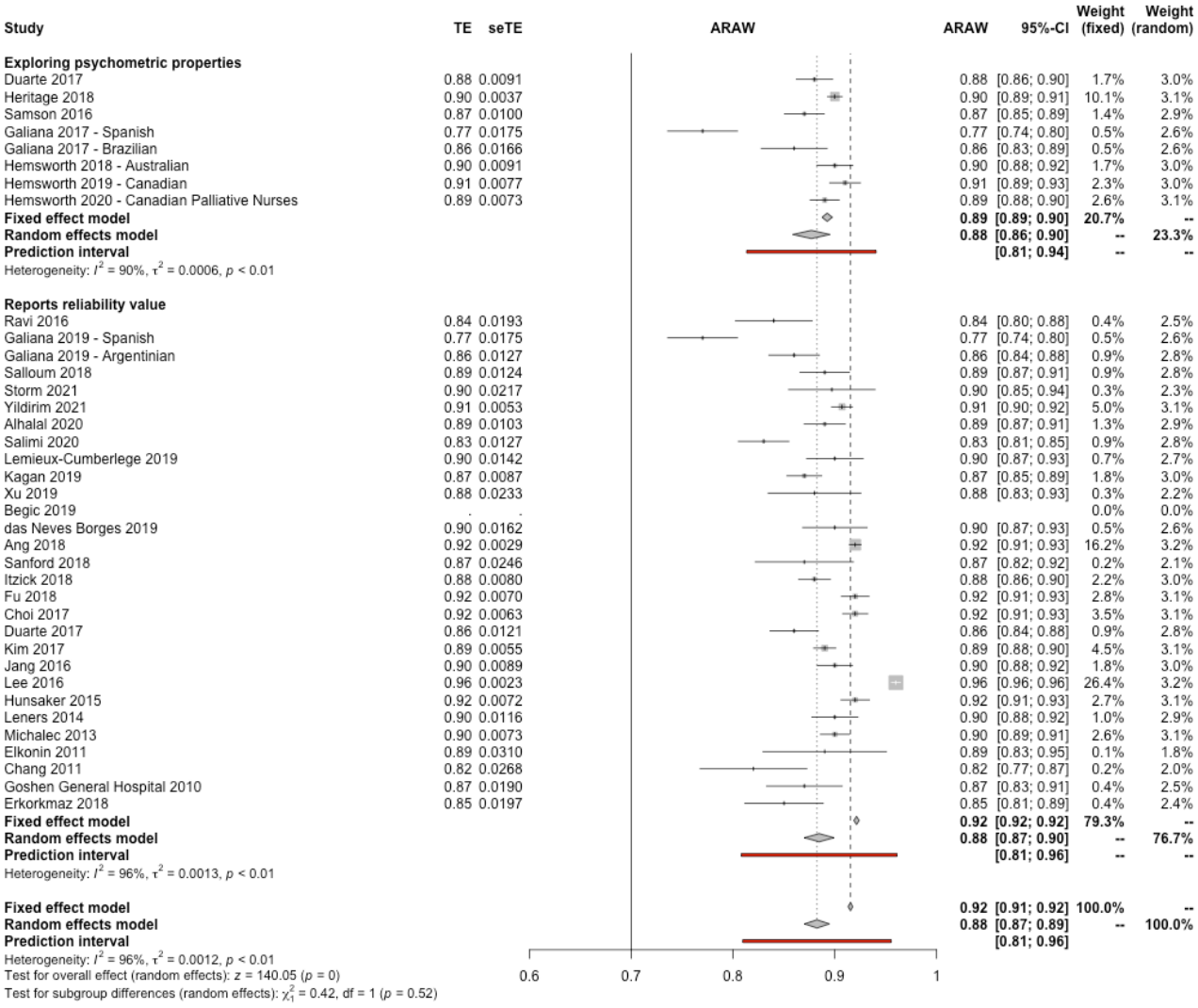
Figure 3.4: Subgroup plot of generalisability bias for the compassion satisfaction subscales



Differences based on the primary aims of the included studies for the compassion satisfaction subscales

The primary studies were divided into (a) those studies that were specifically designed to assess the psychometric properties of the ProQOL scale and (b) those studies that reported psychometric properties but were designed to address a different question. There was no significant difference in the average alpha coefficients for the two types of study design (X^2 = 0.42, p=0.52) (see Figure 3.5).

Figure 3.5. Difference based on the primary aims of the studies



Differences between different versions of the ProQOL compassion satisfaction subscales

The difference in alpha coefficients was calculated for the different versions of the ProQOL scale. No statistically significant difference was observed between the internal reliability of the different versions of the ProQOL compassion satisfaction subscales (X^2 = 7.02, p=0.13).

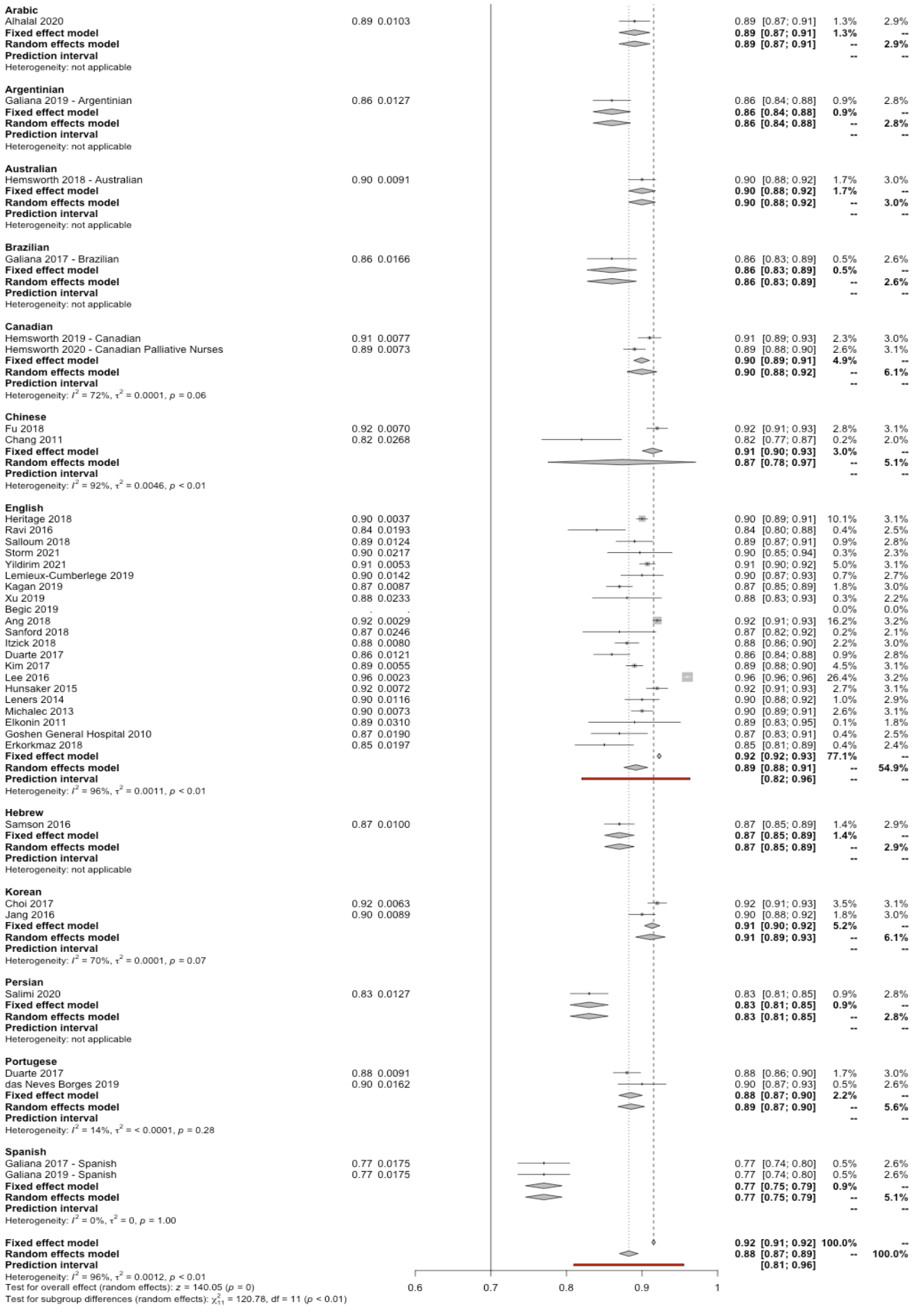
Differences between language versions of the ProQOL compassion satisfaction subscales

The difference in alpha coefficients was calculated for the different language versions of the ProQOL compassion satisfaction subscales (see Figure 3.6)

A statistically significant difference was observed between the internal consistency of the different language versions of the ProQOL compassion satisfaction subscale (X^2 = 120.78, p <0.01). The two language versions that showed a significant difference from the overall mean were the Persian language version of the ProQOL scale with an alpha coefficient of α = 0.83 (95% CI 0.81 to 0.85) and the Spanish language version with an alpha coefficient of α = 0.77 (95% CI of 0.75 to 0.79).

It should be noted that the estimates of internal consistency of these language versions were obtained from only one and two studies, respectively, which both produced lower alpha coefficients than were reported for other language versions. Therefore, it is likely that the average internal consistency score will change with the publication of future studies.

Figure 3.6 – Differences between different language versions of the ProQOL compassion satisfaction subscales

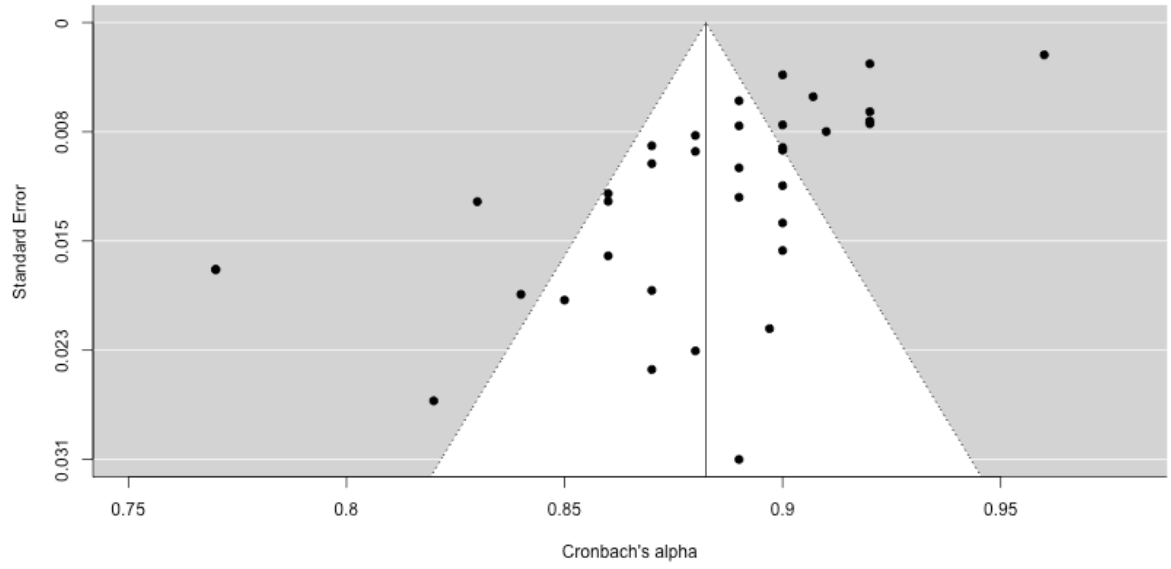


The impact of publication and small study biases for the compassion satisfaction subscales

Publication bias is caused by the tendency for statistically significant results to be published and the reticence to publish papers with non-significant results. Small study bias is the tendency for studies with smaller sample sizes to show greater variability in their measurement of internal consistency. These biases can be identified in a funnel plot, which plots the magnitude of the study's reported alpha coefficients (i.e., the importance of the study in the synthesis) against an estimate of the studies deviation from the meta-analytic average (i.e., the discrepancy of the study within the literature).

If there is an absence of publication bias, the alpha coefficients reported from the studies with smaller sample sizes which show greater variability and will scatter more widely at the bottom of the plot compared to studies from larger samples which will lie closer to the overall meta-analytic effect, creating a symmetrical funnel shape. If there is an absence of studies in the area of the plot associated with small sample sizes and non-significant results, then it is likely there is some publication bias leading to an overestimation of the true effect. The funnel plot is presented in Figure 3.7.

Figure 3.7. funnel plot of the alpha coefficients for the compassion satisfaction subscales. The 95% confidence interval of the expected distribution of alpha coefficients is shown as an inverted "funnel". Points filled in white represent imputed studies using the trim and fill procedure (Duval & Tweedle, 2000).



The high levels of heterogeneity reported above are reflected in the funnel plot by the larger than expected number of studies outside of the 95% confidence interval of the meta-analysis. However, in the area of the plot associated with publication bias (i.e. relatively low alpha coefficients derived from small samples) there is no obvious emission of studies. Therefore, estimation and correction for publication bias is not required.

Discussion

This literature review aimed to explore the reliability of the reported Cronbach's Alpha for each of the three subscales of the ProQOL scale. As there are different versions of the scale as well as different language versions, used across different countries and populations, reliability was examined across the different subscales and versions.

The scale has been revised several times to address and improve psychometric difficulties noted in previous versions. Furthermore, to acknowledge difficulties separating burnout and secondary traumatic stress items have been reworded in different versions, based on recent research to better reflect the core concepts of the subscales. Due to a complex relationship between the subscales and still some ambiguity regarding the relationship between them, it is important to note the ProQOL manual emphasises the lack of a composite score for the three subscales combined. Specifically, there is an identified issue of collinearity between the subscales from the original versions, and repeated adaptations across versions aimed to reduce this. Therefore, in the current analysis, it was deemed appropriate to explore reliability for each subscale for the different ProQOL versions and language versions separately, as suggested by the manual.

This ties in with the literature reporting reliability, specifically internal consistency, of the ProQOL scale, where alpha coefficients are reported separately for each of the subscales. Only three of the forty studies retrieved from the initial search reported alpha values for the ProQOL scale as a whole outcome measure, with thirty-seven studies reporting the subscales separately. This, alongside the original authors acknowledgements of the importance of treating the subscales separately, makes it difficult to ascertain how reliable the ProQOL scale is as a total outcome measure.

Results for subscales

The findings from this review suggest the burnout subscale possessed the lowest reliability of the three subscales (0.48-0.94). The compassion fatigue subscale had better reliability (0.64-0.91), with the compassion satisfaction subscale demonstrating the strongest reliability (0.77-0.96).

Furthermore, the alpha coefficients reported for the subscales suggest some discrepancy between studies. For instance, for the burnout subscale, alpha values range from 0.48 (Michalec et al., 2013) to 0.94 (Alhalal et al., 2020). These studies reported alpha coefficients for the English and Arabic language versions of the ProQOL-5 scale, with large sample sizes of 416 and 255 participants, respectively. The only version of the ProQOL which showed a significantly different alpha coefficient from the overall mean was the ProQOL-21, with no significant difference noted between all other versions. As only one study used this version, it is likely that the average internal consistency score will vary with the publication of future studies. Furthermore, a statistically significant difference was observed between the internal consistency of the different language versions of the ProQOL burnout subscales.

There were a large number of studies that reported acceptable alpha coefficient values >0.7 for different versions of the ProQOL burnout scale and different language versions. However, just less than a quarter of the studies included reported alpha values of <0.7. These studies included different versions of the ProQOL (specifically ProQOL-5, ProQOL, ProQol R-IV and the ProQOL revised) as well as different language versions across a range of professional groups across various countries. This may explain the difference in reliability. For instance, the Chinese language version was used on soldiers in China (Chang & Taormina, 2011); the English version was used on nurses in South Africa (Elkonin & Van der Vyver, 2011), nurses in Turkey (Erkorkmaz et al. 2018), social workers in Israel (Kagan & Itzick 2019, Itzick et al. 2018) and audiologists and speech and language pathologists in India (Ravi, 2016); Galiana et al. (2017) and Galiana et al. (2019) also used palliative care professionals in Spain, Brazil and Argentina whereas Samson (2016) used the Hebrew version on palliative care professionals in Israel.

As there is a wide variation in the range of professional groups who the scale has been administered to, this may be linked to differences in work ethic, expectations and approaches to working with and/or supporting those who have experienced trauma. For instance, nurses are often expected to support those who have experienced trauma as part of their caring roles, although there is acknowledgement that more than a quarter of nurses exceeded the threshold for burnout (Marchand et al., 2015). Military personnel are believed to be more resilient to stress and trauma due to exposure to military training, resulting in lower levels of burnout and increased tolerance to high-intensity stress or screening out inidividuals who demonstrate lower stress tolerance when selecting soldiers (Mohammad, 2012). Nonetheless, evidence suggests that although soldiers are increasingly likely to experience burnout due to demographic factors, job characteristics and personal characteristics including workplace stress, social support, social skills, group culture and individual contributions to their working factions (Wu et al., 2022), they tend to underreport symptoms of burnout (Mohammad, 2012). Other professional groups such as speech and language pathologists are also susceptible to developing burnout due to emotional connections formed with their patients, witnessing depressive symptoms, receiving aggression from patients and pressure for instant improvements from carers and families (Maura et al., 2019). Nonetheless, as there is little research on burnout in audiologists and speech and language pathologists (Brito-Marcelino et al., 2020), this may contribute to differing levels of burnout reported and may change with the publication of future research.

The difference in language versions may also contribute to differences in reported reliability; although they have been adapted to make sense linguistically to their populations, it can be argued that even minor changes can have a significant impact on the interpretation of burnout and does not hold the same value or convey the same meaning as intended in the original English version. This view is supported by Hambleton & De Jong (2003) who suggest translated and adapted instruments will almost always be different to the original language version.

Additionally, differences in cultural values across different countries regarding burnout may also contribute to differences in reported reliability. A wide variation in cultural diversity has significant impacts on health seeking behaviours, attitudes of patients and practitioners as

well as how health and illness are perceived (Gopalkrishnan, 2018). Some evidence suggests chronic stress resulting from perceived threat is one of the biggest causes of burnout (Johnson & Naidoo, 2017) where perceived threat is largely influenced by cultural conditions (Bracken, 2002). Nonetheless, further evidence is needed to provide a better understanding of how burnout develops in different countries and contexts; at present it remains unclear how culture-specific values, behavioural norms and perceptions of the self in relation to one's society may impact the development of burnout (Barker et al., 2021).

Therefore, the reliability of the burnout dimension for the ProQOL scale has some inconsistencies, although this may represent the varied contexts (language, culture, healthcare professions) that it has been used across, as reliability appeared to vary not only between different versions of the ProQOL, but also between different language versions. Until such inconsistencies have been better explored and understood, at present, the result of variable reliability reduces confidence in the reliability and the scale should be approached with some caution particularly when used across different groups, languages and cultures.

In comparison, the compassion fatigue subscale appeared to demonstrate better reliability. Although one study (Alhalal et al., 2020) did not report an alpha value and two studies (Chang & Taormina, 2011; Sanford et al., 2018) reported alpha values of <0.7 (0.68 and 0.649, respectively), all other studies reported alpha values of >0.7, with Lee & Seomun (2016) reporting the highest alpha value (0.91) for this subscale. The only version of the ProQOL which showed a significant difference from the overall mean was the ProQOL-21, with no significant difference noted between all other versions. Importantly, only one study used this version. Furthermore, the only language version that showed a significant difference from the overall mean was the Chinese version which was obtained from only two studies and is likely to change with the publication of future studies. As there is more consistency across reported reliability values and within the "acceptable" range, across different versions and languages, there may be an increased sense of confidence when assuming reliability for this subscale.

Similarly, the alpha values reported for the compassion satisfaction subscale would suggest this subscale has good reliability, with all studies reporting alpha values of >0.7.

Galiana et al. (2017) & Galiana et al. (2019) reported the lowest alpha value (0.77) which was still greater than the acceptable threshold for reliability. Lee & Seomun (2016) reported the highest alpha value of 0.96, reflecting high reliability. All other studies reported alpha values ranging from 0.82-0.92, which is generally deemed to be good and increases confidence when assuming reliability for this subscale. There was no significant difference observed between the internal reliability of the difference versions of the ProQOL subscale. However, a statistically significant difference was observed between the internal consistency of the different language versions of the ProQOL compassion satisfaction subscales, specifically for the Spanish and Persian language versions. This included only three studies and is likely to change with the publication of further studies.

It is important to note the overall quality of the studies based on risk of bias varies greatly. Although the study by Heritage et al. (2018) has a perfect overall quality score of 100% with Duarte (2017), Galiana et al. (2017), Hemsworth et al. (2018) and Samson et al. (2016) also possessing good overall quality (95%, 95%, 95% and 91%, respectively), three studies have a much lower overall quality of 36% (Chang & Taormina, 2011) and 41% (Alhalal et al., 2020 & Choi & Lee, 2017). The quality scores for all other studies ranges between 45-55%. This would suggest interpreting the reported alpha values and reliability of the ProQOL scale with caution as the overall quality of the studies appears to be mixed with some indicating low quality.

When the studies were split into ‘any risk’ (comprising of high risk of bias and unclear risk of bias) and ‘low risk’ studies for each of the risk of bias criteria, there were no significant differences between the weighted average alpha coefficients for the low and high risk of bias studies for any of the risk of bias criteria for the burnout subscale. Selection bias, reporting bias and generalisability bias had only five, three and two studies, respectively, rated as any risk. The other risk of bias criteria comprised a better balance of any risk and low risk studies. Although, as mentioned, this did not result in any significant differences.

For the compassion fatigue subscale, only selection bias evidenced statistically significant differences in estimates of internal reliability for low risk and any risk studies, with only 5 studies being rated as any risk.

Finally, for the compassion satisfaction subscale, only generalisability bias evidenced statistically significant differences between estimates of internal reliability for low risk and any risk studies. There was only one study for this criterion rated as any risk of bias.

Clinical implications

As there were generally not many studies rated as any risk, this makes it difficult to judge the impact of the overall quality estimates for the studies and may reduce confidence in the overall conclusions about the reliability of the subscales across different contexts i.e. healthcare and different countries.

Although the presence of bias for the compassion fatigue and compassion satisfaction subscales did not appear to result in any substantive alterations in the conclusions of this review, it is important to note that the burnout subscale possessed lower reliability and broader range of alpha coefficients reported (0.48-0.94) than the compassion fatigue (0.64-0.91) and compassion satisfaction subscales (0.77-0.96). Importantly, there was no statistically significant difference observed between different versions of the scale.

This is significant when any version of the scale is used as a clinical tool for clinicians or healthcare professionals when exploring any of the subscales. For instance, when assessing levels of burnout in clinical or healthcare staff, the scale may not accurately reflect staff experiences. Their difficulties may not be acknowledged, and they may continue to work with prolonged feelings of stress and burnout which evidence suggests is associated with a number of physical health conditions including pain, fatigue, cardiovascular disorders, respiratory diseases, diabetes, gastrointestinal difficulties and significantly premature death (Salvagioni et al., 2017). These authors further state negative psychological consequences of burnout resulting in the use of psychotropic medication and admission to hospital for mental disorders. This was found to be associated with a 21% increase in work absence days due to sickness, regardless of health status, sociodemographic status and working conditions (Borritz et al., 2006). Therefore, if burnout is not measured accurately, there is a range of potential negative consequences for clinicians.

Additionally, reviews on compassion in healthcare, encompassing compassion fatigue (Sinclair et al., 2016) have supported the notion that the construct has been under studied, with many emerging critiques of compassion fatigue lacking a universal definition (Hofmeyer et al., 2020). Without this, it is difficult to differentiate when staff might be experiencing compassion fatigue, compared to conditions with similar symptoms such as PTSD, disruptions to cognitive abilities, difficult relationship dynamics, emotional distress and physical pain experiences (Sodeke-Gregson et al., 2013). This is likely to impact not only personal relationships, but also their ability to effectively work with colleagues and deliver patient care to a good standard (Collins & Long, 2003); nurses who experience burnout and fatigue lack empathy and compassion, often overlooking patient needs due to feelings of depersonalization, and deliver a reduced standard of care (Bramley & Matiti, 2014).

Due to this lack of universal definition, one can argue the ProQOL scale, regardless of which version is used, does not measure compassion effectively as it lacks construct validity (Ledoux, 2015). Some authors have suggested that studies which have used the ProQOL scale need to be interpreted with caution as some of the core components of compassion are not measured by the ProQOL scale (Sinclair, Raffin-Bouchal, et al., 2017). Nonetheless, the current study found consistently reliable alpha coefficient scores reported for the compassion satisfaction and compassion fatigue subscales, which can increase confidence when using these scales.

As the ProQOL scale is the most widely used tool to explore PQoL, scores across different versions should be interpreted tentatively; due to difficulties measuring each of the subscales accurately, as mentioned above, with scope for overlooking symptoms, the implications of use in clinical populations can be detrimental with potential serious negative consequences. The lack of statistical difference between the reliability of ProQoL versions offers little evidence to support the use of one version over any other. However, given the difficulties and reported conceptual confusions already discussed around some of the core concepts, and the reported attempts of the ProQoL authors to address these in subsequent measures, it seems prudent for researchers to be using the most recent ProQoL measures and scoring guidelines. This would help to ensure that the most-up to date theoretical considerations

of the different aspects of professional quality of life have been appropriately incorporated into the measure being used.

Strengths and weaknesses

A strength of the analysis is that the main text for each study was thoroughly screened, as not all studies reported reliability coefficients in the title or abstract. This ensured appropriate studies were not overlooked in error and all relevant studies that reported a reliability value were included in the analysis. Furthermore, all studies reporting internal consistency and Cronbach's Alpha were included. As this is currently the most widely used index purporting to inform about scale reliability (Raykov & Marcoulides, 2017) in applied research, it can be argued that this makes it easier to compare studies, with increased confidence that they are comparable. Nonetheless, as found in this review, there appears to be a lot of heterogeneity in the alpha coefficients when demonstrating reliability of the different subscales as reported by studies.

The analysis also included studies from different populations, including different countries and different healthcare disciplines. This allowed opportunity to explore different uses of the ProQOL scale and compare reliability in different contexts. Interestingly, there were no significant differences found. This may help explain difficulties ascertaining accurate reliability of the ProQOL scale as using different populations and groups of people will naturally provide different findings; outcome measures tend to not be universally reliable, rather, better suited for use with a particular population (Roach, 2006). As such, reliability of the tool could be better harnessed by researchers and clinicians calculating and using specific reliability estimates for their specific populations of interest.

Overall, most of the studies included in the analysis used adequate sample sizes (n >30) with the overall number of participants >13,000. As a general rule, sample sizes of around 30-50 are deemed sufficient for the Central Limit Theorem (CLT) to hold (Ross, 2017). According to CLT the mean of a sample of data will be closer to the mean of the overall population in question as the sample size increases, notwithstanding the actual distribution of the data. Only one study (Begic et al., 2019) used an insufficient sample size of 27 and another used a sample size of 30 (Elkonin & Van der Vyver, 2011).

Nonetheless, one limitation of the analysis is that studies included were on the basis of reporting a Cronbach's alpha value. Although it is unlikely to have made a significant different in the overall conclusions of this analysis, it may have been useful to also include studies that reported other measures of reliability such a test re-test and inter-rater reliability.

Recommendations for future research

Given the importance of internal consistency, which is the most widely used measure of reliability (DeVellis & Thorpe, 2021), future recommendations include studies reporting an alpha value for the ProQOL scale as a whole outcome measure, in addition to individual subscales. This should be combined with other measures of reliability, which could not be done in this analysis, largely due to time constraints. Therefore, all studies that were retrieved from the literature search could be screened again to include studies that reported test re-test values and inter-rater reliability. This will add to a more comprehensive understanding of reliability when exploring the psychometric properties of the ProQOL scale.

Other future recommendations include further research using different language versions of the tool which may have an impact on the findings of the overall analysis. The current analysis included many studies reporting individual language versions. For instance, only one study reported use of the Arabic language version (Alhalal et al., 2020). If there were additional studies reporting alpha values for the Arabic language version, this is likely to impact the overall scores in the analysis, although at present cannot be predicted whether this will increase or decrease reliability of the scale.

Conclusion

A key finding of this analysis is that further research is needed exploring the psychometric properties of the ProQOL scale, specifically exploring reliability. Only one version demonstrated a significant difference although this version was only used in one study. Most of the studies included in this meta-analysis did not explore reliability of the scale as a whole but did report a Cronbach's Alpha for the three subscales, making it difficult to

determine reliability of the ProQOL scale as a whole, although this is an acknowledged problem of the scale.

References

Adams, R. E., Figley, C. R., & Boscarino, J. A. (2008). The compassion fatigue scale: Its use with social workers following urban disaster. *Research on social work practice, 18*(3), 238-250.

Adriaenssens, J., De Gucht, V., & Maes, S. (2015). Causes and consequences of occupational stress in emergency nurses, a longitudinal study. *Journal of nursing management, 23*(3), 346-358.

Alhalal, E., Alrashidi, L. M., & Alanazi, A. N. (2020). Predictors of patient-centered care provision among nurses in acute care setting. *Journal of nursing management, 28*(6), 1400-1409.

Amjad, S., Abbasi, S., Altaf, U., & Ijaz, F. (2020, 02/13). Mediating Role of Professional Quality of Life Between Emotional Regulation and Psychological Well-Being among Mental Health Professionals. *14*, 101-109. https://doi.org/10.19080/PBSIJ.2020.14.555888

Andersen, J., Papazoglou, K., Nyman, M., Koskelainen, M., & Gustafsberg, H. (2015, 10/15). Fostering Resilience among the Police. *The Journal of Law Enforcement, 5*, 1-13.

Ang, S. Y., Hemsworth, D., Uthaman, T., Ayre, T. C., Mordiffi, S. Z., Ang, E., & Lopez, V. (2018). Understanding the influence of resilience on psychological outcomes—Comparing results from acute care nurses in Canada and Singapore. *Applied Nursing Research, 43*, 105-113.

Ardito, R. B., & Rabellino, D. (2011). Therapeutic alliance and outcome of psychotherapy: historical excursus, measurements, and prospects for research. *Frontiers in psychology, 2*, 270.

Başoğul, C., Baysan Arabaci, L., Mutlu Satıl, E., & Büyükbayram Aslan, A. (2021). Professional values and professional quality of life among mental health nurses: A cross-sectional study. *Nursing & health sciences, 23*(2), 362-371.

Begic, S., Weaver, J. M., & McDonald, T. W. (2019). Risk and protective factors for secondary traumatic stress and burnout among home visitors. *Journal of Human Behavior in the Social Environment, 29*(1), 137-159.

Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *Bmj, 314*(7080), 572.

Borritz, M., Rugulies, R., Christensen, K. B., Villadsen, E., & Kristensen, T. (2006). Burnout as a predictor of self-reported sickness absence among human service workers: prospective findings from three year follow up of the PUMA study. *Occupational and environmental medicine, 63*(2), 98-106.

Bramley, L., & Matiti, M. (2014). How does it really feel to be in my shoes? Patients' experiences of compassion within nursing care and their perceptions of developing compassionate nurses. *Journal of clinical nursing, 23*(19-20), 2790-2799.

Burton, A., Burgess, C., Dean, S., Koutsopoulou, G. Z., & Hugh-Jones, S. (2017). How effective are mindfulness-based interventions for reducing stress among healthcare professionals? A systematic review and meta-analysis. *Stress and Health, 33*(1), 3-13.

Cavanagh, N., Cockett, G., Heinrich, C., Doig, L., Fiest, K., Guichon, J. R., Page, S., Mitchell, I., & Doig, C. J. (2020). Compassion fatigue in healthcare providers: A systematic review and meta-analysis. *Nursing ethics, 27*(3), 639-665.

Chang, K., & Taormina, R. J. (2011). Reduced secondary trauma among Chinese earthquake rescuers: A test of correlates and life indicators. *Journal of Loss and Trauma, 16*(6), 542-562.

Choi, S. H., & Lee, H. (2017, Oct). Workplace violence against nurses in Korea and its impact on professional quality of life and turnover intention. *J Nurs Manag, 25*(7), 508-518. https://doi.org/10.1111/jonm.12488

Cocker, F., & Joss, N. (2016). Compassion Fatigue among Healthcare, Emergency and Community Service Workers: A Systematic Review. *International journal of environmental research and public health, 13*(6), 618. https://doi.org/10.3390/ijerph13060618

Cohen, R. J., Swerdlik, M. E., & Phillips, S. M. (1996). *Psychological testing and assessment: An introduction to tests and measurement*. Mayfield Publishing Co.

Collins, S., & Long, A. (2003). Too tired to care? The psychological effects of working with trauma. *Journal of Psychiatric and Mental Health Nursing, 10*(1), 17-27.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology, 78*(1), 98.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika, 16*(3), 297-334.

Daundasekara, S. S., Arlinghaus, K. R., & Johnston, C. A. (2020). Quality of Life: The Primary Goal of Lifestyle Intervention. *American journal of lifestyle medicine, 14*(3), 267-270. https://doi.org/10.1177/1559827620907309

Davis, S., Lind, B. K., & Sorensen, C. (2013). A comparison of burnout among oncology nurses working in adult and pediatric inpatient and outpatient settings. Oncology nursing forum,

DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage publications.

Drury, V., Craigie, M., Francis, K., Aoun, S., & Hegney, D. G. (2014). Compassion satisfaction, compassion fatigue, anxiety, depression and stress in registered nurses in A ustralia: Phase 2 results. *Journal of nursing management, 22*(4), 519-531.

Duarte, J. (2017). Professional quality of life in nurses: Contribution for the validation of the Portuguese version of the Professional Quality of Life Scale-5 (ProQOL-5). *Análise Psicológica, 35*(4), 529-542.

Duarte, J., & Pinto-Gouveia, J. (2017). Empathy and feelings of guilt experienced by nurses: A cross-sectional study of their role in burnout and compassion fatigue symptoms. *Applied Nursing Research, 35*, 42-47.

Dubale, B. W., Friedman, L. E., Chemali, Z., Denninger, J. W., Mehta, D. H., Alem, A., Fricchione, G. L., Dossett, M. L., & Gelaye, B. (2019). Systematic review of burnout among healthcare providers in sub-Saharan Africa. *BMC public health, 19*(1), 1-20.

Elkonin, D., & Van der Vyver, L. (2011). Positive and negative emotional responses to work-related trauma of intensive care nurses in private health care facilities. *Health SA Gesondheid, 16*(1).

Elliott, T. R., & Gramling, S. E. (1990). Personal assertiveness and the effects of social support among college students. *Journal of counseling psychology, 37*(4), 427.

Erkorkmaz, U., Dogu, O., & Cinar, N. (2018). The relationship between burnout, self-esteem and professional life quality of nurses. *Journal of the College of Physicians and Surgeons Pakistan, 28*(7), 549-553.

Fernando III, A. T., & Consedine, N. S. (2014). Beyond compassion fatigue: the transactional model of physician compassion. *Journal of pain and symptom management, 48*(2), 289-298.

Figley, C. R. (2002). Compassion fatigue: Psychotherapists' chronic lack of self care. *Journal of clinical psychology, 58*(11), 1433-1441.

Figley, C. R. (2013). *Compassion fatigue: Coping with secondary traumatic stress disorder in those who treat the traumatized*. Routledge.

Fu, C. Y., Yang, M. S., Leung, W., Liu, Y. Y., Huang, H. W., & Wang, R. H. (2018). Associations of professional quality of life and social support with health in clinical nurses. *Journal of nursing management, 26*(2), 172-179.

Galiana, L., Arena, F., Oliver, A., Sansó, N., & Benito, E. (2017). Compassion satisfaction, compassion fatigue, and burnout in Spain and Brazil: ProQOL validation and cross-cultural diagnosis. *Journal of pain and symptom management, 53*(3), 598-604.

Galiana, L., Oliver, A., De Simone, G., Linzitto, J. P., Benito, E., & Sansó, N. (2019). A brief measure for the assessment of competence in coping with death: The coping with death scale short version. *Journal of pain and symptom management, 57*(2), 209-215.

Hall, L. H., Johnson, J., Watt, I., Tsipa, A., & O'Connor, D. B. (2016). Healthcare staff wellbeing, burnout, and patient safety: a systematic review. *PloS one, 11*(7), e0159015.

Heeb, J.-L., & Haberey-Knuessi, V. (2014). Health professionals facing burnout: what do we know about nursing managers? *Nursing research and practice, 2014*.

Hemsworth, D., Baregheh, A., Aoun, S., & Kazanjian, A. (2018). A critical enquiry into the psychometric properties of the professional quality of life scale (ProQol-5) instrument. *Applied Nursing Research, 39*, 81-88.

Heritage, B., Rees, C. S., & Hegney, D. G. (2018). The ProQOL-21: A revised version of the Professional Quality of Life (ProQOL) scale based on Rasch analysis. *PloS one, 13*(2), e0193478.

Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L., & Sterne, J. A. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *Bmj, 343*.

Hinderer, K. A., VonRueden, K. T., Friedmann, E., McQuillan, K. A., Gilmore, R., Kramer, B., & Murray, M. (2014). Burnout, compassion fatigue, compassion satisfaction, and secondary traumatic stress in trauma nurses. *Journal of Trauma Nursing| JTN, 21*(4), 160-169.

Hofmeyer, A., Kennedy, K., & Taylor, R. (2020, 2020/04/01/). Contesting the term 'compassion fatigue': Integrating findings from social neuroscience and self-care research. *Collegian, 27*(2), 232-237. https://doi.org/https://doi.org/10.1016/j.colegn.2019.07.001

Horkan, A. M. (2014). Alarm fatigue and patient safety. *Nephrology Nursing Journal, 41*(1), 83-86.

Hundall Stamm, B. (2009). Professional Quality of Life Measure: Compassion, Satisfaction, and Fatigue Version 5 (ProQOL).

Itzick, M., Kagan, M., & Ben-Ezra, M. (2018). Social worker characteristics associated with perceived meaning in life. *Journal of Social Work, 18*(3), 326-347.

Kagan, M., & Itzick, M. (2019). Work-related factors associated with psychological distress among social workers. *European journal of social work, 22*(1), 30-42.

Kim, K., Han, Y., Kwak, Y., & Kim, J.-s. (2015). Professional quality of life and clinical competencies among Korean nurses. *Asian nursing research, 9*(3), 200-206.

Kim, S. (2014). World Health Organization Quality of Life (WHOQOL) Assessment. In A. C. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research* (pp. 7260-7261). Springer Netherlands. https://doi.org/10.1007/978-94-007-0753-5_3282

Kim, S. Y., Park, J. E., Lee, Y. J., Seo, H.-J., Sheen, S.-S., Hahn, S., Jang, B.-H., & Son, H.-J. (2013). Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *Journal of clinical epidemiology, 66*(4), 408-414.

Kim, Y. H., Kim, S. R., Kim, Y. O., Kim, J. Y., Kim, H. K., & Kim, H. Y. (2017). Influence of type D personality on job stress and job satisfaction in clinical nurses: the mediating effects of compassion fatigue, burnout, and compassion satisfaction. *Journal of advanced nursing, 73*(4), 905-916.

Ledoux, K. (2015). Understanding compassion fatigue: Understanding compassion. *Journal of advanced nursing, 71*(9), 2041-2050.

Lee, Y., & Seomun, G. (2016). Role of compassion competence among clinical nurses in professional quality of life. *International Nursing Review, 63*(3), 381-387.

Lynch, S. H., & Lobo, M. L. (2012). Compassion fatigue in family caregivers: a Wilsonian concept analysis. *Journal of advanced nursing, 68*(9), 2125-2134.

Maslach, C. (2003). *Burnout: The cost of caring*. Ishk.

Mathieu, F. (2007). Running on empty: Compassion fatigue in health professionals. *Rehab & Community Care Medicine, 4*, 1-7.

Michalec, B., Diefenbeck, C., & Mahoney, M. (2013). The calm before the storm? Burnout and compassion fatigue among undergraduate nursing students. *Nurse Education Today, 33*(4), 314-320.

Mills, J. D. (2003). SPSS textbooks: A review for teachers. *Statistics Education Research Journal, 2*(2), 59-70.

Nolte, A. G., Downing, C., Temane, A., & Hastings-Tolsma, M. (2017). Compassion fatigue in nurses: A metasynthesis. *Journal of clinical nursing, 26*(23-24), 4364-4378.

Norrman Harling, M., Högman, E., & Schad, E. (2020). Breaking the taboo: eight Swedish clinical psychologists' experiences of compassion fatigue. *International journal of qualitative studies on health and well-being, 15*(1), 1785610.

Nunnally, J. C. (1994). *Psychometric theory 3E*. Tata McGraw-hill education.

O'Connor, K., Neff, D. M., & Pitman, S. (2018). Burnout in mental health professionals: A systematic review and meta-analysis of prevalence and determinants. *European Psychiatry, 53*, 74-99.

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: applications to practice* (Vol. 892). Pearson/Prentice Hall Upper Saddle River, NJ.

Rauvola, R. S., Vega, D. M., & Lavigne, K. N. (2019). Compassion fatigue, secondary traumatic stress, and vicarious traumatization: A qualitative review and research agenda. *Occupational Health Science, 3*(3), 297-336.

Ravi, R., Yerraguntla, K., Gunjawate, D. R., Guddattu, V., & Bellur, R. (2016). Professional quality of life in audiologists and speech language pathologists working in India. *Journal of Workplace Behavioral Health, 31*(3), 162-172.

Raykov, T., & Marcoulides, G. A. (2017, 2019/02/01). Thanks Coefficient Alpha, We Still Need You! *Educational and Psychological Measurement, 79*(1), 200-210. https://doi.org/10.1177/0013164417725127

Roach, K. E. (2006). Measurement of Health Outcomes: Reliability, Validity and Responsiveness. *JPO: Journal of Prosthetics and Orthotics, 18*(6). https://journals.lww.com/jpojournal/Fulltext/2006/01001/Measurement_of_Health_Outcomes__Reliability,.3.aspx

Roney, L. N., & Acri, M. C. (2018). The cost of caring: an exploration of compassion fatigue, compassion satisfaction, and job satisfaction in pediatric nurses. *Journal of Pediatric Nursing, 40*, 74-80.

Ross, S. M. (2017). *Introductory statistics*. Academic Press.

Ruiz-Fernández, M. D., Ramos-Pichardo, J. D., Ibáñez-Masero, O., Carmona-Rega, M. I., Sánchez-Ruiz, M. J., & Ortega-Galán, Á. M. (2021). Professional quality of life, self-compassion, resilience, and empathy in healthcare professionals during COVID-19 crisis in Spain. *Research in Nursing & Health, 44*(4), 620-632.

Salvagioni, D. A. J., Melanda, F. N., Mesas, A. E., González, A. D., Gabani, F. L., & Andrade, S. M. d. (2017). Physical, psychological and occupational consequences of job burnout: A systematic review of prospective studies. *PloS one, 12*(10), e0185781-e0185781. https://doi.org/10.1371/journal.pone.0185781

Samson, T., Iecovich, E., & Shvartzman, P. (2016). Psychometric characteristics of the Hebrew version of the Professional Quality-of-Life Scale. *Journal of pain and symptom management, 52*(4), 575-581. e571.

Sanford, R. L., Cerel, J., & Frey, L. M. (2018). Survivor of Suicide Loss Support Group Facilitators: Do Peers and Professionals Differ? *Social Work With Groups, 41*(4), 306-322.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological assessment, 8*(4), 350.

Sinclair, S., Beamer, K., Hack, T. F., McClement, S., Raffin Bouchal, S., Chochinov, H. M., & Hagen, N. A. (2017). Sympathy, empathy, and compassion: A grounded theory study of palliative care patients' understandings, experiences, and preferences. *Palliative medicine, 31*(5), 437-447.

Sinclair, S., Kondejewski, J., Schulte, F., Letourneau, N., Kuhn, S., Raffin-Bouchal, S., Guilcher, G. M., & Strother, D. (2020). Compassion in pediatric healthcare: A scoping review. *Journal of Pediatric Nursing, 51*, 57-66.

Sinclair, S., Norris, J. M., McConnell, S. J., Chochinov, H. M., Hack, T. F., Hagen, N. A., McClement, S., & Bouchal, S. R. (2016). Compassion: a scoping review of the healthcare literature. *BMC palliative care, 15*(1), 1-16.

Sinclair, S., Raffin-Bouchal, S., Venturato, L., Mijovic-Kondejewski, J., & Smith-MacDonald, L. (2017). Compassion fatigue: A meta-narrative review of the healthcare literature. *International journal of nursing studies, 69*, 9-24.

Slatten, L. A., Carson, K. D., & Carson, P. P. (2011). Compassion fatigue and burnout: What managers should know. *The health care manager, 30*(4), 325-333.

Sodeke-Gregson, E. A., Holttum, S., & Billings, J. (2013). Compassion satisfaction, burnout, and secondary traumatic stress in UK therapists who work with adult trauma clients. *European journal of psychotraumatology, 4*, 10.3402/ejpt.v3404i3400.21869. https://doi.org/10.3402/ejpt.v4i0.21869

Sorenson, C., Bolick, B., Wright, K., & Hamilton, R. (2016). Understanding compassion fatigue in healthcare providers: A review of current literature. *Journal of Nursing Scholarship, 48*(5), 456-465.

Stamm, B. (2010). The concise manual for the professional quality of life scale.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education, 2*, 53-55. https://doi.org/10.5116/ijme.4dfb.8dfd

CHAPTER TWO

Adult mental healthcare staff experiences of psychology-led supervision groups for managing risk within Assertive Outreach Teams (AOT) and Inpatient Rehabilitation teams

Abstract

Background

Clinical supervision within healthcare can be defined as "The formal provision, by approved supervisors, of a relationship-based education and training that is work-focused and which manages, supports, develops and evaluates the work of colleague/s." It is considered an essential component of modern effective health care systems and can be provided in different formats. There is a growing evidence base for the impact of clinical supervision on patient outcomes related to risk; supervisions on how to record information and manage identified risks is a way of improving staff confidence in managing risks and ongoing supervision should be available to assess, formulate and manage risk.

Aim

This study aimed to explore adult mental healthcare staffs experiences of psychology-led supervision groups when managing risk within Assertive Outreach Teams (AOT) and Inpatient Rehabilitation.

Method

Online semi-structured interviews were conducted with eight participants. Data was analysed using Reflexive Thematic Analysis (RTA).

Findings

Three main themes were generated; 'the function and value of group sessions;' 'managing the NHS hierarchy in the group' and 'the experience of responsibility and accountability.' The importance of a team approach was emphasised by all participants and appears to have an impact on individual well-being as well as patient outcomes.

Conclusion

The findings from this study emphasise the importance of clinical supervision groups in meeting a range of different functions, which are highly valued by staff. A team approach increased a sense of cohesion and confidence to manage difficulties and better support patients in their recovery.

Introduction

Clinical supervision within healthcare can be defined as "The formal provision, by approved supervisors, of a relationship-based education and training that is work-focused and which manages, supports, develops and evaluates the work of colleague/s" (Milne, 2007). This author further conceptualises clinical supervision as a space for junior clinicians to receive support and guidance from senior clinicians with more expertise.

(Kilminster et al., 2007) supports this definition, referring to clinical supervision as a formal professional activity to improve individual clinical skills, knowledge and competence, as well as enhanced quality of care to patients, by gaining an insight into supervisor experiences. (Proctor et al., 1988) further suggest clinical supervision has three key functions: 'normative' (organisational and quality control, dealing with professional issues such as codes of practice and restrictions), 'formative' (promoting education and development, examining clinical interventions and development of clinical skills) and 'restorative' (ensuring support for staff, assessing for work distress and burnout). This model is commonly used within the National Health Service (NHS), to ensure staff feel safe, within a protected, confidential environment, supporting their personal and professional development and reflecting on their practice (Commission, 2012). It is considered an essential component of modern effective health care systems (Kilminster & Jolly, 2000).

Clinical supervision aims to promote autonomous decision-making, with a goal of improving professional practice (Bunyan et al., 2017). Important during the process, is valuing the individual's well-being and sense of safety and protection, which can be monitored via reflective processes (Beddoe, 2010) and non-judgemental analysis of one's clinical practice

(Hall & Cox, 2009). The supervision process can also improve positive working relationships and teamwork with other colleagues and help staff to manage work demands required on an individual level, personally, as well as part of a team (Martin et al., 2021). Driscoll (2007) further states a key component of clinical supervision is to enhance learning opportunities, with an experienced supervisor facilitating a space for reflection on clinical experiences.

There are many identified benefits of clinical supervision, with peer support and management of stress and pressure one of the most obvious ones (Brunero & Stein-Parbury, 2008), and proving to be key in preventing burnout (Hagen et al., 2017). Clinical supervision allows a space to consider the emotional effects staff and patients have on each other and what staff can learn from their thoughts, feelings and experiences (Palmer-Olsen et al., 2011). This is crucial as burnout is more prevalent in particular settings such as mental health, where there is a high proportion of patients who self-harm, express suicidal ideation, actively attempt to die by suicide or engage in other risky behaviours which evoke a range of emotions that may be distressing and/or difficult to manage (Hagen et al., 2017). Other benefits of clinical supervision include better staff retention rates and increased job satisfaction due to shared responsibilities and accountability, risk management plans and overall better support for quality improvements (Cutcliffe et al., 2018).

Whilst clinical supervisions share responsibility between the clinical supervisor and supervisee, responsibility for ensuring policies and legislation are discussed and followed appropriately lies with the clinical supervisor (O'Shea et al., 2019). This helps to develop the supervisees skills, knowledge and competence in safety, risk management and better-quality practices (HSE, 2015a). Additionally, it may be cost-effective due to emphasis on compassion, where the supervisor is available to offer support in a proactive manner, maintaining sensitivity when exploring detail, whilst adopting a Socratic approach (Walker & Clark, 1999).

Clinical supervisions can be provided in different formats such as in groups or on a one-to-one basis and can also be conducted in-person or remotely. Increasing evidence suggests group supervisions are preferrable to one-to-one supervisions due to the breadth of experience and reflection available (Mastoras & Andrews, 2011) and are increasingly utilised in various clinical settings including acute care and community settings (Pollock et al., 2017).

Benefits of group supervision include opportunities to critically reflect on case studies, receive ongoing support from peers, reduced stress and burnout, and better quality feedback and validation of contributions (Golia & McGovern, 2015).

Due to positive associations with patient outcomes such as reduced risk of complications and mortality, there is increasing focus on the importance of clinical supervision (Snowdon et al., 2017). Furthermore, although medication related risks are present in both mental and physical healthcare settings (Cullen et al., 2010), risks related to violence, self-harm and suicide are more prevalent in the former (Flewett, 2010).

Rationale

There is some literature looking at clinical supervision, however, this is predominantly focused on the process behind it and the impact on psychological wellbeing such as burnout, stress and managing distressing emotions. Additionally, a large proportion of the literature is focused largely on the viewpoints and experiences of mental health nurses. There is little exploration of how clinical supervision differs across different settings and the viewpoints of other disciplines of mental health care staff in how this is used to manage risk which can present in the form of challenging behaviours including violence and aggression, both physical and verbal, absconding and self-harm (Bowers et al., 2015). This is an important factor as some mental health settings such as crisis teams, Assertive Outreach and inpatient rehabilitation teams support people with complex and severe mental health needs. More than a quarter of AOT patients, for example, were rated as being seriously at risk, or currently in institutional care (Slade et al., 2005). Similarly, Meaden et al. (2014) found 34% and 55% of 80 service users in inpatient rehabilitation settings had a history of physical and verbal aggression incidents in the last 6 months, respectively.

Assertive Outreach was developed in the UK based on the development of Assertive Community Treatment in the USA (Stein & Test, 1980). These teams are designed to focus on the development of engagement with service users with complex mental health difficulties and risk behaviours, many of whom can be ‘hard to reach’ (Priebe et al., 2004; Wright et al., 2003). Indeed, patients under the care of AOTs appeared to be at the severe end of the spectrum for a range of different measures, with 95% deemed to experience psychosis and 30% having had

three or more hospital admissions in the past two years (Schneider et al., 2006). This study also found this population of patients to present with higher risk, compared to people in community mental health care settings, in terms of violence and aggression towards others, past suicide attempts, self-harm and self-neglect, dual-diagnosis and unstable accommodation and/or homelessness.

Recent findings suggests ongoing supervision on how to record information and manage identified risks is a way of improving staff confidence in managing risks and ongoing supervision should be available to assess, formulate and manage risk (Appleby et al., 2019). Team-based assessment and formulations are increasingly perceived as one of the most significant recommendations to ensure best practice within inpatient settings, although differences between teams regarding structure, frequency of meetings and use of theoretical approaches is acknowledged (Raphael et al., 2021). Additionally, there has been an increased recognition of the value of shared formulation in supporting staff teams who work with people with distressing psychosis (Berry et al., 2009), including the management of challenging and risk behaviour using the Shared Assessment, Formulation and Education (SAFE) approach (Meaden et al., 2015; Meaden & Hacker, 2010). This approach supports MDT assessment and care planning practices when assessing and managing risk, by improving communication via shared understanding by aiding the development of clear and realistic goals, focusing on enabling recovery (Meaden et al., 2022). The use of the MDT, combining specialist knowledge and skills, particularly when dealing with complex needs and higher risk, is considered essential in understanding, assessing and managing risks, whilst also ensuring adequate care is delivered to these patients (Mason et al., 2002).

The Health Service Executive (HSE, 2007) guidelines suggest risk management in mental health settings is deemed to be effective only by involving all staff, using an MDT approach. It advises staff to be supported to develop and exercise professional skills such as information sharing to manage risk better, with MDTs jointly develop risk management plans, adopting an open, egalitarian environment, and emphasis on the use of reflective practice. The SAFE approach, with the focus on shared assessment and formulation amongst the MDT, facilitates this.

The Royal College of Psychiatrists (RCP, 2017) also emphasises the use of MDTs when assessing risk, allowing information to be shared from different perspectives, where assessment and formulations guide risk management plans which should be reviewed regularly. Risk assessment tools such as the Historical and Clinical Risk History (HCR-20) are also increasingly used as a clinical framework to guide clinicians to use a multidisciplinary approach to assess and manage risk (Douglas et al., 2013). Importantly, team formulations have demonstrated positive impacts on staff wellbeing (Kramarz et al., 2022), creating a space for team communication and reflections, enabling psychologically informed care plans for managing risk (Cole et al., 2011). Staff morale has also been found to be increased as a result of good communication and relationships within the team, a culture encouraging the availability of support following violent incidences and one where every staff is listened to regardless of seniority (Totman et al., 2011). However, the degree to which certain team members' perspectives may be encouraged or disregarded, may vary largely across teams whereby risk formulations may differ depending on the experiences and perspectives of individual team members, their values and/or dynamics between team members (Lewis & Doyle, 2009).

Overall, team approaches to risk assessment and formulation have been acknowledged as preferable to clinicians working alone, with team supervision and shared formulation deemed better positioned to support the management of risk, particularly in services such as inpatient psychiatric rehabilitation and AOT e.g. (Meaden & Fox, 2015). However, how these approaches are used by the staff that attend these, and how staff experience and make sense of them (e.g. what they find valuable) remains unclear. Therefore, this study hopes to explore this further, specifically looking at ***Adult mental healthcare staff experiences of psychology-led supervision groups within Assertive Outreach and inpatient psychiatric rehabilitation teams in managing risk.***

Method

Recruitment

An anonymous NHS mental health foundation Trust that uses team supervision sessions across various AOT and inpatient rehabilitation teams, led by psychologists, was invited to

take part in the research study. Staff across both these teams work with patient groups with similar presentations that are often complex, vary in risk behaviours and have longer periods of recovery times. These teams also share psychiatric medical staff across both teams, with Psychology providing team supervisions for both multidisciplinary teams (MDT). A clinical psychologist employed by this NHS Trust and known by both teams, informed staff of the study and forwarded study information to staff across both teams, including a research poster. Opportunity sampling was utilised and any interested participants from any discipline (except for clinical psychologists) contacted the researcher directly, providing full informed consent before volunteering to participate in the interviews.

One team of inpatient rehabilitation staff and one team of AOT staff from the NHS Trust invited responded and thus took part. Any member of staff within these team, from any professional discipline, including staff with qualified professional backgrounds such as registered mental health nurses and those with unqualified status such as healthcare assistants attended the supervision groups and were thus eligible to participate. Furthermore, as the supervision groups within these teams were facilitated by lone team psychologists (there was no other psychology involvement in the teams at the time), it was not deemed appropriate to include them in the interviews as, essentially, they would be being asked to comment on the facilitation of their own group supervisions. There were no rewards, compensation or incentives provided for participation.

Model of Supervision Groups

With the AOT, the supervision groups were facilitated on a weekly basis and focused on ‘case-busts.’ This referred to discussing specific cases staff teams were struggling with and which had posed increased risk, helping staff to identify different solutions and plans for managing difficulties, as well as provide a space to offer each other support and reflection.
For inpatient rehabilitation, supervision groups were facilitated fortnightly, also focussing on ‘case-busts’, primarily using the SAFE approach (Meaden et al., 2022) to formulate cases. Both teams used psychological formulation (e.g., the 5Ps model, SAFE approach) to aid reflection on cases to aid team understanding of service user needs, with a focus on risk and problematic behaviour.

Inclusion criteria

Participants were eligible to take part in the study if they met all the following:

- Were aged 18 years and above
- Spoke and understood English without the need for an interpreter
- Were mental health care staff working in an Assertive Outreach Team (AOT) or inpatient rehabilitation setting. This includes both permanent and bank/agency staff.
- Have attended at least 3 psychology-led clinical supervision groups in their current team setting
- Have attended 3 or more psychology-led clinical supervision groups in the last 12 months.

Exclusion criteria

Participants were not able to take part in the study if they:

- Had never attended a psychology-led supervision group
- Were unable to speak English without the help of an interpreter

Design

Data was collected online, via a secure video platform. Eight participants completed semi-structured interviews, lasting between 35-65 minutes. An interview schedule guide consisting of open questions (see Appendix 6) was developed by the researcher and refined based on feedback from staff working in the AOT and inpatient rehabilitation teams. This was used to guide each interview, which focused on how participants used the supervision groups and their experiences of it.

Participants:

Participants were staff working in AOTs or inpatient rehabilitation teams for an NHS Trust. Professional roles varied between participants and included a consultant psychiatrist, a specialist registrar, qualified mental health nurses (x4), an Occupational Therapist (OT) and a healthcare assistant (HCA). The length of time participants had worked within these teams

varied between 9 months to 16 years. The ages of participants and the jobs roles for each pseudo-name is not provided, to support anonymity of participants.

Sample size

For small projects, 6-10 participants are recommended for interviews (Braun & Clarke, 2013). Therefore, a sample size of 8 participants was deemed sufficient for this study.

Analysis

Reflective Thematic Analysis (RTA) (Braun & Clarke, 2021) was chosen for this study. Thematic analysis is a broad approach to analysis used to explore, interpret, and report patterns of meaning across datasets, which are then coded to help develop themes (Clarke & Braun, 2021). It allows the researcher to organise and guide the data, resulting in the development of themes which are relevant to the research question (Clarke & Braun, 2021). Importantly, the researcher is actively involved in the analytic process, making decisions about how codes are grouped into themes, and how these relate to the research question. "Big Q" approaches (Kidder & Fine, 1987) to thematic analysis actively use the researcher as a tool for enhancing the analysis, recognising the importance of understanding and making sense of the data within the context it was conducted (e.g., time, place, local culture) (Coyle, 2007). Given the present research focus on a specific approach to risk management for a group of service users within a specific setting, it was deemed important to use an approach that would facilitate an appreciation of the contextualised nature of the responses being analysed.

Interpretative Phenomenological Analysis (IPA) was considered as it considers the experiences of individual participants and how they make sense of this. However, IPA places a focus on the unique experiences and individual characteristics of participants, developing themes for each participant before grouping them and exploring themes at a group level. This aims to preserve and capture the individual aspects of the experiences relevant to each participant as well as the group commonality and is well suited to questions that prioritise a focus on the experiential (Smith et al, 2022). The present study question is less focussed on the individual participants experiences of the supervision process, and more focussed on how they,

as a group, use this to support their work. Furthermore, IPA focuses on inductive analyses, and attempts to "bracket" presuppositions and theories to allow a focus on the lived experience of the participants (Smith et al., 2022). In the present study, the research question was already contextualised within an existing knowledge of the psychology of team supervision (e.g., the SAFE approach), and so flexibility to incorporate this into the analysis (elements of a 'deductive' approach) was considered beneficial.

RTA is an easily accessible and theoretically flexible interpretative approach to qualitative data analysis, which adds to thematic analysis by facilitating the identification and analysis of patterns or themes in a given data set (Braun & Clarke, 2012). It allows *"the researcher's reflective and thoughtful engagement with their data and their reflexive and thoughtful engagement with the analytic process"* (Braun & Clarke, 2019. p.594). It requires the researcher to critically reflect on themselves, their positioning, and how this may impact the overall research. The analyst is encouraged to acknowledge and value their subjectivity and creativity with the datasets, using this as a resource for analysis rather than a hindrance (Clarke & Braun, 2021). Evidence suggests reflexivity is vital in ensuring a high standard of analysis, and researchers must endeavour to understand their own perspectives (Elliott et al., 1999).

Additionally, RTA enables flexibility when interpreting data, where all knowledge related to the subject matter can be brought together. It also allows for a broad analytical focus which increases flexibility and allows a deeper understanding; for instance, an experiential focus can be used which involves an analysis with the aim of capturing individual experiences and understanding whereas a critical approach involves an RTA which aims to understand meaning around a broader issue or concept (Clarke & Braun, 2021). Importantly, both these approaches could be applied to this research study which focused on group processes and the impact of this when managing risk, rather than the emotional or psychological consequences of attending these groups.

RTA is helpful as it offers different foci of meaning where data can be coded into themes which are either semantic or latent; semantic codes are identified through the explicit or surface meanings of the data and can be described as a descriptive analysis of the data, aimed solely at presenting the content of the data as communicated by the respondent. Latent coding,

however, goes beyond the descriptive level of the data and attempts to identify hidden meanings or underlying assumptions, ideas, or ideologies that may shape or inform the descriptive or semantic content of the data (Byrne, 2022). Both semantic and latent codes are understood to represent the researcher's interpretations of patterns of meaning across the datasets (Byrne, 2022) (see Appendix 7 for coding of a transcript). Themes are then actively produced by organising these codes around a relative core commonality, or 'central organising concept', that the researcher interprets from systematically engaging with the data (Braun & Clarke, 2019) (see Appendix 8). This was particularly important in this study as the aim was to gain a better understanding of not only individual staff experiences, but also shared meanings and ideas within teams when managing risk; RTA offered an avenue for flexible interpretation and meaning to be derived from the data and shed some insight into an understudied area of research.

Throughout the analytic process, 6 key phases (Braun & Clarke, 2021) were used as guidance, aiding the researcher to identify and attend to the important aspects of a thematic analysis. These included:

1. ***Familiarisation with the data*** by thoroughly examining the transcripts and identifying appropriate information that is relevant to the research question

2. Generating initial codes (both semantic and latent). This will help produce succinct, shorthand descriptive or interpretive labels for pieces of information that may be of relevance to the research question

3. Generating themes – combining codes with shared meanings to form themes

4. Reviewing potential themes – reviewing whether themes provide an accurate interpretation of the data in relation to the research question

5. Defining and naming themes - ensuring each theme and sub-theme depicts a coherent and consistent account of the data that is distinct and cannot be told by other themes

6. Producing the report – accurately summarising the data in themes that are relevant to answering the research question, reflecting on changes in codes and themes over the course of the analysis and noting this in informal notes and memos

Ethical considerations

Ethical approval was received by the Health Research Authority (HRA) (see Appendix 3). A favourable opinion was also provided by the University of Birmingham who were acting as sponsor for the study (see Appendix 4) and the host Trust's Research and Innovation (R&I) team. All participants volunteered to take part in the study, without being forced to or coerced. They were mental healthcare staff who are not deemed vulnerable and who had capacity throughout the study. Although they knew the researcher's supervisor, who was the clinical psychologist that facilitated the supervision groups for one of these teams, they were not previously known to the researcher. The group facilitators were not directly involved in the analysis of the study. Participants were informed of their right to terminate the interviews at any time and their right to withdraw their data up to two weeks post-interview date, without any negative repercussions. All participants were given a participant identification number and pseudo-name to protect their identity. This ensured anonymity where the researcher's supervisor could not identify who the participants were. Quotes that have been used refer to the allocated pseudo-name and not participants' real names.

Reflections on researcher's positioning

The researcher is a Trainee Clinical Psychologist, completing this research as part of a doctorate training program. The researcher was able to identify with some of the participants experiences, having worked in acute inpatient settings, where patients posed similar levels of risk and presented with complex needs and comorbidities. Reflection on this illuminated some initial assumptions about what the researcher expected participants to talk about; importantly, it became apparent very early on during the interviews that participants shared insights about a range of topics and experiences that the researcher had not necessarily anticipated.

This reflection resulted in the researcher making a conscious effort to try not to impose their personal views or opinions, or sway participants to answer a certain way. Rather, effort was made to be curious and explorative in questioning, enabling participants to freely discuss their experiences and processes of the supervision groups.

The use of a reflective journal further aided the researcher to reflect on their positioning and how this may have impacted the interviews. A key reflection was on the dual role of the researcher as both a researcher and clinician; although the researcher felt more competent in their role of a clinician, participants appeared to assume they were an expert in the field of research. This is likely to have inhibited participants from openly discussing their experiences; during the first interview, the participant appeared nervous and later disclosed it was because they were worried about 'giving the wrong answers' and commented on the researcher's role as an academic professional. This prompted the researcher to try to build a better rapport with the participants, clarifying their dual role and clinical experiences and being more mindful to the language they were using as well as their tone of voice.

The reflective journal also aided the researcher to notice any personal anxieties they were facing such as limited previous experience using this type of research methodology and time pressures to recruit participants in. Issues such as these were discussed during supervisions and managed appropriately with an experienced clinical psychologist. The reflective journal further allowed the researcher to reflect on their personal/demographic characteristics, ethnic background, cultural and religious beliefs and compare any similarities or differences to that of participants which may have had an impact on how forthcoming participants were during the interviews in sharing their experiences.

Findings

Table 3: describing themes and sub-themes developed from the data

Theme	
1. The function and value of the group sessions: enhancing belonging and team cohesion	Participants who contributed to theme
1.1 Identifying plans and solutions	All
1.2 Seeking and providing support	All
1.3 Sharing ideas and perspectives	All
1.4 Dealing with difficulties and worries	All
2. Managing the NHS hierarchy in the group	All
3. The experience of responsibility and accountability	All

Function and value of the group sessions: Enhancing belonging and team cohesion

All participants spoke about their experiences of the supervision groups and how these were used to achieve a range of functions, including practical and emotional support. There was a strong sense of participants valuing the group, as it served a range of important purposes. As such, a major theme was ***'the function and value of the group sessions'*** and within this, there were four subordinate themes that reflected four different functions. Throughout these themes, it became apparent that a core function of the group was enhancing a sense of cohesion and belonging among the team. This was noted as a thread that connected all of these subthemes, facilitating and enhancing the practical functions of the MDT. Approaching the work as a team enabled participants to feel more supported and less isolated as illustrated below by Deborah.

> *Deborah: "It does help when it's a team approach and you can obviously, you know, have that confidence to say yes, this is something that we agreed as a collective and that's what we're doing."*

One clear function of the group was about ***identifying plans and solutions.*** Identifying plans was important as it helped guide participants on what to do and how to approach various things, including problems, issues or challenges faced by staff across a range of issues, related to both clinical practice as well as relationships with staff and team members. It also helped staff to be more consistent in their approaches, to be in agreement about how to manage various situations. There was a sense of increased confidence, by identifying specific ways of how to approach situations, using practical skills, which seemed to be valued by participants.

> *Cameron: "Having everybody in a meeting that is well organised and has definite kind of formulation, and comes up with definite plans, that means that everybody ends up hopefully doing the same thing, taking the same approach with the patient."*

Here, Cameron identifies how group cohesion is achieved through the sessions ("doing the same thing") and Geraldine below notes the importance of consistency; by identifying plans and solutions in the group, this had an impact on what patients expected from staff, ensuring better understanding between both staff and patients, resulting in a smoother running of the unit.

> *Geraldine: "If you're all on the same page, and you're, you can then develop a kind of consistent approach, so you haven't got somebody accidentally undermining what, what works or, and the patient also knows then eventually what to expect, how our approach is going to be."*

Other important aspects of identifying plans meant staff were able to plan ahead and manage risks better; doing this as a team was deemed to be vital as it helped increase team consensus and relieve feelings of individual pressure. Essentially, it allowed staff to provide person-centred care that was appropriate in meeting the individual needs of that patient with better outcomes.

> *Harriet: I think we're quite good at like discussing okay what are we gonna do, if like you know, he's expressing these thoughts, what should we do now."*

Another function of the supervision groups was around ***seeking and providing support.*** It appeared that seeking support from other members of the team was key to identifying common feelings and experiences and providing a sense of reassurance. The group was crucial in meeting these needs, offering a formal space to do this in, with Elsa describing it as their *"resource"* to *"get the support that I need."* Support from colleagues appeared to help reduce feelings of stress when dealing with risk and isolation experienced as part of this. Importantly, it appeared to have a positive impact on individual well-being. It further helped create a sense of unity between all team members, with Fiona describing it as *"working towards the same sort of ideals and goals,"* enabling staff to feel valued, their contributions appreciated and connected to the wider team.

> *Alfie: "It's a way where people don't feel isolated erm, in the sense of we have to carry the can on their own."*

> *Cameron: "Hopefully kind of gets everybody feeling supported and feeling part of the team and their contribution valued."*

> *Harriet: "So yeah, I think it's more of like the support of the team that that helps me deal with the stress of dealing with risk."*

Support from other team members also seemed to allow staff to feel confident and assured enough to encourage patients to support themselves better. Participants were able to discuss in the supervision groups how they can assist patients to enhance their independent living skills and then empower patients to reintegrate in the community, whilst also better managing potential risks.

> *Alfie: "Work out a way how we can best support that individual so that they can function to the best of abilities in the community while err monitoring the risk that they potentially could be posing"*

> *Deborah: "We sort of talk about, how to go about reducing the risk and how to you know, help them support their physical health better"*

This was seen as enhancing the functions of the team, which was also viewed as enhancing the support for service users. In this way we can see how the team members link the supervision groups to clinical outcomes ("function to the best of their abilities").

Another function of the group was about ***sharing ideas and perspectives.*** Across interviews, participants were vocal in how sharing different ideas helped to provide different perspectives they had not considered previously and explore all options available to them. This was important in identifying alternative solutions and ways of approaching things collectively, enhancing team cohesion in service of the common goal (patient recovery). It helped increase different insights, identifying new information about patients with a better understanding of what might work better or why something might not have worked so well. There was a sense of reassurance that staff were doing the best they can, regardless of the pressures or stresses they were facing, attempting to meet individual patient needs by exploring different possibilities.

> *Alfie: "You feel reassured that you haven't erm, you haven't missed out on exploring different avenues that you possibly haven't thought through while you were absorbed in it."*

> *Cameron: "For an individual patient there will always be things that I don't know. And, erm, and having different staff insights is important. You know the more people who are looking at a patient and thinking about them, the more insight you are going to get."*

> *Elsa: "Offers that opportunity to discuss that as a team because we can all come up with different reasons why and ideas why someone might not be doing something. There might be things that we're missing so it's useful to get together as a team"*

By using a team approach to share ideas and perspectives, this allowed staff to unite as a team to meet the best interests of the patients they were caring for. Staff were able to offer validation which increased a sense of reassurance and confidence in the approach they were

taking. It also allowed different members of the team, comprising of different roles, to share their experiences of working with patients. The importance of this on patient outcomes is reflected in the interview with Fiona, who was able to reflect on the role, how much time they spent with patients and how this affected what they were aware of directly:

> *Fiona: "I don't sit out with the patients all the time so I don't get the same sort of conversations with them. I don't see the same things."*

This suggests the use of the supervision groups unified the team and was helpful in increasing awareness of different views, perspectives and understandings of patients. Importantly, this allowed the team to provide appropriate care to the patients.

The final function of the group identified was around ***dealing with difficulties and worries.*** This was reported to be important as the patients within the teams (Assertive Outreach and Inpatient Rehabilitation) tended to be "*complex patients with quite a lot of challenges" (Cameron),* with severe mental illness, comorbidities, and often high risk. They also tend to be *"fairly difficult to engage, with multiple needs"* (Alfie).

Alfie further added that working with this group of patients, with their difficulties and risk issues, can be quite "traumatic" and have a "personal impact" on staff. It was indicated that it is not unusual for this group of patients to engage in high-risk behaviours or offences, which creates a dilemma; balancing the need to provide care for patients whilst managing personal feelings towards the behaviour or action committed, which sometimes go against team members' own ethical principles.

> *Alfie: "We talk about unconditional positive regards, err, in nursing but sometimes erm when you, when you put that into context, how erm, traumatic those events could be err, it erm, it can have a personal impact on you. Erm, and er the patient is quite complex."*

The supervision groups offered a space to share these worries and helped to reduce anxieties when dealing with difficult situations, with participants seeming to rely on the team,

and in many cases, the facilitators (clinical psychologists) to offer them guidance and a lead in managing the difficulties. The groups were deemed important in maintaining the well-being of individual staff, with support from others and a sense of togetherness counteracting worries and stress. It allowed a forum to discuss things they were struggling with, worries about safety, with opportunities to break issues down into smaller chunks, identifying what the actual risks are and then managing these effectively. Ultimately, this was viewed as having a positive impact on the functions of the team.

> *Deborah: "At times it can be really helpful, er, at times particularly when we've got a difficult scenario or a particular client who's raising a lot of difficult scenarios we get a chance to kind of, just get like having someone to kind of guide us through, er that."*

> *Fiona: "If people are having a particular issue and problem with somebody erm, and then we sort of try and break it down into what the actual risk is"*

The group also allowed a space to identify worries that were common across different staff, which helped participants feels less isolated and reassured that they were not doing something wrong. It created a sense of belonging and togetherness, which was especially crucial for individual staff representing a whole discipline e.g. Occupation Therapists (OT), who did not have other OTs to share concerns with; Elsa reported the group to be helpful in managing worries regarding what to do with patients and advice on how to best support patients in moving forward.

Elsa: "As an OT I think the supervision sessions are particularly beneficial because I don't have any other OT's that I work with, so it gives me that opportunity to feedback anything that I'm worried about or anything that I'm struggling with or anything I'm having challenges with."

Geraldine: "It could be a few other people then turn around and say oh actually I'm having a tough time. So you don't feel so isolated, which is nice."

The sense of togetherness and support from others appears to be crucial in challenging worries felt by individual staff which could otherwise have a significant negative impact on staff mental health and general well-being.

Additionally, participants (including those with very senior roles and responsibilities) seemed to have confidence in facilitators' skill and ability to support assessment, formulation and the development of management plans, which helped to ease their worries and concerns. Again, doing this as a team, and sharing their concerns with other members of the team seemed to increase individual confidence to manage difficulties, whilst also increasing reassurance and validation, particularly when there were common worries and concerns that team members were struggling with.

> *Cameron: "If I'm really worried about a particular patient, and it's complicated, erm, having a team supervision discussion, erm, with psychologists facilitating is the you know, by far and away the best way of making sure that a very rigorous and good erm, sort of erm, risk assessment and risk formulation and risk management plan gets put together."*

The four sub-themes illustrate how the team-based processes of shared risk assessment, formulation and intervention are seen to not only enhance the rigour of plans for service users, but also appear to foster a valued sense of group cohesion and belonging. In summary, this theme highlights how there are various practical principles and activities taking part within supervision groups that support the main functions of the teams such as risk management and care planning. However, a common thread connecting these is the sense of team membership and belonging that is fostered through these group-based supervision processes.

<u>Managing the NHS hierarchy</u>

A second theme was identified around participant experiences of ***Managing the NHS hierarchy in the group.*** There were no subordinate themes identified within this theme. There was a strong sense of a role and/or grade-based hierarchy, and a general acceptance that decisions were made by senior staff e.g. the consultant psychiatrist, clinical psychologist or team manager. Additionally, unqualified staff such as healthcare assistants were perceived as

having less of a voice within the overall team setting (outside of the supervision groups) with minimal authority or power to make decisions. Participants with qualified statuses acknowledged the importance of unqualified staff members' perspectives and input into the group supervision sessions, valuing their contributions and stating they sometimes had a better sense of happenings around the unit:

> *Brad: "Sometimes I know with the HCAs, our Band 3s, that they can be a little bit more, they've almost got a better sense of what's going on cos they're delivering more hands-on care, but I think some of them feel because they haven't got qualified status, or just in general by virtue of their role, they almost feel like they're less entitled to speak out which I think is really unfair."*

More senior members of staff supported a sense of working as a team through encouraging more junior members of staff to voice their opinions and perspectives. They acknowledged unqualified staff members have little say in the overall management of care, and advocated for them to be more involved in decision making, which the group enabled:

> *Harriet: "The HCA's don't always get involved [in clinical reviews], at all, so I think it's a good chance for them to have a say as well."*

Participants recognised the efforts of all staff, including those without qualified status, and attempted to reduce the feelings of disparity or disempowerment by allowing an open space within groups for all staff to have a voice and make shared decisions, which increased a sense of team cohesion. They recognised that there were a lot of things unqualified staff members observed, that were vital in making important decisions related to enhanced patient care and better outcomes and attempted to facilitate an environment that felt comfortable and safe to speak in.

> *Elsa: "It's not just the same people making decisions, that a whole team can kind of come together, including healthcare assistants, because they're the people out on the front line a lot of the time and they're the people that are observing a lot of these issues*

so it's useful to get them involved because there's probably a lot of things that they have observed and we haven't."

Participants with qualified and/or more senior backgrounds, made conscious choices to take more of a "back seat" in sessions and allow other (less senior) staff share their views. Importantly, they also acknowledged the power dynamics and potential difficulties for less senior staff raise difficult issues, where they take an active role in initiating more sensitive or difficult topics for group discussions:

Fiona: "Sometimes it needs somebody to say something cos people are a bit reluctant or find it hard"

Additionally, participants with senior roles were able to reflect on the impact of their role, acknowledging that they *"don't see the same things,"* and that patients *"always act differently in front of a manager."* They attempted to counteract the effects of these by promoting team cohesion by encouraging all staff members, especially HCAs and students, to attend groups and provide their input, with *"no pressure"* on how much or what to contribute. By doing this, they acknowledged the significant contributions and positive impacts of junior staff input on the functions of the unit and improving patient rehabilitation.

It appears that these groups can be successful in achieving this safe space for less senior staff to contribute, to an extent, with Geraldine, an unqualified staff member, reporting:

"Supervision the other day was with our consultant and the unit psychologist, an SHO, another doctor, and they're all lovely but I'm still a band 3 and that can be quite intimidating, but they don't make you feel intimidated."

Geraldine went further on to say the senior members of the group are open to and encourage input from all members of the team. This appeared important in addressing some of the power dynamics and imbalances that typically coincide with the NHS hierarchy, which operates a hierarchical, pyramidal form of leadership (Fernandopulle, 2021) and is linked to

differences in clinical authority, for example between nurses and physicians (Omura et al., 2017) where the latter are seen to be "bottom of the hierarchy" (Weiss et al., 2017).

> *Geraldine:* "*He's more than happy, as is the Consultant actually, to get your perspective."*

Responsibility and accountability

A final theme was identified as ***the experience of responsibility and accountability***. Again, there were no subordinate themes identified. Participant experiences were partly dependent on their roles - for instance, being a qualified nurse or not being a doctor and the responsibilities that were attributed to staff based on their roles. For instance, Cameron had a strong awareness that their senior role within the team significantly increased their perceived responsibility and duties towards the patients and team members:

> *Cameron: "I am responsible for most things on the team but that's what I get paid for."*

Fiona reinforced role-based responsibilities, stating *"named nurses usually update the care plans"* which was again evident in the viewpoints of Geraldine who stated:

> *"If it's something they've agreed. That's, that's normally handed over to the nurse in charge then."*

Additionally, participants' perceptions of their roles, especially as a qualified member of staff, seemed to increase feelings of accountability and a sense of pressure, with fear about how things might be interpreted. This was evident in Elsa's experiences, where they stated:

> "*It's hopefully a good like, forum cos yeah you can feel a bit like you're, like as a nurse, the only one, cos you're giving hand over you can, anything you say is kind of, can influence how it's handed over the next day."*

However, working as part of a team and coming to shared decisions with the wider team, really seemed to counteract these strong feelings and instil a sense of relief. Participants generally agreed that regardless of their role-based responsibilities which were inevitable, the group sessions were useful in creating better team cohesion, reducing feelings of blame or individual accountability and approaching things in a "general way."

> *Elsa: "we wouldn't sort of pinpoint like you did this wrong on this day, so we'd talk more in a general way"*

> *Brad: "I get comfort from a team decision because I'm not taking direct responsibility and potential criticism."*

Similarly, Elsa also talks about the group offering a platform for shared decisions and a team approach, which aids feelings of reassurance and safety and reduces a sense of individual accountability. This perceived support from team members appeared to be important in staff managing feelings of stress and on their overall well-being:

> *"You can kind of see it as a bit of a safety net as well if you're really worried about a kind of, I dunno, a service user's risk or something risky on the unit. It just really offers a nice opportunity to discuss that, erm rather than feeling like decision's have been solely left to yourself."*

Additionally, the impact of a team approach was also perceived to have a significant role in important clinical outcomes and the delivery of better-quality patient care.

> *Harriet: "Rather than having one person make a decision about a service users care, the whole MDT can come to that conclusion together which is better quality care to be honest."*

Discussion

This study aimed to explore the experiences of staff working in adult mental healthcare, specifically AOT and inpatient rehabilitation teams, when using psychology-led supervision groups to support the management of risk. Three main themes were identified including the function and value of the group sessions: enhancing belonging and cohesion, managing the NHS hierarchy in the group and the experience of responsibility and accountability. The importance of working as a team was emphasised by all participants, who referred to the positive impact this had on the functions of the team in rehabilitating patients, specifically shared goals and enhanced consistency, resulting in better quality risk management plans and better patient outcomes.

Findings and implications

Findings from this study emphasise the value of a team approach and a sense of belonging, with dedicated space and opportunity to share ideas and perspectives. This may be perceived as sharing of knowledge and can be referred to as "team members sharing task-relevant ideas, information, and suggestions with each other" (Srivastava et al., 2006) (p.4). Communicating and sharing ideas and information with professional colleagues will inform decisions that are made (Sim et al., 2001) with patient care and safety compromised if this isn't done to an adequate standard (Lin & Hsieh, 2006). Therefore, this emphasises the significance of the function of the supervision groups in sharing perspectives, ideas and relevant information.

The supervision groups also enabled staff to identify plans and solutions for any difficulties encountered, increasing confidence and assurance. Team supervisions evidence support for better communication within teams, shared problems solving and an increased sense of unity (Cross et al., 2010). This ties in with previous findings which suggest team communication improves taskwork and teamwork due to better relationships between team members when approaching problems and solutions (Morgan Jr et al., 1993). It also helps to create better team cohesion, which significantly contributes to enhanced psychological

wellbeing, job satisfaction and engagement (Chan, 2016). Interpersonal communication between colleagues, feeling appreciated and mutual respect impact meaningful work experiences and are closely linked with a sense of enhanced belongingness (Baumeister & Leary, 2017). The importance of belonging is emphasised in studies which have found this to improve staff mental well-being, reducing emotional burnout, depersonalization, lack of self-realisation and psychological demands (Coissard et al., 2017).

Team leaders are perceived to be crucial figures in this process, promoting positive team engagement by encouraging motivation, coordination between team members, effective teamwork and essentially taking a lead in problem solving, which is focused on identifying and implementing plans and solutions to help achieve desired goals (Zaccaro et al., 2001). This was referred to within the supervision groups, where the facilitators were perceived as crucial in supporting staff to identify solutions and specific plans for difficult situations or worries staff were experiencing. This is important as productivity at work is enhanced if there are better relationships between staff and supervisors and good cohesion, even during periods of intense stress (Cummins, 1990). This is particularly relevant to this study as staff working in AOT and inpatient rehabilitation often worked with patients with complex needs and high risk, under pressure to manage risks and ensure safety for both staff and patients.

Support from team members through the group nature of the supervision was also found to help reduce feelings of isolation and worries. Team supervisions can allow staff a safe space to share their difficulties, with shared difficulties and experiences enabling staff to feel more emotionally contained (James, 1994) and better team cohesion. Support from colleagues and immediate supervisors positively impacts individual well-being by decreasing a sense of work overload and workplace stress (Bowling et al., 2015). It also helps to facilitate an environment that feels safe and where members are able to develop emotional connections and a sense of belonging that enables them to communicate worries, fears and anxieties and seek appropriate support (Dunbar & Carter, 2017) This team support is fundamental, reflecting on the findings of other studies that have found that staff in mental health care settings are often faced with challenging patients who may be physically aggressive, where emotional exhaustion experienced was alleviated by peer social support (Jenkins & Elliott, 2004). Likewise, lower levels of support from peers and senior staff can increase the risk of burnout (Westwood et al., 2017). This is particularly relevant to the findings of this study, as staff working in AOTs and

inpatient rehabilitation teams tend to support people with complex needs and difficulties, providing intensive care for people who are difficult to engage and have had multiple hospital admissions, predominately due to severe psychosis (Firn et al., 2018).

Social support has also been linked to an increased sense of social connectedness which can moderate feelings of isolation and burnout, and an overall effective method for reducing stress (Grant & Kinman, 2014).

Based on these findings, a key function of supervision groups appears to be to provide support for staff, ensuring a safe environment where staff feel able to contribute. In order to help create a safe environment, it may be helpful for facilitators to be aware of the hierarchical structure of the NHS, which was a major theme generated from this study. Teams and leaders should consider the impacts of this on less senior and/or unqualified staff, which may impact psychological safety, which is reflected in environments where team members have shared views and goals, with an emphasis on positive risk taking and increasing knowledge (Edmondson, 1999). Psychological safety is enabled in situations where one feels included, safe to learn and contribute and safe to challenge notions and ideas without fear of being humiliated, marginalised or penalised (Clark, 2020). Research on psychological safety also suggests team leaders are responsible for managing power dynamics and supporting inclusion; this plays a key role in facilitating an environment where staff feel safe sharing their experiences, identifying solutions and opportunities to learn from mistakes (Nembhard & Edmondson, 2006). These team dynamics are essential when creating a safe environment for teams as well as individual staff to learn from any mistakes. The power dynamics that coincide with a hierarchical structure may further discourage team members to voice concerns or propose suggestions (Weiss et al., 2017) with team members fearing rejection, punishment, or criticism, from senior staff (Morrison, 2014).

For HCAs specifically, there have been visible concerns when accessing clinical supervision (Long et al., 2014). The Care Quality Commission notes “clinical supervision is often primarily aimed at registered professionals” (Commission, 2013) (p.5). However, HCAs tend to be at increased risk of stress and burnout due to prolonged periods of direct contact with patients, who present with high risk challenging behaviours in psychiatric settings. Therefore, it is essential to enable a safe space, where hierarchies are acknowledged, and

encourage staff from less senior positions to share their worries, difficulties and experiences without fear of criticism or rejection.

Importantly, throughout these themes generated in this study, the importance of working together as a team, with a sense of belonging, was a recurring factor for all participants. This is significant, supporting the notion that clear team aims and objectives enhances the quality of team work and cohesion, where teams are more effective and productive (Cooke & Hilton, 2015). Individual staff are able to improve their own clinical performances, having a positive impact on the quality of care delivered to patients (Kilminster et al., 2007). Moreover, clear communication forms the basis for good team understanding, having the biggest impact on team effectiveness compared to other factors such as individual levels of knowledge (Cooke et al., 2013). Based on these findings, it is important for facilitators and team leaders to support a team approach, where communication between different grades and disciplines of staff is encouraged. Facilitators and leaders should demonstrate competent communication skills by using different methods of communication such as body language, gestures and verbal dialogue, ensure information is clear and concise and remain receptive to hearing the viewpoints of all team members (Shaw, 2005). This has shown to have a direct impact on team member satisfaction (Mikkelson et al., 2015), job performance (Mayfield & Mayfield, 2010) and increased self-efficacy and innovation (Mayfield & Mayfield, 2012). Facilitators should also aim to understand better the communication styles of their team members and be flexible and considerate of these. This will assist them in supporting team members to meet their needs and feel more comfortable to attend and participate during group supervisions.

Finally, the findings emphasise the notion that a team approach and support from team members helps to acknowledge some of the challenges faced by staff and may counteract some of the individual stress. Evidence suggests occupational stress is largely due to intense work pressures, combined with increased responsibilities but reduced authority (Mark & Smith, 2012). Therefore, supervision groups and processes should support shared decision making and responsibility to ease the stress of these.

The risk of moral injury i.e. where individuals adopt negative views of themselves, doubt their competence and experience negative emotions such as shame, is also likely to increase when unsure about what ones responsibilities are (Greenberg et al., 2020). Team

accountability on the other hand, refers to shared decisions and actions, where responsibility lies with the whole team rather than on individuals (Kou & Stewart, 2018). This shared responsibility and a consensus of what is deemed helpful as a team (Mead et al., 2001) positively impacts individuals that are supported and cared for as well as the employing organisation (Salzer et al., 2013). This is crucial, especially regarding risk assessments and formulation, where the quality of these is significantly enhanced when approached, shared and co-constructed with an MDT. Risk assessments form part of the SAFE process (Meaden & Hacker, 2010) in which risk management aims to promote positive risk taking by informing relevant clinicians and stakeholders of potential risk factors, most importantly as a team, where assessments and formulations are completed together and accountability is shared. This allows information to be shared accurately, where risk mitigation plans are clearly outlined, there is more clarity regarding treatment decisions, with an aim of reducing risk in the longer term (Meaden & Hacker, 2010).

Essentially, there are many advantages of sharing responsibilities and accountabilities between team members, particularly within healthcare settings (Babiker et al., 2014) and should be encouraged in all teams, acknowledging different roles and responsibilities that may be ascribed to those and mitigating the effects of these.

Strengths and limitations

A strength of the study is the inclusion of participants from a range of different disciplines, including both qualified and unqualified staff members as previous literature has focused mainly on qualified nursing staff. The length of employment within these teams varied between 9 months to 16 years, providing a range of experiences and involvement as a team member, both within the group supervision sessions as well as outside of the group sessions. This allowed an in-depth analysis, with rich quality data. Furthermore, the inclusion of two different teams (representatives from an AOT and an inpatient rehabilitation team) strengthened the plausibility of the findings and applicability across settings.

Additionally, as an inductive method was adopted, which generally supports the generation of new findings, ideas and theories, this enabled an approach where codes and

themes are solely guided by the data obtained, rather than existing theories or frameworks (Byrne, 2022). This allowed new phenomena to be identified, as communicated by participants regarding their experiences, in an area where very little research has previously been conducted. It has helped identify what can be done to enable supervision groups are effective in meeting the needs of staff, increasing the likelihood of a positive impact on clinical patient care.

Most importantly, the study adhered to numerous aspects of APA style Journal Article Reporting Standards (JARS) for qualitative research (JARS-Qual). JARS-Qual recommends researchers reflect on their approach to the research, justifying the methodology adopted, whilst remaining conscious of their positioning in relation to the research question, which this study adhered to. Additionally, there was the inclusion of appropriate literature to provide a rationale for the research question, information on the recruitment process, methodological integrity, evidence of coding and theme development, excerpts from transcriptions and reflections on the researcher's positioning via the use of a reflective journal.

Nonetheless, some demographic details for participants such as age and ethnicity as recommended by JARS-Qual were not explored, although this was to maintain anonymity. Additionally, consideration should be given to the relationship between the researcher and participants, and ensuring there are no ethical issues which may invalidate the findings (Levitt et al., 2018). Although the clinical psychologist facilitating supervision groups for one of the teams involved was also the supervisor for the researcher conducting this study, they did not have direct involvement in the coding of the data, as this was completed by the primary author. Participants were also anonymised with pseudo-names given to protect their identities, making it unlikely that the supervisor could establish their identities. Furthermore, not all participants were from the supervision groups, with some from a different team completely with a different supervision group facilitator.

Despite these safeguards, the supervisor being known to research participants may have impacted the findings, such as limiting any negative feedback given (such as identifying barriers experienced in the supervision groups and/or resistance to change). Furthermore, participants who volunteered to take part may have done so as they had positive views of the

supervision groups whereas staff with more negative views may not have participated. To manage this in future studies, effort could be made by the primary (neutral) researcher to approach all staff and to ascertain reasons given for non-involvement in the study.

Additionally, the topic guide used in this study focused predominantly on the supervision group processes and how they work in managing risk. The psychological consequences including emotional impact of taking part in these groups and discussing risk was not explored. As a result, it is unclear whether there was an emotional impact of the groups on participants and whether participants felt more or less emotionally disturbed after their participation in these supervision groups, after having discussed risk issues in detail.

Finally, it is also important to note that these findings may not be replicated within a different service. As participants mentioned, within these two teams, there were low rates of staff turnover and participants were familiar with team members, including bank staff who appeared to be 'regular' bank staff. In most mental health settings, there is a continuous staffing gap, with difficulties meeting the rising demand; a high percentage (63%) of mental health professionals reported working shifts where there was a shortage of staff, with an even higher percentage (69%) reporting this to occur most, it not all the time (Mahase, 2020). Additionally, in other mental health settings, for instance community mental health teams, where patients needs and impairments may not be as complex, staff may have different experiences of supervision groups when managing risk.

Recommendations for facilitators of group supervision sessions

Based on the findings of this study, a recommendation would be to suggest regular supervision groups are facilitated across AOT and inpatient rehabilitation teams. These should include a protected forum to discuss any difficulties, worries or anxieties staff may be experiencing. Supervision groups should be prioritised, with allocated time and space to ensure these are able to go ahead. Effort should be made to ensure different disciplines and staff of all grades attend supervision groups. If possible, it should be noted the disciplines and grades of staff who might or might not attend regularly as this may be linked to a perceived hierarchy, a sense of responsibility and/or accountability and the difficult feelings associated with this. It

may also be useful for facilitators and team leaders to acknowledge the hierarchical nature of the NHS and encourage less senior members of teams to attend supervision groups and contribute to whatever extent they feel comfortable and confident in doing so. In order to evaluate these recommendations, staff can complete anonymous surveys to feedback their experiences of the groups after each group and the impact it had, including what went well and what could be improved.

Additionally, emphasis should be placed on a team approach, with shared responsibility in identifying plans and solutions and highlighting the numerous significant benefits of peer support from staff and other team members, both on patient care provided as well as on individual staff wellbeing. As evidence suggests, an MDT approach to clinical risk assessment and risk management is encouraged (Otto & Douglas, 2011); it offers guidance when managing risk (Webster et al., 2013) and is strongly recommended for best practices in managing risk (Health, 2007).

Future research recommendations

A recommendation based on the findings of this study would be to explore staff experiences of psychology-led supervision groups when managing risk in different areas of mental health and across different teams where the needs of patients are different to those in AOT or inpatient rehabilitation. For instance, people in community mental health care settings tend to display less violence and aggression towards others, have fewer past suicide attempts, fewer incidences of self-harm and self-neglect, and more stable housing (Schneider et al., 2006). Staff within different teams and settings may have a different sense of team inclusion and membership which may again impact their experience of such supervision groups. Therefore, it will also be useful to explore whether similar findings may be retrieved from working with different age groups such as Child and Adolescent Mental Health Services (CAMHS) or older adult settings.

Additionally, it would be valuable to explore guidelines for supervisions or different types of supervisions. For instance, some supervisions may be task-orientated where leaders communicate what to do, when to do it and how to do it, supporting their subordinates in

clarifying roles and objectives, identifying specific plans to help achieve desired goals and monitoring performance with an overall aim of increasing personal efficacy (Manyak & Mujtaba, 2013). This study also found some supervisions may be relationship-orientated, with particular emphasis on commitment, teamwork and building trust within teams. This latter type of supervision places emphasis on achieving goals through empowering, supporting and motivating subordinates (Van Wart, 2011), striving to meet the needs of individual employees and ensuring the provision of social and emotional support (Bowers & Seashore, 1966). This has shown a positive impact on group cohesion (Tabernero et al., 2009) which was crucially found to be an important factor in this study. It may also influence staff attributions and levels of containment.

Supervisor leadership styles may also be explored; the supervisor experiences were not included in this study and assumed to be equivalent which is unlikely to be true. There are various leadership styles supervisors may opt for, including delegative (allowing group members to make decisions and failing to provide feedback), autocratic (making decisions without input from team members), democratic (considering and valuing input from all team members, and encouraging everyone to share ideas), transactional (focusing on specific tasks and rewarding or punishing team members to motivate them based on performance results) and transformational (motivating others and enhancing productivity through communication and high visibility). Of these, a democratic leader was found to be the most effective leadership style whereas a delegative leader possesses poor, ineffective leadership skills which are highly dissatisfying for team members (Gadirajurrett et al., 2018). Further exploration into different supervisor leadership styles can shed some insight into the impact of this on staff contributions and experiences of supervision groups.

Conclusion

The findings from this study emphasise the importance of clinical supervision groups in meeting a range of different functions, which are highly valued by staff. These include identifying plans and solutions, seeking and providing support, sharing ideas and perspectives and dealing with difficulties and worries. The hierarchical structure of the NHS needs to be acknowledged in order to provide a safe environment for staff to feel comfortable to share their

difficulties and concerns. This may have an impact on staff experiences of responsibility and accountability which can increase feelings of stress and reduced uncertainty. All participants commented on a team approach helping to counteract any negative feelings, rather aiding feelings of team support and unity which resulted in increased confidence to manage difficulties and better support patients in their recovery.

References:

Adams, R. E., Boscarino, J. A., & Figley, C. R. (2006). Compassion fatigue and psychological distress among social workers: A validation study. *American Journal of orthopsychiatry, 76*(1), 103-108.

Adams, R. E., Figley, C. R., & Boscarino, J. A. (2008). The compassion fatigue scale: Its use with social workers following urban disaster. *Research on social work practice, 18*(3), 238-250.

Adriaenssens, J., De Gucht, V., & Maes, S. (2015). Causes and consequences of occupational stress in emergency nurses, a longitudinal study. *Journal of nursing management, 23*(3), 346-358.

Alhalal, E., Alrashidi, L. M., & Alanazi, A. N. (2020). Predictors of patient-centered care provision among nurses in acute care setting. *Journal of nursing management, 28*(6), 1400-1409.

Amjad, S., Abbasi, S., Altaf, U., & Ijaz, F. (2020, 02/13). Mediating Role of Professional Quality of Life Between Emotional Regulation and Psychological Well-Being among Mental Health Professionals. *14*, 101-109. https://doi.org/10.19080/PBSIJ.2020.14.555888

Andersen, J., Papazoglou, K., Nyman, M., Koskelainen, M., & Gustafsberg, H. (2015, 10/15). Fostering Resilience among the Police. *The Journal of Law Enforcement, 5*, 1-13.

Ang, S. Y., Hemsworth, D., Uthaman, T., Ayre, T. C., Mordiffi, S. Z., Ang, E., & Lopez, V. (2018). Understanding the influence of resilience on psychological outcomes—Comparing results from acute care nurses in Canada and Singapore. *Applied Nursing Research, 43*, 105-113.

Appleby, L., Kapur, N., Shaw, J., Hunt, I., Ibrahim, S., Turnbull, P., Bojanić, L., Rodway, C., Tham, S.-G., & Richards, N. (2019). National Confidential Inquiry into Suicide and Safety in Mental Health: Annual Report: England, Northern Ireland, Scotland and Wales.

Ardito, R. B., & Rabellino, D. (2011). Therapeutic alliance and outcome of psychotherapy: historical excursus, measurements, and prospects for research. *Frontiers in psychology, 2*, 270.

Babiker, A., El Husseini, M., Al Nemri, A., Al Frayh, A., Al Juryyan, N., Faki, M. O., Assiri, A., Al Saadi, M., Shaikh, F., & Al Zamil, F. (2014). Health care professional development: Working as a team to improve patient care. *Sudanese journal of paediatrics, 14*(2), 9-16. https://pubmed.ncbi.nlm.nih.gov/27493399 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4949805/

Barker, G. G., Volk, F., & Peters, C. (2021). Cultural influences on burnout: a Swedish–American comparison. *International Journal of Workplace Health Management*.

Başoğul, C., Baysan Arabaci, L., Mutlu Satıl, E., & Büyükbayram Aslan, A. (2021). Professional values and professional quality of life among mental health nurses: A cross-sectional study. *Nursing & health sciences, 23*(2), 362-371.

Baumeister, R. F., & Leary, M. R. (2017). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Interpersonal development*, 57-89.

Beddoe, L. (2010). Surveillance or reflection: Professional supervision in 'the risk society'. *British journal of social work, 40*(4), 1279-1296.

Begic, S., Weaver, J. M., & McDonald, T. W. (2019). Risk and protective factors for secondary traumatic stress and burnout among home visitors. *Journal of Human Behavior in the Social Environment, 29*(1), 137-159.

Berry, K., Barrowclough, C., & Wearden, A. (2009). A Pilot Study Investigating the Use of Psychological Formulations to Modify Psychiatric Staff Perceptions of Service Users with Psychosis. *Behavioural and cognitive psychotherapy, 37*(1), 39-48. https://doi.org/10.1017/S1352465808005018

Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *Bmj, 314*(7080), 572.

Borritz, M., Rugulies, R., Christensen, K. B., Villadsen, E., & Kristensen, T. (2006). Burnout as a predictor of self-reported sickness absence among human service workers: prospective findings from three year follow up of the PUMA study. *Occupational and environmental medicine, 63*(2), 98-106.

Bowers, D. G., & Seashore, S. E. (1966). Predicting organizational effectiveness with a four-factor theory of leadership. *Administrative science quarterly*, 238-263.

Bowers, L., James, K., Quirk, A., Simpson, A., Stewart, D., & Hodsoll, J. (2015). Reducing conflict and containment rates on acute psychiatric wards: The Safewards cluster randomised controlled trial. *International Journal of Nursing Studies, 52*(9), 1412-1422.

Bowling, N. A., Alarcon, G. M., Bragg, C. B., & Hartman, M. J. (2015). A meta-analytic examination of the potential correlates and consequences of workload. *Work & stress, 29*(2), 95-113.

Bracken, P. (2002). *Trauma: culture, meaning and philosophy*. Whurr Publishers.

Bramley, L., & Matiti, M. (2014). How does it really feel to be in my shoes? Patients' experiences of compassion within nursing care and their perceptions of developing compassionate nurses. *Journal of clinical nursing, 23*(19-20), 2790-2799.

Braun, V., & Clarke, V. (2012). Thematic analysis.

Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners*. sage.

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health, 11*(4), 589-597.

Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative research in psychology, 18*(3), 328-352.

Bride, B. E., Robinson, M. M., Yegidis, B., & Figley, C. R. (2004). Development and validation of the secondary traumatic stress scale. *Research on social work practice, 14*(1), 27-35.

Brito-Marcelino, A., Oliva-Costa, E. F., Sarmento, S. C. P., & Carvalho, A. A. (2020, Dec 11). Burnout syndrome in speech-language pathologists and audiologists: a review. *Rev Bras Med Trab, 18*(2), 217-222. https://doi.org/10.47626/1679-4435-2020-480

Brunero, S., & Stein-Parbury, J. (2008). The effectiveness of clinical supervision in nursing: an evidenced based literature review. *Australian Journal of Advanced Nursing, The, 25*(3), 86-94.

Bunyan, M., Crowley, J., Smedley, N., Mutti, M. F., Cashen, A., Thompson, T., & Foster, J. (2017). Feasibility of training nurses in motivational interviewing to improve patient experience in mental health inpatient rehabilitation: a pilot study. *Journal of psychiatric and mental health nursing, 24*(4), 221-231.

Burton, A., Burgess, C., Dean, S., Koutsopoulou, G. Z., & Hugh-Jones, S. (2017). How effective are mindfulness-based interventions for reducing stress among healthcare professionals? A systematic review and meta-analysis. *Stress and Health, 33*(1), 3-13.

Byrne, D. (2022, 2022/06/01). A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Quality & Quantity, 56*(3), 1391-1412. https://doi.org/10.1007/s11135-021-01182-y

Cavanagh, N., Cockett, G., Heinrich, C., Doig, L., Fiest, K., Guichon, J. R., Page, S., Mitchell, I., & Doig, C. J. (2020). Compassion fatigue in healthcare providers: A systematic review and meta-analysis. *Nursing ethics, 27*(3), 639-665.

Chan, S. (2016). Belonging to a workplace: first-year apprentices' perspectives on factors determining engagement and continuation through apprenticeship. *International Journal for Educational and Vocational Guidance, 16*(1), 9-27.

Chang, K., & Taormina, R. J. (2011). Reduced secondary trauma among Chinese earthquake rescuers: A test of correlates and life indicators. *Journal of Loss and Trauma, 16*(6), 542-562.

Choi, S. H., & Lee, H. (2017, Oct). Workplace violence against nurses in Korea and its impact on professional quality of life and turnover intention. *J Nurs Manag, 25*(7), 508-518. https://doi.org/10.1111/jonm.12488

Clark, T. R. (2020). *The 4 stages of psychological safety: Defining the path to inclusion and innovation*. Berrett-Koehler Publishers.

Clarke, V., & Braun, V. (2021). Thematic analysis: a practical guide. *Thematic Analysis*, 1-100.

Cocker, F., & Joss, N. (2016). Compassion Fatigue among Healthcare, Emergency and Community Service Workers: A Systematic Review. *International journal of environmental research and public health, 13*(6), 618. https://doi.org/10.3390/ijerph13060618

Cohen, R. J., Swerdlik, M. E., & Phillips, S. M. (1996). *Psychological testing and assessment: An introduction to tests and measurement*. Mayfield Publishing Co.

Coissard, F., Ndao, M.-L., Gilibert, D., & Banovic, I. (2017). Relationships at work and psychosocial risk: The feeling of belonging as indicator and mediator. *European Review of Applied Psychology, 67*(6), 317-325.

Cole, S., Johnstone, L., Oliver, D. N., & Whomsley, S. (2011). Good practice guidelines on the use of psychological formulation.

Collins, S., & Long, A. (2003). Too tired to care? The psychological effects of working with trauma. *Journal of Psychiatric and Mental Health Nursing, 10*(1), 17-27.

Commission, C. Q. (2012). *The state of health care and adult social care in England in 2011/12* (Vol. 763). The Stationery Office.

Commission, C. Q. (2013). Supporting information and guidance: Supporting effective clinical supervision. *London: Care Quality Commission*.

Cooke, N. J., Gorman, J. C., Myers, C. W., & Duran, J. L. (2013). Interactive team cognition. *Cognitive science, 37*(2), 255-285.

Cooke, N. J., & Hilton, M. L. (2015). Committee on the Science of Team Science, Board on Behavioral, Cognitive, and Sensory Sciences, Division of Behavioral and Social Sciences and Education, National Research Council. 2015. *Enhancing the effectiveness of team science*.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology, 78*(1), 98.

Coyle, A. (2007). Introduction to qualitative psychological research. *Analysing qualitative data in psychology, 2*, 9-30.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika, 16*(3), 297-334.

Cross, W., Moore, A., & Ockerby, S. (2010). Clinical supervision of general nurses in a busy medical ward of a teaching hospital. *Contemporary Nurse, 35*(2), 245-253.

Cullen, S. W., Nath, S. B., & Marcus, S. C. (2010). Toward understanding errors in inpatient psychiatry: a qualitative inquiry. *Psychiatric quarterly, 81*(3), 197-205.

Cummins, R. C. (1990). Job stress and the buffering effect of supervisory support. *Group & Organization Studies, 15*(1), 92-104.

Cutcliffe, J. R., Sloan, G., & Bashaw, M. (2018). A systematic review of clinical supervision evaluation studies in nursing. *International journal of mental health nursing, 27*(5), 1344-1363.

Daundasekara, S. S., Arlinghaus, K. R., & Johnston, C. A. (2020). Quality of Life: The Primary Goal of Lifestyle Intervention. *American journal of lifestyle medicine, 14*(3), 267-270. https://doi.org/10.1177/1559827620907309

Davis, S., Lind, B. K., & Sorensen, C. (2013). A comparison of burnout among oncology nurses working in adult and pediatric inpatient and outpatient settings. Oncology nursing forum,

DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage publications.

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20V3: Assessing risk for violence: User guide*. Mental Health, Law, and Policy Institute, Simon Fraser University.

Driscoll, J. (2007). Adventures in facilitating group clinical supervision in practice. *Practising Clinical Supervision–A Reflective Approach for Healthcare Professionals*, 53-71.

Drury, V., Craigie, M., Francis, K., Aoun, S., & Hegney, D. G. (2014). Compassion satisfaction, compassion fatigue, anxiety, depression and stress in registered nurses in A ustralia: Phase 2 results. *Journal of nursing management, 22*(4), 519-531.

Duarte, J. (2017, Dec 2017). Professional quality of life in nurses: Contribution for the validation of the Portuguese version of the Professional Quality of Life Scale-5 (ProQOL-5). *Análise Psicológica, 35*(4), 529-542. http://publicacoes.ispa.pt/index.php/ap/article/download/1260/pdf

Duarte, J., & Pinto-Gouveia, J. (2017). Empathy and feelings of guilt experienced by nurses: A cross-sectional study of their role in burnout and compassion fatigue symptoms. *Applied Nursing Research, 35*, 42-47.

Dubale, B. W., Friedman, L. E., Chemali, Z., Denninger, J. W., Mehta, D. H., Alem, A., Fricchione, G. L., Dossett, M. L., & Gelaye, B. (2019). Systematic review of burnout among healthcare providers in sub-Saharan Africa. *BMC public health, 19*(1), 1-20.

[Record #348 is using a reference type undefined in this output style.]

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative science quarterly, 44*(2), 350-383.

Elkonin, D., & Van der Vyver, L. (2011). Positive and negative emotional responses to work-related trauma of intensive care nurses in private health care facilities. *Health SA Gesondheid, 16*(1).

Elliott, R., Fischer, C. T., & Rennie, D. L. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British journal of clinical psychology, 38*(3), 215-229.

Elliott, T. R., & Gramling, S. E. (1990). Personal assertiveness and the effects of social support among college students. *Journal of counseling psychology, 37*(4), 427.

Erkorkmaz, U., Dogu, O., & Cinar, N. (2018). The relationship between burnout, self-esteem and professional life quality of nurses. *Journal of the College of Physicians and Surgeons Pakistan, 28*(7), 549-553.

Fernando III, A. T., & Consedine, N. S. (2014). Beyond compassion fatigue: the transactional model of physician compassion. *Journal of pain and symptom management, 48*(2), 289-298.

Fernandopulle, N. (2021). To what extent does hierarchical leadership affect health care outcomes? *Medical journal of the Islamic Republic of Iran, 35*, 117-117. https://doi.org/10.47176/mjiri.35.117

Figley, C. R. (1995). *Compassion fatigue: Coping with secondary traumatic stress in those who treat the traumatized*. Brunner/Mazel, Publishers.

Figley, C. R. (2002). Compassion fatigue: Psychotherapists' chronic lack of self care. *Journal of clinical psychology, 58*(11), 1433-1441.

Figley, C. R. (2013). *Compassion fatigue: Coping with secondary traumatic stress disorder in those who treat the traumatized*. Routledge.

Figley, C. R., & Stamm, B. H. (1996). Psychometric review of compassion fatigue self test. *Measurement of stress, trauma, and adaptation*, 127-130.

Firn, M., White, S. J., Hubbeling, D., & Jones, B. (2018). The replacement of assertive outreach services by reinforcing local community teams: a four-year observational study. *Journal of Mental Health, 27*(1), 4-9.

Flewett, T. (2010). *Clinical Risk Management: An introductory text for mental health clinicians*. Elsevier Australia.

Fu, C. Y., Yang, M. S., Leung, W., Liu, Y. Y., Huang, H. W., & Wang, R. H. (2018). Associations of professional quality of life and social support with health in clinical nurses. *Journal of nursing management, 26*(2), 172-179.

Gadirajurrett, H., Srinivasan, R., Stevens, J., & Jeena, N. (2018). Impact of leadership on team's performance.

Galiana, L., Arena, F., Oliver, A., Sansó, N., & Benito, E. (2017, Mar 2017). Compassion satisfaction, compassion fatigue, and burnout in Spain and Brazil: ProQOL validation and cross-cultural diagnosis. *Journal of Pain and Symptom Management, 53*(3), 598-604. http://www.jpsmjournal.com/article/S0885392416311988/pdf

Galiana, L., Oliver, A., De Simone, G., Linzitto, J. P., Benito, E., & Sansó, N. (2019). A brief measure for the assessment of competence in coping with death: The coping with death scale short version. *Journal of pain and symptom management, 57*(2), 209-215.

[Record #407 is using a reference type undefined in this output style.]

Golia, G. M., & McGovern, A. R. (2015). If you save me, I'll save you: the power of peer supervision in clinical training and professional development. *The British Journal of Social Work, 45*(2), 634-650.

Gopalkrishnan, N. (2018). Cultural Diversity and Mental Health: Considerations for Policy and Practice. *Front Public Health, 6*, 179. https://doi.org/10.3389/fpubh.2018.00179

Grant, L., & Kinman, G. (2014). The Importance of Resilience for Staff and Students in the 'Helping'Professions: Developing an Emotional Curriculum. *Higher Education Academy: York, UK*.

Greenberg, N., Docherty, M., Gnanapragasam, S., & Wessely, S. (2020). Managing mental health challenges faced by healthcare workers during covid-19 pandemic. *Bmj, 368*.

Hagen, J., Knizek, B. L., & Hjelmeland, H. (2017). Mental health nurses' experiences of caring for suicidal patients in psychiatric wards: an emotional endeavor. *Archives of psychiatric nursing, 31*(1), 31-37.

Hall, L. H., Johnson, J., Watt, I., Tsipa, A., & O'Connor, D. B. (2016). Healthcare staff wellbeing, burnout, and patient safety: a systematic review. *PloS one, 11*(7), e0159015.

Hall, T., & Cox, D. (2009). Clinical supervision: an appropriate term for physiotherapists? *Learning in Health and Social Care, 8*(4), 282-291.

Hambleton, R., & De Jong, J. H. A. L. (2003, 04/01). Advances in Translating and Adapting Educational and Psychological Tests. *Language Testing - LANG TEST, 20*, 127-134. https://doi.org/10.1191/0265532203lt247xx

[Record #392 is using a reference type undefined in this output style.]

Heeb, J.-L., & Haberey-Knuessi, V. (2014). Health professionals facing burnout: what do we know about nursing managers? *Nursing research and practice, 2014*.

Hemsworth, D., Baregheh, A., Aoun, S., & Kazanjian, A. (2018). A critical enquiry into the psychometric properties of the professional quality of life scale (ProQol-5) instrument. *Applied Nursing Research, 39*, 81-88.

Heritage, B., Rees, C. S., & Hegney, D. G. (2018). The ProQOL-21: A revised version of the Professional Quality of Life (ProQOL) scale based on Rasch analysis. *PloS one, 13*(2), e0193478.

Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L., & Sterne, J. A. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *Bmj, 343*.

Hinderer, K. A., VonRueden, K. T., Friedmann, E., McQuillan, K. A., Gilmore, R., Kramer, B., & Murray, M. (2014). Burnout, compassion fatigue, compassion satisfaction, and secondary traumatic stress in trauma nurses. *Journal of Trauma Nursing| JTN, 21*(4), 160-169.

Hofmeyer, A., Kennedy, K., & Taylor, R. (2020, 2020/04/01/). Contesting the term 'compassion fatigue': Integrating findings from social neuroscience and self-care research. *Collegian, 27*(2), 232-237. https://doi.org/https://doi.org/10.1016/j.colegn.2019.07.001

Horkan, A. M. (2014). Alarm fatigue and patient safety. *Nephrology Nursing Journal, 41*(1), 83-86.

Horowitz, M., Wilner, N., & Alvarez, W. (1979). Impact of Event Scale: A measure of subjective stress. *Psychosomatic medicine, 41*(3), 209-218.

Hundall Stamm, B. (2009). Professional Quality of Life Measure: Compassion, Satisfaction, and Fatigue Version 5 (ProQOL).

[Record #346 is using a reference type undefined in this output style.]

Jenkins, R., & Elliott, P. (2004). Stressors, burnout and social support: nurses in acute mental health settings. *Journal of advanced nursing, 48*(6), 622-631.

Johnson, S., & Naidoo, A. (2017). Can evolutionary insights into the brain's response to threat suggest different group interventions for perceived stress and burnout of teachers in high-risk schools? *South African journal of psychology, 47*(3), 401-415.

Kidder, L. H., & Fine, M. (1987). Qualitative and quantitative methods: When stories converge. *New directions for program evaluation, 1987*(35), 57-75.

Kilminster, S., Cottrell, D., Grant, J., & Jolly, B. (2007). AMEE Guide No. 27: Effective educational and clinical supervision. *Medical teacher, 29*(1), 2-19.

Kilminster, S. M., & Jolly, B. C. (2000). Effective supervision in clinical practice settings: a literature review. *Medical education, 34*(10), 827-840.

Kim, K., Han, Y., Kwak, Y., & Kim, J.-s. (2015). Professional quality of life and clinical competencies among Korean nurses. *Asian nursing research, 9*(3), 200-206.

Kim, S. (2014). World Health Organization Quality of Life (WHOQOL) Assessment. In A. C. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research* (pp. 7260-7261). Springer Netherlands. https://doi.org/10.1007/978-94-007-0753-5_3282

Kim, S. Y., Park, J. E., Lee, Y. J., Seo, H.-J., Sheen, S.-S., Hahn, S., Jang, B.-H., & Son, H.-J. (2013). Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *Journal of clinical epidemiology, 66*(4), 408-414.

Kim, Y. H., Kim, S. R., Kim, Y. O., Kim, J. Y., Kim, H. K., & Kim, H. Y. (2017). Influence of type D personality on job stress and job satisfaction in clinical nurses: the mediating effects of compassion fatigue, burnout, and compassion satisfaction. *Journal of advanced nursing, 73*(4), 905-916.

Kou, C.-Y., & Stewart, V. (2018). Group accountability: A review and extension of existing research. *Small Group Research, 49*(1), 34-61.

Kramarz, E., Mok, C. L. M., Westhead, M., & Riches, S. (2022). Staff experience of team case formulation to address challenging behaviour on acute psychiatric wards: a mixed-methods study. *Journal of Mental Health*, 1-12.

Ledoux, K. (2015). Understanding compassion fatigue: Understanding compassion. *Journal of advanced nursing, 71*(9), 2041-2050.

Lee, Y., & Seomun, G. (2016). Role of compassion competence among clinical nurses in professional quality of life. *International Nursing Review, 63*(3), 381-387.

Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., Josselson, R., & Suárez-Orozco, C. (2018). Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The APA Publications and Communications Board task force report. *American psychologist, 73*(1), 26.

Lewis, G., & Doyle, M. (2009, 2009/12/15). Risk Formulation: What are We Doing and Why? *International Journal of Forensic Mental Health, 8*(4), 286-292. https://doi.org/10.1080/14999011003635696

Lheureux, F., Truchot, D., Borteyrou, X., & Rascle, N. (2017). The Maslach Burnout Inventory–Human Services Survey (MBI-HSS): factor structure, wording effect and psychometric qualities of known problematic items. *Le travail humain, 80*(2), 161-186.

Lin, B., & Hsieh, C.-t. (2006). Critical factors for assessing service quality of online pharmacies: a research framework. *International Journal of Electronic Healthcare, 2*(4), 398-414.

Long, C., Harding, S., Payne, K., & Collins, L. (2014). Nursing and health-care assistant experience of supervision in a medium secure psychiatric service for women:

implications for service development. *Journal of Psychiatric and Mental Health Nursing, 21*(2), 154-162.

Lynch, S. H., & Lobo, M. L. (2012). Compassion fatigue in family caregivers: a Wilsonian concept analysis. *Journal of advanced nursing, 68*(9), 2125-2134.

Mahase, E. (2020). Workforce crisis has left mental health staff at "breaking point" as demand rises. *Bmj, 368*, m88. https://doi.org/10.1136/bmj.m88

Manyak, T., & Mujtaba, B. (2013). Task and relationship orientations of Ugandans and Americans. *International Business and Management, 6*(1), 12-20.

Marchand, A., Blanc, M.-E., & Durand, P. (2015). Genre, âge, catégorie professionnelle, secteur économique et santé mentale en milieu de travail: les résultats de l'étude SALVEO. *Canadian journal of public health, 106*(4), e223-e229.

Mark, G., & Smith, A. P. (2012). Occupational stress, job characteristics, coping, and the mental health of nurses. *British journal of health psychology, 17*(3), 505-521.

Martin, P., Lizarondo, L., Kumar, S., & Snowdon, D. (2021). Impact of clinical supervision on healthcare organisational outcomes: A mixed methods systematic review. *PloS one, 16*(11), e0260156. https://doi.org/10.1371/journal.pone.0260156

Maslach, C. (2003). *Burnout: The cost of caring*. Ishk.

Maslach, C., Jackson, S. E., & Leiter, M. P. (1997). *Maslach burnout inventory*. Scarecrow Education.

Mason, T., Williams, R., & Vivian-Byrne, S. (2002). Multi-disciplinary working in a forensic mental health setting: ethical codes of reference. *Journal of psychiatric and mental health nursing, 9*(5), 563-572.

Mastoras, S. M., & Andrews, J. J. (2011). The supervisee experience of group supervision: Implications for research and practice. *Training and Education in Professional Psychology, 5*(2), 102.

Mathieu, F. (2007). Running on empty: Compassion fatigue in health professionals. *Rehab & Community Care Medicine, 4*, 1-7.

Maura, G., Igor, P., Nicola, F., Marco, P., Federico, M., Gabriele, F., Paolo, C., & Marcello, C. (2019). Association between burnout and sense of coherence among speech and language therapists: an exploratory study in Italy. *Acta Bio Medica: Atenei Parmensis, 90*(Suppl 4), 25.

Mayfield, J., & Mayfield, M. (2010). Leader-level influence on motivating language: A two-level model investigation on worker performance and job satisfaction. *Competitiveness Review: An International Business Journal*.

Mayfield, J., & Mayfield, M. (2012). The relationship between leader motivating language and self-efficacy: A partial least squares model analysis. *The Journal of Business Communication (1973), 49*(4), 357-376.

Mead, S., Hilton, D., & Curtis, L. (2001). Peer support: a theoretical perspective. *Psychiatric rehabilitation journal, 25*(2), 134.

Meaden, A., Commander, M., Cowan, C., & Edwards, T. (2014, Dec). Patient engagement and problematic behaviours in nurse-staffed residential rehabilitation units. *Psychiatr Bull (2014), 38*(6), 260-264. https://doi.org/10.1192/pb.bp.113.045252

Meaden, A., Fox, A., & Hacker, D. (2015). Team-Based Cognitive Therapy for distress and problematic behaviour associated with positive symptoms. *Innovations in Psychosocial Interventions for Psychosis: Working with the hard to reach*, 184.

Meaden, A., Fox, A., & Hussain, H. (2022). *Team-Based Shared Formulation for Psychosis: The SAFE Approach*. Taylor & Francis.

Meaden, A., & Hacker, D. (2010). *Problematic and risk behaviours in psychosis: A shared formulation approach*. Routledge.

Michalec, B., Diefenbeck, C., & Mahoney, M. (2013). The calm before the storm? Burnout and compassion fatigue among undergraduate nursing students. *Nurse Education Today, 33*(4), 314-320.

Mikkelson, A. C., York, J. A., & Arritola, J. (2015). Communication competence, leadership behaviors, and employee outcomes in supervisor-employee relationships. *Business and Professional Communication Quarterly, 78*(3), 336-354.

Mills, J. D. (2003). SPSS textbooks: A review for teachers. *Statistics Education Research Journal, 2*(2), 59-70.

Milne, D. (2007). An empirical definition of clinical supervision. *British Journal of Clinical Psychology, 46*(4), 437-447.

Mohammad, A. A. A. (2012). Resilience, burnout, and role stress among military personnel. *Middle East current psychiatry, 19*(2), 123-129.

Morgan Jr, B. B., Salas, E., & Glickman, A. S. (1993). An analysis of team evolution and maturation. *The Journal of General Psychology, 120*(3), 277-291.

Morrison, E. W. (2014). Employee voice and silence. *Annu. Rev. Organ. Psychol. Organ. Behav., 1*(1), 173-197.

Nembhard, I. M., & Edmondson, A. C. (2006). Making it safe: The effects of leader inclusiveness and professional status on psychological safety and improvement efforts in health care teams. *Journal of Organizational Behavior: The International Journal*

of Industrial, Occupational and Organizational Psychology and Behavior, 27(7), 941-966.

Nolte, A. G., Downing, C., Temane, A., & Hastings-Tolsma, M. (2017). Compassion fatigue in nurses: A metasynthesis. *Journal of clinical nursing, 26*(23-24), 4364-4378.

Norrman Harling, M., Högman, E., & Schad, E. (2020). Breaking the taboo: eight Swedish clinical psychologists' experiences of compassion fatigue. *International journal of qualitative studies on health and well-being, 15*(1), 1785610.

Nunnally, J. C. (1994). *Psychometric theory 3E*. Tata McGraw-hill education.

O'Connor, K., Neff, D. M., & Pitman, S. (2018). Burnout in mental health professionals: A systematic review and meta-analysis of prevalence and determinants. *European Psychiatry, 53*, 74-99.

[Record #321 is using a reference type undefined in this output style.]

Omura, M., Maguire, J., Levett-Jones, T., & Stone, T. E. (2017). The effectiveness of assertiveness communication training programs for healthcare professionals and students: A systematic review. *International journal of nursing studies, 76*, 120-128.

Otto, R. K., & Douglas, K. S. (2011). *Handbook of violence risk assessment*. Routledge.

Palmer-Olsen, L., Gold, L. L., & Woolley, S. R. (2011). Supervising emotionally focused therapists: A systematic research-based model. *Journal of marital and family therapy, 37*(4), 411-426.

Pollock, A., Campbell, P., Deery, R., Fleming, M., Rankin, J., Sloan, G., & Cheyne, H. (2017). A systematic review of evidence relating to clinical supervision for nurses, midwives and allied health professionals. *Journal of Advanced Nursing, 73*(8), 1825-1837.

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: applications to practice* (Vol. 892). Pearson/Prentice Hall Upper Saddle River, NJ.

Priebe, S., Fakhoury, W., White, I., Watts, J., Bebbington, P., Billings, J., Burns, T., Johnson, S., Muijen, M., Ryrie, I., & Wright, C. (2004). Characteristics of teams, staff and patients: associations with outcomes of patients in assertive outreach. *British Journal of Psychiatry, 185*(4), 306-311. https://doi.org/10.1192/bjp.185.4.306

Proctor, B., Marken, M., & Payne, M. (1988). Enabling and ensuring. *Supervision: A co-operative exercise in accountability. Leicester, UK: Leicester National Youth Bureau and Council for Education and Training in Youth and Community Work*.

Raphael, J., Hutchinson, T., Haddock, G., Emsley, R., Bucci, S., Lovell, K., Edge, D., Price, O., Udachina, A., & Day, C. (2021). A study on the feasibility of delivering a psychologically informed ward-based intervention on an acute mental health ward. *Clinical psychology & psychotherapy, 28*(6), 1587-1597.

Rauvola, R. S., Vega, D. M., & Lavigne, K. N. (2019). Compassion fatigue, secondary traumatic stress, and vicarious traumatization: A qualitative review and research agenda. *Occupational Health Science, 3*(3), 297-336.

Raykov, T., & Marcoulides, G. A. (2017, 2019/02/01). Thanks Coefficient Alpha, We Still Need You! *Educational and Psychological Measurement, 79*(1), 200-210. https://doi.org/10.1177/0013164417725127

Roach, K. E. (2006). Measurement of Health Outcomes: Reliability, Validity and Responsiveness. *JPO: Journal of Prosthetics and Orthotics, 18*(6). https://journals.lww.com/jpojournal/Fulltext/2006/01001/Measurement_of_Health_Outcomes__Reliability,.3.aspx

Roney, L. N., & Acri, M. C. (2018). The cost of caring: an exploration of compassion fatigue, compassion satisfaction, and job satisfaction in pediatric nurses. *Journal of Pediatric Nursing, 40*, 74-80.

Ross, S. M. (2017). *Introductory statistics*. Academic Press.

Ruiz-Fernández, M. D., Ramos-Pichardo, J. D., Ibáñez-Masero, O., Carmona-Rega, M. I., Sánchez-Ruiz, M. J., & Ortega-Galán, Á. M. (2021). Professional quality of life, self-compassion, resilience, and empathy in healthcare professionals during COVID-19 crisis in Spain. *Research in Nursing & Health, 44*(4), 620-632.

Salvagioni, D. A. J., Melanda, F. N., Mesas, A. E., González, A. D., Gabani, F. L., & Andrade, S. M. d. (2017). Physical, psychological and occupational consequences of job burnout: A systematic review of prospective studies. *PloS one, 12*(10), e0185781-e0185781. https://doi.org/10.1371/journal.pone.0185781

Salzer, M. S., Darr, N., Calhoun, G., Boyer, W., Loss, R. E., Goessel, J., Schwenk, E., & Brusilovskiy, E. (2013). Benefits of working as a certified peer specialist: Results from a statewide survey. *Psychiatric rehabilitation journal, 36*(3), 219.

Samson, T., Iecovich, E., & Shvartzman, P. (2016). Psychometric characteristics of the Hebrew version of the Professional Quality-of-Life Scale. *Journal of pain and symptom management, 52*(4), 575-581. e571.

Sanford, R. L., Cerel, J., & Frey, L. M. (2018). Survivor of Suicide Loss Support Group Facilitators: Do Peers and Professionals Differ? *Social Work With Groups, 41*(4), 306-322.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological assessment, 8*(4), 350.

Schneider, J., Brandon, T., Wooff, D., Carpenter, J., & Paxton, R. (2006). Assertive outreach: policy and reality. *Psychiatric Bulletin, 30*(3), 89-94.

Shaw, K. (2005). Getting leaders involved in communication strategy. *Strategic Communication Management, 9*(6), 14.

Sim, I., Gorman, P., Greenes, R. A., Haynes, R. B., Kaplan, B., Lehmann, H., & Tang, P. C. (2001). Clinical decision support systems for the practice of evidence-based medicine. *Journal of the American Medical Informatics Association, 8*(6), 527-534.

Sinclair, S., Beamer, K., Hack, T. F., McClement, S., Raffin Bouchal, S., Chochinov, H. M., & Hagen, N. A. (2017). Sympathy, empathy, and compassion: A grounded theory study of palliative care patients' understandings, experiences, and preferences. *Palliative medicine, 31*(5), 437-447.

Sinclair, S., Norris, J. M., McConnell, S. J., Chochinov, H. M., Hack, T. F., Hagen, N. A., McClement, S., & Bouchal, S. R. (2016). Compassion: a scoping review of the healthcare literature. *BMC palliative care, 15*(1), 1-16.

Sinclair, S., Raffin-Bouchal, S., Venturato, L., Mijovic-Kondejewski, J., & Smith-MacDonald, L. (2017). Compassion fatigue: A meta-narrative review of the healthcare literature. *International Journal of Nursing Studies, 69*, 9-24.

Slade, J., Schneider, J., & Brandon, T. (2005). Assertive Outreach team and caseload survey. *Mental Health and Learning Disabilities Research and Practice, 2*(2), 78-94.

Slatten, L. A., Carson, K. D., & Carson, P. P. (2011). Compassion fatigue and burnout: What managers should know. *The health care manager, 30*(4), 325-333.

Snowdon, D. A., Leggat, S. G., & Taylor, N. F. (2017). Does clinical supervision of healthcare professionals improve effectiveness of care and patient experience? A systematic review. *BMC Health Services Research, 17*(1), 1-11.

Sodeke-Gregson, E. A., Holttum, S., & Billings, J. (2013). Compassion satisfaction, burnout, and secondary traumatic stress in UK therapists who work with adult trauma clients. *European journal of psychotraumatology, 4*, 10.3402/ejpt.v3404i3400.21869. https://doi.org/10.3402/ejpt.v4i0.21869

Sorenson, C., Bolick, B., Wright, K., & Hamilton, R. (2016). Understanding compassion fatigue in healthcare providers: A review of current literature. *Journal of Nursing Scholarship, 48*(5), 456-465.

Srivastava, A., Bartol, K. M., & Locke, E. A. (2006). Empowering leadership in management teams: Effects on knowledge sharing, efficacy, and performance. *Academy of management journal, 49*(6), 1239-1251.

Stamm, B. (2010). The concise manual for the professional quality of life scale.

Stein, L. I., & Test, M. A. (1980). Alternative to mental hospital treatment: I. Conceptual model, treatment program, and clinical evaluation. *Archives of general psychiatry, 37*(4), 392-397.

Tabernero, C., Chambel, M. J., Curral, L., & Arana, J. M. (2009). The role of task-oriented versus relationship-oriented leadership on normative contract and group performance. *Social Behavior and Personality: an international journal, 37*(10), 1391-1404.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education, 2*, 53-55. https://doi.org/10.5116/ijme.4dfb.8dfd

Totman, J., Hundt, G. L., Wearn, E., Paul, M., & Johnson, S. (2011). Factors affecting staff morale on inpatient mental health wards in England: a qualitative investigation. *BMC psychiatry, 11*(1), 1-10.

Van Wart, M. (2011). Changing dynamics of administrative leadership. *The State of Public Administration: Issues, Challenges, and Opportunities*, 89-107.

Walker, R., & Clark, J. J. (1999). Heading off boundary problems: Clinical supervision as risk management. *Psychiatric Services, 50*(11), 1435-1439.

Webster, C. D., Haque, Q., & Hucker, S. J. (2013). *Violence risk-assessment and management: Advances through structured professional judgement and sequential redirections*. John Wiley & Sons.

Weiss, D., & Marmar, C. (2004). The impact of event scale—revised In: JP Wilson, & TM Keane. *Assessing psychological trauma and PTSD*, 168-189.

Weiss, M., Kolbe, M., Grote, G., Spahn, D. R., & Grande, B. (2017). Why didn't you say something? Effects of after-event reviews on voice behaviour and hierarchy beliefs in multi-professional action teams. *European Journal of Work and Organizational Psychology, 26*(1), 66-80.

West, C. P., Dyrbye, L. N., Satele, D. V., Sloan, J. A., & Shanafelt, T. D. (2012). Concurrent validity of single-item measures of emotional exhaustion and depersonalization in burnout assessment. *Journal of general internal medicine, 27*(11), 1445-1452.

Westwood, S., Morison, L., Allt, J., & Holmes, N. (2017). Predictors of emotional exhaustion, disengagement and burnout among improving access to psychological therapies (IAPT) practitioners. *Journal of Mental Health, 26*(2), 172-179.

[Record #331 is using a reference type undefined in this output style.]

Wright, C., Burns, T., James, P., Billings, J., Johnson, S., Muijen, M., Priebe, S., Ryrie, I., Watts, J., & White, I. (2003). Assertive outreach teams in London: Models of operation: Pan-London Assertive Outreach Study, Part 1. *British Journal of Psychiatry, 183*(2), 132-138. https://doi.org/10.1192/bjp.183.2.132

Wu, L., Chen, Y., Liu, X., Fang, P., Feng, T., Sun, K., Ren, L., Liang, W., Lu, H., Lin, X., Li, Y., Wang, L., Li, C., Zhang, T., Ni, C., & Wu, S. (2022). The influence of job burnout on the attention ability of army soldiers and officers: Evidence from ERP. *Front Neurosci, 16*, 992537. https://doi.org/10.3389/fnins.2022.992537

Zaccaro, S. J., Rittman, A. L., & Marks, M. A. (2001). Team leadership. *The leadership quarterly, 12*(4), 451-483.

CHAPTER 3:

3.1 Press release for literature review:

A systematic literature review exploring the reliability of the subscales of the Professional Quality of Life (ProQOL) Scale

Background:

Professional quality of life (PQoL) refers to both pleasant and unpleasant emotions experienced by professionals who have a caring role and consists of burnout, compassion fatigue, and compassion satisfaction.

Burnout results from long periods of physical and emotional fatigue and can result in feeling frustrated, incapable and disinterested at work. Compassion fatigue is characterised by exhaustion, anger and irritability, and a reduced sense of satisfaction with work. Due to working with patients who experience trauma, frontline workers and nurses working in mental health settings, emergency departments and critical care services are at an increased risk of developing burnout and compassion fatigue.

However, this can be reduced with increased levels of compassion satisfaction which is reflected in positive emotions experienced when supporting patients with difficulties. It coincides with increased job satisfaction and is linked to better psychological well-being.

The most common way of measuring PQoL is using the Professional Quality of Life (ProQOL) Scale which is used across different healthcare settings healthcare and in different countries. However, little is known about the reliability of the ProQOL scale.

Therefore, this review aimed to explore this further, focusing on a Cronbach's Alpha value, the most reported form of reliability. This was done by completing a systematic literature search to explore the reported Cronbach's Alpha value for each of the three subscales by studies that have used different versions of the ProQOL and across different countries.

What did the review find?

The review found thirty-seven studies had reported a Cronbach's Alpha value for each of the three subscales. These studies used five different versions of the ProQOL as well as ten different language versions.

All studies were screened for risk of bias across six different criteria (selection, performance, detection, statistical, reporting and generalisability bias). The overall quality of the studies varied greatly, with most studies obtaining an overall risk quality score ranging from 45-55%. Only one study had a quality rating score of 100%.

Overall, when comparing the studies using statistical procedures generally known as "meta-analysis" the findings suggest that the burnout and compassion fatigue subscales for the ProQOL-21 were the only subscales that showed a significantly different reliability value compared to other versions of the ProQOL. However, these results involved only a small number of studies and therefore would be open to change following publication of further studies.

Additionally, significant differences in reliability were found for different language versions of the ProQOL, however, these only included a small number of studies. For instance, only two studies used the Chinese language version and only one study used the Persian language version which were found to have a statistically significant difference in reliability for the compassion fatigue and compassion satisfaction subscales, respectively.

What does this mean?

The lack of consistent statistical difference between the reliability of ProQoL versions offers little evidence to support the use of one version over any other. Additionally, although some significant differences were found between language versions, these only include a small number of studies and could easily change with the publication of new research.

Further research is needed to clarify reliability of the subscales for the ProQOL scale. However, it is recommended to use the latest version as this better reflects updated considerations to measuring PQoL.

3.2 Press release for empirical study:

Adult mental healthcare staff experiences of psychology-led supervision groups within Assertive Outreach Teams (AOT) and inpatient Rehabilitation teams in managing risk.

Background

Clinical supervision is defined as "the formal provision, by approved supervisors, of a relationship-based education and training that is work-focused and which manages, supports, develops and evaluates the work of colleague/s" (Milne, 2007). It aims to improve individual clinical skills, knowledge and competence, with enhanced quality of patient care.

There are many benefits of clinical supervision, including management of stress and pressure. Other benefits include increased job satisfaction, better risk management plans and overall better support for improving the quality of the service for patients.

Because of the benefits of clinical supervision, there is increased focus on making sure that all staff have access to this part of their working role. Violence, self-harm and suicide are more prevalent in mental health settings. This is important as some mental health teams such as Assertive Outreach Teams (AOT) and inpatient rehabilitation teams treat people with complex and severe mental health needs, with more than a quarter of AOT patients considered seriously at risk.

There is little known about how clinical supervision differs across these settings and the viewpoints of different disciplines of mental health care staff within these team in using supervision groups to manage risk.

What did the study do?

The study explored adult mental healthcare staff experiences of psychology-led supervision groups when managing risk. This was done by conducting online interviews with eight staff members from NHS AOT and inpatient rehabilitation teams, from different professional disciplines, via a secure video link.

What did the study find?

Three main themes came up that were important to the participants. Firstly, participants talked about how supervision sessions enhanced their sense of team belonging and cohesion by helping them think about risk systematically as a team. This included *Identifying plans and solutions*, *Seeking and providing support*, *Sharing ideas and perspectives and Dealing with difficulties and worries.*

Secondly, participants spoke about how it was important for them to ***manage the NHS hierarchy***. Superivison sessions were useful in helping junio staff share their views with senior staff.

Finally, participants identified the usefulness of team supervision groups in facilitating an experience of shared responsibility and accountability. All participants agreed that working as a team helped them manage feelings of stress and pressure, regardless of role, with shared team decisions instilling a sense of relief.

What does this mean?

The findings emphasise the importance of group supervisions when managing risk and providing a safe space for mental healthcare staff to share their difficulties. A team approach should be encouraged with all staff supported to share different ideas and perspectives, with facilitators perceived as crucial in supporting staff to identify solutions and specific plans for difficult situations or worries staff were experiencing.

APPENDICES

Appendix 1: Table summarising individual papers reviewed in the analysis

ProQOL Version	Study name	N	Alpha Coefficient (Burnout)	Alpha Coefficient (Compassion Fatigue)	Alpha Coefficient (Compassion Satisfaction)	Language	Country	Professional Group	Nature of Study
ProQOL-5	Alhalal 2020	255	0.78	DID NOT REPORT	0.89	Arabic	Saudi Arabia	Nurses	Nurses were recruited from 5 hospitals, 3 of which provide secondary and tertiary care and 2 which specialise in maternity and childcare. Acute services provided include intensive care, paediatrics, oncology, burns and emergency care.
ProQOL-5	das Neves Borges 2019	87	0.77	0.82	0.9	Portuguese	Portugal	Nurses	Nurses were recruited from an emergency and urgent care unit, using convenience sampling.
ProQOL-5	Duarte 2017	390	0.86	0.82	0.88	Portuguese	Portugal	Nurses	Nurses from 5 hospitals volunteered to take part and worked in different areas including oncology, surgery, internal medicine, paediatrics, palliative care, emergency, psychiatry, outpatients, family medicine, intensive care and continuity care.
ProQOL-5	Duarte 2017	298	0.74	0.74	0.86	English	Portugal	Nurses	Nurses were recruited from 5 hospitals using convenience sampling. They worked in a range of settings, although these were not specified.
ProQOL-5	Fu 2018	294	0.8	0.72	0.92	Chinese	Taiwan	Nurses	This study recruited only female nurses from one teaching hospital, using convenience sampling.
ProQOL-5	Hemsworth 2018 – Australian	273	0.8	0.82	0.9	English	Australia	Nurses	Nurses were recruited from one tertiary acute hospital. Data used in this study was obtained as part of a larger study exploring well-being of Australian nurses.
ProQOL-5	Hemsworth 2018 – Canadian	303	0.75	0.85	0.91	English	Canada	Nurses	All nurses from one large health centre were invited to take part. They were given the option to complete the questionnaire online or a hard copy. They were given a $5 card which could be used in the cafeteria as a reward for participating.
ProQOL-5	Hemsworth 2018 – Canadian	503	0.74	0.78	0.89	English	Canada	Palliative Care-Workers	Palliative care-workers were contacted via e-mail using the British Columbia Hospice Palliative Care Association and the provincial branch of the Canadian Hospice and Palliative Care Association.

	Palliative Nurses								
ProQOL-5	Hunsaker 2015	278	0.82	0.79	0.92	English	USA	Nurses	Nurses who worked in various emergency departments were recruited using purposive sampling. Contact details of nurses were obtained from the Emergency Nurses Association.
ProQOL-5	Jang 2016	285	0.81	0.8	0.9	Korean	Korea	Nurses	Nurses were recruited from 2 tertiary hospitals and worked with cancer patients at oncology wards, outpatient clinics of internal medicine and chemotherapy clinics. Participants voluntarily agreed to take part in this study.
ProQOL-5	Kim 2017	875	0.73	0.72	0.89	English	Korea	Nurses	Nurses were recruited from a large tertiary hospital using convenience sampling. Minimum age for inclusion was 20years.
ProQOL-5	Lemieux-Cumberlege 2019	112	0.74	0.81	0.9	English	Scotland/ United Kingdom	Homelessness Practitioners	Participants included frontline workers who worked with homeless or vulnerably housed people, including drop-in centres, health services, social care and third sector organisations.
ProQOL-5	Michalec 2013	416	0.48	0.82	0.9	English	USA	Nurses	First, second, third- and fourth-year undergraduate nursing students from a university for nursing students were invited to complete questionnaires. Third and fourth year students were also invited to take part in interviews further exploring burnout and/or compassion fatigue.
ProQOL-5	Ravi 2016	155	0.6	0.73	0.84	English	India	Audiologists/ Speech & Language Pathologists	Participants were professional Audiologists and Speech and Language Pathologists registered under the Indian Speech and Hearing Association. 500 potential participants were randomly contacted, with 155 taking part in the study. They were employed in both academic and clinical settings such as schools, hospitals and private practice including hearing aid dispensing.
ProQOL-5	Salloum 2018	177	0.8	0.85	0.89	English	USA	Child Welfare Workers	Participants were child welfare workers from five private case management agencies that work with children and families referred due to allegations of abuse or neglect.
ProQOL	Ang 2018	1667	0.7	0.8	0.92	English	Singapore/ Canada	Nurses	Nurses from Singapore were recruited from 2 academic medical centres and completed an online survey. Nurses from a large regional health centre in Canada were given the option to complete an online survey or a hard copy. They were also given a $5 card to use in the cafeteria as a reward for participating.
ProQOL	Begic 2019	27	0.8	0.9	DID NOT REPORT	English	USA	Home Visitors	Participants made up the complete number of home visitors within that geographical location. Sampling was not used due to the small

									sample of home visitors within the geographical area. All participants completed 2 structured interviews and 2 quantitative surveys, 6 months apart.
ProQOL	Chang 2011	102	0.65	0.68	0.82	Chinese	China	Soldiers	Participants were Chinese soldiers who were dispatched as front-line rescuers and emergency relief workers after a big earthquake had occurred.
ProQOL	Erkorkmaz 2018	131	0.52	0.71	0.85	English	Turkey	Nurses	Nurses from a Training and Research hospital volunteered to take part in the study.
ProQOL	Galiana 2017 – Spanish	385	0.53	0.78	0.77	Spanish	Spain	Palliative Care Professionals	Palliative Care Professionals including doctors, nurses, psychologists, nursing assistants and social workers volunteered to complete an anonymous online survey. They worked within a range of palliative care settings including in hospitals, at home, social health care centre, hospice and paediatric palliative care.
ProQOL	Galiana 2017 – Brazilian	161	0.65	0.77	0.86	Brazilian	Brazil	Palliative Care Professionals	Palliative Care Professionals including doctors, nurses, psychologists, nursing assistants and social workers volunteered to complete an anonymous online survey. They worked within a range of palliative care settings including in hospitals, at home, oncology unit, intensive treatment unit and paediatric palliative care.
ProQOL	Galiana 2019 - Spanish	385	0.54	0.78	0.77	Spanish	Spain	Palliative Care Professionals	Palliative Care Professionals were identified through the Spanish Society for Palliative Care and the Pallium Latin-American Institute. Participants included doctors, nurses, psychologists, nursing assistants and social workers who volunteered to participate.
ProQOL	Galiana 2019 - Argentinian	273	0.65	0.77	0.86	Argentinian	Argentina	Palliative Care Professionals	Palliative Care Professionals were identified through the Argentinian Association for Palliative Medicine and Care. Participants included doctors, psychologists, nurses, social workers, occupational therapists and nursing assistants who volunteered to participate.
ProQOL	Kagan 2019	494	0.64	0.77	0.87	English	Israel	Social Workers	Social Workers included in this study worked in Welfare, Healthcare, Community and Correctional settings. The sample comprised of mainly females (88.1%) who volunteered to take part.
ProQOL	Lee 2016	680	0.92	0.91	0.96	English	South Korea	Nurses	Nurses in this study were recruited from 3 tertiary hospitals using convenience sampling and worked in Intensive Care, Emergency Care and General Wards.
ProQOL	Leners 2014	168	0.77	0.85	0.9	English	USA	Military Healthcare Professionals	Participants in this study included doctors, nurses and advanced practice nurses working in the Navy, Air Force, Army and Marines who were identified at a convention of the American Military

									Surgeons of the United States (AMSUS) and volunteered to participate.
ProQOL	Salimi 2020	400	0.78	0.79	0.83	Persian	Iran	Nurses	Nurses working in critical care units across 8 educational hospitals volunteered to participate. Majority (92.2%) of participants were female and worked in intensive care, coronary care, neonatal intensive care and paediatric intensive care.
ProQOL	Samson 2016	377	0.69	0.82	0.87	Hebrew	Israel	Palliative Care Professionals	Participants included doctors, nurses and social workers working in palliative care settings, specifically home and hospital-based hospice units. A control group included primary healthcare providers.
ProQOL	Sanford 2018	64	0.767	0.649	0.87	English	USA	Peers or Professional	This study included facilitators of groups for survivors of suicide loss who were either peers (i.e. a suicide loss survivor), professionals with backgrounds in mental health, including Psychology, Counselling, Social Work and Psychiatric Nursing or a Peer Professional.
ProQOL	Storm 2021	52	0.824	0.734	0.897	English	USA	Nurses	Convenience sampling was used to recruit critical care nurses from eight intensive care units and four step-down units of three selected medical centres in Philadelphia.
ProQOL	Xu 2019	61	0.79	0.79	0.88	English	USA	Social Workers	Social Workers were randomly selected from the state Board of Social Work Examiners' address list. They were working in various areas including administration/management, community/organising/advocacy and training/consultation.
ProQOL Revised	Itzick 2018	501	0.63	0.76	0.88	English	Israel	Social Workers	The head of social services at the Ministry of Welfare and the Ministry of Health were asked to distribute the questionnaires to managers in different organizations, in which social workers practice. It is not specified which areas participating social workers worked in but majority (88.1%) were female.
ProQOL Revised	Choi 2017	358	0.76	0.82	0.92	Korean	Korea	Nurses	Nurses were recruited form 3 hospitals using convenience sampling to explore the impact of experiencing workplace violence, specifically verbal abuse, physical threats and physical violence from patients, parents/families, doctors, nurses or other. They areas/settings nurses worked in was not specified.
ProQOL R-IV	Elkonin 2011	30	0.69	0.8	0.89	English	South Africa	Nurses	Nurses were recruited using convenience sampling from 3 privately owned intensive care units, with majority (93.3%) of participants being female.
ProQOL R-IV	Goshen General	106	0.7	0.75	0.87	English	USA	Nurses	Nurses in this study were recruited using convenience sampling and worked in different areas including home care, emergency

	Hospital 2010								department, intensive care, progressive care, oncology and medical-surgical units. 71 participants also provided qualitative data exploring trigger situations and coping strategies.
ProQOL R-IV	Yildirim 2021	697	0.728	0.861	0.907	English	Turkey	Nurses	Nurses were recruited from tertiary university hospitals using convenience sampling. Majority of participants (91.1%) were female and worked in different areas including surgical department, internal medicines department, daily treatment services, intensive care unit and emergency services.
ProQOL-21	Heritage 2018	1615	0.8	0.84	0.9	English	Australia	Nurses	Nurses in this study were employed in hospitals, both in the private and public sector, including some who worked in the aged care sector.

Appendix 2: Risk of bias criteria

Domain	Details	Risk of Bias
Selection bias	Selection bias occurs when there is a systematic difference between the characteristics of those selected for the study and those who are not. Have the selection method and characteristics of participants been described adequately?	High Risk – No method of how participants were selected, nor characteristics of participants are described. Unclear Risk – The characteristics of the study population are not clearly or fully reported. This includes age range, education years, socioeconomic status, ethnicity, where participants were recruited from and how they were recruited. Low Risk - The characteristics of the study population are clearly described and without evidence of bias.
Performance Bias	Performance bias refers to systematic differences between/within groups in the participants motivation to complete the study.	High Risk – The study does not report levels of confidentiality and anonymity. It is not clear if participants were rewarded for their participation (e.g., motivation to respond in a certain way). It is unclear how much information was provided to the participant prior to taking part in the study Unclear Risk – The study does not report levels of confidentiality and anonymity. It is not clear if participants were rewarded for their participation (e.g., motivation to respond in a certain way). It is unclear how much information was provided to the participant prior to taking part in the study Low Risk - Study reports level of confidentiality and anonymity. Participants were not rewarded for their participation in the study. Information and procedures are provided in a way that does not differentially motivated participants
Detection bias	Was the ProQOL scale delivered in its original format? Was the scoring of the test completed as per the author's recommendations?	High Risk – Major alterations to the test, including wording and/or scoring matrix. Combined with or amalgamated with a different test. Unclear Risk – Minor changes made to the wording of questions; changes made to the scoring; changes to questions due to translation. Reporting only a subset of test scores. Low Risk - Test administered in its original or agreed format and scored following the recommended matrix.
Statistical bias	Bias resulting from the (inappropriate) statistical treatment of the data.	High Risk – Study does not report a Cronbach's Alpha value. A variation or alternative value is provided in place of a Cronbach's Alpha from which an estimate of Cronbach's Alpha cannot be derived. Unclear risk - Study does not report a Cronbach's Alpha value. A variation or alternative value is provided in place of a Cronbach's Alpha from which an estimate of Cronbach's Alpha can be derived Low Risk - Analysis using Cronbach alpha
Reporting bias	Is there evidence of selective outcome reporting? Are there measures that have not been reported in the results that have been mentioned in the method section?	High risk - Reported only a subsample of results and/or only significant results. Did not report on entire sample. Data does not appear to be accurately reported (e.g., final values are suspect, or data is reported in a manner requiring reconstruction or transformation). Unclear risk - Did not report Cronbach's Alpha for *all of the subtests* of the ProQOL scale. Low risk - Full sample size reported. Reported results for all subtests of the ProQOL scale.
Generalisability	Can the results be applied to other populations groups or settings based on the sample used?	High risk – Small sample with or without idiosyncratic features (<20 per group). Unclear risk - Sufficient sample for generalisation but with some idiosyncratic feature (> 20 per group). Sample taken from only one population group (i.e., students) with attempts to generalise to entire population. Low risk- Sufficient sample for generalisation and representative of target population (>20 per group)

Appendix 3: Health Research Authority (HRA) ethical approval documentation

Ymchwil Iechyd a Gofal Cymru
Health and Care Research Wales

NHS
Health Research Authority

[redacted]

Email: approvals@hra.nhs.uk
HCRW.approvals@wales.nhs.uk

10 August 2021

Dear Miss [redacted]

HRA and Health and Care Research Wales (HCRW) Approval Letter

Study title:	**Adult healthcare staff experiences of psychology-led supervision groups within Assertive Outreach Teams (AOT) and inpatient Rehabilitation teams in managing risk**
IRAS project ID:	**292171**
Protocol number:	**ERN_21-0357**
Sponsor	**University of Birmingham**

I am pleased to confirm that **HRA and Health and Care Research Wales (HCRW) Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications received. You should not expect to receive anything further relating to this application.

Please now work with participating NHS organisations to confirm capacity and capability, in line with the instructions provided in the "Information to support study set up" section towards the end of this letter.

How should I work with participating NHS/HSC organisations in Northern Ireland and Scotland?
HRA and HCRW Approval does not apply to NHS/HSC organisations within Northern Ireland and Scotland.

If you indicated in your IRAS form that you do have participating organisations in either of these devolved administrations, the final document set and the study wide governance report

Appendix 4: University of Birmingham sponsorship and favourable opinion letter

UNIVERSITY OF BIRMINGHAM
FINANCE OFFICE

[redacted]
School of Psychology
University of Birmingham

Tuesday, 6 July 2021

[redacted]

Project Title: **Healthcare staff experiences of psychology lead supervision groups within Assertive Outreach Teams (AOT) and inpatient Rehabilitation teams in managing risk**
IRAS ID: **292171**
Sponsor Reference: **RG_21-037**
UoB Ethics Reference: **ERN_21-0357**

Under the requirements of UK Policy Framework for Health and Social Care Research, the University of Birmingham agrees to act as Sponsor for this project. Sponsorship is subject to you obtaining a favourable ethical opinion, HRA approval and NHS R&D management approval where appropriate.

As Chief Investigator, you must ensure that local study recruitment does not commence until all applicable approvals have been obtained. Where a study is or becomes multi-site you are responsible for ensuring that recruitment at external sites does not commence until local approvals have been obtained.

Following receipt of all relevant approvals, you should ensure that any subsequent amendments are notified to the Sponsor, REC, HRA and relevant NHS R&D Office(s), and that an annual progress report is submitted to the Sponsor, REC and NHS R&D departments where requested.

Please ensure you are familiar with the University of Birmingham Code of Practice for Research (http://www.birmingham.ac.uk/Documents/university/legal/research.pdf) and any appropriate College or School guidelines.

Finally please contact researchgovernance@contacts.bham.ac.uk should you have any queries.

You may show this letter to external organisations.

Yours sincerely

[redacted]
Head of Research Governance and Integrity

Appendix 5: Structured Meta-analysis Plan

Centre for Applied Psychology Meta-Analysis Strategy

Transformation of effects for calculations and back transformation for presentation

The event rates and relative risk estimates in primary study were log transformed prior to numerical synthesis however, unless otherwise indicated, the values presented in tables and figures have been back-transformed to their original format for clarity of presentation.
For meta-analysis of alpha coefficient then the transformation reported by Bonett (2002, 2010) is most commonly used to correct for issues relating to normalisation and variance stabilisation.

The omnibus test

The omnibus test can be calculated using either the fixed effects or the random effects models. Under the fixed-effect model we assume that the true effect size for all studies is identical, and the only reason the effect size varies between studies is sampling error (error in estimating the effect size). Therefore, when assigning weights to the different studies we can largely ignore the information in the smaller studies since we have better information about the same effect size in the larger studies. It makes sense to use the fixed-effect model if two conditions are met. First, we believe that all the studies included in the analysis are functionally identical (i.e., all studies have a uniformly excellent methodology). Secondly, our goal is to compute the common effect size for the identified population, and not to generalize to other populations. In point of fact, effects in psychological studies are likely to vary as a result of a number of uncontrolled factors (e.g., the distribution methodological weakness across studies, uncontrolled moderators, natural variation in the effect that is being measured).

In contrast, under the random-effects model the goal is not to estimate one true effect but to estimate the mean of a distribution of possible effects (which may show true variation due to the idiosyncratic characteristics of the individual or the unique circumstances of the intervention or exposure). Since each study provides information about a different effect size,

we want to be sure that all these effect sizes are represented in the summary estimate. This means that we cannot discount a small study by giving it a very small weight (the way we would in a fixed-effect analysis). The estimate provided by that study may be imprecise, but it is information about an effect that no other study has estimated. By the same logic we cannot give too much weight to a very large study (the way we might in a fixed-effect analysis). Our goal is to estimate the mean effect in a range of studies, and we do not want that overall estimate to be overly influenced by any one of them. When the researcher has gathered data from studies that had been undertaken by researchers operating independently (and will therefore show different methodological strengths and weaknesses), it would be unlikely that all the studies are functionally equivalent. Typically, the participants and/or interventions in these studies would have differed in ways that would have impacted on the results, and therefore we should not assume a common effect size. Therefore, in in the case of the current review the random-effects model is more easily justified than the fixed-effect model.

The DerSimonian and Laird method is the simplest and most commonly used method for calculating the between studies variation (tau) for fitting the random effects model.

Handling problematic variance

Defining problematic variance

An effect is considered heterogeneous if it presents with variation from the meta-analysis synthesis that cannot be attributed to true variation in the distribution population effect. Heterogeneity can result from methodological variation in the studies, measurement error or uncontrolled individual difference factors within the body of literature. Higgins I^2 is a commonly used measure of heterogeneity, with greater values of I^2 indicating variation in effect that cannot be attributed to true variation in the distribution of effect in the population. As there is considerable variation in methodologies of the primary studies that was used to calculate the meta-analytic synthesis, problematic heterogeneity was defined as a Higgins I^2 value greater than 75%.

Estimation of unexplained variance due to methodological factors and uncontrolled covariates

If problematic heterogeneity is observed then a leave-one-out analysis will be conducted to identify primary studies that exert a disproportionately influential effect on the meta-analytic synthesis. Any such study will be reviewed with regard to the possibility of exclusion due to risk of bias.

In addition, subgroup analyses and meta regression will be used to attempt to identify the source or sources of problematic heterogeneity and the attenuated estimate of the synthesis will be reported.

The quality effects model

In the random effects model the precision of an effect is usually estimated as a function of the sample size from which the effect is derived. The quality effects model (Doi & Thalib, 2008) extends the random effects model by explicitly including rating of methodological quality in addition to the size of the sample in the estimation of precision. In this review the quality effects model was calculated using the total score from the risk of bias ratings reported in section XX. The quality effects model can be interpreted as the meta-analytic synthesise that would have been obtained had all of the studies been of the same methodological quality as the best study in the review. Accordingly, the quality effects model provides a measure of attrition attributable to methodological variation.

Identifying Influential Studies

To examine whether any particular study or studies are exerting a disproportionately high influence on the overall meta-analytic effect, a “one left out” procedure was conducted. This procedure identifies individual studies with a disproportionate influence on the quantitative synthesis, by observing the impact of removing each study in turn. If omitting a study results in an effect that lies outside of the 95% CI for the complete meta-analysis then that study is deemed to have a disproportionate influence and is remove from the omnibus test.

Identifying Publication Bias and Small Study Effects

For outcomes with a sufficient number of primary studies, publication bias and small study effects will be identified through visual and statistical inspection of the funnel plot. A funnel plot is a scatterplot of the effects from against a measure of study precision. It is used primarily as a visual aid for detecting systematic heterogeneity.

In the absence of publication bias, it is assumed that studies with high precision will be plotted near the average (i.e., the meta analytic synthesis), and studies with low precision will be spread evenly on both sides of the average, creating a roughly funnel-shaped distribution where the distance from the average is inversely proportionate to the precision of the study. A symmetric inverted funnel shape arises from a 'well-behaved' data set, in which publication bias is unlikely whereas deviation from this shape can indicate publication bias especially if there is an absence of studies in the region associated with small samples sizes and non-significant effects.

If publication bias is identified, then a trim and fill procedure (Duval & Tweedle, 2000a; Duval & Tweedle, 2000b) will be undertaken. The trim and fill procedure builds on the assumption that publication bias would lead to an asymmetrical funnel plot. Trim and fill procedure uses an iterative algorythm to remove the most extreme small studies from the side of the funnel plot associated with positive effects, re-computing the effect size at each iteration until the funnel plot is symmetric about the (corrected) effect size. In theory, this will yield an unbiased estimate of the effect size. While this trimming yields the adjusted effect size, it also reduces the variance of the effects, yielding a too narrow confidence interval. Therefore, the algorithm then adds the original studies back into the analysis, and imputes a mirror image for each on the side of the funnel plot associated with negative effects.

In addition, the fail-safe N will also be calculated (Rosenthal, 1979). The fail-safe N is an estimation of the number of missing studies that would need to be retrieved for the effect to be no longer significant. If this number is large (relative to the number of primary studies in the meta-analysis) then the omnibus test can be considered robust to the effects of publication bias.

Planned Contrasts

Where specific a priori hypothesis made been posited, then sub-group analysis will be conducted for categorical moderators and meta-regression will be calculated for continuous moderators.

Analysis of Sub-groups

Where categorical moderators are considered then summary effects and associated heterogeneity measures will be calculated for each of the sub-groups. The significance of the difference between the sub-groups will be evaluated by comparison of their 95% confidence intervals.

Potential moderators of the effect will be explored using a series of subgroup analyses. The significance of sub-group differences will be evaluated using the Q statistic, which may be view as an extension of analysis of variance. The Q statistic is calculated by summing the within-studies variation (the weighted sum of squares of all of the studies within a subgroup about the mean of the subgroup) across all subgroups and the subtracting this from the total variance (i.e., the weighted sum of squares between all of the studies and the overall grand mean). The resulting Q statistic therefore represents the ., the weighted sum of squares attributable to between studies variation and conforms to a chi-squared distribution (Borenstein, 2009). A 95% confidence interval for each subgroup will be used to determine the significance of the pairwise differences between the sub-groups.

Meta-Regression

Meta-regression is similar to simple regression, in that the effects of the primary studies are predicted according to the values of one or more explanatory variables. However, larger studies have more influence on the relationship than smaller studies, since studies are weighted by the precision of their respective effect estimate. The explanatory variables are typically characteristics of studies or participants that might influence the size of intervention effect.

Appendix 6:

Interview Topic Guide

Adult mental healthcare staff experiences of team supervision sessions within Assertive Outreach Teams (AOT) and inpatient Rehabilitation teams for managing risk
Version: 2.0
IRAS ID: 292171
07.05.2021

Opening questions:

1. What is your job role?

   What does this involve?

   Which team do you work in?

   How long have you worked in this team?

Interview question:

1. **Can you tell me about the team supervision sessions you have in your team (the ones led by psychology)?**

   Potential prompts:

   - What format do they take?
   - What do you understand these groups are for?
   - If you had to explain to someone what these supervision groups are like, how would you describe it? What would you say?

2. **What do these supervision sessions involve?**

   Potential prompts:

   - How frequently do you have these supervision sessions?
   - How long do they last?
   - Who facilitates them?
   - Has it always been this person or has this varied? Does this make a difference? How? In what ways?
   - Who attends? How is this decided upon? (Is there a choice?)
   - Does attendance make any difference to the group at all? If so, what?
   - What sort of things are spoken about or discussed? How is this decided?
   - Is there a format or specific framework used to guide these supervision groups?

3. **Moving now to think about risk. How do you use these supervision sessions in managing risk?**

   Potential prompts:

   - How do you decide what is discussed?
   - What sorts of risk issues are discussed?
   - What sorts of risks are spoken about more often?
   - What sorts of risks are spoken about less often?
   - What factors do you think may influence the types of risks that are discussed in the group?
   - What do you hope to get out of these supervision groups?
   - To what extent are these met? How are these met / not met?
   - To what extent and in what way do you participate in these groups?
   - Is there anything else you would like to mention here?

4. **What effects (if any) do you think the supervision groups have when managing risk?**

   Key areas to consider:

   - On clinical practice (e.g. dynamics and relationships with other staff and colleagues or work with clients)
   - Knowledge and ability to manage or deal with risk? (think about the types and levels of risk situations, clients or incidences that are discussed)
   - Your levels of stress regarding dealing with risk?
   - Your confidence in managing risk?
   - Can you give me an example of applying things discussed/reflected on in group supervisions within your clinical practice, related to managing risk**?**

5. **Thinking more about communication now**

   Key areas to consider:

   - How do you communicate outcomes to the team, considering risk factors? Particularly to those staff members that were not present for the session (thinking about shift patterns)?

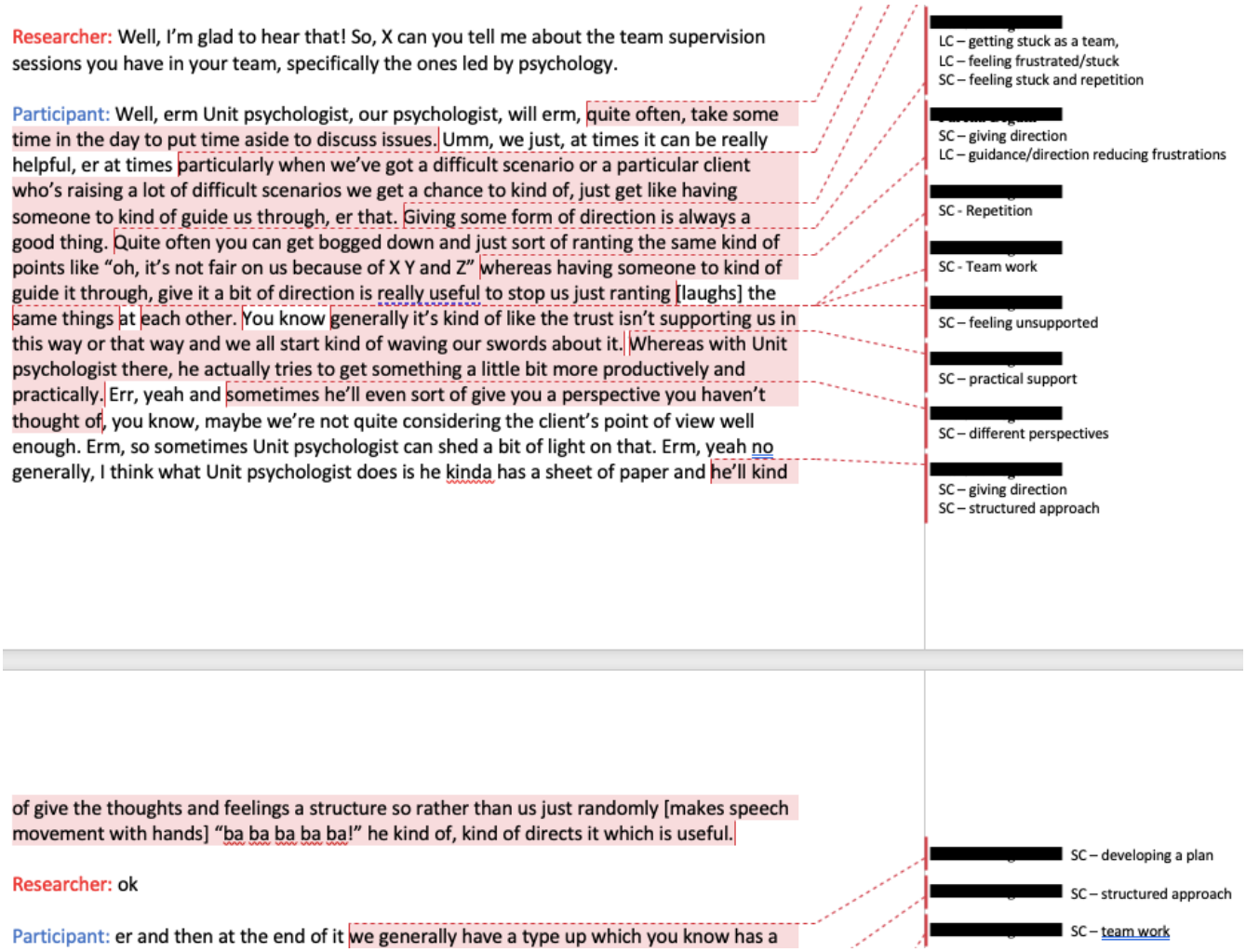
- How do you review the outcomes of what has been discussed? Does this impact your ability to manage risk? If yes, how?

- What, if any, are the barriers to accessing these supervision groups? How have these been managed? What has helped facilitate access to them?
- If you could change anything about these supervision groups in terms of managing risk, what would it be and why?

6. **Is there anything else you would like to tell me about your experiences of taking part in the team supervision sessions regarding managing risk?**

Thank you for taking part in this interview.

Appendix 7 - transcript excerpt demonstrating semantic and latent coding



Appendix 8: Table describing theme development

Themes	Interview							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1 The function and value of the group sessions								
1.1. Identifying plans and solutions	"Try to work out a plan err in what way we can co-ordinate their care so that the patient could err benefit" "Try to figure out what level of risk err the patient poses and erm how to medicate them. How to avoid it, what level of contact maintained with the patient. So it's also an opportunity to review the care plan that we have for the patient."	"We need plans above and beyond what's written on section papers and leave prescriptions" "People want to discuss and hear the plans for a particularly risky person" "With a bit of discussion we were at least able to know what we shouldn't be doing, which erm was yeah vital"	"It does kind of give you insight into kind of the static factors and the dynamic factors. Erm, and should hopefully generate a sort of sensible management plan" "Having everybody in a meeting that is well organised and has definite kind of formulation, and comes up with definite plans, that means that everybody ends up hopefully doing the same thing, taking	"You know using multipronged approach to look at the issue, really, and erm how best to manage it" "It's helpful because obviously you help to manage the patient in the long run" "And it's not just down to you, the medic to you know, come up with	"We'll kind of summarise what's being discussed as well. Erm, and then that will kind of lead to the action plan" "It kind of allows us to analyse that situation, think about whether we'd actually be able to deal with that or whether we need to think about alternative placements for them" "Another area that often comes up	"We came up with a plan in terms of how we were going to manage that for the best for them" "Talking through with the team to try and come to some sort of solution or action plan" "The group came up with a plan in terms of how we can manage both parties and erm we contacted	"We would go in and we would discuss around how we can maybe come up with a plan of the next step in that patient's care or how best to approach something and who might do it" "We agreed, we would therefore, monitor the front door, so then we had kind of erm, a casual rota… and that helped	"I think last week they might have had one where they discussed someone's, [relapse prevention plan] so like creating that plan, so yeah. Everyone sort of feeds into it" "I think we're quite good at like discussing okay what are we gonna do, if like you know, he's expressing these thoughts, what should we do now"

	“It’s an opportunity to pause a clock and reflect, focus, erm on a particular patient. And er, what we expect to come out of it, update the plan of action er in what we we going to move forward to try to support the patient” “It gives us a fair framework, a fair idea, err how the care plan, what should the care plan reflect.” “So as a team, we work out the solution”	“I almost feel comforted that there’s a firm plan that covers different eventualities” “I felt we came out with a plan or do come out with a plan ones whereas with this other chap’s, we, it’s very relaxing but err, I didn’t quite feel like we had an end result from” “I really like things that enable me to erm, know practically how to approach very specific situations” “Some people didn’t like that plan at all. Erm, but it was a team decision”	the same approach with the patient” “It’s hopefully easier to deliver because everybody knows what is expected, what the plan is"	a solution. Everybody would put their heads together and yeah, it works well”	around erm substance misuse as well, risks around that and how we might manage that. Erm, do we look at that service users leave, do we need to change that service users leave, how can we minimise the risks around that.” “I suppose it allows us to think about recommendations around how we can reduce risk, how we can reflect as a team, it, I suppose it enables us to have discussions around whether this service user is suitable to be here so we kind of look at the environment of team a and you know the, the potential for us	safeguarding, spoke about things with safeguarding, and we contacted various people.”	manage him, we were all on the same page” “If you're all on the same page, and you’re, you can then develop a kind of consistent approach, so you haven't got somebody accidentally undermining what, what works or, and the patient also knows then eventually what to expect, how our approach is going to be.” “You can say what you feel and it's like a problem-solving group.” “You can go away and think oh, actually if I do this better, we agreed this	

					dealing with aggressive behaviour"		and you've got a kind of a team backing and that feels quite nice."	
1.2 Seeking and providing support	"It's a chance to provide supervision where we can all support each other" "It's a way where people don't feel isolated erm, in the sense of we have to carry the can on their own." "Work out a way how we can best support that individual so that they can function to the best of abilities in the community while err monitoring the risk that they potentially could be posing"	"If a decision is worrying me, if I have a particular chap, usually I get to run it past someone. Not always but during the week. So generally, you get support" "To go form that environment, where you have no support, no kind of team discussion, no open forum, to have that, I'm just infinitely grateful really" "Usually there is one thing most people want to talk about. And usually it doesn't even need that much discussion because most people are like yeah, fair	"It can be about just how the, the unit is running, what the sort of erm dynamics are like, erm, kind of on, on the team, how we maintain consistency and erm, provide support to staff as well because it can be quite draining, err, working with some of the patients that we have considering erm, the slow progress that we make." "Hopefully kind of gets everybody feeling supported and feeling part of the team and their contribution valued" "Just working in that sort of silo with your own patients and not really talking about	"We sort of talk about, how to go about reducing the risk and how to you know, help them support their physical health better." "We're able to come up with ideas on how to help each other without, you know, anybody feeling like they're being judged or, you know, or they're not doing what they're supposed to be doing or they're complaining or	"It's my resource that I use to get the support that I need" "We can think about as a team, you know who we can bring into those discussions and is there any special service we can think about to increase our knowledge of best supporting that person" "I think it's around support and getting a team perspective on things"	"I think it makes everybody a bit more cohesive" "I'm sure it reduces stress quite, as I say, the more sort of cohesive a team that you've got, the better it is." "We try and sort of make them welcome and say, you know, there's no pressure on you, you don't have to talk but if you just come in you might find there's something you've got to say" "Everybody is on the same	"You can normally deal with most things if you're supporting one another" "Mostly it's to do with we pick a patient and then, and then discuss how we're coping with that particular patient and how we can best support that patient, and staff to be honest." "It could be a few other people then turn around and say oh actually I'm having a tough time. So you don't feel so	"So yeah, I think it's more of like the support of the team that that helps me deal with the stress of dealing with risk."

		enough, that is like the thing we should talk about at the moment or that is the person we should concentrate on”	them much to other people, is really stressful. Err, but you don't get that sort of in AOT and it's even less if you're having good regular supervision” “By virtue of trying to involve everybody, so everybody feels included, that hopefully you know sets things up for everybody to kind of, whatever plan you come up with.”	anything like that” “I would say it brings the team together in a way”		page, working towards the same sort of ideals and goals.”	isolated, which is nice.” “I feel comfortable going in that group. Erm, and you can, no one’s gonna judge you. You can say what you feel” “It feels like quite a nice space to share it in because the unit psychologist is very welcoming. You don't feel like you're saying something daft” “Sometimes it's like a majority rule so if someone you know, like says I don't think we need to talk about that person, what about this	

							person and someone else might say yeah actually that's a good point” “If you're all erm, you're all in agreement on that, then I think that, that's better care. It’s continuity of care, isn’t it? You know, you’re going to get a better outcome.”	
1.3 Sharing ideas and perspectives	“I think it does, it does, in the sense of we try to be creative. Obviously as I said earlier, different agencies are involved and try to work out ways erm, erm, other creative ways in what we, we can manage the situation better.”	“The fact that this can all be you know a big melting pot of ideas and feelings, particularly for Dr X to hear, that’s really useful” “This is the only time we’re actually sensibly kind of, more just one person to another sharing our feelings and plans”	“If you compare that to one person thinking what's the best thing to do, is so much better having lots of people contribute. Cos, cos you could just get a whole range of ideas” “People will attend because I think they value it and it is useful, particularly when people are	“What we do is to come together, put our heads together and try to come up with a better way of dealing with the situation” “Mostly we come up with brilliant ideas and they are helpful.	“We can all come up with different reasons why and ideas why someone might not be doing something.” “We have a lot of nurses who all work together, erm, so they can kind of bounce ideas off of each other”	“You get so many different ideas and you've talked it through, we go through the pros and cons of anything” “We wouldn't have considered that, I don't think, or even sort of known much about how to go about	“Because of the psychology input, erm, and the psychiatrist input, you can kind of thrash out some ideas a bit better, and then we can put it into practice” “You kind of come out sometimes renewed with kind of a better	“I think we're quite good at bouncing ideas off each other” “It just gets different ideas thrown in so erm, yeah. Erm, so then if there's something I haven't thought about and I go to the ward review, it's useful for that.”

	“You feel reassured that you haven’t erm, you haven’t missed out on exploring different avenues that you possibly haven’t thought through while you were absorbed in it “Everybody providing their input from different err disciplines” “A psychologist there, err gives erm a also a psychological perspective of a patient” “So as a team, we work out the solution, err everybody providing their input from different err disciplines”	“Sometimes he’ll even sort of give you a perspective you haven’t thought of, you know, maybe we’re not quite considering the client’s point of view well enough”	feeling burnt out with a particular patient and they want to kind of contribute, trying to do something differently, or better” “Just that kind of open forum, erm, just does generate things that I wouldn’t have thought of” “You get everybody's perspective, everybody can contribute, is encouraged to contribute. Erm, and that's really helpful because you just get everybody’s views and the more people who are thinking about a, a case, the better really” “For an individual patient there will always be things that I don't know. And, erm, and having	Usually practical ideas we come up with erm that we hadn’t thought about before the supervision” “Yeah, definitely, to come up with ideas that one person wouldn't usually come up with so yeah it's definitely useful” “You tend to have erm, a wider view, other people’s perspectives, bringing in other ideas and erm, you kind of gain, you know, that kind of knowledge of how to you know, widen your	“It's useful to bring it to the team and discuss that altogether because you get different perspectives from different members of mdt as well” “It's opened my eyes to different tools we can use to manage risk as well.” “Different professionals coming forward, and you know, kind of showing their own perspectives and bringing forward their own skills and knowledge is really useful” “Offers that opportunity to discuss that as a team because we can all come up with different reasons why and	doing that, if we hadn't done the supervision group.” “It's just a given time for everybody just to give an opinion if they want or have one, and just give different views” “I don't sit out with the patients all the time so I don't get the same sort of conversations with them. I don't see the same things” “It's really good to have everybody’s view. Sometimes it's quite surprising the difference that you see”	idea of how you’re going to approach something or it might be something that you definitely didn't think of” “Just having somebody say well I would remove yourself, or you could try this, you could try that, you go away with a kind of a fresher feeling” “He's more than happy, as is the consultant actually, to get your perspective on how to manage patients they might not have seen or observed.” “Yeah, different approaches because I might be doing	“We kind of just like brainstorm, I suppose, see what, if anyone’s got any extra ideas or, which I think is quite good” “You feel like if something comes up then you can always bounce ideas, like, which I tend to do anyway and that's something I've always done” “We get students involved as well so I think it's useful sometimes to have like an outsider’s perspective or someone that's just come to the service.” “I think just seeing other people’s perspective. Erm, I think as a nurse and someone in charge a lot, like it is good to get feedback of what

			different staff insights is important. You know the more people who are looking at a patient and thinking about them, the more insight you are going to get" "Everybody regardless of their role erm, has some insight or some contribution." "We also have quite a lot of staff members, erm, who contribute to the multidisciplinary team, so it's, it's really important to have the opportunity to kind of get together and think about the patients"	perspective really"	ideas why someone might not be doing something. There might be things that we're missing so it's useful to get together as a team"		something that is exacerbating something and I'm not sure about how to tweak my approach, or I actually think my approach is quite good, and then someone else might think well actually it's rubbish and vice versa. So we can kind of talk about that"	other people are thinking" "Those kind of discussions can help with leading into ward review discussions because, you know it's only like two members of the team and then the doctor sometimes so it's good to, to have everyone's opinion, definitely"
1.4 Dealing with difficulties and worries	"We talk about unconditional positive regards, err, in nursing but sometimes erm when you, when	"At times it can be really helpful, er at times particularly when we've got a difficult scenario or a particular	"We have these complex patients with quite a lot of challenges"	"We discuss cases where we're struggling with to manage, either risks or	"As an OT I think the supervision sessions are particularly beneficial because I don't have any	"If people are having a particular issue and problem with somebody erm, and then	"If you've got a rough time with somebody, you might just wanna go in and say look, I'm	"We just discuss, sort of, erm, any issues we've got as a team"

	you put that into context, how erm, traumatic those events could be err, it erm, it can have a personal impact on you. Erm, and er the patient is quite complex" "Especially within an assertive outreach team where patients tend to be with us who are, erm, fairly difficult to engage, with multiple needs"	client who's raising a lot of difficult scenarios we get a chance to kind of, just get like having someone to kind of guide us through, er that." "When people worry about safety, erm, then that makes, that makes the discuss naturally gravitate towards that person"	"when somebody is starting to become unwell or they've used a load of drugs or they've refused their medication. Erm, erm, and and it tends to kind of follow on if we're sort of worried that the risks might increase, that's when we have those sort of discussions again. "If I'm really worried about a particular patient, and it's complicated, erm, having a team supervision discussion, erm, with psychologists facilitating is the you know, by far and away the best way of making sure that a very rigorous and good erm, sort of erm, risk assessment and risk formulation and risk management plan gets put together."	certain behavioural issues" "brainstorm on possible, you know, causes or what might be perpetuating the problem, what might be precipitating the problem" "It's usually in relation to you know, a particular behaviour that they've adopted that we think isn't, you know, isn't good for their health. Could be their mental health, could be their physical health" "Oh, certainly [increases confidence]	other OT's that I work with so it gives me that opportunity to feedback anything that I'm worried about or anything that I'm struggling with or anything I'm having challenges with" "Particularly within my care as an OT, sometimes I come across issues, erm, within my kind of work life and I don't really know how to kind of move forward with those service users." "Providing opportunity to analyse and explore difficult scenarios that are coming up in the unit."	we sort of try and break it down into what the actual risk is"	struggling with this particular patient, they won't engage with me, or my approach, I don't know what to do" "If we've got a patient, who, that we're struggling with, with interactions or we are very concerned about, we can have a discussion around that." "If we're having any problems, we kind of air those." "We're all worried about her, her physical health and her mental health. She doesn't want us to help."	

				because obviously you've discussed it in a group using a team approach to, you know, do whatever you're doing"				
2. Managing the NHS hierarchy in the group								
No sub-themes	"Obviously the consultant psychiatrist is not with the patient 24 hours, err, our input is very important in that process. Err, input from all angles really. From an OT perspective, from a social worker perspective, from a nursing perspective, from a support worker perspective, from safeguarding perspective, erm, and psychological perspective"	"Sometimes I know with the HCAs, our band 3s, that they can be a little bit more, they've almost got a better sense of what's going on cos they're delivering more hands-on care, but I think some of them feel because they haven't got qualified status, or just in general by virtue of their role, they almost feel like they're less entitled to speak out which I think is really unfair"	"Even though students are kind of learning, be by virtue of the fact that they are on shift and are spending time with the patients, they give their opinions fairly well"	"It might have been discussed with the consultant or with the team manager and then we'll agree that yeah, we should bring it to supervision"	"It's not just the same people making decisions, that a whole team can kind of come together, including healthcare assistants, because they're the people out on the front line a lot of the time and they're the people that are observing a lot of these issues so it's useful to get them involved because there's probably a lot of things that they have	"They always act differently in front of a manager as well."	"Supervision the other day was with our consultant and the unit psychologist, an SHO, another doctor, and they're all lovely but I'm still a band 3 and that can be quite intimidating but they don't make you feel intimidated"	"The HCA's don't always get involved in ward reviews, at all, so I think it's a good chance for them to have a say as well" "It's in person and I think if there's any doctors that want to do it virtually, they can log on" "It's like a whole team sort of thing so we get the doctors involved as well, erm, if they're around. So, yeah, no, everyone

	"The whole team are involved, picking and choosing, err. Who is involved." "We value every team member in our team. Erm, you see erm. Occupational therapists, social worker, er and the doctors and the students also. Student psychologist, student doctors. Student nurses also"	"In this one meeting, it felt like, a little bit like it all came to a head in a very kind of professional but subtle kind of like comments being made and it kind of felt a little bit like a power struggle" "People of course can get very passionate when they feel like it comes to a particular area of the team where they feel is needing more help or isn't being recognised enough. I mean I'm very passionate about nursing. I say we do all the work and deserve all the credit [laughs]			observed and we haven't"			gets involved and has a say in what they think could improve things."
3. The experience of responsibility and accountability								
No sub-themes	"We do try to have the care co-	"There's never been an element of	"I am glad when we have done a	"My role would be to highlight if	"It is reassuring knowing that I	"It's shared ownership"	"If it's something they've agreed.	"It's hopefully a good like, forum

	ordinator present because they tend to know a bit more than us about the patient." "Sometimes if you practice independently, er, you might have to take a fair amount of stress personally and it might impact your health, your level of functioning. But when err, it becomes a team decision, err, I guess you feel that it is not only you" "Think it makes you feel, erm, less worried because it's not a personal decision, it's a collective team decision"	oh you've failed or you have to take accountability" "I get comfort from a team decision because I'm not taking direct responsibility and potential criticism" "It's nice to share difficult decisions, naturally."	good job with the risk assessment. I'm glad it has been done cos part of the problem you know, if something does go wrong, people always ask, you know, what was the risk assessment." "If it's the kind of standard Rio risk assessment, which isn't brilliant, that's usually the care coordinator. If it's a case bust outcome and or it's been done as a sort of START or a very specific tool then, that tends to be **the** unit psychologist" "I am responsible for most things on the team but	there were any risks, erm to highlight if there were any issue with medication that could compound the risks" "Then we were able to, you know, agree, you know, that we did it in their best interest so. But if it were to be one person that came up with the idea, you might struggle within yourself to, you know, to go through with it because you might feel oh am I being paternalistic towards the patient or am I mothering them in a way" "It does help when it's a team approach and you	can bring that forward and get a team perspective on it so I'm not left to make decisions on my own." "You can kind of see it as a bit of a safety net as well if you're really worried about a kind of, I dunno, a service user's risk or something risky on the unit. it just really offers a nice opportunity to discuss that, erm rather than feeling like decision's have been solely left to yourself." "Rather than having one person make a decision about a service users care. the whole mdt can come to that conclusion together which is	"The named nurses usually update the care plans."	That's, that's normally handed over to the nurse in charge then" "The nurse in charge, the named nurse would. They would, they would maybe alter how an approach is to be made on a particular patient, they would alter that in their risk."	cos yeah you can feel a bit like you're, like as a nurse, the only one, cos you're giving hand over you can, anything you say is kind of, can influence how it's handed over the next day" "The minutes for the team supervisions are put on the shared drive so it's sort of on the person to, on the individual to look at that themselves." "We'd talk about it in a general way, we wouldn't sort of pinpoint like you did this wrong on this day, so we'd talk more in a general way"

			that's what I get paid for” “And certainly, you know, involving a whole team helps to share out the stress.”	can obviously, you know, have that confidence to say yes, this is something that we agreed as a collective and that's what we're doing”	better quality care to be honest”			