An Exploration of Risk Factors and Consequences of Occupational Burnout in Mental Health Professionals

Ben Davis

A thesis submitted in partial fulfilment of the requirements for the award of Doctor of Clinical Psychology at the University of Sheffield

The University of Sheffield
Faculty of Science
Clinical Psychology Unit
Department of Psychology

Submission Date: June 2020

This page is intentionally left blank

Declaration

I declare that this work has not been submitted for any other degree at the University of Sheffield, or any other institution. The work presented is original and all other sources have been referenced accordingly.

Structure and Word Counts

Literature review

Without references and tables: 7,455

With references and tables: 12,420

Research report

Without references and tables: 7,989

With references and tables: 12,975

Overall Abstract

Occupational burnout is an established problem in the caring professions and there is a high prevalence amongst mental health professionals. Burnout can occur in any work context and has two key aspects. The first, emotional exhaustion, relates to feeling emotionally drained and worn out by one's work, potentially to the extent that individuals find it harder to manage work pressures and are negatively impacted outside of work as well. The second, disengagement, relates to losing interest in one's work, and potentially disconnecting from colleagues and patients. Models of burnout commonly think about this problem as an imbalance of demands and effort compared to resources and reward. Higher levels of occupational stress and workload increase risk of burnout and higher levels of professional support and job satisfaction reduce this risk. Burnout has been shown to have detrimental effects on physical and mental health in healthcare professionals and also to negatively impact on service delivery.

Another reported consequence is increased job turnover, and this was examined in a systematic review and meta-analysis of mental health professionals. Twenty-three eligible studies were reviewed, sixteen of which were included in a meta-analysis. The latter found a significant, moderate, positive association between burnout and turnover intention. Differences in how studies were performed appeared to account for the variability in the size of association which supports confidence in the findings. The other articles supported the positive association between burnout and turnover, however, as the data was largely correlational and measured at single time-points, we

cannot conclude that burnout causes turnover and additional research with better design is required to investigate this possibility.

The research study investigated risk factors for burnout in mental health professionals by collecting measures of burnout on seven occasions over 6-months. Personal characteristics such as age and gender, job characteristics such as role and service were collected at the start of the study, as well as several questionnaire-based measures, previously shown to be associated with levels of burnout. The study developed models of burnout aimed to support predicting which individuals might be at higher risk of burnout over time and found a number of factors significantly associated with burnout.

Overcommitment and workload-related stress were associated with higher risk of exhaustion. Higher levels of job autonomy and self-efficacy were associated with lower risk of exhaustion. Stress related to organisational processes (e.g. poor management and supervision) was associated with higher risk of disengagement with higher levels of autonomy and job satisfaction associated with lower levels of disengagement. Work-family conflict and overtime were also associated with increased exhaustion indicating work-life balance impacts burnout. Supervisor and colleague support were associated with lower burnout levels. We concluded that interventions directed at both organisational and individual levels, to increase job autonomy and self-efficacy and reduce overcommitment and poor work-life balance, may be more effective than more commonly applied stress management interventions to reduce burnout.

Acknowledgments

I would firstly like to express my gratitude to my supervisors, Dr Jaime Delgadillo and Professor Michael Barkham for their fabulous support and advice during this project. In addition, the support of the collaborators and Research and Development teams at the different research sites has made this project possible alongside, of course, all the participants who completed several surveys over a number of months. Last, but not least, I would like to thank my fiancé Michelle for her wonderful support, patience, and belief in me.

Table of Contents

This remained of this page is intentionally left blank

Section 1 Literature Review

Mental Health staff burnout and staff turnover: A systematic review and meta-analysis

Abstract

Objectives: Several studies have identified associations between burnout and staff turnover in healthcare professionals. This systematic review and meta-analysis aimed to investigate this association specifically in mental health professionals to investigate whether higher levels of burnout are associated with increased staff turnover.

Method: Systematic searches of six databases were performed using search terms relating to burnout and staff turnover in healthcare professionals. Articles which included analysis of associations between quantitative measures of burnout and turnover or turnover intent in mental health professionals were included. Random effects meta-analysis was performed on reported and transformed *r* correlation coefficients for burnout associations with turnover intent. A subgroup analysis was conducted on methodological factors. A narrative synthesis was performed on other studies meeting criteria. Risk of bias evaluation was performed on all studies included in the synthesis.

Results: Twenty-three peer-reviewed studies were reviewed. Sixteen were included in the meta-analysis. Most studies were classified as having low risk of bias. Omnibus meta-analysis found a significant, moderate, positive association between burnout and turnover intention, $r = 0.43$ (95% CI [.38,.48] $p<.0001$). Methodological differences appeared to account for a significant proportion of heterogeneity. Narrative synthesis supported the positive association between burnout and turnover.

Conclusion: A significant, moderate positive association between burnout and turnover intention is reported with additional evidence to include actual turnover, although these findings are mainly from correlational data. Further research using longitudinal or interventional design is required to develop understanding

of any potential causal relationship between burnout and measures related to staff turnover.

Practitioner Points:

- Mental health service managers should be aware of the potentially greater risks of their team members seeking alternative employment due to burnout with potentially negative implications to service delivery if staff are not adequately supported.
- Mental health service retention programmes should consider the impact of burnout on staff turnover and co-ordinate accordingly with occupational health and wellbeing teams to consider appropriate initiatives to support staff and reduce turnover and its potentially negative consequences.

Introduction

Retention of healthcare staff has been acknowledged as a global problem with an estimated shortfall of 12.9 million skilled health professionals by 2035 (WHO, 2013). In the UK, there is an acknowledged staffing crisis with over 100,000 staff vacancies (Buchan et al., 2019). This relates to more staff leaving their professions or taking early retirement and lower levels of new professionals, and includes increases in newly qualified staff leaving (Health Education England, 2014). A recent survey of UK mental health professionals reported that nearly half of respondents were thinking about leaving their work in mental health, with nearly one third of those actually planning to do so (UNISON, 2017). Mental health nurses are also significantly more likely to be seeking new jobs than nurses in other roles (Marangozov et al., 2017). These intentions to leave are also matched by high levels of actual turnover in mental health services (Brabson et al., 2020; Bukach et al., 2017; Woltmann et al., 2008). In addition, recruitment into psychiatry and mental health nursing roles has not kept up with demand, as evidenced by negative net employment rates in the UK, despite training courses being oversubscribed (NHS England, 2017). In 2017, NHS Improvement was asked by the UK Government to deliver a programme aimed at reducing clinical staff turnover in mental health staff services as well as in general nursing (NHS Improvement, 2019).

High staff turnover is associated with significant financial costs. These involve the cost of recruitment, temporary replacement with expensive agency staff, training replacement staff, and costs of loss in productivity (Hayes et al., 2006). Estimates of actual costs vary between professions and levels of seniority with estimates of the cost of replacing nurses ranging between 31% and 131% of an individual’s annual salary (Halter et al., 2017); these costs are in-line with general employment turnover cost estimates (Allen, 2008). In total, staff turnover may cost a minimum of five percent of services' annual budgets (Waldman et al., 2010). Evaluating the impact of these financial costs directly is understandably difficult; however, studies of the impact of financial pressures on services report negative impacts on patient care (e.g. delays and reduced access to services) as well as additional stress on staff, both of which are related to staff reductions (Robertson et al., 2017). Lower staffing levels have also been associated with poor patient outcomes and increased clinical errors (Frijters et al., 2007).

The direct negative impact of staff turnover on quality of care is often assumed, but may more often reflect opinion rather than evidence as research in this area is limited (Bae et al., 2010; Hayes et al., 2006; Shields & Ward, 2001; West, 2018). One study reported no effects of turnover on service delivery as measured by cancellations and staff absence rates (Gray et al., 1996). A study of mental health clinical outcomes showed no impact of number of staff leaving per month (Brandt et al., 2016). However, higher staff-patient ratios were associated with better outcomes and staff stability with better outcomes for patients starting treatment, although effect sizes were very small. Their results may indicate that staffing consistency is more important when patients first engage with a service which may relate to factors such as development of trust with clinical professionals (Birkhäuer et al., 2017).

Studying factors which contribute to actual staff turnover is challenging due the requirement to perform longitudinal studies and subsequent high loss to follow-up rates (Ball & Pike, 2009). Working with healthcare managers to capture cross-sectional turnover data during staff notice periods may, however, be a feasible approach (Mccarthy et al., 2017). The majority of studies investigating staff turnover measure turnover intent instead of actual turnover. Turnover intent relates to thoughts and/or plans about leaving either a role or profession (Vandenberg & Nelson, 1999) and is considered the most reliable predictor of actual staff turnover (Hayes et al., 2006; Yamazakia & Petchdee, 2015; Zaheer et al., 2019).

Turnover intent has been described as a process triggered by negative psychological responses to jobs or organisations, and in healthcare professionals it has been associated with factors such as job dissatisfaction, occupational stress and bullying (Hayes et al., 2012; Mccarthy et al., 2002; Takase, 2010). However, seeking new experiences and skills was given as the main reason for job change by over half of the participants in one large study of UK nurses, indicating that staff changes frequently relate to professional development. The same study did however find that stress, workload and dissatisfaction did account for a significant minority (≈30%) of job changes (Ball & Pike, 2009) indicating that negative aspects of job experience have a significant part to play in staff turnover.

Models of turnover intent have incorporated a variety of constructs such as perspectives of organisational culture, work identity-threat, organisational citizenship and work alienation (Bothma & Roodt, 2013). One prominent model of turnover intent is the Job-Demands Resources model and studies using this model commonly infer turnover intent as the result of occupational burnout due to job demands and lack of resources (Bakker & Demerouti, 2007; Bothma & Roodt, 2013). There is compelling evidence that high levels of occupational burnout are associated with turnover across healthcare professionals as a whole, as observed in a large number of studies (Hoff et al., 2019; Karakachian & Colbert, 2019; Lo et al., 2018; Nantsupawat et al., 2017). However, there is a paucity of research evaluating this relationship in mental health professionals despite the high levels of burnout commonly seen in this population (Morse et al., 2012; Pines & Maslach, 1978). On the basis that mental health staff retention has been specifically highlighted in government policy, understanding the link between burnout and staff turnover may be a key factor in developing effective retention strategies going forward.

Burnout has been most commonly measured using the Maslach Burnout Inventory (MBI) which was first developed for healthcare professionals, although other versions now exist for other populations (Maslach, 1982; Poghosyan et al., 2009) . Factor analysis of burnout constructs has revealed three key dimensions; emotional exhaustion, depersonalisation and personal accomplishment. Emotional exhaustion manifests as feelings of being overwhelmed and fatigued by one's role. Depersonalization represents diminished emotional connection and distance towards one's own work or clients. Personal accomplishment describes the estimation of one's capability and proficiency when working with people and this can be reduced as part of burnout (Maslach, 1982). More recently, the Oldenburg Burnout Inventory (OLBI) was developed to address some of the methodological limitations of the MBI in relation to item design but also underlying factors (Demerouti et al., 2001). The authors propose that burnout is a two-dimensional construct consisting of emotional exhaustion and disengagement, and that reduced personal accomplishment is a consequence rather than a component of burnout. This proposal of a two-factor structure of burnout has been supported by factor analysis of the three most commonly used burnout measures (Maroco & Campos, 2012). A number of influential models exist in relation to causal and protective factors related to burnout. These relate to demands and effort being out of balance with rewards and resources which leads to increased risk of burnout (Bakker & Demerouti, 2017; Siegrist et al., 2014).

Unlike burnout, turnover related measures are more varied, inconsistent and often poorly validated (Bothma & Roodt, 2013). Staff turnover is an objective measure commonly reported as a percentage related to total staffing numbers and usually expressed as a rate (e.g. per annum). Turnover intent, despite its frequent investigation, is commonly measured using unvalidated single items, and the use of more than three-item scales is limited (Martin, 2007). Longer validated measures do exist, although these appear to be more commonly used outside of healthcare research, and there may be face-validity issues relating to items that measure associated constructs such as job satisfaction (Bothma & Roodt, 2013).

The current review aimed to assess the strength of evidence for a positive association between levels of burnout in mental health professionals and staff turnover and to report a summary estimate of the strength of this relationship via meta-analysis. It also aimed to investigate sources of heterogeneity within studies investigating this relationship.

Methods

A systematic review protocol was pre-registered on the Center for Open Science OSF platform before the database searches were performed:

https://mfr.de-1.osf.io/render?url=https://osf.io/jfwpg/?direct%26action=download.

Searches

Literature database searches were performed using Ovid PsycINFO, Ovid MEDLINE, Scopus, Web of Science, Embase and CINAHL platforms in January 2020. Search terms (related to burnout, turnover intention and healthcare professionals) were used to capture a superset of all related studies involving healthcare professionals (Appendix A). Where possible, searches were limited to articles published in the English language with no restrictions on publication date. Backward and forward citation searches were completed on full-text articles which met inclusion criteria.

Inclusion and exclusion criteria

Studies were eligible for inclusion if they met the following criteria: quantitative studies including mental health care professionals; associations of burnout and staff turnover- related measures were reported (observational) or where burnout and turnover-related outcomes were reported (interventional); publication in a peer reviewed journal.

Studies were excluded as follows: significant levels of non-clinical professionals within the sample or where they could not be differentiated from clinical professionals; where the study design was not appropriate to support a meta-analysis (e.g. qualitative or case studies); full-text articles not written in English; conference proceedings; studies of healthcare students; care home and school settings.

Data Extraction and Transformation

Database searches were merged using Mendeley Desktop software (version 1.19.4) and duplicates were removed. Titles and abstracts of retrieved study articles were screened to exclude those not meeting inclusion criteria. Full texts of remaining search results were then reviewed against the eligibility criteria. Study characteristics were captured and coded using the Cochrane Collaboration Data Collection form (Higgins and Green, 2011). Correlation coefficients for burnout and staff turnover-related measures were extracted along with sample size to enable effect sizes to be compared. Studies that reported other measures of association between burnout and turnover intent were included to maximise inclusion of relevant data. Transformation of these was applied with careful consideration due to potential issues combining derived metrics from different statistical variables (Lipsey & Wilson, 2001). Additional parameters included standardised regression β-coefficients and analysis of variance (ANOVA) *F* scores (extracted or calculated), where these could be appropriately transformed to *r* (Cohen, 1988; Peterson & Brown, 2005; Rosenthal, 1994; Lakens, 2013; Table 1).

Where sub-sample *r* values were reported, these were combined into a single summary statistic. Where multiple correlation coefficients for burnout and intention to leave were reported for the same sample, the mean was calculated as appropriate. First authors or authors listed for correspondence were contacted via email where clarification of study data was required.

Risk of bias assessment

Risk of bias was assessed using a modified version of the Quality Assessment Tool for Observational Cohort and Cross-sectional Studies, which standardised scores for cross-sectional and longitudinal studies (Hagger et al. 2017). A random sample (50%) of included studies was assessed by an independent researcher and the intraclass correlation coefficient was calculated using a random effects model as an index of inter-rater reliability. Inconsistencies were discussed and resolved without the need for moderation.

Meta-analysis

A Meta-analysis effect size integration was performed on reported or calculated *r* values (e.g., *r* values transformed from other statistical tests) to estimate the summary relationship between levels of burnout and intent to leave (ITL). The random-effects method was utilised to account for between-study differences such as sample population and different measures used (Riley et al., 2011). R statistical software (www.R-project.org, version 3.6.3) was used with a number of meta-analysis specific packages (Harrer et al., 2019; Schwarzer et al., 2015; Viechtbauer, 2020) Q and I^2 statistics were calculated using the Q-test and heterogeneity was considered meaningful if this was significant based on a larger than normal alpha value (due to the Q-test's poor detection of true heterogeneity; α =.1; Higgins et al., 2003) or if I^2 was greater than 75% (Deeks et al., 2019). Publication bias was evaluated by assessing funnel plot asymmetry using Egger's test and a rank correlation test, as well as the "trim and fill" method (Begg & Mazumdar, 1994; Duval & Tweedie, 2000; Egger et al., 1997). An updated Fail-safe N method was also used to estimate the number of non-significant findings required to render the summary statistic as below significance (Rosenberg, 2005).

Analysis of subgroups or subsets:

Categorical moderator analyses were not performed based on the number of samples (>20) recommended for this type of analysis (Rubio-Aparicio et al., 2017). Heterogeneity was investigated using subgroup analyses of study characteristics; including risk of bias category, burnout measure, and studies using reported versus transformed *r* values. A *uniform method comparison* was also performed, comparing the most frequent study design (i.e. cross-sectional studies reporting correlation coefficients using the MBI burnout measure).

Results

Search Results

Database searches resulted in 7,246 publications identified from the 6 databases searched. De-duplication resulted in a total of 3,797 articles for title and abstract review. Initial screening led to 3,444 articles being excluded (Figure 1). Secondary screening of the remaining 353 articles led to 36 articles being included for full-text review with 2 additional articles resulting from hand searches. Out of these 38 studies, 15 further studies were excluded leaving 23 studies for risk of bias assessment. Of these, 16 studies met criteria for meta-analysis, and 7 were only included in the narrative review.

Study Characteristics

Measures

Study and sample characteristics are summarised in Table 1. Most studies used one of several versions of the Maslach Burnout Inventory (MBI; Maslach & Leiter, 2016) either in full or using selected subscales. Other burnout measures were the Oldenburg Burnout Inventory (n = 1; Demerouti, Bakker, Nachreiner, & Schaufeli, 2001), Copenhagen Burnout Inventory (n = 1; Kristensen et al., 2005*) and a scale developed prior to the MBI (n = 1; Pines & Kafry, 1978). In addition, two studies used a single-item burnout measure. Most studies examined associations between emotional exhaustion subscales with turnover intention, and two studies only examined associations between overall burnout level and turnover intention. Nine studies published depersonalisation or disengagement associations, with only seven of these reporting data suitable for meta-analysis. Four studies included personal accomplishment comparisons. Due to the low numbers of studies, these latter two variables were not included in the meta-analysis.

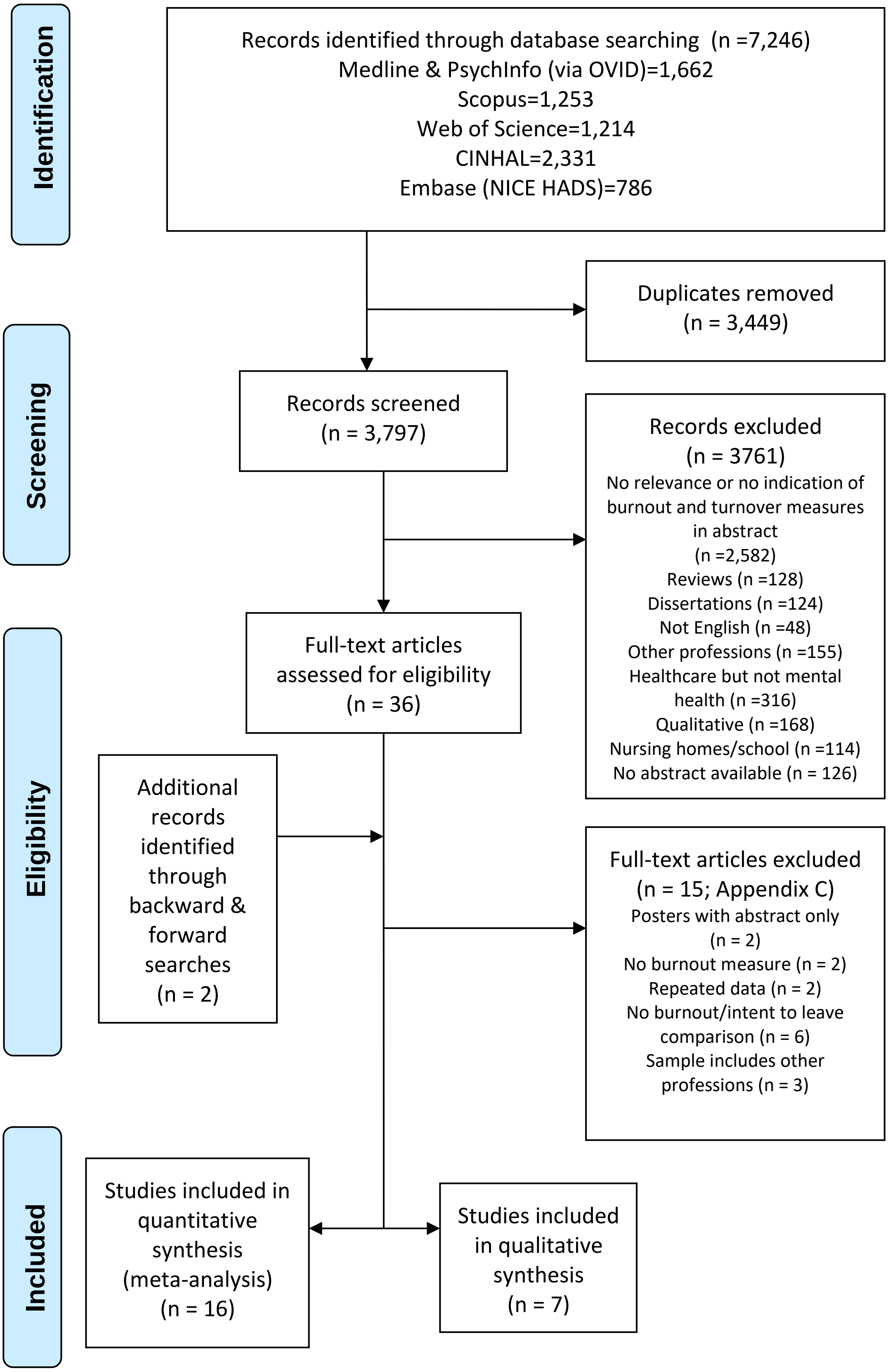

Figure 1. PRISMA flow diagram (Moher et al., 2009).

Table 1. Summary of study characteristics for burnout and turnover measures or intention to leave (ITL)

Author (year)	Study design	Population	Size	Percent Female	Age	Country	Burnout measure	Turnover measure or Intent-to-Leave (ITL)	Correlation between burnout and ITL	Other results
Beidas et al. (2016)[d]	Longitudinal	Mental health clinicians and supervisors	149	77	39.8	USA	Therapist Background Questionnaire 1-item	Actual turnover after 12 months		OR=0.85 for burnout with staff leaving
Blankertz et al. (1997)	Cross sectional	Rehabilitation mental health workers	848	70	37	USA	MBI	1-item, intent to leave in next 2 years	Spearman *rho* correlation EE=.39*** DP=.15** PA=-.21**	
Ducharme et al. (2008)	Cross sectional	Substance abuse counsellors	1869	61	45	USA	MBI-EE	3-item (Walsh et al., 1985)	EE: r=.41[e]	Standardised β= 0.361**
Fukui et al. (2019)	Intervention	Community mental health clinicians	195	80	40	USA	MBI-EE	2-item - considered in last 6 months, likelihood of leaving next 6 months	At baseline r=.65**	Standardised β=0.62**
Garcia et al. (2015)[d]	Cross sectional	Non-prescribing PTSD clinicians for veterans	138	50	52	USA	MBI	1-item, intent to leave in next 2 years		Non-standardised Beta regression coefficient DP=.067** EE=.041(p=.07)
Garcia et al. (2014)	Cross sectional	Veteran association psychiatrists	109	50	51.7	USA	MBI	1-item, intent to leave in next 2 years	EE: r=.48[c] DP: r=.48[c] PA: r=-.36[c]	ANOVA F(4,122)[c] PA=4.44 EE=9.13 DP=9.38
Geurts et al. (1998)	Cross sectional	Mental health care professionals	208	57	37	Netherlands	MBI-EE & DP	4-item bespoke measure of ITL	EE: r=.44*** DP: r=.27***	
Green et al. (2013)	Cross sectional	Community CAMHS workers	388	81	36	USA	MBI-EE	5-item measure of ITL	r=.44**	
Kim (2016)	Cross sectional	Music therapists	163	97	35.3	Korea	MBI	3 item ITL scale (Nissly et al., 2005)	r=.52**	
Knudsen et al. (2006)	Cross sectional	Substance abuse counsellors	817	59	43	USA	MBI-EE	3-item based on Walsh et al., (1985)	EE: r=.30[e]	Standardised β = 0.247***
Knudsen et al. (2008)	Cross sectional	Substance abuse counsellors	823	69	44	USA	MBI-EE	3-item based on Walsh et al. (1985)	EE: r=.32[e]	Standardised β = 0.272***

Author (year)	Study design	Population	Size	Percent Female	Age	Country	Burnout measure	Turnover measure or Intent-to-Leave (ITL)	Correlation between burnout and ITL	Other results
Raquepaw et al. (1989)	Cross sectional	Psychologists and social worker psychotherapists	68	42	NS	USA	MBI	How likely to leave in next 5 years	EE: r=.32[c] DP: r=.45[c]	[c]ANOVA of burnout with ITL category F(2,65) EE=3.62* DP=8.29** Converted to eta^2
Roche et al. (2013)[d]	Cross sectional	Drug and alcohol workers	179	70	mode 40-49	Australia	MBI-EE	4-item (O'Driscoll & Beehr, 1994)		Unstandardized Beta ranges from 0.145** to 0.252** in different regression models
Roncalli et al. (2016)[d]	Cross sectional	Clinical Psychologists	77	77	38	Ireland	MBI	Having left a CMHT in last 3 years with reasons		Mean burnout for group that has left role versus group that has stayed
Salyers et al. (2011)[d]	Intervention	Community mental health clinicians	84	87	NS	USA	MBI	2-item - considered in last 6 months, likelihood of leaving next 6 months		At 6-week follow-up burnout (EE & DP) significantly reduced Cohen's d =0.65 & 0.43 No effect on ITL
Salyers et al. (2015)	Cross sectional	Community mental health service employees	113	83	NS	USA	MBI	2-item - considered in last 6 months, likelihood of leaving next 6 months	† ITL past 6 months EE: r=.58*** PA: r=-.14 DP: r=.42*** ITL next 6 months EE: r=.51*** PA r=-.05 DP: r=.35*** Mean EE r=.55	

Author (year)	Study design	Population	Size	Percent Female	Age	Country	Burnout measure	Turnover measure or Intent-to-Leave (ITL)	Correlation between burnout and ITL	Other results
Scanlan et al. (2019)	Cross sectional	Mental health practitioners	277	63	mode 31-40	Australia	OLBI	3-item ITL scale	EE: r=.42*** DIS: r=.55***	
Trudeau et al. (2001)[d]	Cross sectional	Therapists Psychologists Psychiatrist & Social workers in mental health	382	23-60[a]	43.8-47.14[a]	USA	MBI	2-item scale, intention to *stay*		r = -.23*** for intent to stay
Van Bogaert et al. (2013)[d]	Cross sectional	Psychiatric hospital nurses and care workers	357	89	35.9	Belgium	MBI	ITL nursing profession in next 12 months yes/no		OR for no ITL with mean DP 0.81**
Von Hippel et al. (2019)	Cross sectional	Mental health community workers	349	76	mode 31-39	Australia	Copenhagen Burnout Inventory (Kristensen et al., 2005)	2-item ITL scale for both organisation and profession (Boroff & Lewin, 1997)	† ITL organisation r=.28*** ITL profession r=.36*** Mean r=.32	
Weinberg et al. (1983)	Cross sectional	Mental health clinical professionals & administration support (22%)	416	62	NS	USA	9-item scale (Pines & Kafry, 1978)	1-item "how soon would you leave your job for one with the same money"	r= .45***	Turnover rates in high vs. low burnout organisations* High burnout organisation=49% Low burnout organisation=17%
Yanchus et al. (2017)	Cross sectional	Psychologist, psychiatrists, nurses & social workers in veteran association mental health	7586	66	mode 30-39	USA	1-item "*I feel burned out from my work*"	turnover intention (ITL) and turnover plan as 2 separate measures	† Kendall *tau* ITL [a,b]N:r=.49*** SW:r=.44*** PsyL: r=.46*** PsyI: r=.47*** Mean r=.46***	

Author (year)	Study design	Population	Size	Percent Female	Age	Country	Burnout measure	Turnover measure or Intent-to-Leave (ITL)	Correlation between burnout and ITL	Other results
Yanos et al. (2019)	Longitudinal	Mental health service staff including 2 administration staff	64	78	38.6	USA	MBI	2-item considered in last 6 months, likelihood of leaving next 6 months	† ITL last 6 months EE:*r*=.45*** DP:*r*=-.25 PA: *r*=-.10 ITL next 6 months EE: *r*=.42** DP: *r*=.30* PA: *r*=-.26 Mean *r*=.44	

Note: NS - not specified;† combined *r* using Fisher's *z* transformation or mean for same sample; [a] Reported for different professions; [b] N=Nurse; SW=Social Worker; PsyL=Psychologist; PsyI=Psychiatrist; [c] ANOVA performed or reported and eta^2 effect sizes converted to *r;* [d] narrative synthesis only; [e] transformed β coefficient;

* *p*<.05, ** *p*<.01, *** *p*<.001; MBI- Maslach Burnout Inventory (Maslach, 1982); EE- Emotional exhaustion subscale; DP- Depersonalisation subscale; PA- Personal accomplishment subscale. OLBI- Oldenburg Burnout Inventory (Demerouti et al., 2001), DIS-Disengagement subscale; ITL – intention to leave.

Measures of turnover varied between studies and ranged from short, validated measures of turnover intent (e.g. Walsh, Ashford, & Hill, 1985; O'Driscoll & Beehr, 1994), single or two-item measures capturing intent to leave between 6-months to 5-years ahead, intent to leave profession (yes or no), intent to stay and actual turnover.

A variety of other measures related to burnout and turnover intent were reported including measures of job satisfaction, organisational climate, co-worker and supervisor support, autonomy, distributive and procedural justice, occupational stress and workload related variables.

Study Setting, Population and Design

Studies were conducted in the USA (n = 16), Europe (n = 3), Australia (n = 3) and Asia (n = 1). Several studies recruited unspecified groups of mental health service employees working in inpatient and community mental health services, in substance misuse services and in learning disability services. Other studies specified roles including occupational therapists, psychologists, psychiatrists and mental health nurses and social workers. One study recruited music therapists through a national organisation (Kim, 2016).

Most studies were cross-sectional (n = 20). Only three eligible studies had a longitudinal design, one of which was interventional (Beidas et al., 2016; Fukui et al., 2019; Yanos et al., 2020).

Risk of bias

Standardised risk of bias assessments are summarised in Table 2. Ratings for cross-sectional and longitudinal studies ranged from 3.84-8.13 out of a maximum score of 10. Scores above 6 were classed as indicative of high methodological quality, and below this cut-off as low methodological quality (Hagger et al., 2017; see Appendix B for all items). Most studies (n = 13/16) selected for meta-analysis were above the cut-off score of 6. Four out of seven studies, included for narrative synthesis only, reached this cut-off.

The total sample size for included studies was 15,659 with 19 of the 23 studies including over 100 participants. Whilst a small number of studies acknowledged that the exploratory nature of the study meant that power calculations were not possible, mostly this subject was not addressed. The majority of studies were unlikely to be underpowered based on the variables studied with some very large sample sizes of over 1000 participants. The utility of power calculations for exploratory studies has also been questioned (Jones et al., 2003). Another area where quality scores were often marked down, was around explicit mention of ethical approval, although this may in some cases be inferred from the process of studies being awarded grants from bodies which undertake an ethical review process. Several studies highlighted that participation was voluntary but informed consent was frequently not described. This maybe common practice with studies of healthcare employees and whilst not ideal from an ethical perspective, it may have limited impact on the study results.

The Intraclass Correlation Coefficient based on 50% of studies compared was .89, 95% CI [.68, .96] which is indicative of high inter-rater reliability (Cicchetti, 1994).

Meta-analysis

Standardisation of statistical data

Most studies eligible for the meta-analysis reported correlations (n = 11). Of these, 9 were Pearson's *r*, with two studies reporting Spearman's *rho* or Kendall's *Tau*. Three studies with standardized β values from regression analysis were included and converted to *r* using the guidelines from Peterson & Brown, (2005). Studies which reported both *r* and β indicated that these were either very similar or β was lower than *r.* Two *eta-squared* (η^2) values were calculated and converted to *r* using published

Table 2. Quality and risk of bias assessment (based on Hagger et al. 2017).

Article	Research Question	Population	Participation	Outcome defined	Inclusion/Exclusion Applied	Ethics	Informed Consent	Drop-out analysis	Sample Size to IV ratio >=10	Power calc	Longitudinal	Outcome measured > 4 weeks	Independent variables defined and valid	Dependent variables defined and valid	Loss to follow up <=20%	Control for confounding variables	Score	Quality
Beidas et al. (2016)*	Y	Y	Y	Y	Y	Y	Y	N	Y	NS	Y	Y	N	Y	Y	Y	8.1	High
Blankertz et al. (1997)	Y	Y	N	Y	Y	NS	NS	NA	Y	NS	N	NA	Y	Y	NA	Y	6.2	High
Ducharme et al. (2008)	Y	Y	Y	Y	Y	NS	NS	NA	Y	NS	N	NA	Y	Y	NA	Y	6.9	High
Fukui et al. (2019)	Y	Y	NS	Y	Y	Y	Y	Y	Y	NS	Y	Y	Y	N	N	NS	6.9	High
Garcia et al. (2014)*	Y	Y	N	Y	Y	Y	Y	NA	Y	NS	N	NA	Y	N	NA	Y	6.9	High
Garcia et al. (2015)	Y	Y	N	Y	Y	Y	Y	NA	Y	NS	N	NA	Y	NS	NA	Y	6.9	High
Geurts et al. (1998)	Y	Y	Y	Y	Y	NS	Y	NA	Y	NS	N	NA	Y	Y	NA	Y	7.7	High
Green et al. (2013)	Y	N	Y	Y	NS	Y	Y	NA	Y	NS	N	NA	Y	Y	NA	Y	6.9	High
Kim (2016)	Y	Y	N	Y	Y	NS	NS	NA	Y	NS	N	NA	Y	Y	NA	Y	6.2	High
Knudsen et al. (2006)	Y	Y	Y	Y	Y	NS	Y	NA	Y	NS	N	NA	N	N	NA	Y	6.2	High
Knudsen et al. (2008)	Y	Y	Y	Y	Y	Y	Y	NA	Y	NS	N	NA	N	N	NA	Y	6.9	High
Raquepaw et al. (1989)	Y	Y	N	Y	Y	NS	NS	NA	Y	NS	N	NA	N	N	NA	Y	4.6	Low
Roche et al. (2013)*	Y	Y	N	Y	NS	Y	NS	NA	Y	NS	N	NA	Y	Y	NA	Y	6.2	High
Roncalli et al. (2016)*	Y	Y	N	Y	Y	NS	NS	NA	Y	NS	N	NA	Y	Y	NA	Y	6.2	High
Salyers et al. (2011)*	Y	Y	N	Y	Y	NS	NS	N	Y	NS	Y	Y	N	N	Y	N	5.0	Low
Salyers et al. (2015)	Y	N	Y	Y	NS	Y	NS	NA	Y	NS	N	NA	NS	NS	NA	N	3.8	Low
Scanlan et al. (2019)	Y	Y	N	Y	NS	Y	Y	NA	Y	NS	N	NA	Y	Y	NA	Y	6.9	High
Trudeau et al. (2001)*	N	Y	N	N	NS	NS	NS	NA	Y	NS	N	NA	Y	Y	NA	Y	3.8	Low
Van Bogaert et al. (2013)*	Y	Y	Y	N	NS	Y	NS	NA	Y	NS	N	NA	Y	N	NA	Y	5.4	Low
Von Hippel et al. (2019)	Y	Y	N	Y	NS	Y	Y	NA	Y	NS	N	NA	Y	Y	NA	Y	6.9	High
Weinberg et al. (1983)	Y	Y	Y	Y	NS	NS	NS	NA	Y	NS	N	NA	N	N	NA	N	3.8	Low
Yanchus et al. (2017)	Y	Y	Y	Y	Y	Y	NS	NA	Y	NS	N	NA	N	N	NA	Y	6.2	High
Yanos et al. (2020)	Y	N	Y	Y	NS	Y	NS	N	Y	NS	Y	Y	Y	Y	Y	Y	6.9	High

Notes: Y- present, N- not present, NS- not specified, NA- not applicable. Y scores one point all others zero. Longitudinal scores divided by 16, cross-sectional divided by 13 and scored out of 10. 6 or more is rated as high otherwise low. * narrative only. Items relating to longitudinal studies only in blue text.

ANOVA F scores (Raquepaw & Miller, 1989) or from a one-way ANOVA performed on a set of mean burnout values for each of five levels of turnover intent (Garcia et al. 2014).

One study (Yanchus et al., 2017) reported *r* values by different professions (nurses, psychiatrists, psychologists & social workers) and these sub-sample effect-sizes were pooled using Fisher's *z* summary converted back to *r* to be incorporated into the meta-analysis. Another study reported two *r* values for intent to leave the organisation and intent to leave the profession (von Hippel et al., 2019) and these were pooled using the same method. Where separate measures of intent to leave versus plans to leave were recorded as separate items the former was put forward into the meta-analysis as being the most consistent construct in intention to leave measures.

Omnibus analysis – Relationship of Emotional exhaustion with Turnover intention

The effect sizes from all of the studies included in the omnibus meta-analysis (n=16) were in the moderate-to-large range (r = .32 to .65) according to Cohen's criteria (1988), indicating positive and significant correlations between burnout or emotional exhaustion with turnover intention. The pooled correlation coefficient using a random effects model was 0.43 (95% CI [.38,.48] p<.0001) indicating that levels of burnout or emotional exhaustion are moderately correlated with turnover intention (see figure 2). The I^2 statistic was 80.6% (95% CI [69.5%; 87.7%]) and the Q test was highly significant ($Q(15) = 77.51$, $p < .0001$) indicating that the level of study heterogeneity was both large and statistically significant.

The sample size range for the different studies was very large (53-7586) with the largest sample (n=7586; Yanchus et al., 2017) greater than all the other

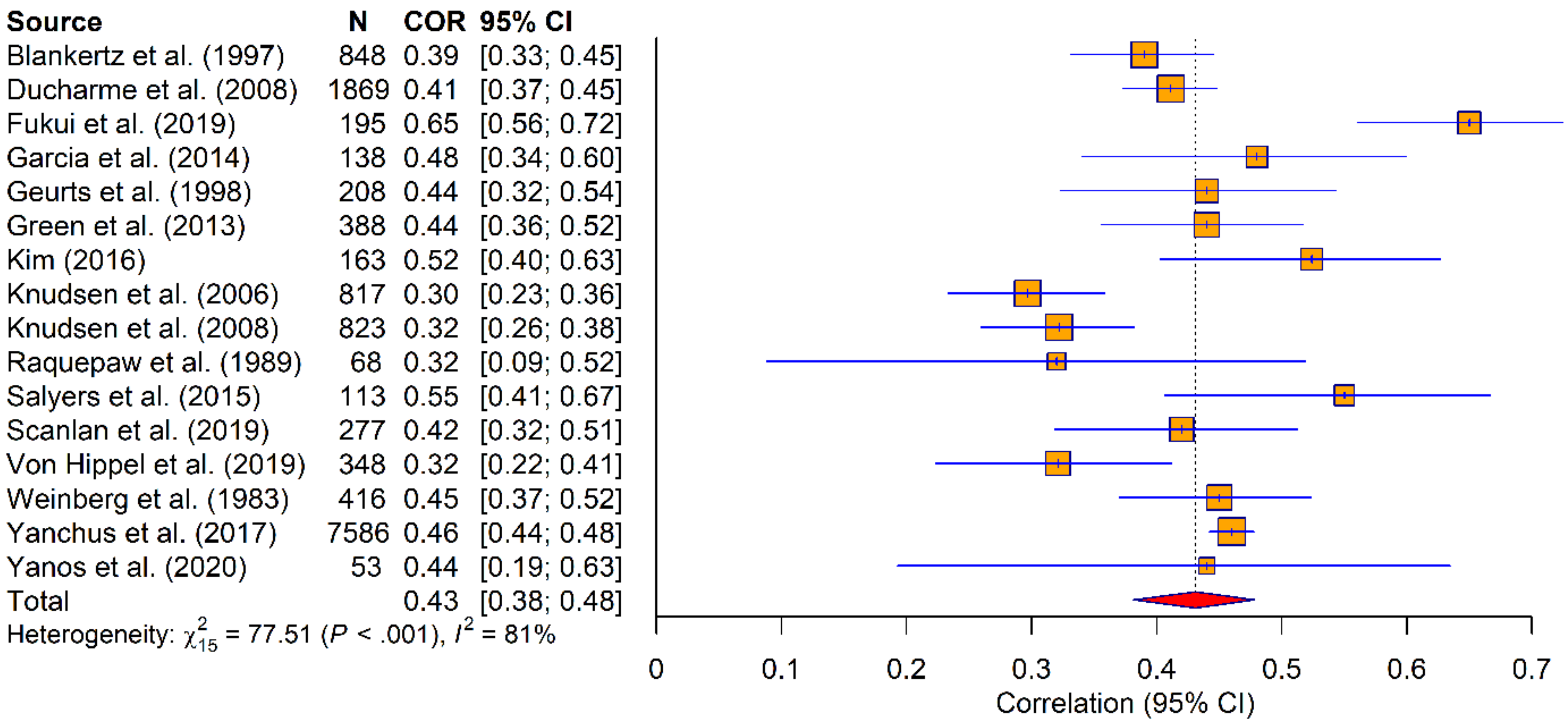

Figure 2. Random effects meta-analysis: correlation (*r*) between burnout and turnover intent. Squares show individual *r* values with 95% CI (blue lines). Smaller squares represent smaller sample sizes. Dotted vertical line – pooled effect size with 95% CI (red diamond).

studies' summed sample size (n=6724). Influence analysis using a leave-one-out method (Viechtbauer & Cheung, 2010) resulted in k-1 pooled effect sizes (.42-.44) which were very close to the main effect size including all studies, and I^2 statistics ranged between 72.7% to 80.7% indicating that no single study overly influenced the summary effect, or contributed to heterogeneity more than other studies to a significant degree as part of the omnibus analysis. This study used a single item relating to burnout which has limited demonstration of validity although a similar item has been previously demonstrated large correlations with MBI exhaustion and depersonalisation subscales (Beidas et al., 2016).

Publication bias

A visual examination of the funnel plot (Figure 3) indicated potential for reporting bias due to some evidence of asymmetry. However, the absence of augmented data points in the trim and fill modelling (Duval & Tweedie, 2000) provides evidence of low risk of publication bias alongside Egger's test which was not significant (p = .69; Egger et al., 1997). In addition a rank correlation test for asymmetry was not significant (Kendall's *tau* = .15, p = 0.45; Begg & Mazumdar, 1994) and a weighted fail-safe measure indicated that the number of non-significant studies needed to make the result of the summary analysis non-significant is 14,441 (Rosenberg, 2005). Overall, these analyses indicated there is a low risk of publication bias.

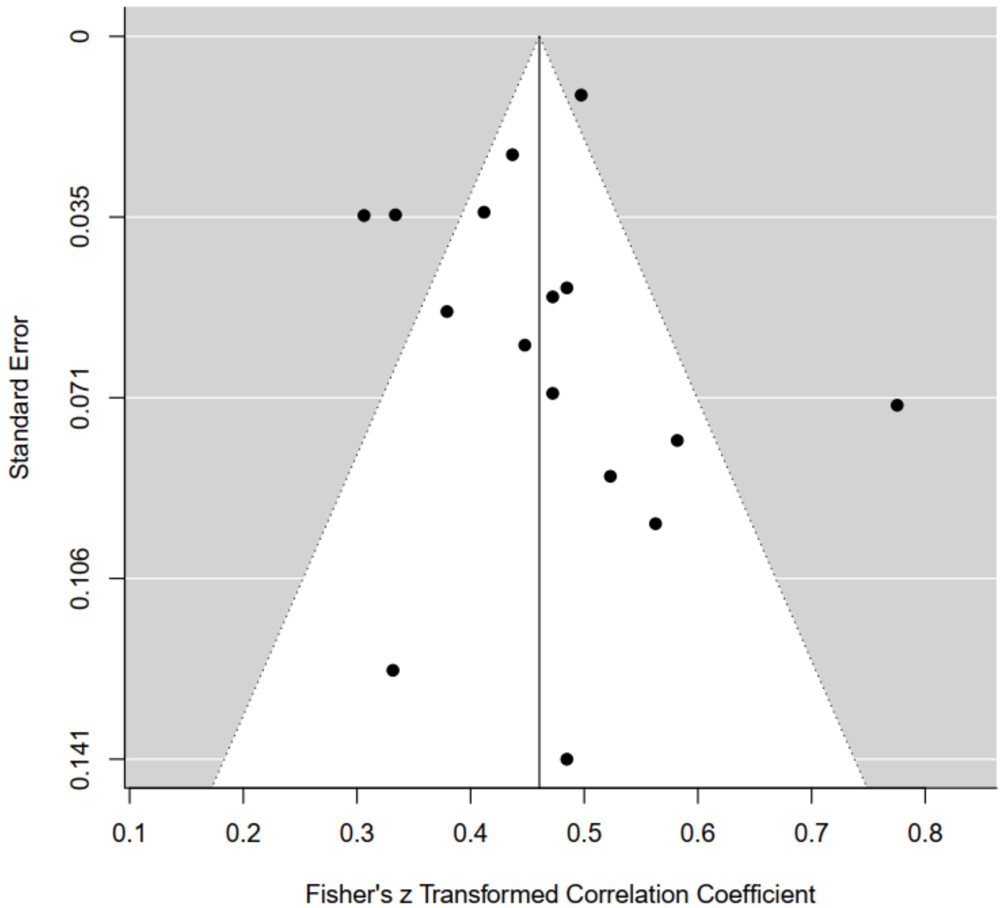

Figure 3. Funnel plot to examine publication bias after trim and fill modelling

Subgroup Analysis

Comparative subgroup analyses using the Q test for subgroup difference (Borenstein & Higgins, 2013) were completed to investigate how methodological differences might influence summary effects sizes and measures of heterogeneity. Comprehensive subgroup analysis results are detailed in Table 3.

High quality (r =.43, 95% CI [0.37; 0.48], p<.0001) and low quality (r =.45, 95% CI [0.32; 0.57], p<.0001) subgroups had summary effects sizes that were very close to the omnibus value and were not significantly different between groups (p=.69). However, the I^2 statistic was markedly lower in the low-quality subgroup (42% versus 84%) and the Q test for heterogeneity was only significant in the high-quality subgroup (p<.0001 versus p=.18).

Table 3. Subgroup analysis of study quality and methodological factors

Moderator	*r* [95% CI]	Q between groups	N	k	I^2	Q group heterogeneity
Quality						
High	.43*** [.37, .48]	.15	13,713	13	84%	73.5***
Low	.45*** [.32, .57]		597	3	42%	3.5
Analysis						
Correlation	.46*** [.40, .51]	4.84*	10,595	11	70%	33.6**
Other	.36*** [.29, .43]		3,715	5	73%	15.0**
Burnout Measure						
MBI	.44*** [.37, .50]	.14	5,683	12	80%	53.5***
Other	.42*** [.36, .48]		8,627	4	68%	9.4*
Full Method Analysis						
Not Uniform	.42*** [.34,.49]	.50	12,537	10	87%	70.3***
Uniform	.45*** [.39,.50]		1,773	6	30%	7.2

p <.0001***, *p* <.005**, *p* <.05*

Subgroup analysis of studies which reported correlation values had a pooled effect size similar to the omnibus analysis (r =.46, 95% CI [0.40, 0.51], p<.0001). This was significantly higher (p=.028) than the subgroup effect size for studies where effect sizes were converted to r from reported standardised β or calculated partial eta^2 values. The summary effect size for this subgroup was markedly lower than the omnibus effect size (r =.36, 95% CI [0.29, 0.43], p<.0001). Heterogeneity indicators were, however, similar to both omnibus analysis and each other with the correlations subgroup a little lower for I^2 (70% versus 73%). Both Q tests of heterogeneity were significant (p<.005).

All but one study was observational with one study having an interventional design with Pearson's r being measured at baseline prior to intervention (Fukui et al., 2019). The study's effect size (r =.65) was outside both omnibus and subgroup confidence intervals by the largest margin. Single outlier analysis using 1-sided Grubb's test identified this study as a significant outlier (p<.05; Grubbs, 1969). Whilst subgroup analysis for heterogeneity could not be completed, the Q test for between subgroup effects was highly significant (p<0.0001).

Study burnout measures were examined in two subgroups as either a version of the MBI or other measure (short scales or single items relating to burnout). Subgroup effect summaries were similar, both between subgroups, and when compared to the omnibus effect summary with the MBI subgroups effect size slightly larger (r =.44, 95% CI [0.37, 0.50], p<.0001) compared to non-MBI measures (r =.42, 95% CI [0.36, 0.48], p<.0001). These effect sizes were not statistically different (p=0.76) however the I^2 statistic was lower in the non-MBI subgroup (68% versus 80%). The Q test for homogeneity was significant in both the MBI group and non-MBI group (p<.0001 versus p <.05)

Finally, a uniform method subgroup analysis was completed for studies which reported correlations, using an observational design and an MBI-based measure of occupational burnout. Again, effects sizes were very similar between subgroups and the omnibus analysis with uniform subgroup effect size (r =.45, 95% CI [0.39, 0.50], p<.0001) being not significantly higher (p=.48) than the non-uniform subgroup (r =.42, 95% CI [0.34, 0.49], p<.0001). However, the uniform method subgroup had a much lower I^2 statistic compared to the high value for the non-uniform subgroup (30% versus 87%). Q test heterogeneity significance aligned with the I^2 statistic for these two subgroups with only the non-uniform subgroup reaching significance (p=.21 versus p<.0001).

Narrative Synthesis of Data Not Suitable for Meta-analysis

Due to issues with study design and data analysis, seven studies were excluded from the meta-analysis. Two studies (Garcia et al., 2015; Roche et al., 2013) presented non-standardised regression coefficients which could not be transformed to *r*. One study measured actual staff turnover 12-months after burnout measures were administered (Beidas et al., 2016) and another study compared burnout with staff who had left a team in the last three years with those who had not (Roncalli & Byrne, 2016). Another study measured the impact of an intervention on burnout and turnover intention but did no direct comparison (Salyers et al., 2011) and one study measured intention to stay rather than intention to leave (Trudeau et al., 2001). Finally, one study analysed no intent to leave as a binary variable with job satisfaction and quality of care as multiple outcomes in a single mixed-effects model (Van Bogaert et al., 2013). As all these studies results related to the original research question, a narrative synthesis was performed rather than bias the results by leaving out relevant study material. This was combined with depersonalisation and personal accomplishment reports and other relevant data.

The two studies which published unstandardized regression coefficients both found significant or close to significant positive relationships between emotional exhaustion and turnover intention (Garcia et al., 2015; Roche et al., 2013). In addition, a significant positive relationship between depersonalisation and turnover intention was revealed (Roche et al., 2013). Staff who had left a mental health team withing the last 3 years did not have significantly different levels of emotional exhaustion or depersonalisation compared to staff who stayed in the same team. Staff who left because of dissatisfaction or stress levels had signifciantly higher levels of emotional exhaustion compared to staff who left for other reasons such as for a better job or location (Roncalli & Byrne, 2016).

A significant negative relationship between intent-to-stay and burnout was reported by Trudeau et al., (2001) and lower depersonalisation was shown to predict no intention to leave the nursing profession (Van Bogaert et al., 2013). Increased burnout has also been shown to predict staff leaving their post over a 12-month period (Beidas et al., 2016) and organisations with high levels of burnout had significacntly higher rates of turnover (Weinberg et al., 1983). Finally, an intervention designed to reduce levels of burnout showed a signficant and moderate effect on reduction of burnout (emotional exhaustion and depersonalisation) at 6-weeks follow-up. Intent to leave was not significantly different at follow-up (Salyers et al., 2011). This was the only study in both meta-analysis and narritive synthesis to not show a positive relationship between a dimension of burnout and turnover or turnover intention on the basis that one might expect reductions in burnout to parallel reductions in turnover intent.

Nine studies reported significant associations between depersonalisation or disengagement and turnover intent. Three of these studies reported smaller, positive coefficient values compared to associations with emotional exhaustion (Blankertz & Robinson, 1997; Geurts et al., 1998; Salyers et al., 2015), with three reporting larger associations (Garcia et al., 2015; Raquepaw & Miller, 1989; Scanlan & Still, 2019) and one almost identical (Garcia et al., 2014). One study reported inconsistent and opposing associations of depersonalisation with two measures of turnover intent (Yanos et al., 2020).

Personal accomplishment was negatively associated with turnover intention in four studies, but only two studies showed significant correlations which were small-to-moderate in size (Blankertz & Robinson, 1997; Garcia et al., 2014).

Discussion

This review aimed to provide a comprehensive synthesis of empirical studies investigating the relationship between occupational burnout and staff turnover in mental health professionals. The meta-analysis provides robust evidence indicating that a moderate positive association between levels of emotional exhaustion and turnover intention may exist in this population. All 16 studies included in the meta-analysis reported moderate-to-large associations between burnout and turnover intent, with 14 of these specifically related to emotional exhaustion. Two additional studies showed significant positive associations between burnout and turnover intent; and intent to stay was associated with lower levels of burnout in two studies. In addition, higher burnout was associated with higher actual turnover in three studies. Only one study showed no association between burnout and turnover, although this was not measured directly.

Positive correlations of the depersonalisation/disengagement component of burnout with turnover intention were also reported, however these were more variable and less consistent and too small in number to include in the meta-analysis. A low number of small negative associations of personal accomplishment with turnover intent were reported which aligns with its opposite subscale score direction as part of the MBI (Maslach 1982).

The omnibus meta-analysis had high levels of heterogeneity which might indicate that the results should be interpreted with caution. However, sensitivity analysis of methodological sub-groups indicated that these accounted for a significant proportion of the heterogeneity, with transformed effect-sizes having significantly smaller *r* coefficients and studies with the most similar methodology having low or potentially unimportant levels of heterogeneity (Deeks et al., 2019; Higgins, 2003).

A significant majority of studies were evaluated as having low risk of bias (high-quality), and there was no significant difference between summary statistics for low- and high-quality studies from sub-group analysis.

Together, the meta-analysis and narrative synthesis triangulate the findings that higher burnout is associated with increased turnover intent in mental health staff, and to a much more limited degree that turnover intent may be reflected in actual staff turnover.

The summary correlation coefficient reported in this meta-analysis is consistent with wider studies of burnout and turnover intention in healthcare professionals (Hazell, 2010; Meeusen et al., 2011; Tziner et al., 2015).

Strengths and limitations

This is the first systematic review and meta-analysis investigating the relationship between burnout and staff turnover specifically in mental health professionals. Strengths of this review include searching across six databases without date restrictions with additional forward/backward citation searches. The search terms were also designed to capture a broad base of healthcare related studies to minimise the risk of relevant studies being missed. In addition, authors were contacted to clarify data to maximise the inclusion of articles. Non-English articles and non-peer reviewed data, such as dissertations were, however, excluded. The review meets PRISMA reporting standards (Moher et al., 2009; see Appendix D) including the protocol being made publicly available online ahead of searches; and including a risk of bias evaluation which included independent validation. There were however some limitations of the review which should be considered when interpreting the results.

We included studies with standardised beta coefficients using a simple conversion based on comparative study with correlations (Peterson & Brown, 2005). Additionally, two ANOVA *F* scores were converted via eta-squared (one reported, one calculated from mean/standard deviation). Subgroup analysis of these against studies which reported correlation coefficients showed a significantly lower summary metric which may have led to an underestimation of the overall association being evaluated. However, this did not appear to be significant, with only .03 points difference between the omnibus and correlation only summary value. Alternatively, these coefficients might represent a more accurate smaller strength of association as other factors were controlled for or included in models and where both correlation and beta-coefficients were reported, the latter were lower. This is unlikely to be the case for the ANOVA transformations however as effect size calculated from *F* can have a tendency to be overestimated (Hullett & Levine, 2003).

Spearman's *rho* and Kendall's *tau* were assumed to be equivalent to Pearson *r*, for the purposes of the meta-analysis based on verbal advice from the authors organisations statistical support team. However, this may not have been appropriate due to the ranked non-parametric calculation for these coefficients. Assuming these tests were used to treat scale-based data as ordinal, then these may be seen as approximately equal, although overall non-parametric correlation may be prone to be more bias than parametric tests (Jamieson, 2004; Rodgers & Nicewander, 1988; Zimmerman et al., 2003).

The data included in the meta-analysis mostly pertained to the emotional exhaustion dimension of burnout as several studies only measured this dimension. A smaller number of studies reported depersonalisation or disengagement coefficients which could have been averaged with exhaustion to create a more global burnout correlation, and subsequently the overall association of burnout with turnover intention. However, this might have introduced additional heterogeneity and evaluating the impact of separate burnout dimensions may be useful for future research to develop effective burnout interventions targeting staff turnover issues.

The meta-analysis was only able to include measures of turnover intention and there is a known discrepancy between intentions and actual behaviours (Ajzen, 1991). As previously described, turnover intention has been evaluated as a good predictor of actual turnover, however not all studies support this (Cohen et al., 2016). Subsequently, the summary effect size may be smaller when considering actual turnover. However, the discrepancy between attitudes and behaviours has been shown to be reduced when attitudes are derived from direct experience, which supports the strong association between intent to leave and actual turnover due to occupational factors (Fazio et al., 1978). In addition, turnover intent is linked to lower organisational commitment which in turn relates to lower staff performance and lower quality of service provision to patients, so may have relevance to services in its own right (West & Dawson, 2012).

The variety of measures used to capture turnover intent may have influenced the reliability of the findings particularly with the range of time periods used (6-months to five years) which may impact the relationship with actual turnover, with longer time periods being less predictive (Hom et al., 1979). Turnover intention for role and profession was not differentiated with studies measuring one or the other or both. Internal staff turnover may be seen as less damaging or even positive as a whole within an organisation as staff move roles to support professional development compared to the “wastage” of staff leaving the profession (Ball & Pike, 2009; Frijters et al., 2007). In that regard, controlling for turnover intent that was not associated with an assumed response to negative occupational factors could provide a more accurate measure of the association between burnout and turnover. This is supported by the finding that staff leaving due to adverse job factors showed higher levels of burnout compared to other reasons for leaving (Roncalli & Byrne, 2016).

Additional sub-group analyses of turnover intention measures may have further explained heterogeneity in the studies analysed, and potentially provided additional confidence in the findings. However, the degree of variability meant that grouping the studies in any meaningful way was problematic and several groups would potentially have had a single study.

Most of the studies were evaluated as having high methodological quality and low risk of bias, however there may have been some shortcomings with the process of evaluation. The quality assessment tool selected, whilst previously validated, was not a standard tool and was selected so that cross-sectional studies would not be overly penalised as is commonly the case with tools such as the frequently used tool developed by Downs and Black (Downs & Black, 1998; Neild, 2018; Protogerou & Hagger, 2019). Most data extracted for the meta-analysis were correlations, which were often calculated as a precursor for more sophisticated analysis such as structural equation modelling. These were all standard rather than partial correlations so no other variables were controlled for even though the studies main analysis often accounted for potential confounding variables. Arguably, one could subsequently classify the majority of the studies as low quality, which would then potentially reduce confidence in the overall summary finding. However, it would seem to be better practice to acknowledge the well-known limitations of inference from a correlation co-efficient as a measure of association, rather than dismiss the findings due to concerns related to risk of bias. An example of where quality assessment can be misleading was observed with an interventional study which due to its experimental and longitudinal nature scored the second highest number of total points. However, the data extracted from the study was a baseline correlation and so technically no different to a cross-sectional study. The score standardisation approach between cross-sectional and longitudinal studies was able to account for this. This perhaps highlights the difficulty of selecting a fit-for-purpose quality assessment tool and interpretation of risk of bias assessment should be considered in the context of the types of measures being reviewed (Protogerou & Hagger, 2019).

The number of studies included in the meta-analysis was relatively low, and while no evidence of publication bias was observed, suppression of non-significant findings was not explicitly evaluated (Peters et al., 2008; Rubio-Aparicio et al., 2017). The fail-safe N was large which might indicate the results of the meta-analysis were robust, however this statistic is sensitive to non-publication bias and should be treated with caution (Petticrew & Roberts, 2008). However, if the results of publication bias analysis are taken at face value, alongside an explanation of heterogeneity due to methodological rather than sample considerations, this study's findings may be generalisable to a wider population of mental health professionals based on the different professionals and organisational contexts included in the review.

Implications for theory, practice, and future research

Due to the nature of observational and largely cross-sectional methods applied in the studies reviewed, there is limited scope to infer mechanisms of the relationship between burnout and staff turnover from a correlation-based summary statistic. However, the results of the only study which did not provide evidence for an association between burnout and staff turnover, may have implications for both theory and practice. This was an interventional study aimed at reducing levels of burnout (Salyers et al., 2011). The intervention (BREATHE) aimed to develop stress/job management skills in participants. Burnout levels were significantly reduced with a moderate effect size; however, levels of turnover intent were not affected, which does not support a causal link between burnout and turnover intent. This might indicate that burnout and turnover intent occur in parallel, potentially explained by a shared mediating factor related to job demands/resources. Job satisfaction is a potential candidate for such a mediating factor and has been proposed to mediate between job conditions and burnout (Spector, 1997). It has also has been shown to have stronger associations with turnover intent compared to burnout in mental health professionals and also to act as a mediator between career development needs not being met and turnover intention, relating to staff leaving for better opportunities (Ball & Pike, 2009; Rahman & Syahrizal, 2018; Yanchus et al., 2017). The fact that the BREATHE intervention showed no changes in job satisfaction would support this explanation. This perspective, however, goes against the finding that investigation of the Job Demands-Resources model revealed evidence for a strong link between job resources and turnover intent without any significant influence of job satisfaction (Bakker et al., 2003). This study involved call-workers though, and the lack of involvement of job satisfaction might relate to known socio-economic and cultural differences of reward preferences (Slocum, 1971; Yeh, 2015).

One alternative model of job satisfaction, with both direct and indirect effects on turnover intent, the latter mediated by emotional exhaustion, has been proposed (Yanchus et al., 2017). It may be the case that once turnover intent has been established, reducing levels of exhaustion alone will not reverse increased turnover intent without improvements in job satisfaction and other factors. This might be supported by the finding that lower job satisfaction, but not emotional exhaustion, has been associated with turnover plans (Yanchus et al., 2017). This finding should perhaps be interpreted with caution however, as this study used a single item measure for job satisfaction, although similar direct and indirect effects of job characteristics on turnover intent have been proposed (Knudsen et al., 2008). The complexity of interplay between these factors is highlighted by a perspective that occupational stress increases burnout which then decreases job satisfaction with all three factors affecting turnover intent both directly and indirectly (Tziner et al., 2015).

The additional component of depersonalisation/disengagement should perhaps also be considered as this may be a consequence of longer term emotional exhaustion, and subsequently increase turnover intent through additional mechanisms (Bakker et al., 2003; Rastogi et al., 2018; Rogala et al., 2016). Subsequently, interventions which address the two core dimensions of burnout alongside job satisfaction may have greater potential for reducing staff turnover related to high levels of staff burnout. This might include, in addition to individually directed interventions to reduce the impact of job demands on exhaustion, organisational interventions that support key aspects of job satisfaction such as autonomy, supervisor support, and team cohesion (Dreison et al., 2018; Fields, 2002; Nagy, 2002; O'Connor et al., 2018).

Further research is recommended capturing longitudinal data which includes burnout, job satisfaction and turnover intent. More sophisticated analysis, such as cross-lagged panel analysis, may help establish whether burnout and intent to leave co-occur as a result of other factors or whether there is indeed a causal mechanism linking burnout to increased turnover intent. Incorporation of turnover intent measures as part of studies of interventions to reduce burnout might help further inform on any causal link between burnout and turnover.

Measures of turnover intent need to be well validated, and to incorporate cognitions, plans and behaviours related to turnover and ideally focused on the short term for accuracy. Capturing intent to leave at both organisation and profession levels may have utility in developing better understanding of these processes to both prevent costly staff turnover and professional “wastage” in the context of a crisis of staff shortages due to unfilled posts.

The World Health Organisation has recommended that policy makers employ “radical measures” to ensure retention of healthcare workers (WHO, 2013). However, the current NHS Improvement guide to improve retention of clinical staff only highlights examples of rewards and benefits approaches to retention and downplays health and wellbeing to an extent. Furthermore, corporate re-branding is highlighted to support staff engagement, yet the national NHS brand was voted the most relevant brand in 2019, so one might speculate these approaches may have limited value (NHS Improvement & NHS England, 2019; Prophet, 2019). In addition, the associated workforce health and wellbeing framework, infers that burnout is a problem related to the individual – contrary to the widely evidenced view of burnout as a problem created by organisations (Brown & Quick, 2013; Moss, 2019). Retention programmes that highlight the potential importance of addressing burnout at both organisational and individual levels should be considered for the mental health sector.

Conclusions

The results of this systematic literature review and meta-analysis indicate the existence of a significant moderate, positive association between burnout and turnover intention, which in turn may be linked to actual turnover. These findings should be interpreted within the context of well-established limitations of simple correlational association as well as the studies focus on intentions rather than actual turnover based the research available to review (Asamoah, 2014). Additional longitudinal and interventional research examining this relationship may help understand the role of burnout in the current staffing crisis in mental health services.

References

Studies prepended with an asterisk were included in the review.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Allen, D. G. (2008). Retaining Talent: A guide to analyzing and managing employee turnover. *Society for Human Resource Mangement Foundation*, 57. http://www.shrm.org/about/foundation/research/documents/retaining talent- final.pdf

Asamoah, M. K. (2014). Re-examination of the limitations associated with correlational research. *Journal of Educational Research and Reviews*, *2*, 45–52.

Bae, S. H., Mark, B., & Fried, B. (2010). Impact of nursing unit turnover on patient outcomes in hospitals. *Journal of Nursing Scholarship*, *42*(1), 40–49. https://doi.org/10.1111/j.1547-5069.2009.01319.x

Bakker, A. B., & Demerouti, E. (2007). The Job Demands-Resources model: State of the art. *Journal of Managerial Psychology*, *22*(3), 309–328. https://doi.org/10.1108/02683940710733115

Bakker, A. B., & Demerouti, E. (2017). Job demands-resources theory: Taking stock and looking forward. *Journal of Occupational Health Psychology*, *22*(3), 273–285. https://doi.org/10.1037/ocp0000056

Bakker, A., Demerouti, E., & Schaufeli, W. (2003). Dual processes at work in a call centre: An application of the Job Demands-Resources model. *European Journal of Work and Organizational Psychology*, *12*(4), 393–417. https://doi.org/10.1080/13594320344000165

Ball, J., & Pike, G. (2009). *Past imperfect, future tense Nurses' employment and morale in 2009*. Royal College of Nursing.

Begg, C. B., & Mazumdar, M. (1994). Operating Characteristics of a Rank Correlation Test for Publication Bias. *Biometrics*, *50*(4), 1088. https://doi.org/10.2307/2533446

*Beidas, R. S., Marcus, S., Wolk, C. B., Powell, B., Aarons, G. A., Evans, A. C., Hurford, M. O., Hadley, T., Adams, D. R., Walsh, L. M., Babbar, S., Barg, F., & Byron, M. (2016). A prospective examination of clinician and supervisor turnover within the context of implementation of evidence-based practices in a publicly-funded mental health system. *Administration and Policy in Mental Health and Mental Health Services Research*, *43*(5), 640–649. https://doi.org/http://dx.doi.org/10.1007/s10488-015-0673-6

Birkhäuer, J., Gaab, J., Kossowsky, J., Hasler, S., Krummenacher, P., Werner, C., & Gerger, H. (2017). Trust in the health care professional and health outcome: A meta-analysis. *PLoS ONE*, *12*(2). https://doi.org/10.1371/journal.pone.0170988

*Blankertz, L. E., & Robinson, S. E. (1997). Turnover intentions of community mental health workers in psychosocial rehabilitation services. *Community Mental Health Journal*, *33*(6), 517–529. https://doi.org/10.1023/A:1025000703487

Borenstein, M., & Higgins, J. P. T. (2013). Meta-Analysis and Subgroups. *Prevention Science*, *14*(2), 134–143. https://doi.org/10.1007/s11121-013-0377-7

Boroff, K. E., & Lewin, D. (1997). Loyalty, voice, and intent to exit a union firm: A conceptual and empirical analysis. *Industrial and Labor Relations Review*, *51*(1), 50. https://doi.org/10.2307/2525034

Bothma, C. F. C., & Roodt, G. (2013). The validation of the Turnover Intention Scale. *SA Journal of Human Resource Management*, *11*(1), 1–12. https://doi.org/10.4102/sajhrm.v11i1.507

Brabson, L. A., Harris, J. L., Lindhiem, O., & Herschell, A. D. (2020). Workforce turnover in community behavioral health agencies in the USA: A systematic review with recommendations. *Clinical Child and Family Psychology Review*, Advanced online publication. https://doi.org/10.1007/s10567-020-00313-5

Brandt, W. A., Bielitz, C. J., & Georgi, A. (2016). The impact of staff turnover and staff density on treatment quality in a psychiatric clinic. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00457

Buchan, J., Charlesworth, A., Gershlick, B., & Seccombe, I. (2019). *A critical moment: NHS staffing trends, retention and attrition*. The Health Foundation.

Bukach, A. M., Ejaz, F. K., Dawson, N., & Gitter, R. J. (2017). Turnover among community mental health workers in Ohio. *Administration and Policy in Mental Health and Mental Health Services Research*, *44*(1), 115–122. https://doi.org/10.1007/s10488-015-0706-1

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Cohen, G., Blake, R. S., & Goodman, D. (2016). Does turnover intention matter? Evaluating the usefulness of turnover intention rate as a predictor of actual turnover rate. *Review of Public Personnel Administration*, *36*(3), 240–263. https://doi.org/10.1177/0734371X15581850

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.

Deeks, J. ., Higgins, J. P. T., & Altman, D. G. (2019). Analysing data and undertaking meta-analyses. In J. P. . Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions* (6th ed.). www.training.cochrane.org/handbook

Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The Job Demands-Resources Model of Burnout. *The Journal of Applied Psychology*, *86*(3), 499–512.

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health*, *52*(6), 377–384. https://doi.org/10.1136/jech.52.6.377

Dreison, K. C., Luther, L., Bonfils, K. A., Sliter, M. T., McGrew, J. H., & Salyers, M. P. (2018). Job burnout in mental health providers: A meta-analysis of 35 years of intervention research. *Journal of Occupational Health Psychology*, *23*(1), 18–30. https://doi.org/10.1037/ocp0000047

*Ducharme, L. J., Knudsen, H. K., & Roman, P. M. (2008). Emotional exhaustion and turnover intention in human service occupations: The protective role of coworker support. *Sociological Spectrum*, *28*(1), 81–104. https://doi.org/http://dx.doi.org/10.1080/02732170701675268

Duval, S., & Tweedie, R. (2000). A nonparametric “trim and fill” method of accounting for publication bias in meta-analysis. *Journal of the American*

Statistical Association, *95*(449), 89–98. https://doi.org/10.1080/01621459.2000.10473905

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple , graphical test measures of funnel plot asymmetry. *Bmj*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Fazio, R. H., Zanna, M. P., & Cooper, J. (1978). Direct experience and attitude-behavior consistency: an information processing analysis. *Personality and Social Psychology Bulletin*, *4*(1), 48–51. https://doi.org/10.1177/014616727800400109

Fields, D. (2002). *Taking the measure of work: A guide to validated scales for organizational research and diagnosis* (pp. 82–83). SAGE Publications, Inc. https://doi.org/10.4135/9781452231143

Frijters, P., Shields, M. A., & Price, S. W. (2007). Investigating the quitting decision of nurses: Panel data evidence from the British National Health Service. *Health Economics*, *16*(1), 57–73. https://doi.org/10.1002/hec.1144

*Fukui, S., Wu, W., & Salyers, M. P. (2019). Impact of supervisory support on turnover intention: The mediating role of burnout and job satisfaction in a longitudinal study. *Administration and Policy in Mental Health and Mental Health Services Research*, *46*(4), 488–497. https://doi.org/10.1007/s10488-019-00927-0

*Garcia, H. A., McGeary, C. A., Finley, E. P., Ketchum, N. S., McGeary, D. D., & Peterson, A. L. (2015). Burnout among psychiatrists in the Veterans Health Administration. *Burnout Research*, *2*(4), 108–114. https://doi.org/10.1016/j.burn.2015.10.001

*Garcia, H. A., McGeary, C. A., McGeary, D. D., Finley, E. P., & Peterson, A. L. (2014). Burnout in veterans health administration mental health providers in posttraumatic stress clinics. *Psychological Services*, *11*(1), 50–59. https://doi.org/10.1037/a0035643

*Geurts, S., Schaufeli, W., & De Jonge, J. (1998). Burnout and intention to leave among mental health-care professionals: A social psychological approach. *Journal of Social and Clinical Psychology*, *17*(3), 341–362. https://doi.org/http://dx.doi.org/10.1521/jscp.1998.17.3.341

Gray, A. M., Phillips, V. L., & Normand, C. (1996). The costs of nursing turnover: Evidence from the British National Health Service. *Health Policy*, *38*(2), 117–128. https://doi.org/10.1016/0168-8510(96)00854-8

*Green, A., Miller, E., & Aarons, G. (2013). Transformational leadership moderates the relationship between emotional exhaustion and turnover intention among community mental health providers. *Community Mental Health Journal*, *49*(4), 373–379. https://doi.org/10.1007/s10597-011-9463-0

Grubbs, F. E. (1969). Procedures for detecting outlying observations in samples. *Technometrics*, *11*(1), 1–21. https://doi.org/10.1080/00401706.1969.10490657

Hagger, M. S., Koch, S., Chatzisarantis, N. L. D., & Orbell, S. (2017). The Common Sense Model of self-regulation : Meta-analysis and test of a process model. *Psychological Bulletin*, *143*(11), 1117–1154.

Halter, M., Boiko, O., Pelone, F., Beighton, C., Harris, R., Gale, J., Gourlay, S., & Drennan, V. (2017). The determinants and consequences of adult nursing staff turnover: a systematic review of systematic reviews. *BMC Health Services Research*, *17*(1), 824. https://doi.org/https://dx.doi.org/10.1186/s12913-017-2707-0

Harrer, M., Cuijpers, P., Furukawa, T., & Daniel, D. (2019). *Doing Meta-Analysis in R*. Github.

Hayes, L. J., O'Brien-Pallas, L., Duffield, C., Shamian, J., Buchan, J., Hughes, F., Laschinger, H. K. S., & North, N. (2012). Nurse turnover: A literature review – An update. *International Journal of Nursing Studies*, *49*(7), 887–905. https://doi.org/10.1016/j.ijnurstu.2011.10.001

Hayes, L. J., O'Brien-Pallas, L., Duffield, C., Shamian, J., Buchan, J., Hughes, F., Spence Laschinger, H. K., North, N., & Stone, P. W. (2006). Nurse turnover: A literature review. *International Journal of Nursing Studies*, *43*(2), 237–263. https://doi.org/10.1016/j.ijnurstu.2005.02.007

Hazell, K. W. (2010). *Job stress, burnout, job satisfaction, and intention to leave among registered nurses employed in hospital settings in the state of Florida.* [Doctoral dissertation, Lynn University]. http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=109853581&site=ehost-live

Health Education England. (2014). *Growing Nursing Numbers: Literature review on nurses leaving the NHS*. 1–28. https://doi.org/10.1016/j.ghir.2018.05.002

Higgins, J. P. T. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hoff, T., Carabetta, S., Collinson, G. E., T., H., S., C., & G.E., C. (2019). Satisfaction, burnout, and turnover among nurse practitioners and physician assistants: A review of the empirical literature. *Medical Care Research and Review*, *76*(1), 3–31. https://doi.org/10.1177/1077558717730157

Hom, P. W., Katerberg, R., & Hulin, C. L. (1979). Comparative examination of three approaches to the prediction of turnover. *Journal of Applied Psychology*, *64*(3), 280–290. https://doi.org/10.1037/0021-9010.64.3.280

Hullett, C. R., & Levine, T. R. (2003). The overestimation of effect sizes from f values in meta-analysis : The cause and a solution. *Communication Monographs*, *70*(1), 52–67. https://doi.org/10.1080/715114664

Jamieson, S. (2004). Likert scales: How to (ab)use them. In *Medical Education* (Vol. 38, Issue 12, pp. 1217–1218). Med Educ. https://doi.org/10.1111/j.1365-2929.2004.02012.x

Jones, S. R., Carley, S., & Harrison, M. (2003). An introduction to power and sample size estimation. In *Emergency Medicine Journal* (Vol. 20, Issue 5, pp. 453–458). British Association for Accident and Emergency Medicine. https://doi.org/10.1136/emj.20.5.453

Karakachian, A., & Colbert, A. (2019). Nurses' moral distress, burnout, and intentions to leave: An integrative review. *Journal of Forensic Nursing*, *15*(3), 133–142. https://doi.org/10.1097/JFN.0000000000000249

*Kim, Y. (2016). Music therapists' job demands, job autonomy, social support, and their relationship with burnout and turnover intention. *Arts in Psychotherapy*, *51*, 17–23. https://doi.org/10.1016/j.aip.2016.08.001

*Knudsen, H. K., Ducharme, L. J., & Roman, P. M. (2006). Counselor emotional exhaustion and turnover intention in therapeutic communities. *Journal of Substance Abuse Treatment*, *31*(2), 173–180. https://doi.org/http://dx.doi.org/10.1016/j.jsat.2006.04.003

*Knudsen, H. K., Ducharme, L. J., & Roman, P. M. (2008). Clinical supervision, emotional exhaustion, and turnover intention: A study of substance abuse treatment counselors in the Clinical Trials Network of the National Institute on Drug Abuse. *Journal of Substance Abuse Treatment*, *35*(4), 387–395. https://doi.org/http://dx.doi.org/10.1016/j.jsat.2008.02.003

Kristensen, T. S., Borritz, M., Villadsen, E., & Christensen, K. B. (2005). The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. *Work and Stress*, *19*(3), 192–207. https://doi.org/10.1080/02678370500297720

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*(NOV), 1–12. https://doi.org/10.3389/fpsyg.2013.00863

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. In *PLoS Medicine* (Vol. 6, Issue 7). PLoS Med. https://doi.org/10.1371/journal.pmed.1000100

Lipsey, M., & Wilson, D. (2001). *Practical meta-analysis*. Sage. https://idostatistics.com/lipsey-wilson-2001-practical-meta-analysis-2001/

Lo, D., Wu, F., Chan, M., Chu, R., & Li, D. (2018). A systematic review of burnout among doctors in China: A cultural perspective. In *Asia Pacific Family Medicine* (Vol. 17, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s12930-018-0040-3

Marangozov, R., Huxley, C., Manzoni, C., & Pike, G. (2017). *Royal College of Nursing Employment Survey 2017*. *December*, 50–81. https://www.rcn.org.uk/professional-development/publications/pdf-007076

Maroco, J., & Campos, J. A. D. B. (2012). Defining the student burnout construct: A structural analysis from three burnout inventories. *Psychological Reports*, *111*(3), 814–830. https://doi.org/10.2466/14.10.20.PR0.111.6.814-830

Martin, A. (2007). *Employee perceptions of organisational commitment, job satisfaction and turnover intentions in a post-merger institution*. Johannesburg.

Maslach, C. (1982). *Burnout: The cost of caring*. ISHK.

Maslach, C., & Leiter, M. (2016). Understanding the burnout experience : Recent research and its implications for psychiatry. *World Psychiatry*, *15*(2), 103–111. https://doi.org/10.1002/wps.20311

McCarthy, G., Tyrrell, M. P., & Cronin, C. (2002). *National study of turnover in nursing and midwifery*. Department of Health and Children for the Study of the Nursing and Midwifery Resource

Meeusen, V. C. H., Van Dam, K., Brown-Mahoney, C., Van Zundert, A. A. J., & Knape, H. T. A. (2011). Understanding nurse anesthetists' intention to leave their job: How burnout and job satisfaction mediate the impact of personality and workplace characteristics. *Health Care Management Review*, *36*(2), 155–163. https://doi.org/10.1097/HMR.0b013e3181fb0f41

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*, *339*(7716), 332–336. https://doi.org/10.1136/bmj.b2535

Morse, G., Salyers, M. P., Rollins, A. L., Monroe-DeVita, M., & Pfahler, C. (2012). Burnout in mental health services: A review of the problem and its remediation. *Administrative Policy Mental Health*, *39*(5), 341–352. https://doi.org/10.1007/s10488-011-0352-1.Burnout

Nagy, M. S. (2002). Using a single-item approach to measure facet job satisfaction. *Journal of Occupational and Organizational Psychology*, *75*(1), 77–86. https://doi.org/10.1348/096317902167658

Nantsupawat, A., Kunaviktikul, W., Nantsupawat, R., Wichaikhum, O.-A., Thienthong, H., & Poghosyan, L. (2017). Effects of nurse work environment on job dissatisfaction, burnout, intention to leave. *International Nursing Review*, *64*(1), 91–98. https://doi.org/http://dx.doi.org/10.1111/inr.12342

Neild, L. (2018). *Investigating the relationship between shame and obsessive-compulsive disorder.* [Doctoral dissertation, University of Sheffield]. http://etheses.whiterose.ac.uk/id/eprint/21184

NHS England. (2017). *Stepping forward to 2020/21: The mental health workforce plan for England*. papers3://publication/uuid/FA79A696-FFC6-4A51-AF17-AC794E34554D

NHS Improvement. (2019). *The national retention programme: Two years on.* NHS Improvement.

Nissly, J., Mor Barak, M., & Levin, A. (2005). Stress, social support, and workers' intentions to leave their jobs in public child welfare. *Administration in Social Work*, *29*(1), 79–100.

O'Connor, K., Muller Neff, D., & Pitman, S. (2018). Burnout in mental health professionals: A systematic review and meta-analysis of prevalence and determinants. *European Psychiatry*, *53*, 74–99. https://doi.org/10.1016/j.eurpsy.2018.06.003

O'Driscoll, M. P., & Beehr, T. A. (1994). Supervisor behaviors, role stressors and uncertainty as predictors of personal outcomes for subordinates. *Journal of Organizational Behavior*, *15*(2), 141–155. https://doi.org/10.1002/job.4030150204

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology*, *61*(10), 991–996. https://doi.org/10.1016/j.jclinepi.2007.11.010

Peterson, R. A., & Brown, S. P. (2005). On the use of beta coefficients in meta-analysis. *Journal of Applied Psychology*, *90*(1), 175–181. https://doi.org/10.1037/0021-9010.90.1.175

Petticrew, M., & Roberts, H. (2008). Systematic reviews in the social sciences: A practical guide. In *Systematic Reviews in the Social Sciences: A Practical Guide*. Blackwell Publishing Ltd. https://doi.org/10.1002/9780470754887

Pines, A., & Kafry, D. (1978). Occupational tedium in the social services. *Social Work*, *23*(6), 499–507. https://doi.org/10.1093/sw/23.6.499

Pines, A., & Maslach, C. (1978). Characteristics of staff burnout in mental health settings. *Hospital and Community Psychiatry*, *29*(4), 233–237. https://doi.org/10.1176/ps.29.4.233

Poghosyan, L., Aiken, L. H., & Sloane, D. M. (2009). Factor structure of the Maslach burnout inventory: An analysis of data from large scale cross-sectional surveys of nurses from eight countries. *International Journal of Nursing Studies*, *46*(7), 894–902. https://doi.org/10.1016/j.ijnurstu.2009.03.004

Protogerou, C., & Hagger, M. S. (2019). A case for a study quality appraisal in survey studies in psychology. *Frontiers in Psychology*, *9*, 1–7. https://doi.org/10.3389/fpsyg.2018.02788

Rahman, A., & Syahrizal, S. (2018). Effect of compensation and career development on turnover intention: Job satisfaction as a mediation variable. *Advances in Economics, Business and Management Research*, *64*, 471–477. https://doi.org/10.2991/piceeba2-18.2019.25

*Raquepaw, J. M., & Miller, R. S. (1989). Psychotherapist burnout: A componential analysis. *Professional Psychology: Research and Practice*, *20*(1), 32–36. https://doi.org/10.1037/0735-7028.20.1.32

Rastogi, A., Pati, S. P., Krishnan, T. N., & Krishnan, S. (2018). Causes, contingencies, and consequences of disengagement at work: An integrative literature review. In *Human Resource Development Review* (Vol. 17, Issue 1). https://doi.org/10.1177/1534484317754160

Riley, R. D., Higgins, J. P. T., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *BMJ*, *342*(7804), 964–967. https://doi.org/10.1136/bmj.d549

Robertson, R., Wenzel, L., Thompson, J., & Charles, A. (2017). *Understanding NHS financial pressures: How are they affecting patient care?* The King's Fund. https://www.kingsfund.org.uk/publications/understanding-nhs-financial-pressures

*Roche, A. M., Duraisingam, V., Trifonoff, A., & Tovell, A. (2013). The health and well-being of Indigenous drug and alcohol workers: Results from a national Australian survey. *Journal of Substance Abuse Treatment*, *44*(1), 17–26. https://doi.org/10.1016/j.jsat.2012.01.009

Rodgers, J. L., & Nicewander, A. W. (1988). Thirteen ways to look at the correlation coefficient. *American Statistician*, *42*(1), 59–66. https://doi.org/10.1080/00031305.1988.10475524

Rogala, A., Shoji, K., Luszczynska, A., Kuna, A., Yeager, C., Benight, C. C., & Cieslak, R. (2016). From exhaustion to disengagement via self-efficacy change: Findings from two longitudinal studies among human services workers. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.02032

*Roncalli, S., & Byrne, M. (2016). Relationships at work, burnout and job satisfaction: A study on Irish psychologists. *Mental Health Review Journal*, *21*(1), 23–36. https://doi.org/10.1108/MHRJ-01-2015-0002

Rosenberg, M. S. (2005). The file-drawer problem revisited: A general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution*, *59*(2), 464–468. https://doi.org/10.1111/j.0014-3820.2005.tb01004.x

Rosenthal, R. (1994). Parametric measures of effect size. In J. C. Cooper, H., Hedges, L.V., & Valentine (Ed.), *The Handbook of Research Synthesis* (pp. 231–244). Russell Sage Foundation.

Rubio-Aparicio, M., Sánchez-Meca, J., López-López, J. A., Botella, J., & Marín-Martínez, F. (2017). Analysis of categorical moderators in mixed-effects meta-analysis: Consequences of using pooled versus separate estimates of the residual between-studies variances. *British Journal of Mathematical and Statistical Psychology*, *70*(3), 439–456. https://doi.org/10.1111/bmsp.12092

*Salyers, M. P., Fukui, S., Rollins, A. L., Firmin, R., Gearhart, T., Noll, J. P., Williams, S., & Davis, C. J. (2015). Burnout and self-reported quality of care in community mental health. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(1), 61–69. https://doi.org/10.1007/s10488-014-0544-6

*Salyers, M. P., Hudson, C., Morse, G., AL, R., Monroe-DeVita, M., Wilson, C., Freeland, L., Salyers, M. P., Hudson, C., Morse, G., Rollins, A. L., Monroe-DeVita, M., Wilson, C., & Freeland, L. (2011). BREATHE: A pilot study of a one-day retreat to reduce burnout among mental health professionals. *Psychiatric Services*, *62*(2), 214–217. https://doi.org/10.1176/ps.62.2.pss6202_0214

*Scanlan, J. N., & Still, M. (2019). Relationships between burnout, turnover intention, job satisfaction, job demands and job resources for mental health personnel in an Australian mental health service. *Bmc Health Services Research*, *19*. https://doi.org/10.1186/s12913-018-3841-z

Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). Meta-Analysis with R: Ch 8 - Network Meta-Analysis. *Meta-Analysis with R*, *November 2015*, 187–216. https://doi.org/10.1007/978-3-319-21416-0

Shields, M. A., & Ward, M. (2001). Improving nurse retention in the National Health Service in England: The impact of job satisfaction on intentions to quit. *Journal of Health Economics*, *20*(5), 677–701. https://doi.org/10.1016/S0167-6296(01)00092-3

Siegrist, J., Li, J., & Montano, D. (2014). Psychometric properties of the effort-reward imbalance questionnaire. *Duesseldorf University*, 1–14. https://doi.org/10.1007/s00420-009-0460-3.

Slocum, J. W. (1971). Do cultural differences affect job satisfaction? *Academy of Management Proceedings*, *1*, 136–143. https://doi.org/10.5465/ambpp.1971.4980955

Spector, P. E. (1997). *Job satisfaction: Application, assessment, causes, and consequences*. Sage.

Takase, M. (2010). A concept analysis of turnover intention: Implications for nursing management. *Collegian*, *17*(1), 3–12. https://doi.org/10.1016/j.colegn.2009.05.001

*Trudeau, L. S., Russell, D. W., de la Mora, A., & Schmitz, M. F. (2001). Comparisons of marriage and family therapists, psychologists, psychiatrists, and social workers on job-related measures and reactions to managed care in Iowa. *Journal of Marital and Family Therapy*, *27*(4), 501–507. https://doi.org/http://dx.doi.org/10.1111/j.1752-0606.2001.tb00343.x

Tziner, A., Rabenu, E., Radomski, R., & Belkin, A. (2015). Work stress and turnover intentions among hospital physicians: The mediating role of burnout and work satisfaction. *Revista de Psicologia Del Trabajo y de Las Organizaciones*, *31*(3), 207–213. https://doi.org/10.1016/j.rpto.2015.05.001

UNISON. (2017). Struggling to cope. In *Mental health staff and services under pressure*. https://www.ageuk.org.uk/our-impact/policy-research/struggling-to-cope/

*Van Bogaert, P., Clarke, S., Wouters, K., Franck, E., Willems, R., & Mondelaers, M. (2013). Impacts of unit-level nurse practice environment, workload and burnout on nurse-reported outcomes in psychiatric hospitals: A multilevel modelling approach. *International Journal of Nursing Studies*, *50*(3), 357–365. https://doi.org/10.1016/j.ijnurstu.2012.05.006

Vandenberg, R. J., & Nelson, J. B. (1999). Disaggregating the motives underlying turnover intentions: When do intentions predict turnover behavior? *Human Relations*, *52*(10), 1313–1336. https://doi.org/10.1023/A:1016964515185

Viechtbauer, W. (2020). *Package 'metafor.' 1*.

Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, *1*(2), 112–125. https://doi.org/10.1002/jrsm.11

*von Hippel, C., Brener, L., Rose, G., & von Hippel, W. (2019). Perceived inability to help is associated with client-related burnout and negative work outcomes among community mental health workers. *Health and Social Care in the Community*. https://doi.org/10.1111/hsc.12821

Waldman, J. D., Kelly, F., Arora, S., & Smith, H. L. (2010). The shocking cost of turnover in health care. *Health Care Management Review*, *35*(3), 206–211. https://doi.org/10.1097/HMR.0b013e3181e3940e

Walsh, J. P., Ashford, S. J., & Hill, T. E. (1985). Feedback Obstruction: The Influence of the Information Environment on Employee Turnover Intentions. *Human Relations*, *38*(1), 23–46. https://doi.org/10.1177/001872678503800102

*Weinberg, S., Edwards, G., & Garove, W. E. (1983). Burnout among employees of state residential facilities serving developmentally disabled persons. *Children and Youth Services Review*, *5*(3), 239–253. https://doi.org/http://dx.doi.org/10.1016/0190-7409%2883%2990029-4

West, M. (2018). *The health care workforce in England*. The King's Fund.

West, M., & Dawson, J. (2012). *Employee engagement and NHS performance*. The King's Fund.

WHO. (2013). *Human resources for health: foundation for Universal Health Coverage and the post-2015 development agenda: report of the Third Global Forum on Human Resources for Health, 10-13 November 2013, Recife, Brazil.* World Health Organization.

Woltmann, E. M., Whitley, R., McHugo, G. J., Brunette, M., Torrey, W. C., Coots, L., Lynde, D., & Drake, R. E. (2008). The role of staff turnover in the implementation of evidence-based practices in mental health care. *Psychiatric Services*, *59*(7), 732–737. https://doi.org/10.1176/appi.ps.59.7.732

Yamazakia, Y., & Petchdee, S. (2015). Turnover intention, organizational commitment, and specific job satisfaction among production employees in

Thailand. *Journal of Business and Management*, *4*(4), 22–38. https://doi.org/10.12735/jbm.v4i4p22

*Yanchus, N. J., Periard, D., & Osatuke, K. (2017). Further examination of predictors of turnover intention among mental health professionals. *Journal of Psychiatric and Mental Health Nursing*, *24*(1), 41–56. https://doi.org/http://dx.doi.org/10.1111/jpm.12354

*Yanos, P. T., DeLuca, J. S., Salyers, M. P., Fischer, M. W., Song, J., & Caro, J. (2020). Cross-sectional and prospective correlates of associative stigma among mental health service providers. *Psychiatric Rehabilitation Journal*, *43*(2), 85–90. https://doi.org/http://dx.doi.org/10.1037/prj0000378

Yeh, H. J. (2015). Job demands, job resources, and job satisfaction in East Asia. *Social Indicators Research*, *121*(1), 47–60. https://doi.org/10.1007/s11205-014-0631-9

Zaheer, S., Ginsburg, L., Wong, H. J., Thomson, K., Bain, L., & Wulffhart, Z. (2019). Turnover intention of hospital staff in Ontario, Canada: Exploring the role of frontline supervisors, teamwork, and mindful organizing. *Human Resources for Health*, *17*(1), 66. https://doi.org/10.1186/s12960-019-0404-2

Zimmerman, D., Zumbo, B., & Williams, R. (2003). Bias in estimation and hypothesis testing of correlation. *Psicológica (Valencia, Ed. Impr.)*, 133–158.

Appendix A - Search Terms

OVID and CINHAL
("care coordinator*" OR "case worker*" OR clinician* OR counselor* OR counsellor* OR nurse* OR midwife* OR psychiatrist* OR psychologist* OR therapist* OR "wellbeing practitioner*" OR PWP OR GP OR doctor* OR "general practitioner*" OR surgeon* OR *ologist$ OR p*ediatrician* OR geriatrician* OR physician* OR registrar* OR resident* OR consultant* OR "medical staff" OR "support worker*" OR "health care staff" OR "health care worker$" OR "health care professional$") AND (burnout OR "emotional exhaustion" OR depersonali*ation OR ((occupational OR job OR work) AND stress)) AND (turnover OR "turnover intention*" OR "intention to leave" OR "intention to stay" OR "intention to quit" OR "wish* to leave" OR "willingness to leave" OR resign OR "staff retention" OR "job retention" OR "personnel retention")

SCOPUS
TITLE-ABS-KEY (("care coordinator*" OR "case worker*" OR clinician* OR counselor* OR counsellor* OR nurse* OR midwife* OR psychiatrist* OR psychologist* OR therapist* OR "wellbeing practitioner*" OR pwp OR gp OR doctor* OR "general practitioner*" OR surgeon* OR *ologist* OR p*ediatrician* OR geriatrician* OR physician* OR registrar* OR resident* OR consultant* OR "medical staff" OR "support worker*" OR "health care staff" OR "health care worker*" OR "health care professional*") AND (burnout OR "emotional exhaustion" OR depersonali*ation OR ((occupational OR job OR work) AND stress)) AND (turnover OR "turnover intention*" OR "intention to leave" OR "intention to stay" OR "intention to quit" OR "wish* to leave" OR "willingness to leave" OR resign OR "staff retention" OR "job retention" OR "personnel retention")) AND (LIMIT-TO (LANGUAGE , "English"))

WOS
TOPIC: ((((((((((((((((((((((((((((((((("care coordinator?" OR "case worker*") OR clinician*) OR counselor*) OR counsel?or*) OR psychiatrist*) OR psychologist*) OR therapist*) OR "wellbeing practitioner*") OR PWP) OR GP) OR doctor) OR "general practitioner*") OR surgeon*) OR nurse*) OR midwife*) OR *ologist) OR p*ediatrician*) OR geriatrician*) OR physician*) OR registrar*) OR resident*) OR consultant*) OR "medical staff") OR "support worker") OR "health care staff") OR "health care worker?") OR "health care professional?") AND ((burnout OR "emotional exhaustion") OR depersonali*ation) OR ((occupational OR job OR work) AND stress)) AND ((((((((((turnover OR "turnover intention*") OR "intention to leave") OR “intention to stay”) OR "intention to quit") OR "wish* to leave") OR "willingness to leave") OR resign) OR "staff retention") OR "job retention") OR "personnel retention"))
Refined by: LANGUAGES: (ENGLISH)

Appendix B – Risk of Bias and Quality Tool

Table 1. Quality assessment tool items from Hagger et al., 2017.

No.	Description
1	Was a specific research question, hypothesis, objective or prediction of the study clearly stated?
2	Was the population clearly specified and defined (e.g., population, condition, location, date and time)?
3	Was the participation rate (i.e., proportion of eligible persons invited to participate that agreed to do so) of eligible persons at least 50%?
4	Was the primary outcome clearly defined?
5	Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?
6	Was the study approved by a relevant institutional review board or research ethics committee?
7	Were participants provided with details of the study prior to data collection and required to provide their consent (e.g., by signing a form)?
8	Was the final sample size representative of the population from which the participants were drawn (characteristics compared between those that remained and drop-outs)?
9	Was the ratio of participants to the number of independent variables appropriate (>=10)?
10	Was a statistical power analysis conducted to establish the target sample size a priori?
11	Did the study include longitudinal follow-up of outcomes?
12	Were the outcome follow-up measures assessed more than once over time? Did the study collect long-term (>4 weeks) follow-up measures of the outcomes?
13	Were the independent variables clearly defined, valid, reliable, and implemented consistently across all study participants? In the case of self-report measures, study-specific reports of reliability (e.g., internal consistency test, test-retest reliability) were expected.
14	Were the dependent variables clearly defined, valid, reliable, and implemented consistently across all study participants? In the case of self-report measures, study-specific reports of reliability (e.g., internal consistency test, test-retest reliability) were expected.
15	Was loss to follow-up after baseline 20% or less?
16	Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposures and outcomes?

Hagger, M. S., Koch, S., Chatzisarantis, N. L. D., & Orbell, S. (2017). The Common Sense Model of self-regulation : Meta-analysis and test of a process model. *Psychological Bulletin*, *143*(11), 1117–1154.

Appendix C - Excluded Full-Text Articles	**Reason**
Acker, G. (2009). Job stress associated with managed care mental health services among social workers in the USA. *European Psychiatry*, *24*.	Poster
Acker, G. M. (2018). Self-care practices among social workers: do they predict job satisfaction and turnover intention? *Social Work in Mental Health*, *16*(6), 710–724. https://doi.org/10.1080/15332985.2018.1494082	No burnout measure
Baldwin, A., & Mendoza, N. S. (2013). Exploring the relationships between demographics, emotional exhaustion, and depersonalization among substance abuse counselors. *Alcoholism: Clinical and Experimental Research*, *37*.	Poster
Borritz, M., Rugulies, R., Villadsen, E., Mikkelsen, O. A., Kristensen, T. S., & Bjorner, J. B. (2006). Burnout among employees in human service work: Design and baseline findings of the PUMA study. *Scandinavian Journal of Public Health*, *34*(1), 49–58. https://doi.org/10.1080/14034940510032275	Mixed population
Hernandez, W., Yanchus, N. J., & Osatuke, K. (2018). Evolving the JD-R model: The moderating effects of job resources and burnout taxonomies. *Organization Development Journal*, *36*(1), 31–53.	Mixed population
Jambrak, J., Deane, F. P., & Williams, V. (2014). Value motivations predict burnout and intentions to leave among mental health professionals. *Journal of Mental Health*, *23*(3), 120–124. https://doi.org/10.3109/09638237.2013.869576	No burnout vs. ITL* comparison
Kim, H., & Lee, S. Y. (2009). Supervisory communication, burnout, and turnover intention among social workers in	Mixed population

health care settings. *Social Work in Health Care*, *48*(4), 364–385. https://doi.org/10.1080/00981380802598499	
Koeske, G., & Kirk, S. (1995). The effect of characteristics of human service workers on subsequent morale and turnover. *Administration in Social Work*, *19*(1), 15–31. https://doi.org/10.1300/J147v19n01_02	No burnout vs. ITL comparison
Luther, L., Gearhart, T., Fukui, S., Morse, G., Rollins, A. L., & Salyers, M. P. (2017). Working overtime in community mental health: Associations with clinician burnout and perceived quality of care. *Special Issue: Disability Policy Research*, *40*(2), 252–259. https://doi.org/http://dx.doi.org/10.1037/prj0000234	No burnout vs. ITL comparison
Matthiesen, S. B., & Dyregrov, A. (1988).*Arbeidsbelastninger Knyttet Til Skolepsykologisk Arbeid: Burnout Og Jobbtilfredshet Sett i Forhold Til PPT-Ansattes Planer Om a Skifte Jobb.*, *25*(1), 27–39.	Not English and wrong population
Orkibi, H. (2016). Highly artistic-social personalities buffer the effects of burnout on career commitment. *Arts in Psychotherapy*, *50*, 75–83. https://doi.org/10.1016/j.aip.2016.06.006	No burnout vs. ITL comparison
Pelletier, L. R., Vincent, C., Woods, L., Odell, C., & Stichler, J. F. (2019). Effectiveness of a Psychiatric-Mental Health Nurse Residency Program on Retention. *Journal of the American Psychiatric Nurses Association*, *25*(1), 66–75. https://doi.org/10.1177/1078390318807968	No burnout vs. ITL comparison
Salyers, M. P., Rollins, A. L., Kelly, Y. F., Lysaker, P. H., & Williams, J. R. (2013). Job satisfaction and burnout among VA and community mental health workers. *Administration and Policy in Mental Health and Mental Health Services Research*, *40*(2), 69–75. https://doi.org/10.1007/s10488-011-0375-7	No burnout vs. ITL comparison

Scanlan, J. N., & Still, M. (2013). Job satisfaction, burnout and turnover intention in occupational therapists working in mental health. *Australian Occupational Therapy Journal*, *60*(5), 310–318.	Same data as other included study
Van Bogaert, P., Mondelaers, M., Clarke, S., & Willems, R. (2013). Nurse practice environment, workload, burnout, job outcomes, and quality of care in psychiatric hospitals: A structural equation model approach. *Journal of Advanced Nursing*, *69*(7), 1515–1524. https://doi.org/http://dx.doi.org/10.1111/jan.12010	Same data as other included study
Yanchus, N. J., Periard, D., Moore, S. C., Carle, A. C., & Osatuke, K. (2015). Predictors of job satisfaction and turnover intention in VHA mental health employees: A comparison between psychiatrists, psychologists, social workers, and mental health nurses. *Human Service Organizations Management, Leadership and Governance*, *39*(3), 219–244. https://doi.org/10.1080/23303131.2015.1014953	No burnout measure

*ITL- Intention to leave

Appendix D – PRISMA Checklist

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	i
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	7
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	8
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	8
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	9
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	8
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	9, 10
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	9, 26
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	9
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	9
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	10
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	10
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	11

Page 1 of 2

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	11
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	11
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	13
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	14-17
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	20
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	22
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	21
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	21-23
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	24
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	28, 34
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	29
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	37
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	NA

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

Page 2 of 2

Section 2 Research Report

Development of a Prognostic Model of Mental Health Staff Burnout Using Machine Learning

Abstract

Objectives: High levels of burnout (emotional exhaustion/disengagement) have been acknowledged as a problem in mental health professionals for several decades and a variety of personal and organisational factors have been associated with increased risk of burnout based on largely cross-sectional research. This study aimed to develop a prognostic model for burnout from longitudinal data as well as inform on novel interventions for supporting burned-out staff.

Methods: Occupational burnout and job satisfaction measures were collected on a monthly basis from a diverse sample of 287 mental health professionals for six-months after a battery of measures, previously associated with burnout were administered. Demographics and role characteristics information was also collected. Analysis involved cross-sectional main effects logistic regression, machine learning variable selection techniques and prognostic model development using multi-level modelling of longitudinal data.

Results: Overcommitment and workload-related stress were associated with higher risk of exhaustion. Higher levels of job autonomy and self-efficacy were associated with lower risk of exhaustion. Stress related to organisational processes (e.g. poor management and supervision) was associated with higher risk of disengagement with higher levels of autonomy, job satisfaction and conscientiousness associated with lower levels of disengagement. Work-family conflict and overtime were also associated with increased exhaustion indicating work-life balance impacts burnout. Supervisor and colleague support were associated with lower burnout levels.

Conclusions: Multi-dimensional interventions at both organisational and individual levels, directed at increasing autonomy and self-efficacy and reducing

overcommitment and poor work-life balance may be more effective than more commonly applied stress management interventions to reduce burnout.

Practitioner points:

- Emotional exhaustion is commonplace in mental health professionals.
- Disengagement may be less frequent but previous research indicates this negatively impacts service-user outcomes.
- Increased job autonomy may protect against both emotional exhaustion and disengagement.
- Increasing staff self-efficacy and reducing overcommitment may help reduce emotional exhaustion.
- Improving supervisor support relating to autonomy and encouraging good work-life balance may help support staff with burnout.
- Clinical Psychology may have an important role as leaders, supervision providers, and deliverers of systemic and induvial level interventions.

Introduction

The phenomenon of staff burnout has been associated with caring professions since early investigations over 30 years ago (Maslach, 1982). In mental health staff, prevalence estimates for high levels of burnout vary between 21% to 67% across a range of professionals involved in patient care (Morse et al., 2012). A recent meta-analysis reported that 40% of participants had high levels of burnout (O'Connor et al., 2018).

Research into the consequences of burnout has tended to focus on the impact on the professional, with several studies demonstrating a negative impact on psychological well-being and physical health (Salvagioni et al., 2017). Organisational impact, such as increased absenteeism and turnover intention, has also been reported (Lee et al., 2011; Salvagioni et al., 2017). In wider research, a meta-analysis of objective measures of performance found small to moderate correlations between emotional exhaustion and job performance (Taris, 2006). However, limited research has focused on the direct impact of staff burnout on mental health service users, although the aforementioned organisational consequences of burnout have been associated with reduced clinical performance (Rollins et al., 2010). A recent study has addressed this by examining the relationship between therapists' burnout level and patients' clinical outcomes in psychological therapy, showing that higher therapist burnout was associated with poorer treatment outcomes (Delgadillo et al., 2018).

Maslach (1982) devised the most influential definition of burnout, describing it as a syndrome with three dimensions involving emotional exhaustion, depersonalisation and reduced personal accomplishment (Demerouti et al., 2001). Emotional exhaustion refers to being overextended and depleted of emotional resources or empathy. Depersonalisation is a process of developing cynical and detached attitudes towards the recipients of professional care and/or professional colleagues. Finally, reduced personal accomplishment refers to negative evaluations of self in relation to one's ability to perform one's role effectively (Demerouti et al., 2001; Maslach & Leiter, 2016).

A wide range of factors have been associated with high levels of burnout in mental health staff. These can be categorised into organisational factors such as excessive workload and time pressure as well as lack of supervision, peer support and autonomy, and these may be stronger predictors of burnout compared to personal characteristics of staff (Morse et al. 2012). However, several studies have highlighted the importance of personal attributes relating to burnout including personality traits, self-efficacy, and individual contextual factors such as work-family conflict and social support (O'Connor et al., 2018; Simionato & Simpson, 2018). In addition, studies of therapists have shown over-involvement with clients to be associated with increased burnout (Lee et al., 2011) and this may be more generalised to studies which show that over-commitment to job roles relates to increased burnout (Avanzi et al., 2014; Jachens et al., 2018). Several demographic variables such as age, gender and ethnicity have also been shown to relate to levels of burnout (O'Connor et al., 2018; Westwood et al., 2017). Finally, attitudes towards one's role measured as job satisfaction (which may be interpreted as attitudes towards job demands and rewards) have been shown to relate to burnout and service user outcomes (Delgadillo et al., 2018).

A number of theoretical models have attempted to capture the interplay between job stresses and burnout, such as the Job Demands-Resources model which has been commonly used in burnout research (Chirico, 2016). This framework has evolved to capture organisational and personal factors alike, and proposes a complex relationship between job demands, and both job and personal resources that impact on motivation and strain and subsequent job performance (Bakker & Demerouti, 2017). It posits that job and personal resources are buffers to dealing with job demands, indicative that resources can be seen as mediators of the relationship between increased job demands and burnout (Bakker & Demerouti, 2017; Xanthopoulou et al., 2007). Moreover, there is evidence to indicate cyclical relationships within this model; for example, that increased burnout leads to decreased job resources (Demerouti et al., 2009). Occupational stress is commonly associated with burnout, however these are different constructs (Bamber, 2006). Job demands can be seen as potential occupational stressors, with occupational stress being a reaction to an imbalance between demands and organisational or personal resources (e.g. high workload and not enough time or having low self-efficacy for working with severely ill clients). Burnout is considered a potential longer-term consequence of chronic occupational stress (Chirico, 2016; Ruotsalainen et al., 2015). Occupational stress involves adverse physiological and psychological consequences (e.g. increased levels of stress hormones or emotional distress) which in turn can lead to physical and mental health difficulties in addition to burnout (Quick & Henderson, 2016). Studies of occupational stress in mental health workers have identified a number of key sources of stress including workload, professional conflict, lack of professional support, role ambiguity, job insecurity and client difficulties (Prosser et al., 1997). Perhaps unsurprisingly, interventions aimed to reduce burnout have commonly targeted occupational stress, for example through stress management workshops (Dreison et al., 2018). Whilst these types of interventions have moderate effect sizes on levels of occupational stress, they only have small or no effect on levels of burnout, further highlighting these as associated but separate constructs (Dreison et al., 2018; Ruotsalainen et al., 2015). In this regard, the need for further research to support the design of more effective interventions for burnout has been emphasised (Ahola et al., 2017; Dreison et al., 2018).

The majority of studies of burnout in mental health staff have used a cross-sectional design (Ballenger-Browning et al., 2011) and it has been acknowledged that cross-sectional predictors do not uniformly correlate with longitudinal predictors (Robins et al., 2017). Further longitudinally designed burnout research has been widely advocated (Delgadillo et al., 2018; Maslach & Leiter, 2016; Steel et al., 2015). Such a design seems prudent in light of these findings as well as the potential for forward and reverse relationships of burnout predictors (Bakker & Demerouti, 2017).

A recent study has investigated the Job Demands-Resources model in a sample of mental health professionals using such a longitudinal design (Scanlan & Still, 2019). However, whilst the study's findings were consistent with previous research, and support the application of the Job Demands-Resources model to burnout in mental health professionals, the measures used were unvalidated and largely using single-items per construct, leading to the need for cautious interpretation of their results.

Predicting Burnout

The vast majority of correlational studies have not enabled the development of prognostic models that may enable early interventions to prevent the well-known adverse consequences of occupational burnout. The application of sophisticated analytic techniques, such as machine learning and longitudinal time-series analyses, could propel the field forward to enable early detection and remediation of occupational burnout in mental health professionals. Statistical and machine learning data analysis approaches have historically been seen as distinctive, with the former used for hypothesis testing and theoretical inference, and the latter more commonly associated with predicting future outcomes, with a “black box” approach that lacks the requirement for understanding underlying mechanisms (Bzdok et al., 2018). Statistical methods analyse full data sets, whereas supervised machine learning approaches use techniques such as cross-validation and train-test paradigms. The latter enables estimates of predictive performance using randomly selected holdout data not seen as part of the learning process which can support generalisability as well as highlighting difficulties with overfitting (Pham & Triantaphyllou, 2008).

Machine learning methods are becoming increasingly more visible with utility in theoretical understanding such as variable selection techniques (Azodi et al., 2020; Zou & Hastie, 2005). Conversely, complex statistical techniques that create hierarchical regression models are being utilised for prognostic analysis (Riley et al., 2019).

Variable selection approaches are used when there are several potential predictors, and they enable the removal of redundant or ‘noise’ variables (Galvão & Araújo, 2009; Weisberg, 2005). Statistical approaches to variable selection such as backwards stepwise regression perform less well as the number of variables increases, and may result in important variables being lost (Smith, 2018). Larger numbers of variables also reduce the degrees of freedom for multiple regression approaches leading to significant loss of statistical power (De Gooijer, 2017). In contrast, machine learning techniques cope much better with large numbers of parameters and can model simple and higher-order interactions without the explicit and complex process of defining these using regression techniques (Varian, 2014).

The current study aimed to evaluate organisational and personal predictors of burnout, informed by the Job Demands-Resources model, in a sample of mental health professionals using a longitudinal design and applying cutting-edge time-series and machine learning analyses. The study incorporated multiple factors that have previously been shown to relate to burnout using validated measures from studies of mental health staff and more generic populations. Prognostic model development has the potential to help professionals to self-monitor and to seek support when job demands may have a deleterious impact on their wellbeing and performance, and subsequently this may reduce the impact of clinical staff burnout on service user outcomes.

Aims

The primary aim was to develop a prognostic model of mental health staff burnout using longitudinal data and a range of factors previously associated with staff burnout. The analysis plan aimed to incorporate cross-validated machine learning techniques and multi-level modelling of longitudinal data with the aim of developing a model of that can predict occupational burnout in mental health staff.

A secondary aim was to identify salient factors associated with burnout using machine learning variable reduction techniques and main-effects analysis to inform the design of novel interventions for burnout within mental health service settings.

Methods

Design

A longitudinal observational design was employed with seven monthly online surveys (referred to as Time-1 to Time-7) via the Qualtrics platform (https://www.qualtrics.com). Survey design was adjusted after being piloted at an initial participating site.

Participants

Participants were clinical staff working in six mental health NHS Foundation Trusts across the United Kingdom, covering diverse regions of England including the North West, Yorkshire and Humber and London. Full time students and staff working in physical health services were excluded.

Participant Characteristics

Participants mean age was 42-years (*SD*=11.25), and most were female (79.4%), from a white ethnic background (89.2%), educated to degree level (94.4%) and working in adult (82.6%), community (69.0%), and mental health (63.1%) services. The most frequent professional roles were CBT therapists (17.8%), nurses (18.8%), psychological wellbeing practitioners (16.4%), counsellors (7.3%) and clinical psychologists (7%). Descriptive statistics for participants' (n=287) demographic and work-related variables, Time-1 measures, and comparisons for study completers (n=151) versus drop-outs (n=136) are shown in Appendix A. There were no significant differences between participant characteristics, workload variables or scale measures for completers and drop-outs, indicating no systematic or characteristic related attrition from the study.

Measures and Independent Variables

Participants provided demographic and job characteristics information including role, service, caseload, hours overtime worked and supervision provision. A nine-measure battery was administered at Time-1 as below. Burnout and job satisfaction measures were repeated monthly and the occupational stress measure at Time-7. Higher scores represent greater levels of measured constructs. Cronbach-α values are reported in Table 1 (see Appendix B for variables collected).

Burnout

The Oldenburg Burnout Inventory (OLBI; Demerouti et al. 2001) is a 16-item measure of emotional exhaustion and disengagement dimensions of burnout with improved item design and factorial structure compared to the frequently used Maslach Burnout Inventory (Demerouti et al., 2001; Poghosyan et al., 2009).

Occupational Stress

The Mental Health Professional Stress Scale (MHPSS; Cushway et al., 1996) is a 42-item measure which captures sources of work pressures in healthcare settings. The following sub-scales were administered: workload; client-related difficulties; organisational structures and processes; relationship/conflicts with other professionals; lack of resources.

Social Support

The Social Support Scale (SSS; House & Wells, 1978) is a 6-item measure that captures the levels of work-related support from colleagues, supervisors, spouse/partner and friends/family.

Table 1. Psychometric measure scoring and Cronbach-α values

Subscale	*Item Score Range	Cronbach-α (previous studies)	Cronbach-α (current study)
Oldenburg Burnout Inventory (**OLBI**; Demerouti et al. 2001)			
Exhaustion	1-4	.82	.76
Disengagement	1-4	.83	.80
Mental Health Professional Stress Scale (**MHPSS**; Cushway et al., 1996)			
Workload	0-3	.80	.83
Client-related difficulties	0-3	.80	.73
Organisational structures and processes	0-3	.78	.84
Relationships and conflicts with other professionals	0-3	.79	.79
Lack of resources	0-3	.81	.68
Social Support Scale (**SSS**; Hamaideh, 2011; Jenkins & Elliott, 2004)			
	0-3	>= .84	.89
General Self-efficacy Scale (**GSES**; Shoji et al., 2016)			
	1-4	.75 - .93	.89
Work-family Conflict Scale (**WFCS**; Netemeyer et al., 1996)			
	1-7	.88 - .89	.93
Job Diagnostic Survey (Fields, 2002)			
Autonomy	1-7	.68 - .77	.78
Job Discrepancy and Satisfaction Scale (**JDSS**; Delgadillo et al., 2018)			
	1-4	.75	.77
Overcommitment (Avanzi et al., 2014)			
	1-4	>= .80	.85
Big Five Inventory-10 (Lovik et al., 2017)			
Extraversion	1-5	.42	.64
Agreeableness	1-5	.09	.41
Conscientiousness	1-5	.39	.42
Neuroticism	1-5	.62	.55
Openness	1-5	.28	.23

See reference next to measure name for previous Cronbach-α values article.
*All scale/sub-scale scores are presented as mean of item scores.

Self-Efficacy

The General Self-Efficacy Scale (GSES; Schwarzer & Jerusalem, 1995) is a 10-item scale designed to capture an individual's beliefs about their ability to cope with challenges.

Work-Family Conflict

The Work-Family Conflict Scale (WFCS; Netemeyer et al., 1996) is a 5-item scale designed to capture conflict between work roles and family responsibilities.

Autonomy

The Job Diagnostic Survey (Hackman & Oldham, 1974) contains four items relating to job autonomy, i.e. the level of discretion to choose how to do work.

Job Satisfaction

The Job Discrepancy and Satisfaction Scale (JDSS; Nagy, 2002) is an 8-item scale which captures how satisfied an individual is with their role, including salary, promotion and supervision.

Overcommitment

The Overcommitment subscale of the Effort-Reward Imbalance Questionnaire (Siegrist et al., 2014) is a 6-item scale that captures an individual's ability to separate professional roles from personal life.

Personality

The Big Five Inventory-10 (BFI-10; Rammstedt & John, 2007) is a 10-item scale derived from the 44-item measure of personality capturing Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness. This measure is suitable where limited time is a significant factor and investigations of personality are not the main focus of the study (Rammstedt & John, 2007). An additional item is available to improve the validity of the Agreeableness sub-scale making this 11-items in total.

Job Changes

Participants who registered significant role changes rated whether their workload, responsibility and hours had increased, decreased, or stayed the same. Summed scores represented a simple job-demands change score

Procedure

The study was advertised in participating sites using the same method, including internal Communications emails, emails from individual team managers and principal investigators (Appendix C). Three sites recruited across all clinical services with the others recruiting within Improving Access to Psychological Therapies (IAPT) teams.

Participants accessed the initial survey via site-specific hyperlinks. Study participant information (Appendix D) was provided and consent sought electronically (Appendix E) with eligibility screening. Consenting participants provided an email address to enable monthly survey links to be emailed. Demographic, role, service information and psychometric measures data were collected. Short, monthly surveys were completed via links sent using automated emails. Additional reminder emails were sent after 7-10 days if required. The surveys were launched across participating sites between April-August 2019 and recruitment closed in September 2019. Longitudinal data collection ended in March 2020.

Ethical Considerations and Governance

Ethical approval for the study was obtained from the NHS Health Research Authority via proportionate review (Appendix F).

Power Calculations and Sample Size Considerations

Minimum sample sizes were estimated (G*Power software version-3.1.9.2) based in previous reported R^2 values of >=.25 (f^2=0.33) (Cohen, 1992;Kilfedder et al., 2001; Müller et al., 2018; Westwood et al., 2017). A cautious, minimum sample size of 203 for 36 independent variables was estimated (f^2=0.15;α=.05;80% power). When adjusted for planned 80%-20% train-test paradigms, 10-fold cross-validation, and maximum attrition of 60% based on previous studies, a minimum sample size of 708 was required to retain power (see Appendix G; Robins et al., 2017). This represented a recruitment rate of approximately 28% which was feasible based on initial site staff numbers and previous recruitment to burnout studies (Kilfedder et al., 2001; Westwood et al., 2017).

Data Analysis

Data was analysed using R statistical software (www.R-project.org, version 3.6.3) and Statistical Package for the Social Sciences (SPSS; IBM Corporation, version 25).

Missing Data

Analysis of repeated measures data using Little's test indicated that this was not missing at random, $\chi^2(2)=11.1$, $p<.005$. To allow for the data being missing not at random, a pattern mixture model approach with multi-level multiple imputation was used (Grund et al., 2016; Son et al., 2012). Multiple imputation by chained equations analysis (MICE; van Buuren & Groothuis-Oudshoorn, 2011) was used to impute personality data that were missing at random.

Variable Reduction

The R SuperLearner package was used to compare a set of different machine learning approaches to build classification models of burnout at Time-1 (cross-sectional) and Time-7 (longitudinal) using Time-1 independent variables. For the Time-7 models, OLBI-exhaustion, and OLBI-disengagement at Time-1 were included as predictors. The machine learning algorithms were selected as recommended by the SuperLearner package authors (Eric et al., 2019). These were as follows: Elastic-net regression; Random Forest decision trees; Bayesian additive regression trees (BART); Support vector machine with sequential minimal optimisation; neural networks (Friedman et al., 2019; Kapelner & Bleich, 2016; Karatzoglou et al., 2019; Liaw & Wiener, 2018; Ripley, 2020). Logistic regression and mean models were used as baseline comparators. Previous studies classified high burnout using OLBI-exhaustion and OLBI-disengagement cut-off scores of ≥2.25 and ≥2.1 respectively (Westwood et al., 2017). Nearly 80% of the sample were classified as burned-out using these cut-offs. Instead, cut-offs were calculated as mean, plus one standard deviation providing cut-offs for OLBI-exhaustion and OLBI-disengagement of >3.26 and >2.87 respectively. This approach is more conservative, sensitive to the characteristics of the specific study sample, and has shown significant effects of staff burnout on patient outcomes (Delgadillo et al., 2018). SuperLearner machine learning was configured to perform 10-fold cross-validation (repeated training/testing on randomly selected subsets of the full data sample; Rodríguez et al., 2010). Stratification balanced proportions of burned-out classification within each fold. Interval and ratio variables were mean-centred and standardized and categorical variables transformed into multiple binary dummy variables (n-1 per variable). SuperLearner was configured to maximise the Receiver Operating Characteristic area under curve (ROC AUC) and models not significantly different from the best performing model with moderate or above AUC scores (>.70) were selected (Akobeng, 2007). The packages underlying SuperLearner were used to perform a train-test paradigm with a 2:1 sample ratio respectively. Synthetic minority oversampling (SMOTE) balanced the dataset to approximately equal proportions of burned-out classification enabling more balanced models (Chawla et al., 2011). Model comparison used AUC and other performance metrics (Parikh et al., 2008).

Consensus selection

Mean variable importance score, weighted by AUC above chance, was calculated for 100 model instances. Variables scoring above mean importance scores underwent a consensus vote. Two additional feature selection algorithms were used from the same machine learning approaches, giving a maximum of five votes per variable (Bleich et al., 2014; Kursa & Rudnicki, 2010). Independent variables with a consensus vote of three or more were selected for further analysis.

Main Effects Analysis

Main effects analysis of Time-1 consensus-selected variables was performed using conventional backward stepwise removal of non-significant variables from binomial logistic regression models for OLBI-exhaustion and OLBI-disengagement classification.

Prognostic Multi-Level Modelling

Variables selected by the consensus method for Time-7 were entered as fixed effects into longitudinal multi-level models (MLM) for OLBI-exhaustion and OLBI-disengagement. Variable slopes and intercepts with first order autoregressive residuals were modelled using a log-linear time series. Lower -2log-likelihood was evaluated as representing better goodness-of-fit.

Predictor variables (from the prior step) that were no longer statistically significant in the MLM were removed, until a parsimonious model was obtained.

The equation derived from the final model was then tested as a classifier against completed Time-7 data. Performance was compared to the best modelling method from the previous stage, trained with the same set of consensus-selected variables, including Time-1 OLBI-exhaustion and OLBI-disengagement scores.

Results

Participants

The survey was viewed by 454 staff across six sites. Numbers of consenting participants who completed data at Time-1 and met inclusion criteria are detailed in Figure 1. Overall, the total sample of consenting and eligible participants included in analysis was *n*=287. The drop-out rate of participants who did not complete Time-7 measures was 47.4% with 151 participants completing these (Figure 1).

Missingness

Most missing data related to participants' drop-out by not completing follow-up surveys with 54.4% of participants missing one or more. Nine participants did not fully complete the personality items with one participant not answering any of the questions. Twelve job satisfaction measures were not completed.

Participant Burnout

Of the participant and role characteristics (e.g. gender, education, role, service information), only ethnicity and gender showed any relationship with burnout levels with males having higher levels of OLBI-disengagement than females (Table 2). There was, however, no significant difference in gender for participants classified as burned-out or not, although a higher proportion of males were classified as burned-out compared to females. Participants from an ethnic minority background had significantly higher levels of OLBI-exhaustion and OLBI-disengagement and greater proportions of burned-out classification compared to participants with a white ethnic background. OLBI scores were not significantly different between different sites.

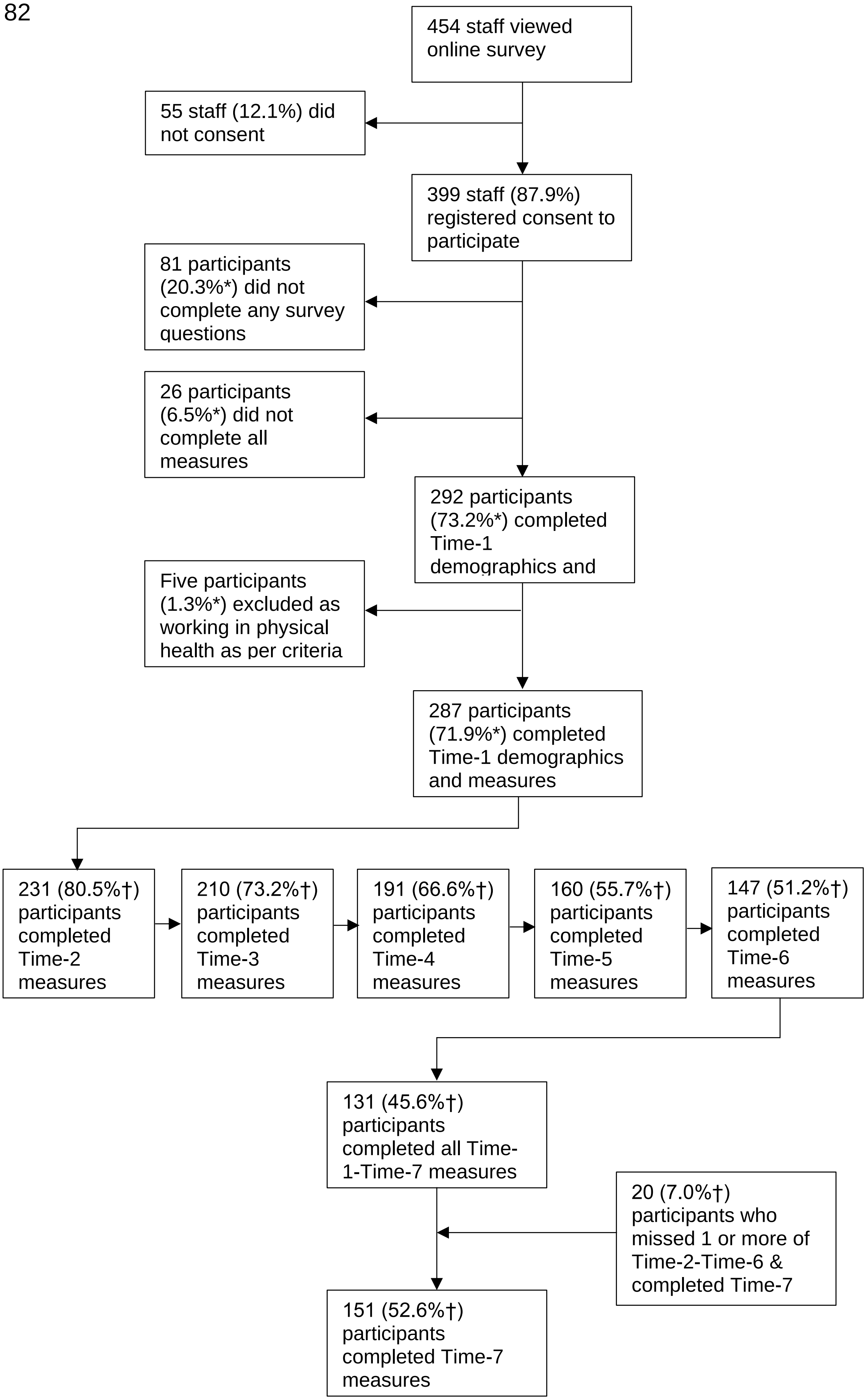

Figure 1. Recruitment and participant flow diagram.
* Percentage of consenting participants. † Percentage of participants meeting inclusion criteria. Drop-out rates were not significantly different between sites, $\chi^2(5,438)=3.00$, $p=.70$

Table 2. Burnout measure mean scores (*SD*) and burnout classification percentages (*n*) at Time-1 for gender and ethnicity

	Gender			Ethnicity		
	Male (n=59)	Female (n=228)	Statistic (*F* or χ^2)	White (n=256)	Other (n=31)	Statistic (*F* or χ^2)
OLBI-exhaustion	2.76 (0.47)	2.78 (0.48)	0.05, *p*=.82	2.75 (0.47)	3.00 (0.57)	7.37, *p*=.007
OLBI-disengagement	2.52 (0.49)	2.39 (0.45)	4.14, *p*=.043	2.39 (0.45)	2.60 (0.55)	5.97, *p*=.015
Burned-out classification	30.5% (18)	20.6% (47)	2.62, *p*=.106	20.3% (52)	41.9% (13)	7.38, *p*=.007

OLBI; Oldenburg Burnout Inventory subscales

The percentage of participants classified as burned out at Time-1 was 22.6% and at Time-7 was 26.4%. More participants were classified as burned-out due to disengagement than exhaustion at Time-1 and Time-7 (Table 3).

Table 3. Burnout subscale classification proportions for participants classified as burned-out at Time-1 and Time-7

Time Point	OLBI Exhaustion	OLBI Disengagement	Both OLBI dimensions
Time-1	20.0%	52.3%	27.7%
Time-7	15.0%	50.0%	35.0%

OLBI; Oldenburg Burnout Inventory subscales

Longitudinal Relationships

JDSS had moderate-to-large negative correlations with OLBI-exhaustion and OLBI-disengagement at every time point (Table 4). In addition, there were small-to-moderate correlations in JDSS score differences between time points (e.g. JDSS at Time-2 minus JDSS at Time-1) and OLBI score difference between the same time points (δ-OLBI) (Table 4). There was a large correlation between JDSS change and OLBI change between Time-1 and Time-7, *r*=.50, *p*<.0001, indicating a potentially causal relationship between burnout and job-satisfaction.

Table 4. Longitudinal job satisfaction correlations with burnout subscales (Time-1-Time-7) for absolute and change between time-point scores

	Time						
	1 (n=287)	2 (n=227)	3 (n=208)	4 (n=189)	5 (n=158)	6 (n=144)	7 (n=151)
OLBI-exhaustion	-.46	-.36	-.46	-.40	-.44	-.46	-.48
OLBI-disengagement	-.61	-.53	-.59	-.60	-.62	-.57	-.58
δ-OLBI vs. δ-JDSS*	*na*†	-.27	-.35	-.29	-.22	-.29	-.43

All correlations were significant at $p < .0001$; *measures changes between time points; †time point differences calculated from Time-2 onwards. OLBI-Oldenburg Burnout Inventory; JDSS-Job Discrepancy and Satisfaction Scale

There was also a large significant correlation between change in total MHPSS score and change in total OLBI score between Time-1 and Time-7, r=.50, p<.0001. Correlations with OLBI-exhaustion (r=.47, p<.0001) and OLBI disengagement (r=.44, p<.0001) were very similar. This indicates occupational stress impacts both exhaustion and disengagement burnout dimensions.

Whilst overall levels of burnout were stable at a sample level, the number of participants classified as burned-out increased in size by 33% from Time-1(n=30) to Time-7(n=40) in the 151 participants who completed the study. In the same group, 50% (n=20) of those classified as burned-out at Time-7 had not been classified as burned out at Time-1.

Thirty-two participants reported that their roles had changed significantly in one or more of the monthly surveys. Most changes (64.3%) were classed as an increase in job demands (i.e. contracted hours, workload, or responsibility). No significant correlation between changes in job demands and δ-OLBI$_{(\text{Time-7-Time-1})}$ scores was observed, *tau b*(30)=.06, p=.68.

Cross-Sectional Associations Between Burnout and Independent Variables at Time-Point One

At Time-1 there were several significant correlations between independent variables and burnout measures as well as between different independent variables (Figure 2). Correlations are positive unless otherwise stated. Of note were large correlation coefficients (r >.5; Cohen, 1988) for JDSS (negative), MHPSS workload, WFCS and OCI with OLBI-exhaustion and moderate (r >.3) correlations for the remaining MHPSS sub-scales, neuroticism, GSES and autonomy (latter two both negative) with OLBI-exhaustion. There were also large intercorrelations between MHPSS domains.

There was a large negative correlation between JDSS and OLBI-disengagement, and moderate negative correlations for autonomy and GSES with OLBI-disengagement. MHPSS workload and organisational stress subscales had moderate positive correlations with OLBI-disengagement as did OCI. GSES and SSS supervisor support had negative moderate correlations with OLBI-disengagement. There were large correlations between OLBI-exhaustion and OLBI-disengagement as expected.

Machine Learning Analysis Using Cross-Sectional Data at Time-Point One

SuperLearner classification algorithms trained using all Time-1 predictor variables (no OLBI scores; no missing data, n=277) against the burnout outcome resulted in three machine learning algorithms performing best, based on highest AUC scores (the null hypothesis that the specific learner was the best performer was tested, Table 5). The Random Forest algorithm had the highest AUC score followed closely by Elastic-net and BART.

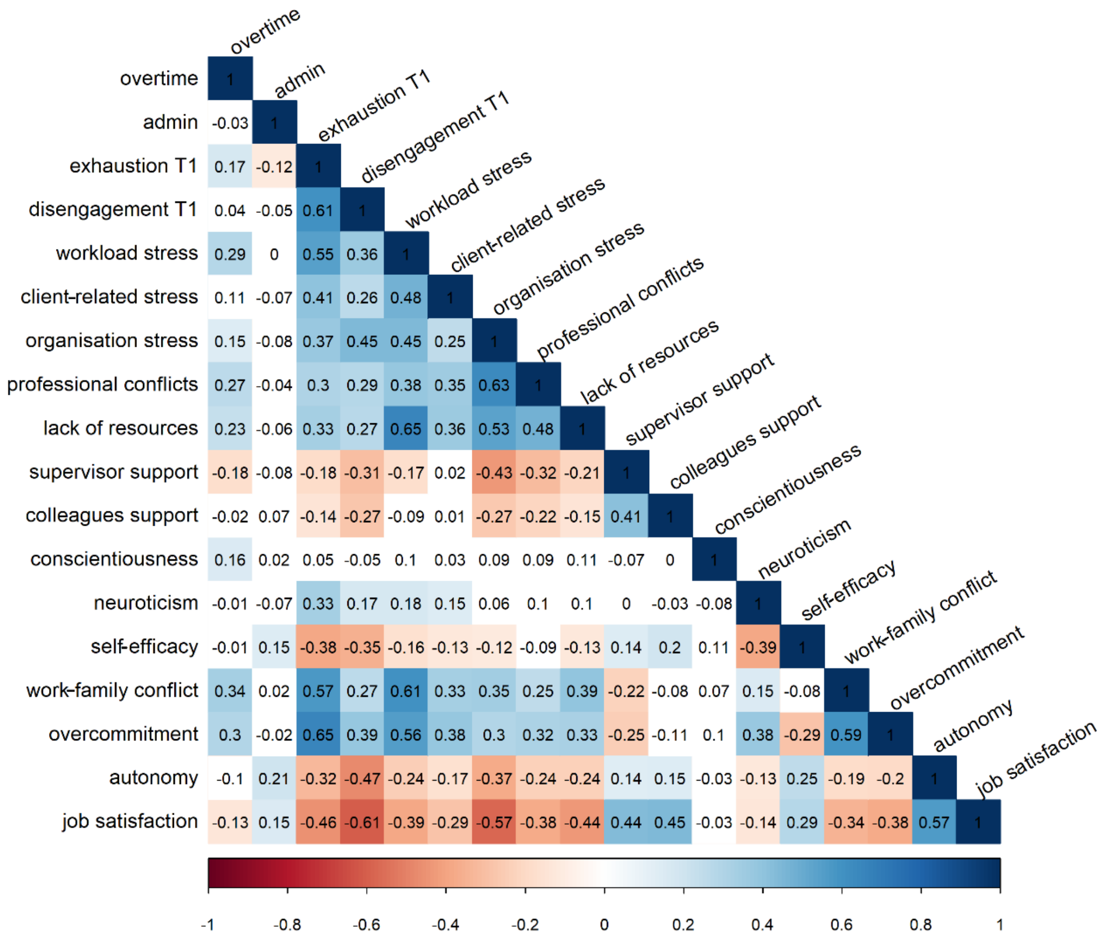

Figure 2. Correlation matrix of independent and burnout variables.
Variables with no significant correlations with burnout (either directly or indirectly) are removed. Stronger colours represent larger correlations, blue = +ve, red= -ve; white background = not significant, *p*>=.05.

Inclusion of imputed missing data for 14 data points in the BF-10 personality scale data (n=287) produced very similar AUC scores for the same three best performing algorithms, topped by BART (Table 6).

Table 5. SuperLearner algorithm AUC performance on complete data (n=277)

Algorithm	AUC	SE	95% CI	*p**
Mean	.50	.07	[.36; .64]	<.0001
Neural network	.54	.07	[.40; .67]	<.0001
Logistic regression	.67	.04	[.59; .75]	<.001
Support vector machine	.75	.03	[.69; .82]	<.05
BART	.79	.03	[.73; .85]	.26**
Elastic-net	.80	.03	[.73; .86]	.39**
Random Forest	.81	.03	[.75; .86]	.50**

*1-tailed test for a null hypothesis that the model is the best performer; **null hypothesis retained at α=0.05. AUC=Area under Receiver Operator curve; SE=Standard error; CI=Confidence intervals. BART=Bayesian additive regression tree

Table 6. SuperLearner algorithm AUC performance on imputed data (n=287)

Algorithm	AUC	SE	95% CI	*p**
Mean	.50	.07	[.36; .64]	<.0001
Neural network	.63	.06	[.52; .74]	<.001
Logistic regression	.67	.04	[.59; .74]	<.0001
Support vector machine	.76	.03	[.70; .82]	<.05
Elastic-net	.79	.03	[.74; .85]	.14**
Random Forest	.81	.03	[.76; .86]	.34**
BART	.82	.03	[.77; .87]	.50**

*1-tailed test for a null hypothesis that the model is the best performer; **null hypothesis retained at α=0.05. AUC=Area under Receiver Operator curve; SE=Standard error; CI=Confidence intervals. BART=Bayesian additive regression tree

Similar classification performance was seen using these three algorithms independent of the SuperLearner framework (Figure 3), which enabled the most accurate classifiers to be used as part of the consensus variable selection process.

Models trained with the raw training set were strongly biased towards specificity rather than sensitivity (Figure 3) indicating that the majority classifier (i.e. not burned-out) biased model learning. A minority oversampling technique which synthesised additional burned-out data using a nearest-neighbours approach, resulted in near-equal proportions of burned-out and not-burned-out cases (51.7% burned-out; n=317). This resulted in large increases in sensitivity with a smaller cost of decreased specificity for the best performing algorithm, BART (Figure 3). Oversampled models had lower accuracy and Positive Predictive Values (PPV) due to increases in false positives; however, Negative Predictive Values (NPV) were increased with reduced false negatives.

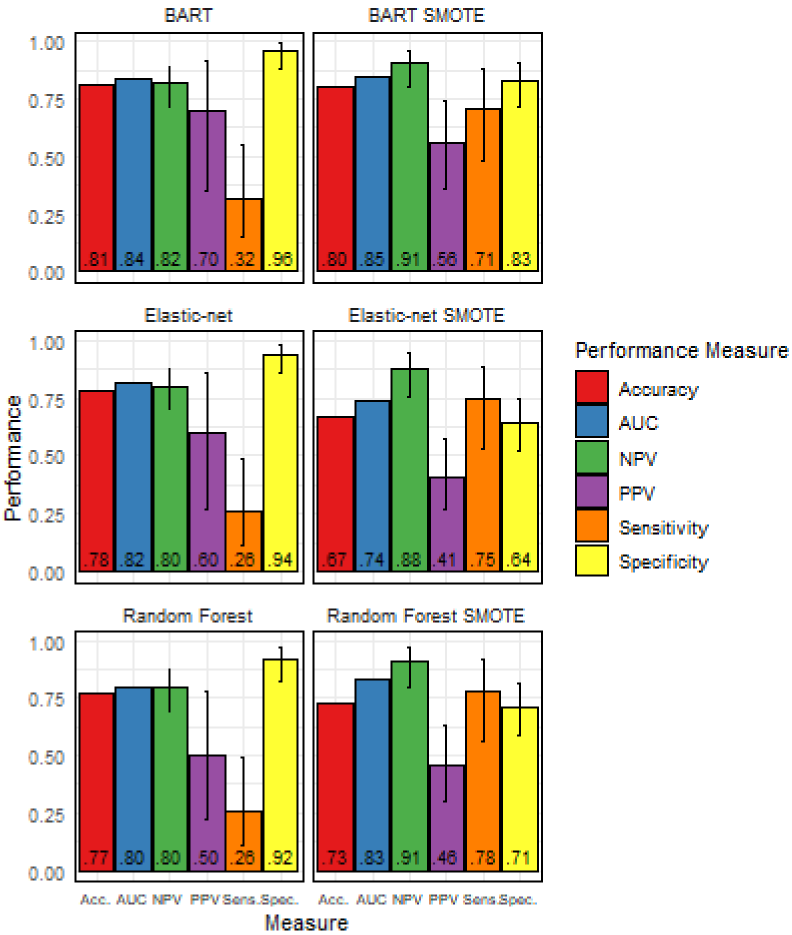

Figure 3. Algorithm performance measures with raw and minority oversampled data
BART- Bayesian additive regression trees. SMOTE-Synthetic minority oversampling technique. AUC-area under Receiver Operating Characteristic curve. PPV-Positive predictive value. NPV-Negative predictive value.

Main Effects Analysis Using Logistic Regression

Selection of the most important variables from these three optimised models in combination with two of their feature selection methods identified 14 independent Time-1 variables with a consensus vote (Table 7). This variable reduction enabled main effects analysis using logistic regression for OLBI-exhaustion (Table 8) and OLBI-disengagement (Table 9) and resulted in seven significant predictors of Time-1 burnout. Autonomy appeared in both models, with workload-related stress, GSES and overcommitment predicting OLBI-exhaustion and organisation-related stress, JDSS and conscientiousness predicting OLBI-disengagement. The final parsimonious model built using these seven independent variables had the highest sensitivity of all models (.92) with moderate specificity (.71; Figure 4)

Table 7. Consensus variable selection for Time-1 predictors with more than one vote

Variable	Random Forest Variable Importance	Random Forest Feature Selection	BART Variable Importance	BART Permute Selection	Elastic-net Variable Importance	Consensus Vote Score*
JDSS	✓(1)	✓	✓(2)	✓	✓(1)	5
Autonomy	✓(2)	✓	✓(1)	✗	✓(2)	4
MHPSS	✓(4)	✓	✓(18)	✗	✓(17)	4
Organisation-related stress†	✓(3)	✓	✓(3)	✗	✓(3)	4
Overcommitment	✓(5)	✓	✓(9)	✓	✓(5)	5
Workload-related stress†	✓(6)	✓	✓(6)	✗	✗	3
GSES	✓(9)	✓	✓(5)	✓	✓(9)	5
Client-related stress†	✓(10)	✓	✓(11)	✗	✓(7)	4
SSS	✓(11)	✓	✓(21)	✗	✗	3
Admin time (hrs)	✓(18)	✗	✓(15)	✗	✓(11)	3
Lack of resources†	✓(16)	✓	✓(9)	✗	✗	3
Neuroticism	✓(13)	✗	✓(12)	✗	✓(13)	3
Conscientiousness	✓(14)	✗	✓(4)	✗	✓(10)	3
WFCS	✓(8)	✓	✗	✗	✓(15)	3
Months in role	✓(23)	✗	✗	✗	✓(14)	2

* Variables selected with consensus majority score of 3 or more. ✓-variable identified, ✗-variable not identified, number in parenthesis-importance rank where relevant. JDSS-Job Discrepancy Satisfaction Scale; MHPSS-Mental Health Professional Stress Scale; GSES-General Self-efficacy Scale; SSS-Social Support Scale; WFCS-Work-Family Conflict Scale

Table 8. Parsimonious binomial logistic regression model for OLBI-exhaustion burned-out classification

	Estimate	Std. Error	z value	*p*
Intercept	-4.2238	0.5767	-7.324	<.0001
Workload-related stress†	1.5897	0.4429	3.589	<.001
GSES	-0.9105	0.3088	-2.949	.003
Overcommitment	0.9367	0.3726	2.514	.011
Autonomy	-0.9645	0.2868	-3.364	<.001

† Mental Health Professional Stress Scale (MHPSS) sub-scale; GSES-General Self-efficacy Scale

Table 9. Parsimonious binomial logistic regression model for OLBI-disengagement burned-out classification

	Estimate	Std. Error	z value	*p*
(Intercept)	-2.0582	0.2259	-9.112	<.0001
Organisation-related stress †	0.5374	0.2077	2.588	.010
Autonomy	-0.6469	0.2029	-3.189	.001
JDSS	-0.5673	0.2514	-2.256	.024
Conscientiousness	-0.3393	0.1702	-1.993	.046

† Mental Health Professional Stress Scale (MHPSS) sub-scale; JDSS-Job Discrepancy Satisfaction Scale

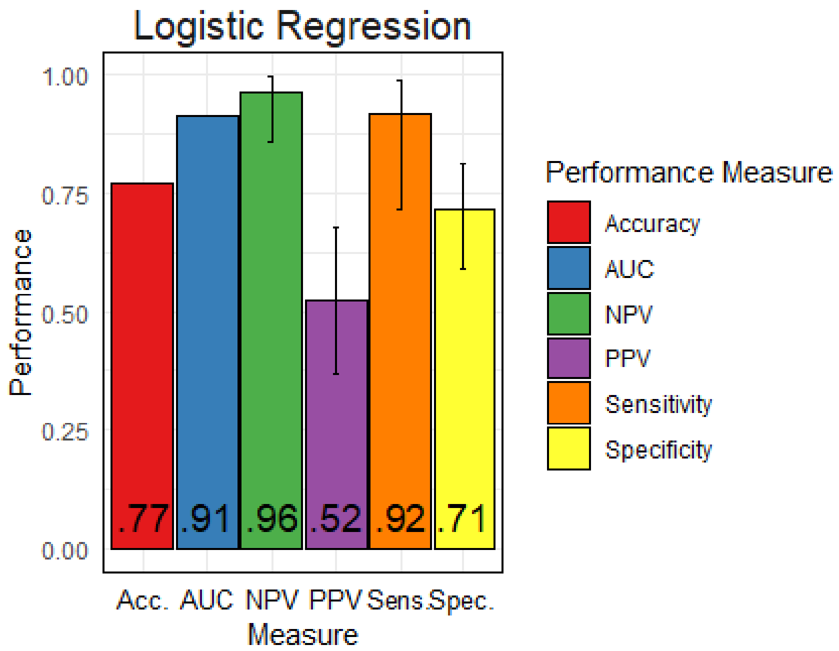

Figure 4. Logistic regression model performance for parsimonious model independent variables. AUC-area under Receiver Operating Characteristic curve. PPV-Positive predictive value. NPV-Negative predictive value.

Prognostic Model Performance for Time-7 Burnout Prediction using Time-1 Measures

At Time-7 there were several significant correlations between Time-1 candidate predictor variables and OLBI-exhaustion and OLBI-disengagement (Figure 5). JDSS and GSES had moderate negative correlations with both OLBI-exhaustion and OLBI-disengagement. Neuroticism, WFCS and OCI had moderate correlations with OLBI-exhaustion but not with OLBI-disengagement. MHPSS workload, client difficulties and resources subscales had moderate correlations with OLBI-exhaustion. There were large correlations between Time-1 and Time-7 OLBI scores. The effect of gender seen at Time-1 was also significant at Time-7 with higher levels of OLBI-disengagement in males (M=2.64,SD=0.52) compared to females (M=2.41,SD=0.52), F=(1,149)=4.47, p=.036.The proportion of males classified as burned-out was also significantly higher (42.9% vs. 22.8%), χ^2(1,N=151)=4.73, p=.03. No significant differences were seen relating to ethnicity and burnout measures at Time-7. MHPSS scores measured at Time-7 resulted in very similar correlation coefficients with OLBI-exhaustion and OLBI-disengagement as seen at Time-1 indicating a consistent relationship with burnout over time.

Longitudinal Imputation.

Missing data, imputed using a longitudinal multi-level model (MLM), was evaluated using randomly removed data (30%) at Time-7 for OLBI-exhaustion and OLBI-disengagement. This method of imputation led to only 29% of removed data being accurately imputed based on burnout classification.

Imputed data was also a poor substrate for machine learning models with the highest performing classifier (SVM), only attaining an AUC of .58 (Figure 6.) which was not significantly above chance (i.e. .5) to consider for further modelling analysis. On this basis, all subsequent analyses were carried out using the sample of cases with complete data, since imputed data was error-prone, and no significant baseline differences between completers vs. dropouts were found.

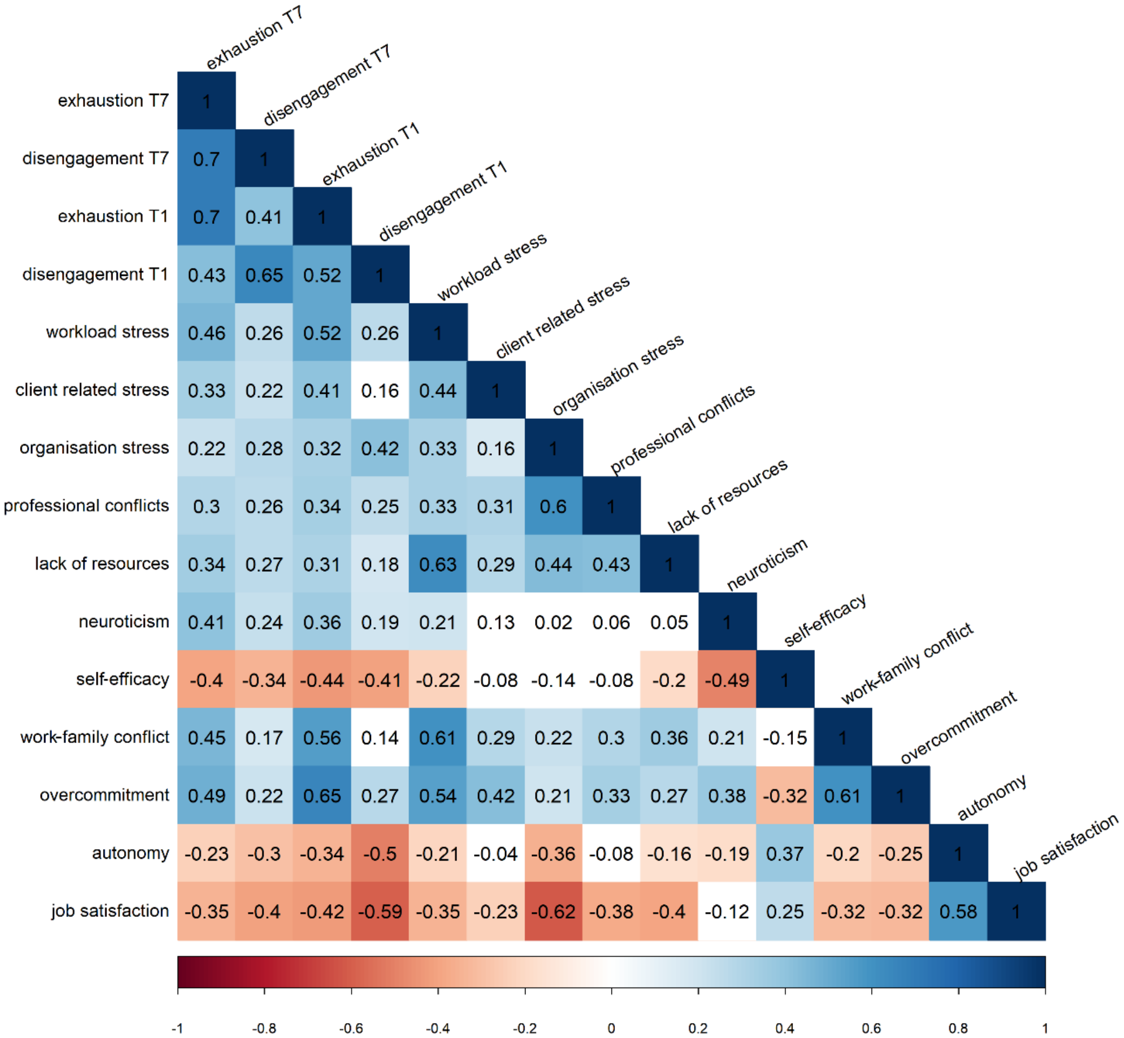

Figure 5. Correlation matrix of predictor and Time-7 burnout
Predictors with no significant correlations are removed. Stronger colours represent larger correlations, blue = +ve, red= -ve; white background = not significant, *p*>=.05.

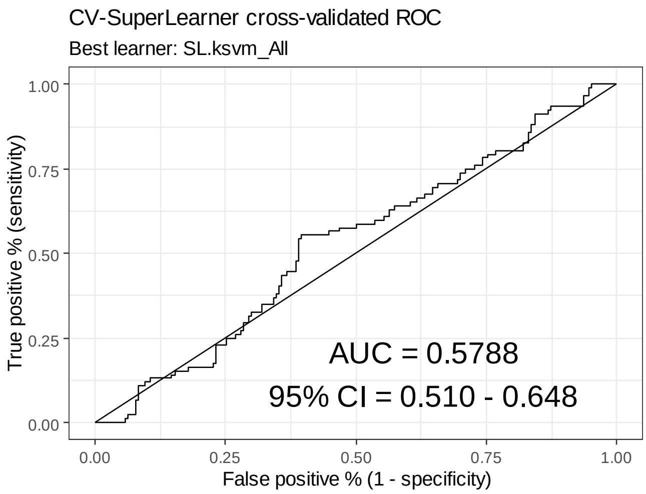

Figure 6. ROC curve for SuperLearner imputed data set best model – support vector machine

SuperLearner Model Ranking for Time-7 Burnout Classifiers

The variable selection process was repeated for Time-7 burnout classification models. The BART classifier model was again the best performer (Figure 7) followed by Elastic-net and Random Forest models with AUC scores >.70 (Table 10).

Table 10. SuperLearner classification scores for imputed BF-10 data set (n=151)

Algorithm	AUC	SE	95% CI	*p**
Mean	.50	.09	[.32; .68]	<.001
Neural network	.58	.07	[.45; .72]	<.01
Logistic regression	.64	.07	[.50; .79]	.07**
Support vector machine	.70	.05	[.61; .79]	.14**
Random Forest	.71†	.05	[.62; .80]	.18**
Elastic-net	.75†	.04	[.66; .84]	.48**
BART	.75†	.05	[.66; .84]	.50**

*1-tailed test for a null hypothesis that the model is the best performer; **null hypothesis retained at α=0.05. † AUC above cut-off. AUC=Area under Receiver Operating Characteristic curve; SE=Standard error; CI=Confidence intervals. BART=Bayesian additive regression tree

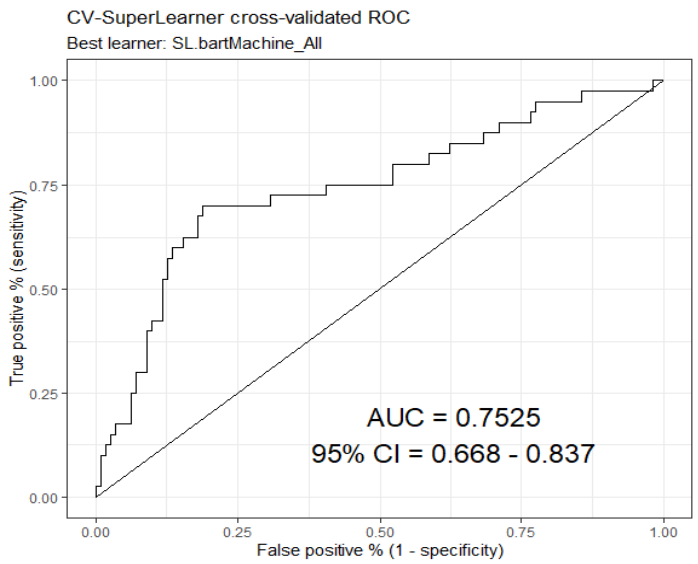

Figure 7. ROC curve for SuperLearner BART model complete data set (n=151)

In a sensitivity analysis, the cross-validated SuperLearner training, using a non-negative binomial likelihood maximization method (LeDell et al., 2016) as an alternative optimiser for binary outcomes, revealed the same order of performance metrics as the AUC optimizer. This indicates that the primary findings were robust and stable.

BART Model Optimisation

There were similar findings to the Time-1 analysis with class imbalanced raw data favouring specificity rather than sensitivity. The AUC value (.81) was higher than the SuperLearner best performing model (Figure 8). Increasing proportions of samples classified as burned-out using a majority undersampling technique produced high sensitivity and low specificity models with similar AUC values (Figure 9). Minority oversampling balanced the proportions of burned-out cases and produced more balanced models with moderate levels of sensitivity and specificity and good AUC values (Figure 8).

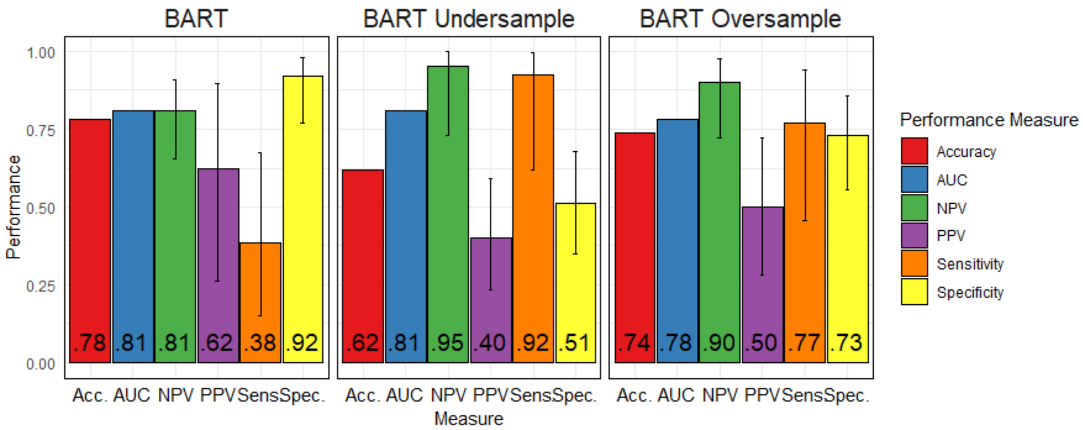

Figure 8. BART performance measures with raw, under-sampled and over-sampled data. AUC-area under Receiver Operating Characteristic curve. PPV-Positive predictive value. NPV-Negative predictive value.

Variable Selection for the longitudinal prediction of Burnout at Time-7

Repeating the consensus approach to variable selection identified eight predictor variables measured at Time-1, with a majority vote out of 12 candidate variables that obtained more than one vote (Table 11). All variable importance methods with ranked scores resulted in OLBI-disengagement as the most important variable above OLBI-exhaustion. An example variable importance plot is shown in Figure 9.

Table 11. Consensus variable selection for Time-1 predictors with more than one vote for burnout prediction at Time-7

Variable	Random Forest Variable Importance	Random Forest Feature Selection	BART Variable Importance	BART Permute Selection	Elastic-net Variable Importance	Consensus Score*
JDSS	✓(5)	×	✓(10)	✓	×	3
Autonomy	✓(4)	✓	✓(1)	✓	✓(12)	5
MHPSS	✓(10)	✓	✓(11)	✓	×	4
Overcommitment	✓(20)	✓	×	×	×	2
GSES	✓(7)	✓	✓(5)	✓	✓(15)	5
Client-related stress	✓(11)	✓	×	×	✓(7)	3
Neuroticism	✓(6)	✓	✓(3)	✓	✓(6)	5
Openness	✓(24)	×	×	✓	×	2
WFCS	✓(3)	✓	✓(6)	×	✓(5)	4
Children	×	×	✓(8)	×	✓(14)	2
Caseload	✓(8)	×	✓(7)	×	×	2
Community Services	×	×	✓(2)	✓	✓(2)	3

* Variables selected with consensus majority score of 3 or more. ✓-variable identified, ×-variable not identified, number in parenthesis-importance rank where relevant. JDSS-Job Discrepancy Satisfaction Scale; MHPSS-Mental Health Professional Stress Scale; GSES-General Self-efficacy Scale; WFCS-Work-Family Conflict Scale

Prognosis Using Reduced Variables with Multi-level Modelling and BART

The seven variables selected by consensus were then included in longitudinal multi-level models and subsequently optimised to develop a prognostic model of burnout. For OLBI-exhaustion, two variables were retained with significant predictive value in addition to Time-1 OLBI-exhaustion; and for the OLBI-disengagement prediction model only OLBI-disengagement at Time-1 was retained (Table 12).

Table 12. Final Multi-Level Models for OLBI-exhaustion and OLBI-disengagement

Variable	Estimate	SE	*p* value
OLBI-exhaustion			
Intercept	.511	.119	< .0001
Time(log)	-.005	.013	.70
OLBI-exhaustion Time-1	.814	.030	< .0001
JDSS Time-1	-.046	.024	.048
Neuroticism	.036	.014	.009
OLBI-disengagement			
Intercept	.264	.067	< .0001
Time(log)	.025	.013	.062
OLBI-disengagement Time-1	.886	.027	< .0001

Both models included a first-order autoregressive covariance matrix which was significant (p <.0001). SE-standard error.

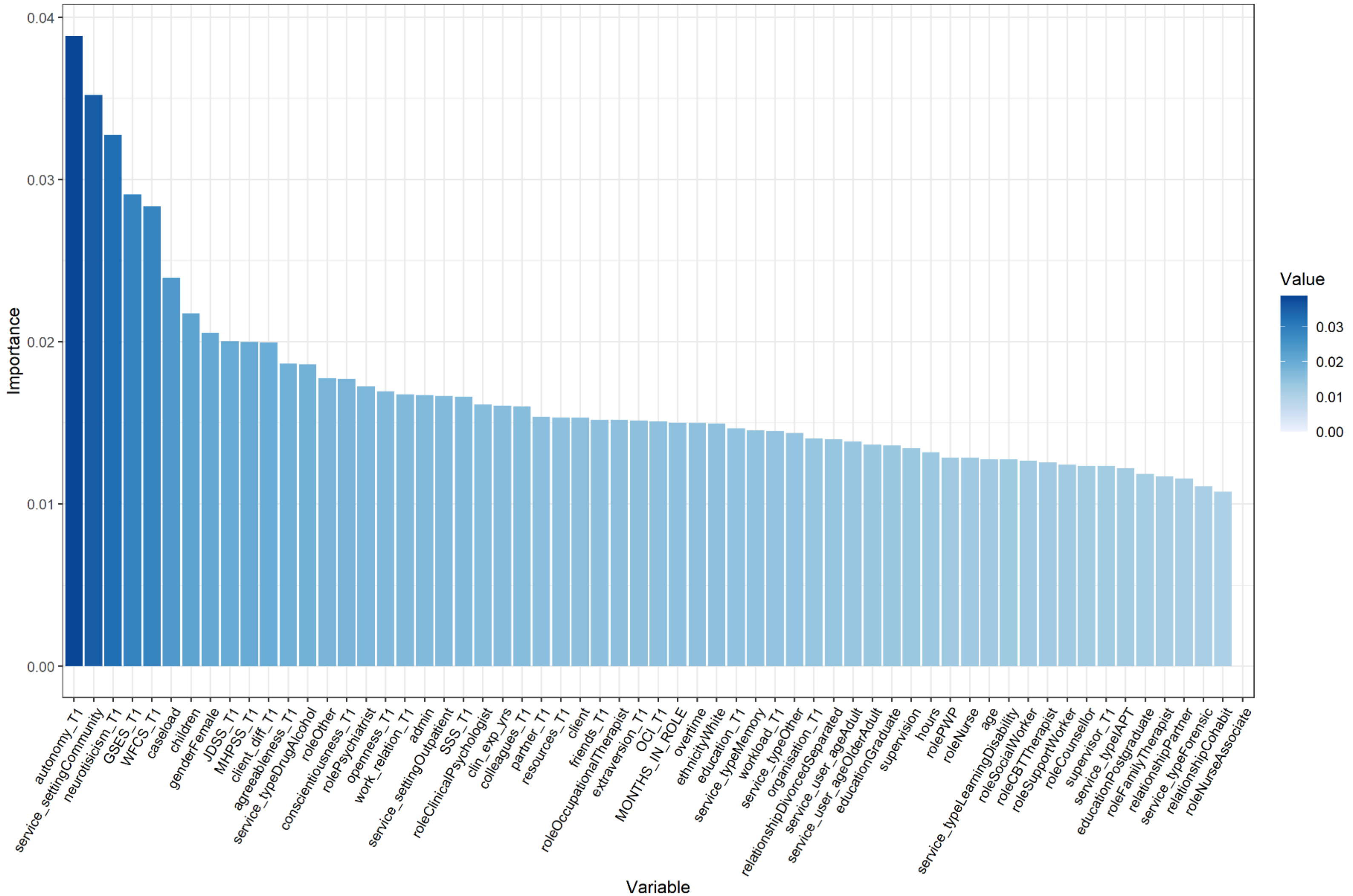

Figure 9. Example variable importance plot (BART) showing model variable importance (y-axis) against predictor variable (x-axis)

Model performance was tested by applying the multi-level models to the complete data set with burnout classifications at Time-7 (n=151). Sensitivity was low (.40) with high specificity (.93), with the ability to identify burned-out participants below chance (Figure 10). BART was also trained using the same consensus-selected variables and the resulting model had a good AUC (.81), almost identical accuracy, significantly higher sensitivity (.71) and a much smaller drop in specificity (.81).

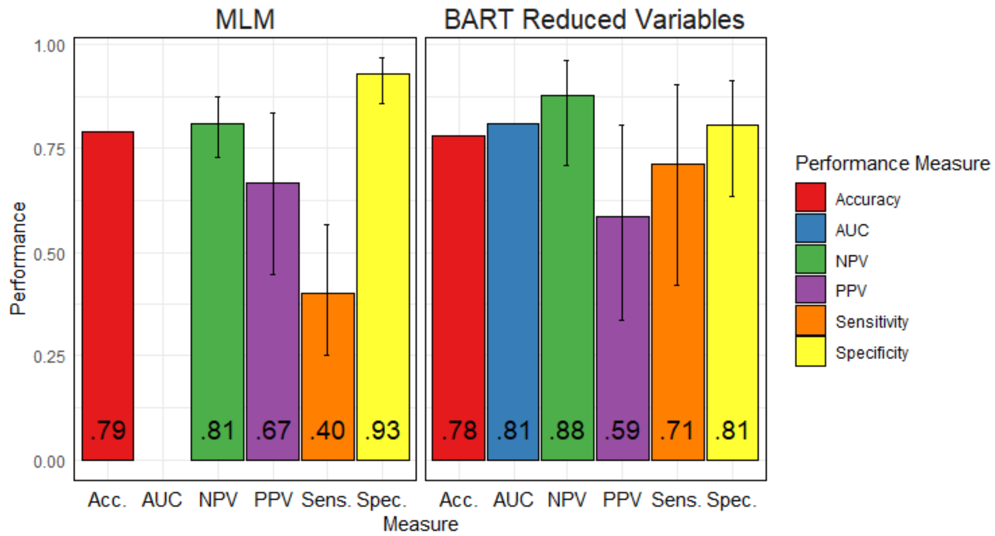

Figure 10. Comparison of MLM and BART model performance
AUC-area under receiver operating characteristic curve. PPV-Positive predictive value. NPV-Negative predictive value.

Overfitting, No-Information-Rate Tests and Sensitivity Analysis

The importance of the train-test paradigm was demonstrated by training the BART classifier with 10 random variables (zero mean-centred with normal distribution) against burnout classification at Time-7 (n=151). This created a close to No-Information-Rate model at the default .5 classification probability cut-off (Figure 11). Optimising this model for best cut-off (.68), with sensitivity and specificity weighted equally, resulted in a trained model that appears highly accurate. In contrast, binomial logistic regression creates a poor model in these conditions with AUC=.39 and non-significant random variable *p* values ranging from .4 to .93.

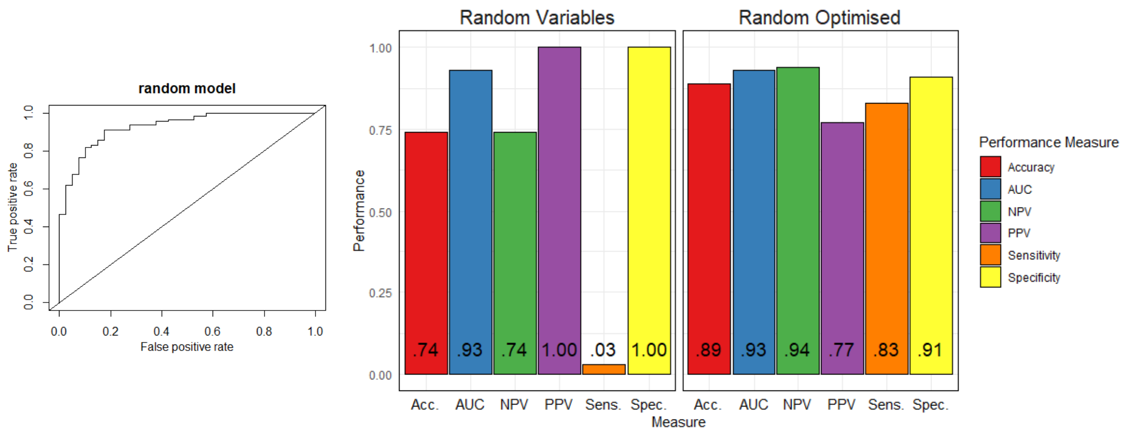

Figure 11. BART model overfitting with random variables. Left-to-right: Model ROC curve; No-Information-Rate model at default cut-off; Optimised random model.
AUC-area under receiver operating characteristic curve. PPV-Positive predictive value. NPV-Negative predictive value.

When the random variable BART model was repeated using the train-test paradigm, the performance against holdout data was close to chance (AUC =.47, sensitivity=.50, specificity=.65). However, when trained against a single random variable, the test performance appeared potentially above chance (AUC=.69, sensitivity =.85, specificity =.47). The random variable was tested for any relationship with burnout with no correlation between the variable and OLBI (*r*= -.01, *p*=.89) or the binary classifier (*r*= -.06, *p*=.49). There was also no significant group effect of the classifier with the variable using Kruskal-Wallis rank test, H(1)=.27, *p*=.60. The same paradigm using a uniform, rather than normally distributed random variable, resulted in expected chance level performance (AUC=0.45, sensitivity=.46, specificity=.51). This analysis indicates that the BART algorithm may have the potential to falsely identify normally distributed predictor variables by chance and that performance metrics may be overestimated. We also observed that changes to the random seed used in model training was able to influence the performance of trained models using the holdout data set indicating that the learning algorithm was sensitive to which data appeared in training and testing samples, as well as potentially creating more accurate models randomly. To test this, models were trained by iterating through 50 different random seed values (1-50) using a single random variable predictor. This produced 8 models which performed above chance (AUC > .60; range=.63-.73), however, the mean AUC for all random variable models was .50 (*SD*=.10; range=.31-.73). Repeating the same paradigm using the reduced variable set gave a mean AUC of .80 (*SD*=.06; range=.65-.93) which was significantly higher, $t(98)=17.61$, $p<.0001$. This indicates that the predictive power of BART machine models was not a random artifact.

Summary Network Model

An interpretive network model based on all the above findings is presented below (Figure 12).

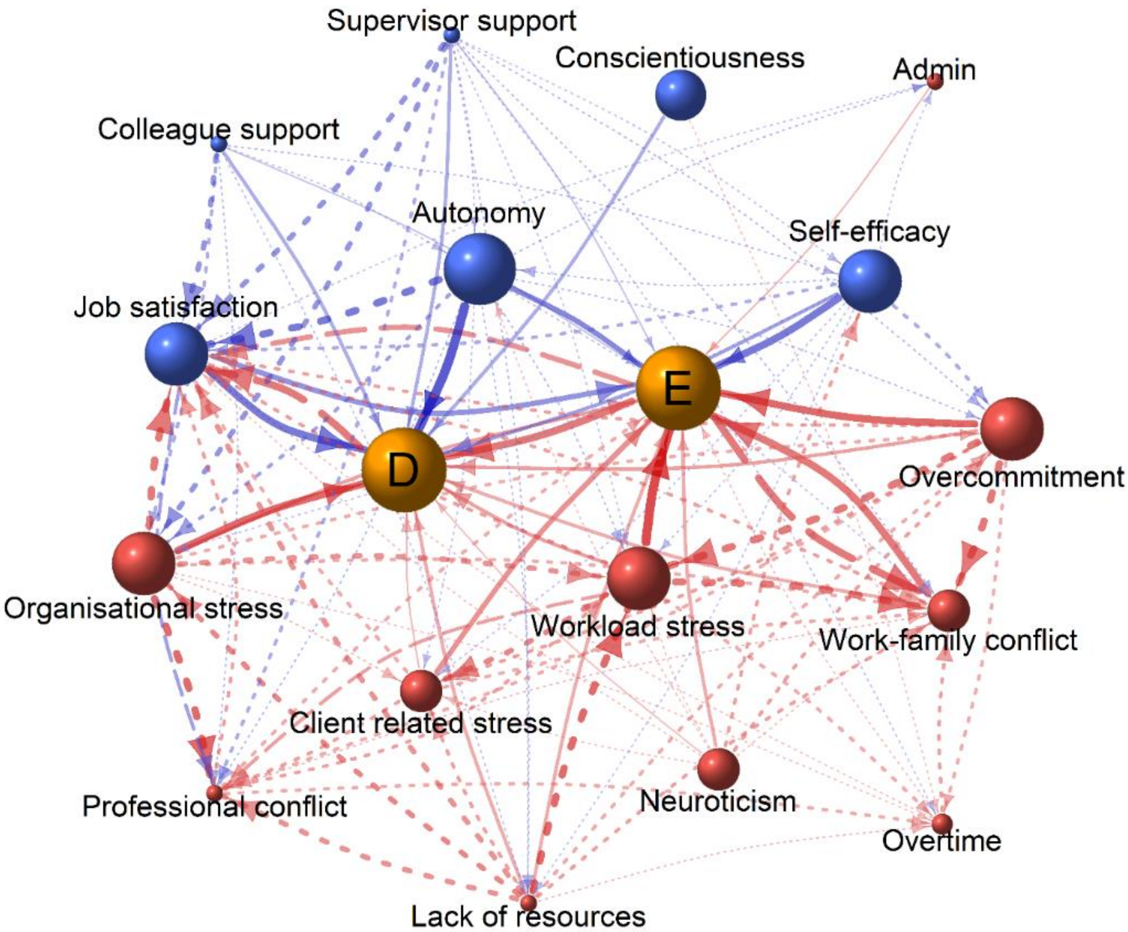

Figure 12. Summary network model showing contributing (red), and protective factors (blue) related to exhaustion (E) and disengagement (D). Node size represents construct importance. Edge width/colour depth represent relationship strength. Dotted lines represent correlations. Solid lines represent relationships from further analysis. Edge directions are theoretically informed with long-dashed lines representing potentially bi-directional relationship directions.

Discussion

The study's primary aim was to develop a prognostic model of burnout in mental health staff to support identification of risk of long-term burnout. The secondary aim was to identify salient factors involved in mental health staff burnout via variable reduction techniques to inform intervention research.

The study design incorporated several independent variables that have previously been shown to relate to burnout levels in mental health staff and other settings (Fields, 2002; O'Connor et al., 2018). This study's results are largely consistent with previous findings having replicated many associations in expected directions.

Cross-sectional main effects analysis showed that overcommitment and workload-related stress were associated with higher risk of exhaustion. Higher levels of job autonomy and self-efficacy were associated with lower risk of exhaustion. Stress related to organisation and management processes (e.g. poor management and supervision) was associated with higher risk of disengagement with higher levels of autonomy, job satisfaction and conscientiousness associated with lower levels of disengagement. Different dimensions of occupational stress influenced different aspects of burnout with job autonomy as a potentially key protective factor for both exhaustion and disengagement.

Individual Characteristics

Individual factors including age, and length of service showed no significant relationship with burnout which is consistent with variable findings in previous research (O'Connor et al., 2018). Male gender and an ethnic minority background were associated with higher levels of burnout, and whilst similar findings have been reported preciously, findings are inconsistent (Angermeyer et al., 2006; O'Connor et al., 2018; Billings et al., 2003; Nelson et al., 2009). As participants numbers with these characteristics were small, and proportionally more ethnic minority background participants identified as male (close to significance), these findings should be interpreted cautiously, although an important interaction between gender and ethnicity may exist.

Personality traits and related personal characteristics of self-efficacy and overcommitment were significantly associated with burnout and each other. Higher levels of self-efficacy were associated with lower levels of burnout and neuroticism, and lower levels of neuroticism were associated with lower levels of burnout. This is consistent with a study in health professionals which found that self-efficacy may protect against stress-induced burnout in individuals with higher neuroticism (Yao et al., 2018).

Overcommitment was associated with higher burnout in medical and psychiatric nurses, however, as far as the author is aware, the current study is the first to describe this finding across different mental health professionals (Schulz et al., 2009). Overcommitment was also associated with neuroticism, consistent with previous research (Vearing & Mak, 2007). The current study showed no significant correlations of conscientiousness with burnout; however, it did identify this trait as having a significant main effect with disengagement at Time-1. Other studies have shown that conscientiousness has a stronger negative relationship with disengagement compared to exhaustion (Robins et al., 2017). Previous research has demonstrated multivariate interactions with these burnout relevant constructs (Yao et al., 2018). This area of research may benefit from future analysis and studies which incorporate mediation and moderator analysis, for example examining whether overcommitment mediates the relationship between neuroticism and burnout. As BART outperformed regression techniques, potentially via modelling higher-order interactions, using additional tools to deconstruct models might support development of theory as well as practical prognostic value (Goldstein et al., 2015; Kapelner & Bleich, 2016).

Work-family conflict was strongly associated with levels of exhaustion and to a lesser degree with disengagement in line with previous studies in physical health nurses (Leineweber et al., 2014; Wang et al., 2012). Overcommitment was also related to work-family conflict as well as hours of overtime, indicating that poor work-life balance and its potentially negative impact on family relationships may be an important predictor for burnout (Kinman & Jones, 2008).

Job Characteristics

Increased workload has been consistently associated with higher levels of burnout (O’Connor et al., 2018). However, objective measures of workload, such as caseload, have inconsistent findings (Galeazzi et al., 2004; Hamaideh, 2011). Subjective perceptions of high workload are more reliably linked with increased risk of burnout (Prosser et al., 1997). The current study found similar results with no associations between caseload, client-facing hours, or working hours with burnout levels; however, there was a small correlation between levels of overtime and exhaustion. Perceptions of high workload were, however, associated with higher levels of burnout, with workload-related stress more strongly related with exhaustion than disengagement, in agreement with previous findings (Evans et al., 2006).

Participants who reported role changes role did not show any relationship between changes in job-demands and burnout levels. Increases in both workload and responsibility was most frequently reported, so the balance of demands and rewards may have remained constant (e.g. increases in workload versus salary, novelty/interest, autonomy) (Basińska & Wilczek-Rużyczka, 2013). In addition, burnout trajectories after changes in role tend to be slow and so potentially were not detectable in this study (Dunford et al., 2012).

In line with previous findings, lower levels of job satisfaction where associated with higher levels of burnout, particularly disengagement (O'Connor et al., 2018). Changes in job satisfaction over time also correlated with changes in burnout levels indicating a possible causal relationship. A causal chain from job conditions to job satisfaction and burnout has previously been proposed, with exhaustion occurring as a result of undertaking an unsatisfying job (Spector, 1997). The strong negative relationship between occupational stress and job satisfaction that we observed is compatible with this theoretical perspective, as well as the finding that job satisfaction was retained in the longitudinal model of OLBI-exhaustion. However, there is also evidence for emotional exhaustion partially mediating the negative effect of job demands on job satisfaction indicating a potentially bidirectional relationship between burnout and job satisfaction (Mijakoski et al., 2015). Cross-lagged panel analysis of the current data might help inform on this relationship (Allen, 2017).

Higher levels of autonomy were associated with lower levels of exhaustion and disengagement. This is perhaps the most consistent finding of burnout studies in mental health professionals (O'Connor et al., 2018). In addition, increased autonomy was associated with increased job satisfaction which may have a mediating or even moderating role with burnout (Özbağ & Ceyhun, 2014). The job satisfaction measure used in this study has an extrinsic satisfaction focus and included an item related to autonomy which may account for part of the association measured, although the satisfaction items are inherently more subjective and so measure different constructs (Fields, 2002).

Organisational Factors

Organisational sources of occupational stress including lack of resources and conflicts with colleagues were associated with higher levels of burnout, in line with previous studies (O'Connor et al., 2018). Occupational stress appeared to be a strong predictor of burnout with changes in stress strongly correlating with changes in burnout over time. This finding aligns with a number of well-regarded models of burnout (Demerouti et al., 2001; Maslach & Leiter, 2016; Siegrist et al., 2004).

No associations between amount of supervision and burnout were observed, in contrast to a previous study which compared regular versus infrequent supervision (Sherring & Knight, 2009). Only five percent of participants in the current study reported having no or infrequent supervision, so the sample characteristics may have impacted this finding. However, supervisor support was associated with lower levels of burnout indicating that the quality of supervisor support rather than supervision time may be more important. Supervisor support was associated with increased job satisfaction, autonomy and self-efficacy and reduced organisational stressors and may indicate a specific mechanism to support staff from becoming burned-out. Support from colleagues showed a similar pattern of relationships with occupational stressors and disengagement, supporting previous findings that team culture may have an important role in burnout processes (Willard-Grace et al., 2014).

Strengths and Limitations

Key strengths of the study were the inclusion of six different sites across four regions and inclusion of a variety of clinical mental health professions across multiple service settings. The subsequent diverse nature of the sample studied may support generalisability of the findings. Also, a broad range of independent variables were investigated to identify key factors based on previous research into burnout in mental health professionals. A sophisticated analytical process was designed and deployed using multiple variable selection techniques. These were designed to detect both main effects and interactions and enabled triangulation of important variables using cross-sectional and longitudinal findings as well as supporting the development of a prognostic model and informing on targets for interventional research.

The high drop-out rate led to underpowered analysis at Time-7 potentially requiring cautious interpretation of the findings and their generalisability. In addition, participation rates were low, although these were difficult to measure accurately. However, *post hoc* power calculations indicated high power achieved (97.9%) based on supplemental analysis (see Appendix G) and encouragingly, the variable selection process at Time-1 (which did reach the required power) had almost complete cross-over with variables selected at Time-7. There was also no evidence of bias due to drop-out from sample characteristics analysis. Reducing questionnaire burden (e.g. fewer time points or planned missing data design) and more direct recruitment activities might support larger sample sizes in future research (Graham et al., 2006).

Previous studies have acknowledged the potential difficulties in using classification compared to regression techniques based on loss of power and information (Delgadillo & Gonzalez Salas Duhne, 2020). OLBI exhaustion scores are consistently higher compared to disengagement scores in mental health professionals (Delgadillo et al., 2018; Rogala et al., 2016; Rzeszutek & Schier, 2014; Westwood et al., 2017). The divide was greater in the current sample indicating that participants were more exhausted than disengaged compared to other samples. Subsequently, the cut-off strategy may have caused bias in burnout classification towards disengagement. An alternative classification strategy based on previously reported scores might have been more appropriate. However, disengagement rather than exhaustion has been shown to have negative impacts on clinical outcomes so our classification may be more relevant to the negative consequences of staff burnout on service users (Delgadillo et al., 2018). Furthermore, high levels of exhaustion compared to disengagement are not uncommon in mental health staff, so identifying staff with unusually high levels of exhaustion, and therefore greater risk of negative consequences to wellbeing and *presenteeism*, could also be of benefit (Johnson et al., 2018; Salvagioni et al., 2017; Steel et al., 2015).

The short big-five questionnaire had low internal consistency potentially reducing confidence in the findings related to personality constructs. However, achieving a large Cronbach-α is an acknowledged difficulty for very short scales measuring broad constructs, even if construct representation is high (Ziegler et al., 2014). Significant findings were consistent with theory and previous research, and as personality is seen as a largely stable construct, the intent for this measure to inform theory rather than intervention design was supported (Griffin et al., 2017).

Evaluation of job changes did not utilise a validated measure and only captured changes in demands, not rewards. However, this was not a key focus of the study rather a sanity check to evaluate whether job changes might interfere with burnout prediction and interpretation of these findings should be done cautiously.

Implications

This study's findings are congruent with the Job-Demands Resources model of occupational burnout when applied to mental health professionals (Scanlan & Still, 2019). Job demands (e.g. workload) and job resources (e.g. supervisor support) respectively contribute to and protect against burnout in mental health professionals. The results highlight the complex interplay between job characteristics (e.g. autonomy), personal characteristics (e.g. personality, self-efficacy & overcommitment) and external factors (e.g. work-family conflict) and demonstrate the close links between occupational stress, job satisfaction and burnout (Figure 12). Together, these findings may explain why stress management-focused interventions have low effect sizes as there may be a requirement for more individualised, multi-dimensional interventions to effectively reduce staff burnout.

Most organisational directed interventions in mental health services have utilised staff training to support skill development for working with clients which have led to small effect sizes to reduce burnout. Enhancing clinical supervision, improving team cohesion and job redesign have not shown any significant benefits (Dreison et al., 2018). Individually directed interventions for burnout show larger effects sizes in comparison, however these effect sizes are still small (Dreison et al., 2018). These include stress management and mindfulness-based approaches which have been shown to have greater effects on stress reduction than burnout, further supporting the need for more sophisticated interventions (Ruotsalainen et al., 2015). Two candidate "targets" for individualised burnout intervention revealed by our study are self-efficacy and overcommitment.

Self-efficacy is considered a modifiable and contextual personal characteristic relating to an individual’s beliefs about their ability to deal effectively with challenges (Bandura et al., 1977). The association of lower self-efficacy with burnout may represent exhaustion reducing self-efficacy which then subsequently increases disengagement, however wider research has concluded that self-efficacy has a protective role against burnout (Amiri et al., 2019; Rogala et al., 2016; Rubio et al., 2015). Clinical training interventions may increase self-efficacy in a narrow context, however, cognitive interventions directed at increasing more general aspects of self-efficacy (as measured in the current study) might have the potential for reducing burnout (Bresó et al., 2011; Resnick, 2018). Addressing supervision quality may also support increased self-efficacy based on previous research (Cashwell et al., 2008).

Overcommitment involves an individual’s excessive striving in combination with a strong need for approval in an occupational context, but may also be related to organisational culture (Kinman, 2016; Siegrist et al., 2004). Addressing overcommitment may reduce burnout via a multivariate effect involving workload-related stress, overtime and work-family conflict (Figure 13). Line-managers may be able to promote healthier work-life balance, and addressing cognitive features of overcommitment may support positive change, however entrenched cognitions may require specialist interventions (Bamber, 2006). An intervention informed by Effort-Reward Imbalance theory (Siegrist et al., 2014), using both psychodynamic and CBT-based approaches, showed long-term reductions in levels of overcommitment and increased effort-reward ratio (Li et al., 2017). Addressing cognitive flexibility using approaches such as Acceptance and Commitment Therapy may also be of benefit (Lloyd et al., 2013).

The current study supports a conclusion of a recent meta-analysis which highlights the need for organisational interventions to promote professional autonomy to address burnout (O'Connor et al., 2018). Promoting autonomy by replacing traditional command-and-control leadership with compassionate and collective leadership has been shown to improve staff wellbeing (West, 2017). Supervisor support in relation to promotion of autonomy may be a key factor to support staff make independent decisions (Upenieks, 2003). Interventions which include collaboration with workplace representatives to address specific causes of stress, may be more effective at addressing depersonalisation aspects of burnout by increasing job control, which is a construct related to autonomy (Ganster, 2011; Hätinen et al., 2007).

Conclusion

This study highlights that levels of burnout in mental health professionals are high, further evidencing the need for effective interventions to support both staff wellbeing and service quality. This study provides further evidence for the need for organisational change to promote autonomy within staff roles as well as evidencing overcommitment and self-efficacy as potential targets for interventions in mental health professionals experiencing burnout. These findings are currently being used to inform the design of a novel burnout intervention being developed within the NHS.

The Clinical Psychology profession may be best placed to support innovation within this field due to its leadership roles, understanding of organisational, systemic and individual psychology and day-to-day involvement in providing supervision and facilitating reflective practice across different professions. This aligns with the British Psychological Society Charter committed to support staff wellbeing (BPS, 2016).

References

Ahola, K., Toppinen-Tanner, S., & Seppänen, J. (2017). Interventions to alleviate burnout symptoms and to support return to work among employees with burnout: Systematic review and meta-analysis. Burnout Research, 4, 1–11. https://doi.org/10.1016/j.burn.2017.02.001

Akobeng, A. K. (2007). Understanding diagnostic tests 3: Receiver operating characteristic curves. Acta Paediatrica, International Journal of Paediatrics, 96(5), 644–647. https://doi.org/10.1111/j.1651-2227.2006.00178.x

Allen, M. (2017). Cross-Lagged Panel Analysis. The SAGE Encyclopedia of Communication Research Methods, 313–314. https://doi.org/10.4135/9781483381411.n117

Amiri, M., Vahedi, H., Mirhoseini, S. R., Eghtesadi, A. R., & Khosravi, A. (2019). Study of the relationship between self-efficacy, general health and burnout among Iranian health workers. In Osong Public Health and Research Perspectives (Vol. 10, Issue 6, pp. 359–367). Korea Centers for Disease Control and Prevention. https://doi.org/10.24171/j.phrp.2019.10.6.06

Angermeyer, M. C., Bull, N., Bernert, S., Dietrich, S., & Kopf, A. (2006). Burnout of caregivers: A comparison between partners of psychiatric patients and nurses. Archives of Psychiatric Nursing, 20(4), 158–165. https://doi.org/10.1016/j.apnu.2005.12.004

Avanzi, L., Zaniboni, S., Balducci, C., & Fraccaroli, F. (2014). The relation between overcommitment and burnout: Does it depend on employee job satisfaction? Anxiety, Stress and Coping, 27(4), 455–465. https://doi.org/10.1080/10615806.2013.866230

Azodi, C. B., Tang, J., & Shiu, S. H. (2020). Opening the black box: Interpretable machine learning for Geneticists. Trends in Genetics, 36(6), 442–455. https://doi.org/10.1016/j.tig.2020.03.005

Bakker, A. B., & Demerouti, E. (2017). Job demands-resources theory: Taking stock and looking forward. Journal of Occupational Health Psychology, 22(3), 273–285. https://doi.org/10.1037/ocp0000056

Ballenger-Browning, K. K., Schmitz, K. J., Rothacker, J. A., Hammer, P. S., Webb-Murphy, J. A., & Johnson, D. C. (2011). Predictors of Burnout Among Military Mental Health Providers. Military Medicine, 176(3), 253–260. https://doi.org/10.7205/MILMED-D-10-00269

Bamber, M. R. (2006). CBT for occupational stress in health professionals: Introducing a schema-focused approach. In CBT for Occupational Stress in Health Professionals: Introducing a Schema-Focused Approach. Routledge. https://doi.org/10.4324/9780203966037

Bandura, A., Adams, N. E., & Beyer, J. (1977). Cognitive processes mediating behavioral change. Journal of Personality and Social Psychology, 35(3), 125–139. https://doi.org/10.1037/0022-3514.35.3.125

Basińska, B. A., & Wilczek-Rużyczka, E. (2013). The role of rewards and demands in burnout among surgical nurses. International Journal of Occupational Medicine and Environmental Health, 26(4), 593–604. https://doi.org/10.2478/s13382-013-0129-8

Billings, J., Johnson, S., Bebbington, P., Greaves, A., Priebe, S., Muijen, M., Ryrie, I., Watts, J., White, I., & Wright, C. (2003). Assertive outreach teams in London: Staff experiences and perceptions - Pan-London Assertive

Outreach Study, Part 2. British Journal of Psychiatry, 183, 139–147. https://doi.org/10.1192/bjp.183.2.139

Bleich, J., Kapelner, A., George, E. I., & Jensen, S. T. (2014). Variable selection for bart: An application to gene regulation. Annals of Applied Statistics, 8(3), 1750–1781. https://doi.org/10.1214/14-AOAS755

BPS. (2016). Charter for Psychological Staff Wellbeing and Resilience. https://www.healthcareconferencesuk.co.uk/news/charter-for-psychological-staff-wellbeing-and-resilience

Bresó, E., Schaufeli, W. B., & Salanova, M. (2011). Can a self-efficacy-based intervention decrease burnout, increase engagement, and enhance performance? A quasi-experimental study. Higher Education, 61(4), 339–355. https://doi.org/10.1007/s10734-010-9334-6

Bzdok, D., Altman, N., & Krzywinski, M. (2018). Points of Significance: Statistics versus machine learning. Nature Methods, 15(4), 233–234. https://doi.org/10.1038/nmeth.4642

Cashwell, T. H., Dooley, K., & Dooley, K. (2008). The impact of supervision on counselor self-efficacy. The Clinical Supervisor, 5223, 37–41. https://doi.org/10.1300/J001v20n01

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2011). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, 321–357. https://doi.org/10.1613/jair.953

Chirico, F. (2016). Job stress models for predicting burnout syndrome: A review. Annali Dell’Istituto Superiore Di Sanita, 52(3), 443–456. https://doi.org/10.4415/ANN_16_03_17

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum.

Cohen, J. (1992). A power primer. Psychological Bulletin, 112(1), 155–159. http://www.ncbi.nlm.nih.gov/pubmed/19565683

Cushway, D., Tyler, P. A., & Nolan, P. (1996). Development of a stress scale for mental health professionals. The British Journal of Clinical Psychology, 35, 279–295. http://www.ncbi.nlm.nih.gov/pubmed/8773803

De Gooijer, J. G. (2017). Elements of Nonlinear Time Series Analysis and Forecasting. Springer.

Delgadillo, J., & Salas, D. P. G. (2020). Targeted prescription of cognitive-behavioral therapy versus person-centered counseling for depression using a machine learning approach. Journal of Consulting and Clinical Psychology, 88(1), 14–24. https://doi.org/10.1037/ccp0000476

Delgadillo, J., Saxon, D., & Barkham, M. (2018). Associations between therapists’ occupational burnout and their patients’ depression and anxiety treatment outcomes. Depression and Anxiety, September 2017, 1–7. https://doi.org/10.1002/da.22766

Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The Job Demands-Resources Model of Burnout. The Journal of Applied Psychology, 86(3), 499–512.

Demerouti, E., Le Blanc, P. M., Bakker, A. B., Schaufeli, W. B., & Hox, J. (2009). Present but sick: A three-wave study on job demands, presenteeism and burnout. Career Development International, 14(1), 50–68. https://doi.org/10.1108/13620430910933574

Dreison, K. C., Luther, L., Bonfils, K. A., Sliter, M. T., McGrew, J. H., & Salyers, M. P. (2018). Job burnout in mental health providers: A meta-analysis of 35 years of intervention research. Journal of Occupational Health Psychology, 23(1), 18–30. https://doi.org/10.1037/ocp0000047

Dunford, B. B., Shipp, A. J., Wayne Boss, R., Angermeier, I., & Boss, A. D. (2012). Is burnout static or dynamic? A career transition perspective of employee burnout trajectories. Journal of Applied Psychology, 97(3), 637–650. https://doi.org/10.1037/a0027060

Eric, A., Ledell, E., Kennedy, C., & Lendle, S. (2019). Package ‘ SuperLearner .’ https://cran.r-project.org/web/packages/SuperLearner/

Evans, S., Huxley, P., Gately, C., Webber, M., Mears, A., Pajak, S., Medina, J., Kendall, T., & Katona, C. (2006). Mental health, burnout and job satisfaction among mental health social workers in England and Wales. British Journal of Psychiatry, 188(1), 75–80. https://doi.org/10.1192/bjp.188.1.75

Fields, D. (2002). Taking the measure of work: A guide to validated scales for organizational research and diagnosis (pp. 82–83). SAGE Publications, Inc. https://doi.org/10.4135/9781452231143

Friedman, J., Hastie, T., Tibshirani, R., Balasubramanian, N., & Simon, N. (2019). Lasso and Elastic-Net regularized generalized linear models (3.0-2). CRAN.

Galeazzi, G. M., Delmonte, S., Fakhoury, W., & Priebe, S. (2004). Morale of mental health professionals in community mental health services of a Northern Italian province. Epidemiologia e Psichiatria Sociale, 13(3), 191–197. https://doi.org/10.1017/S1121189X00003456

Galvão, R. K. H., & Araújo, M. C. U. (2009). Variable Selection. In S. D. Brown, R. Tauler, & B. Walczak (Eds.), Comprehensive chemometrics (3rd ed., Vol. 3, pp. 233–283). Elsevier. https://doi.org/10.1016/B978-044452701-1.00075-2

Ganster, D. (2011). Autonomy and Control. In S. L. Hurrell, Jr., Joseph J., Levi, Lennart, Murphy, Lawrence R., Sauter (Ed.), Encyclopaedia of occupational health and safety. International Labor Organization.

Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. Journal of Computational and Graphical Statistics, 24(1), 44–65. https://doi.org/10.1080/10618600.2014.907095

Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. Psychological Methods, 11(4), 323–343. https://doi.org/10.1037/1082-989X.11.4.323

Griffin, R. W., Phillips, J., & Gully, S. M. (2017). Organizational behavior : managing people and organizations (12th ed.). Cengage Learning.

Grund, S., Lüdtke, O., & Robitzsch, A. (2016). Multiple Imputation of Multilevel Missing Data: An Introduction to the R Package pan. SAGE Open, 6(4). https://doi.org/10.1177/2158244016668220

Hackman, J. R., & Oldham, G. R. (1974). The Job Diagnostic Survey: An instrument for the diagnosis of jobs and the evaluation of job redesign projects. In Organizational Effectiveness Research Programs, Department of Admnnistrative Sciences Yale University (Issue May).

Hamaideh, S. H. (2011). Burnout, social support, and job satisfaction among jordanian mental health nurses. Issues in Mental Health Nursing, 32(4), 234–242. https://doi.org/10.3109/01612840.2010.546494

Hätinen, M., Kinnunen, U., Pekkonen, M., & Kalimo, R. (2007). Comparing two burnout interventions: Perceived job control mediates decreases in burnout. International Journal of Stress Management, 14(3), 227–248. https://doi.org/10.1037/1072-5245.14.3.227

House, J. ., & Wells, J. S. (1978). Occupational stress, social support, and health. In A. McLean, G. Black, & M. Colligan (Eds.), Reducing occupational stress: Proceedings of a conference (pp. 8–29). National Institute of Occupational Health and Safety.

Jachens, L., Houdmont, J., & Thomas, R. (2018). Effort-reward imbalance and burnout among humanitarian aid workers. Disasters. https://doi.org/10.1111/disa.12288

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, J., Hall, L. H., Berzins, K., Baker, J., Melling, K., & Thompson, C. (2018). Mental healthcare staff well-being and burnout: A narrative review of trends, causes, implications, and recommendations for future interventions. International Journal of Mental Health Nursing, 27(1), 20–32. https://doi.org/10.1111/inm.12416

Kapelner, A., & Bleich, J. (2016). bartMachine: Machine learning with bayesian additive regression trees. Journal of Statistical Software, 70(2013). https://doi.org/10.18637/jss.v070.i04

Karatzoglou, A., Smola, A., & Hornik, K. (2019). Kernel-Based Machine Learning Lab.

Kilfedder, C. J., Power, K. G., & Wells, T. J. (2001). Burnout in psychiatric nursing. Journal of Advanced Nursing, 34(3), 383–396. https://doi.org/10.1046/j.1365-2648.2001.01769.x

Kinman, G. (2016). Effort–reward imbalance and overcommitment in UK academics: implications for mental health, satisfaction and retention. Journal of Higher Education Policy and Management, 38(5), 504–518. https://doi.org/10.1080/1360080X.2016.1181884

Kinman, G., & Jones, F. (2008). Effort-reward imbalance, over-commitment and work-life conflict: Testing an expanded model. Journal of Managerial Psychology, 23(3), 236–251. https://doi.org/10.1108/02683940810861365

Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the boruta package. Journal of Statistical Software, 36(11), 1–13. https://doi.org/10.18637/jss.v036.i11

LeDell, E., Van Der Laan, M. J., & Peterson, M. (2016). AUC-maximizing ensembles through metalearning. International Journal of Biostatistics, 12(1), 203–218. https://doi.org/10.1515/ijb-2015-0035

Lee, J., Lim, N., Yang, E., & Lee, S. M. (2011). Antecedents and consequences of three dimensions of burnout in psychotherapists: A meta-analysis. Professional Psychology: Research and Practice, 42(3), 252–258. https://doi.org/10.1037/a0023319

Leineweber, C., Westerlund, H., Chungkham, H. S., Lindqvist, R., Runesdotter, S., & Tishelman, C. (2014). Nurses’ practice environment and work-family conflict in relation to burn out: A multilevel modelling approach. PLoS ONE, 9(5). https://doi.org/10.1371/journal.pone.0096991

Li, J., Riedel, N., Barrech, A., Herr, R. M., Aust, B., Mörtl, K., Siegrist, J., Gündel, H., & Angerer, P. (2017). Long-term effectiveness of a stress management intervention at work: A 9-year follow-up study based on a randomized wait-list controlled trial in male managers. BioMed Research International, 2017. https://doi.org/10.1155/2017/2853813

Liaw, A., & Wiener, M. (2018). Breiman and Cutler's Random Forests for classification and regression. In CRAN (4.6-14). https://doi.org/10.1023/A

Lloyd, J., Bond, F. W., & Flaxman, P. E. (2013). The value of psychological flexibility: Examining psychological mechanisms underpinning a cognitive behavioural therapy intervention for burnout. Work and Stress, 27(2), 181–199. https://doi.org/10.1080/02678373.2013.782157

Lovik, A., Verbeke, G., & Molenberghs, G. (2017). Evaluation of a very short test to measure the big five personality factors on a Flemish sample. Journal of Psychological and Educational Research, 25(2), 7–17.

Maslach, C. (1982). Burnout: The cost of caring. ISHK.

Maslach, C., & Leiter, M. (2016). Understanding the burnout experience : Recent research and its implications for psychiatry. World Psychiatry, 15(2), 103–111. https://doi.org/10.1002/wps.20311

Mijakoski, D., Karadzinska-Bislimovska, J., Basarovska, V., Stoleski, S., & Minov, J. (2015). Burnout and work demands predict reduced job satisfaction in health professionals working in a surgery clinic. Macedonian

Journal of Medical Sciences, 3(1), 166–173. https://doi.org/10.3889/oamjms.2015.020

Morse, G., Salyers, M. P., Rollins, A. L., Monroe-DeVita, M., & Pfahler, C. (2012). Burnout in mental health services: A review of the problem and its remediation. Administrative Policy Mental Health, 39(5), 341–352. https://doi.org/10.1007/s10488-011-0352-1.Burnout

Müller, G., Brendel, B., Freude, G., Potter, G., Rose, U., Burr, H., Falkenstein, M., & Martus, P. (2018). Work-related determinants of burnout in a nationally representative sample of German employees: Results from the study on mental health at work. Journal of Occupational and Environmental Medicine, 60(7), 584–588. https://doi.org/10.1097/JOM.0000000000001328

Nagy, M. S. (2002). Using a single-item approach to measure facet job satisfaction. Journal of Occupational and Organizational Psychology, 75(1), 77–86. https://doi.org/10.1348/096317902167658

Nelson, T., Johnson, S., & Bebbington, P. (2009). Satisfaction and burnout among staff of crisis resolution, assertive outreach and community mental health teams: A multicentre cross sectional survey. Social Psychiatry and Psychiatric Epidemiology, 44(7), 541–549. https://doi.org/10.1007/s00127-008-0480-4

Netemeyer, R. G., Boles, J. S., & Mcmurrian, R. (1996). Development and validation of Work-Family Conflict and Family-Work Conflict scales. Journal of Applied Psychology, 81(4), 400–410. https://doi.org/10.1037/0021-9010.81.4.400

O’Connor, K., Muller Neff, D., & Pitman, S. (2018). Burnout in mental health professionals: A systematic review and meta-analysis of prevalence and determinants. European Psychiatry, 53, 74–99. https://doi.org/10.1016/j.eurpsy.2018.06.003

Özbağ, G. K., & Ceyhun, G. Ç. (2014). The impact of job characteristics on burnout: The mediating role of work family conflict and the moderating role of job satisfaction. International Journal of Academic Research in Management, 3(3), 291–309. www.elvedit.com

Parikh, R., Mathai, A., Parikh, S., Sekhar, G. C., & Thomas, R. (2008). Understanding and using sensitivity, specificity and predictive values. Indian Journal of Ophthalmology, 56(1), 45–50. https://doi.org/10.4103/0301-4738.37595

Peterson, U., Demerouti, E., Bergström, G., Samuelsson, M., Åsberg, M., & Nygren, Å. (2008). Burnout and physical and mental health among Swedish healthcare workers. Journal of Advanced Nursing, 62(1), 84–95. https://doi.org/10.1111/j.1365-2648.2007.04580.x

Pham, H. N. A., & Triantaphyllou, E. (2008). The impact of overfitting and overgeneralization on the classification accuracy in data mining. In Soft Computing for Knowledge Discovery and Data Mining (pp. 391–431). Springer US. https://doi.org/10.1007/978-0-387-69935-6_16

Poghosyan, L., Aiken, L. H., & Sloane, D. M. (2009). Factor structure of the Maslach burnout inventory: An analysis of data from large scale cross-sectional surveys of nurses from eight countries. International Journal of Nursing Studies, 46(7), 894–902. https://doi.org/10.1016/j.ijnurstu.2009.03.004

Prosser, D., Johnson, S., Kuipers, E., Szmukler, G., Bebbington, P., & Thornicroft, G. (1997). Perceived sources of work stress and satisfaction among hospital and community mental health staff, and their relation to mental health, burnout and job satisfaction. Journal of Psychosomatic Research, 43(1), 51–59. https://doi.org/10.1016/S0022-3999(97)00086-X

Quick, J. C., & Henderson, D. F. (2016). Occupational stress: Preventing suffering, enhancing wellbeing. International Journal of Environmental Research and Public Health, 13(5). https://doi.org/10.3390/ijerph13050459

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. Journal of Research in Personality, 41(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Resnick, B. (2018). Theory of Self-Efficacy. In M. J. Smith & P. R. Liehr (Eds.), Middle range theory for nursing. Springer Publishing Company.

Riley, R. D., van der Windt, D., Croft, P., & Moons, K. G. M. (2019). Prognosis Research in Healthcare: Concepts, Methods, and Impact. Oxford University Press.

Ripley, B. (2020). Package ‘ nnet ’ (7.3.13). CRAN.

Robins, T. G., Roberts, R. M., & Sarris, A. (2017). Understanding how personality impacts exhaustion and engagement: The role of job demands, and job and personal resources as mediators. Australian Psychologist, 53, 253–262. https://doi.org/10.1111/ap.12310

Rodríguez, J. D., Pérez, A., & Lozano, J. A. (2010). Sensitivity analysis of k-fold cross validation in prediction error estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(3), 569–575. https://doi.org/10.1109/TPAMI.2009.187

Rogala, A., Shoji, K., Luszczynska, A., Kuna, A., Yeager, C., Benight, C. C., & Cieslak, R. (2016). From exhaustion to disengagement via self-efficacy change: Findings from two longitudinal studies among human services workers. Frontiers in Psychology, 6. https://doi.org/10.3389/fpsyg.2015.02032

Rollins, A. L., Salyers, M. P., Tsai, J., & Lydick, J. M. (2010). Staff turnover in statewide implementation of ACT: relationship with ACT fidelity and other team characteristics. Administration and Policy in Mental Health, 37(5), 417–426. https://doi.org/10.1007/s10488-009-0257-4

Rubio, C., Osca, A., Recio, P., Urien, B., & Peiró, J. M. (2015). Work-family conflict, self-efficacy, and emotional exhaustion: A test of longitudinal effects. Revista de Psicologia Del Trabajo y de Las Organizaciones, 31(3), 147–154. https://doi.org/10.1016/j.rpto.2015.06.004

Ruotsalainen, J. H., Verbeek, J. H., Mariné, A., & Serra, C. (2015). Preventing occupational stress in healthcare workers. In Cochrane Database of Systematic Reviews (Issue 4). John Wiley and Sons Ltd. https://doi.org/10.1002/14651858.CD002892.pub5

Rzeszutek, M., & Schier, K. (2014). Temperament traits, social support, and burnout symptoms in a sample of therapists. Psychotherapy, 51(4), 574–579. https://doi.org/10.1037/a0036020

Salvagioni, D. A. J., Melanda, F. N., Mesas, A. E., González, A. D., Gabani, F. L., & De Andrade, S. M. (2017). Physical, psychological and occupational consequences of job burnout: A systematic review of prospective studies. PLoS ONE, 12(10), 1–29. https://doi.org/10.1371/journal.pone.0185781

Scanlan, J. N., & Still, M. (2019). Relationships between burnout, turnover intention, job satisfaction, job demands and job resources for mental health personnel in an Australian mental health service. Bmc Health Services Research, 19. https://doi.org/10.1186/s12913-018-3841-z

Schulz, M., Damkroger, A., Heins, C., Wehlitz, L., Lohr, M., Driessen, M., Behrens, J., & Wingenfeld, K. (2009). Effort-reward imbalance and burnout among German nurses in medical compared with psychiatric hospital settings. Journal of Psychiatric and Mental Health Nursing, 16(3), 225–233. https://doi.org/http://dx.doi.org/10.1111/j.1365-2850.2008.01355.x

Schwarzer, R., & Jerusalem, M. (1995). Measures in health psychology: A user's portfolio. (& M. J. J. Weinman, S. Wright (ed.); pp. 35–37). NFER-NELSON.

Sherring, S., & Knight, D. (2009). An exploration of burnout among city mental health nurses. British Journal of Nursing, 18(20), 1234–1240. https://doi.org/10.12968/bjon.2009.18.20.45114

Shoji, K., Cieslak, R., Smoktunowicz, E., Rogala, A., Benight, C. C., & Luszczynska, A. (2016). Associations between job burnout and self-efficacy: a meta-analysis. Anxiety, Stress, & Coping, 29(4), 367–386. https://doi.org/10.1080/10615806.2015.1058369

Siegrist, J., Li, J., & Montano, D. (2014). Psychometric properties of the effort-reward imbalance questionnaire. Duesseldorf University, 1–14. https://doi.org/10.1007/s00420-009-0460-3.

Siegrist, J., Starke, D., Chandola, T., Godin, I., Marmot, M., Niedhammer, I., & Peter, R. (2004). The measurement of effort-reward imbalance at work:

European comparisons. Social Science and Medicine, 58(8), 1483–1499. https://doi.org/10.1016/S0277-9536(03)00351-4

Simionato, G. K., & Simpson, S. (2018). Personal risk factors associated with burnout among psychotherapists: A systematic review of the literature. Journal of Clinical Psychology, January 2017, 1–26. https://doi.org/10.1002/jclp.22615

Smith, G. (2018). Step away from stepwise. Journal of Big Data, 5(1), 1–12. https://doi.org/10.1186/s40537-018-0143-6

Son, H., Friedmann, E., & Thomas, S. A. (2012). Application of pattern mixture models to address missing data in longitudinal data analysis using spss. Nursing Research, 61(3), 195–203. https://doi.org/10.1097/NNR.0b013e3182541d8c

Spector, P. E. (1997). Job satisfaction: Application, assessment, causes, and consequences. Sage.

Steel, C., Macdonald, J., Schroder, T., & Mellor-clark, J. (2015). Exhausted but not cynical: Burnout in therapists working within Improving Access to Psychological Therapy Services. Journal of Mental Health, 8237(1), 33–37. https://doi.org/10.3109/09638237.2014.971145

Taris, T. W. (2006). Is there a relationship between burnout and objective performance? A critical review of 16 studies. Work and Stress, 20(4), 316–334. https://doi.org/10.1080/02678370601065893

Upenieks, V. V. (2003). The interrelationship of organizational characteristics of magnet hospitals, nursing leadership, and nursing job satisfaction. Health

Care Manager, 22(2), 83–98. https://doi.org/10.1097/00126450-200304000-00002

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. Journal of Statistical Software, 45(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Varian, H. R. (2014). Big data: New tricks for econometrics. Journal of Economic Perspectives, 28(2), 3–28. https://doi.org/10.1257/jep.28.2.3

Vearing, A., & Mak, A. S. (2007). Big five personality and effort-reward imbalance factors in employees' depressive symptoms. Personality and Individual Differences, 43(7), 1744–1755. https://doi.org/10.1016/j.paid.2007.05.011

Wang, Y., Chang, Y., Fu, J., & Wang, L. (2012). Work-family conflict and burnout among Chinese female nurses: The mediating effect of psychological capital. BMC Public Health, 12(1), 915. https://doi.org/10.1186/1471-2458-12-915

Weisberg, S. (2005). Applied Linear Regression: Third Edition. In Applied Linear Regression: Third Edition. Wiley Blackwell. https://doi.org/10.1002/0471704091

West, M. (2017). Caring to change - how compassionate leadership can stimulate innovation in healthcare. In The Kings Fund. https://doi.org/10.1080/10714421.2011.597235

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. https://doi.org/10.1080/09638237.2016.1276540

Willard-Grace, R., Hessler, D., Rogers, E., Dubé, K., Bodenheimer, T., & Grumbach, K. (2014). Team structure and culture are associated with lower burnout in primary care. Journal of the American Board of Family Medicine, 27(2), 229–238. https://doi.org/10.3122/jabfm.2014.02.130215

Xanthopoulou, D., Bakker, A. B., Demerouti, E., & Schaufeli, W. B. (2007). The role of personal resources in the job demands-resources model. International Journal of Stress Management, 14(2), 121–141. https://doi.org/10.1037/1072-5245.14.2.121

Yao, Y., Zhao, S., Gao, X., An, Z., Wang, S., Li, H., Li, Y., Gao, L., Lu, L., & Dong, Z. (2018). General self-efficacy modifies the effect of stress on burnout in nurses with different personality types. BMC Health Services Research, 18(667), 1–9. https://doi.org/https://doi.org/10.1186/s12913-018-3478-y

Ziegler, M., Kemper, C. J., & Kruyen, P. (2014). Short scales - Five misunderstandings and ways to overcome them. Journal of Individual Differences, 35(4), 185–189. https://doi.org/10.1027/1614-0001/a000148

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society. Series B: Statistical Methodology, 67(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

Appendix A - Participant Characteristics and T1 Measures

Table 1. Participant Characteristics and Job Roles (n=287)

Characteristic	Category	Mean or %	SD or n
Age		42.34	11.25
Gender		79.4% female	228
Children	none	47.7%	137
	one	17.8%	51
	two	27.9%	80
	more than two	6.6%	19
Ethnicity	White	89.2%	256
	Mixed	3.1%	9
	Asian	2.8%	8
	Black	2.1%	6
	Chinese	0.3%	1
	Other	2.4%	7
Education	GCSE	1.4%	4
	A Level	2.4%	7
	NVQ	1.4%	4
	Undergrad	15.3%	44
	PG Dip	46%	132
	Masters	20.6%	59
	Doctorate	12.5%	36
	Apprentice	0.3%	1
Relationship	Single	17.1%	49
	Partner	29.6%	85
	Cohabit	47.4%	136
	Divorced/Separated	5.9%	17
Service setting	Inpatient	13.6%	39
	Outpatient	17.4%	50
	Community	69.0%	198
Service Type	Mental Health	63.1%	181
	Drug & Alcohol	3.1%	9
	Learning Disability	2.1%	6
	Forensic	0.7%	3
	IAPT	23.7%	68
	Memory	0.7%	3
	Other	5.9%	17
Service User Age	Child/Adolescent	7.7%	22
	Adult	82.6%	237
	Older Adult	9.8%	28
Role	Care Coordinator	3.1%	9
	Clinical Psychologist	7.0%	20
	CBT Therapist	17.8%	51
	Family Therapist	0.7%	2
	Nurse	18.8%	54
	Occupational therapist	3.5%	10
	Psychiatrist	5.6%	16
	PWP	16.4%	47
	Social Worker	1.0%	3
	Support Worker	1.4%	4
	Other	15.7%	46
	Nurse Associate	1.0%	3
	Counsellor	7.3%	21

Table 2. Participant Work Characteristics and Time-1 Measures

Type	**Variable**	**Mean or %**	**Sd or n**
Work characteristics	Clinical experience (yrs)	13.93	10.0
	Hrs per week	33.87	6.80
	Weekly overtime (hrs)	2.32	4.75
	Supervision (hrs)	1.81	7.50
	Weekly admin (hrs)	9.91	7.17
	Client facing (hrs)	16.72	8.36
	Client caseload	31.29	47.79
	Months in current role	65.24	67.27
Measures	OLBI Exhaustion	2.78	0.48
	OLBI Disengagement	2.41	0.46
	JDSS	2.42	0.58
	Autonomy	4.72	1.35
	OCI	2.43	0.66
	WFCS	4.32	1.62
	GSES	2.99	0.45
	SSS	1.92	0.56
	MHPSS	1.35	0.52
	Extraversion	3.38	1.02
	Agreeableness	4.06	0.65
	Openness	3.57	0.89
	Conscientiousness	4.27	0.75
	Neuroticism	2.90	0.96
Burnout		22.6%	65

Table 3. Means (Standard deviation), Percent Frequencies (n) and *p* values* for demographic, workload, predictor and outcome variables of participants who dropped out (n=136) and participants who completed T7 measures (n=151).

Participant Characteristics							
		Mean/SD or %/n				**test statistic**	**significance**
		Time-1 (n=136)		**Time-7 (n=151)**		***t, c2, U***	***p***
Age		41.46	12.1	43.13		-1.25	0.21†
Gender female		77.2%	105	81.5%	123	0.79	.37
Children	none	45.6%	62	49.7%	75	9971	0.65
	one	21.3%	29	14.6%	22		
	two	24.3%	33	31.1%	47		
	more than two	8.8%	12	4.6%	7		
Ethnicity	White	87.5%	119	90.7%	137	2.91	.71
	Mixed	3.7%	5	2.6%	4		
	Asian	3.7%	5	2%	3		
	Black	2.9%	4	1.3%	2		
	Chinese	0%	0	0.7%	1		
	Other	2.2%	3	2.6%	4		
Education	GCSE	1.5%	2	1.3%	2	3.20	.87
	A Level	2.2%	3	2.6%	4		
	NVQ	.7%	1	2%	3		
	Undergrad	16.2%	22	14.6%	22		
	PG Dip	48.5%	66	43.7%	66		
	Masters	19.1%	26	21.9%	33		
	Doctorate	11.0%	15	13.9%	21		
	Apprentice	.7%	1	0%	0		
Relationship	Single	16.2%	22	17.9%	27	.74	.86
	Partner	31.6%	43	27.8%	42		
	Cohabit	47.1%	64	47.7%	72		
	Divorced or Separated	5.1%	7	6.6%	10		
Service setting	Inpatient	14.0%	19	13.2%	20	.07	.97
	Outpatient	16.9%	23	17.9%	27		
	Community	69.1%	94	68.9%	104		
Service Type	Mental Health	63.2%	86	62.9%	95	2.17	.90
	Drug & Alcohol	2.9%	4	3.3%	5		
	Learning Disability	2.9%	4	1.3%	2		
	Forensic	1.5%	2	0.7%	1		
	IAPT	22.8%	31	24.5%	37		
	Memory	1.5%	2	0.7%	1		
	Other	5.1%	7	6.6%	10		
Service User Age	Child/Adolescent	6.6%	9	8.6%	13	.72	.70
	Adult	84.6%	115	80.8%	122		
	Older Adult	8.8%	12	10.6%	16		

†- Leven's test of equality of variances $p<.05$, equal variances not assumed
**p* values are calculated using the t-tests except for categorical data where the χ^2 test was used and for ordinal data where the Mann-Whitney U test was used.

		Participant Characteristics					
		Mean / SD or % / n				test statistics	significance
		Time 1 (n=136)		Time 7 (n=151)		*t, c2, U*	*p*
Role	Care Coordinator	2.9%	4	3.3%	5	17.88	.12
	Clinical Psychologist	5.1%	7	8.6%	13		
	CBT Therapist	19.9%	27	15.9%	24		
	Family Therapist	0%	0	1.3%	2		
	Nurse	17.6%	24	19.9%	30		
	Occupational therapist	2.2%	3	4.6%	7		
	Psychiatrist	6.6%	9	4.6%	7		
	PWP	19.9%	27	13.2%	20		
	Social Worker	1.5%	2	0.7%	1		
	Support Worker	2.2%	3	0.1%	1		
	Other	16.2%	22	15.9%	24		
	Nurse Associate	2.2%	3	0%	0		
	Counsellor	3.7%	5	11.3%	17		
Work characteristics	Clinical experience (yrs)	13.26	9.78	14.54	10.21	-1.09	0.28
	Hrs per week	34.65	6.74	33.16	6.79	1.86	0.06
	Weekly overtime (hrs)	2.56	5.68	2.11	3.71	0.81	0.42
	Supervision (hrs)	1.48	5.86	2.11	8.72	-.70	.48
	Weekly admin (hrs)	9.51	7.69	10.27	6.67	-.89	.37
	Client facing (hrs)	16.90	9.53	16.55	7.16	.34	.73†
	Client caseload	30.50	40.93	32.01	53.35	-.27	.79
	Months in current role	65.56	64.28	64.95	70.06	.08	.94
Measures	OLBI Exhaustion	2.82	.50	2.74	0.47	1.44	.15
	OLBI Disengagement	2.45	.47	2.38	0.46	1.21	.23
	JDSS	2.41	.62	2.43	0.54	-.39	.70
	Autonomy	4.63	1.46	4.79	1.24	-1.04	.30†
	OCI	2.46	0.66	2.39	0.65	.94	.35
	WFCS	4.30	1.70	4.32	1.56	-.10	.92
	GSES	2.97	0.44	3.02	0.45	-.98	.33
	SSS	1.92	0.56	1.92	0.55	-.08	.94
	MHPSS	1.38	.55	1.33	0.49	.93	.36
	Extraversion	3.47	1.02	3.291	1.03	1.46	.45
	Agreeableness	4.11	0.70	4.014	0.60	1.28	.20†
	Openness	3.64	0.85	3.513	0.93	1.16	.25
	Conscientiousness	4.25	0.77	4.277	0.74	-.26	.78
	Neuroticism	2.94	0.94	2.867	0.97	.65	.52
Classed as burned out at T1		25.74%	35	19.87%	30	1.41	.24

†- Leven's test of equality of variances $p<.05$, equal variances not assumed
*p values are calculated using the t-tests except for categorical data where the χ^2 test was used and for ordinal data where the Mann-Whitney U test was used.

Appendix B – Demographic Data and Psychometric Measures

Age (Yrs)
Gender (Male/female/Other-please specify/Prefer not to say)
Ethnicity (White/Mixed/Asian/Black/Chinese/Other)
Children (0,1,2,>2)
Education (GCSE/CSE/O Level/, A-Level, NVQ/College, Apprenticeship, Undergrad, PG Diploma, Masters, Doctorate or Equivalent)
Relationship (Single; Partner; Cohabiting with Spouse/Partner; Divorced/Separated; Widowed)

Experience (Yrs)
Time in current role (Yrs and Months)
Role (psychiatrist, psychologist, CBT therapist, PWP, family therapist, nurse, care-coordinator, social worker, support worker, occupational therapist, other)
Service (Inpatient/Community, CAMHS, OA, Adult, Specialist -e.g. CERT, ACCESS, EIS)
Job Status (Permanent, Contract more than 12 months, Contract less than 12 months, Trainee/Full time student)
Service Type (Physical health, Mental Health, Drug and Alcohol, Learning Disability, Forensic, IAPT, Memory, Palliative Care, Other)
Service User Age Group (Child/Adolescent, Adult, Older Adult)
Clinical Setting (Inpatient, Outpatient, Community)

*Contracted hours per week (Hrs)
*Hrs overtime (Hrs)
*Hrs in supervision (Hrs)
*Hrs doing admin (clinical notes, letters, other) (Hrs)
*Hrs client contact (Hrs)
*Caseload (Number)

*Total hours etc. for staff working for more than one organisation.

Oldenburg Burnout Inventory

Oldenburg Burnout Inventory

Instruction: Below you find a series of statements with which you may agree or disagree. Using the scale, please indicate the degree of your agreement by selecting the number that corresponds with each statement

	Strongly agree	Agree	Disagree	Strongly disagree
1. I always find new and interesting aspects in my work.	1	2	3	4
2. There are days when I feel tired before I arrive at work.	1	2	3	4
3. It happens more and more often that I talk about my work in a negative way.	1	2	3	4
4. After work, I tend to need more time than in the past in order to relax and feel better.	1	2	3	4
5. I can tolerate the pressure of my work very well.	1	2	3	4
6. Lately, I tend to think less at work and do my job almost mechanically.	1	2	3	4
7. I find my work to be a positive challenge.	1	2	3	4
8. During my work, I often feel emotionally drained.	1	2	3	4
9. Over time, one can become disconnected from this type of work.	1	2	3	4
10. After working, I have enough energy for my leisure activities.	1	2	3	4
11. Sometimes I feel sickened by my work tasks.	1	2	3	4
12. After my work, I usually feel worn out and weary.	1	2	3	4
13. This is the only type of work that I can imagine myself doing.	1	2	3	4
14. Usually, I can manage the amount of my work well.	1	2	3	4
15. I feel more and more engaged in my work.	1	2	3	4
16. When I work, I usually feel energized.	1	2	3	4

Note. Disengagement items are 1, 3(R), 6(R), 7, 9(R), 11(R), 13, 15. Exhaustion items are 2(R), 4(R), 5, 8(R), 10, 12(R), 14, 16. (R) means reversed item when the scores should be such that higher scores indicate more burnout.

Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The Job Demands-Resources Model of Burnout. The Journal of Applied Psychology, 86(3), 499–512.

Mental Health Professionals Stress Scale

SOURCES OF PRESSURE AT WORK

The following have been found to be sources of pressure at work in health care.
Please respond by circling the numbers which represent the extent to which each item applies to you (i.e. represents a source of pressure at work for you).

		Does not apply to me		Does apply to me	
1	Too much work to do……………………………………	0	1	2	3
2	Ending treatment with clients/patients ………………	0	1	2	3
3	Lack of support from management…………………	0	1	2	3
4	Conflict with other professionals e.g. doctor, nurse…	0	1	2	3
5	Lack of adequate staffing…………………………….	0	1	2	3
6	Feeling inadequately skilled for dealing with emotional needs of clients/patients…………………..	0	1	2	3
7	Not enough time with family…………………………	0	1	2	3
8	Too many different things to do ……………………..	0	1	2	3
9	Dealing with death or suffering……………………….	0	1	2	3
10	Relationship with line manager……………………….	0	1	2	3
11	Conflicting roles with other professionals…………….	0	1	2	3
12	Lack of financial resources for training courses/workshops…………………………………….	0	1	2	3
13	Uncertainty about own capabilities……………………	0	1	2	3
14	Inability to separate personal from professional role…	0	1	2	3
15	Not enough time to complete all tasks satisfactorily….	0	1	2	3
16	No change or slowness of change in clients/patients….	0	1	2	3
17	Communications and flow of information at work……	0	1	2	3
18	Working in a multidisciplinary team………………….	0	1	2	3
19	Shortage of adequate equipment/supplies……………..	0	1	2	3
20	Feeling inadequately skilled for working with difficult clients/patients………………………………..	0	1	2	3

		Does not apply to me		Does apply to me	
21	Taking work home…………………………………..	0	1	2	3
22	Too many clients/patients………………………………	0	1	2	3
23	Difficult and/or demanding clients or patients…………	0	1	2	3
24	Poor management and supervision…………………….	0	1	2	3
25	Criticism by other professional e.g. doctor, nurse……....	0	1	2	3
26	Lack of adequate cover in potentially dangerous Environment……………………………………………	0	1	2	3
27	Doubt about the efficacy of therapeutic endeavours……	0	1	2	3
28	Relationship with spouse/partner affects work………….	0	1	2	3
29	Working too long hours………………………………….	0	1	2	3
30	Physically threatening clients/patients…………………...	0	1	2	3
31	The way conflicts are resolved in the organisation……..	0	1	2	3
32	Lack of emotional support from colleagues…………….	0	1	2	3
33	Inadequate clerical /technical back-up………………….	0	1	2	3
34	Keeping professional/clinical skills up to date………….	0	1	2	3
35	Work emphasises feelings of emptiness and/or isolation.	0	1	2	3
36	Not enough time for recreation…………………………	0	1	2	3
37	Managing therapeutic relationships…………………….	0	1	2	3
38	Organisational structure and policies……………………	0	1	2	3
39	Difficulty of working with certain colleagues…………..	0	1	2	3
40	Poor physical working conditions……………………….	0	1	2	3
41	Fear of making a mistake over a client/patient's treatment	0	1	2	3
42	Inadequate time for friendships/social relationships.	0	1	2	3

Cushway, D., Tyler, P. A., & Nolan, P. (1996). Development of a stress scale for mental health professionals. The British Journal of Clinical Psychology, 35 (Pt 2), 279–95.

Social Support Scale

This part of the questionnaire deals with your present job and life-situation. People around us (both on and off the job) sometimes are very supportive and helpful and sometimes hinder or offer little or no support in our work. This section asks how people around you affect you in such matters. Please circle the response to each question as to how true the statement is concerning the person or persons indicated.

1. How much can each of these people be relied on <u>when things get tough at work?</u>

		Not at all	A little	Some-what	Very Much	
a	Your immediate supervisor	0	1	2	3	
b	Other people at work	0	1	2	3	
c	Your spouse/partner	0	1	2	3	N/A
d	Your friends and relatives	0	1	2	3	

2. How much is each of the following people <u>willing to listen to your work-related problems?</u>

		Not at all	A little	Some-what	Very Much	
a	Your immediate supervisor	0	1	2	3	
b	Other people at work	0	1	2	3	
c	Your spouse/partner	0	1	2	3	N/A
d	Your friends and relatives	0	1	2	3	

3. How much is each of the following people helpful to you in getting your job done?

		Not at all	A little	Some -what	Very Much	
a	Your immediate supervisor	0	1	2	3	
b	Other people at work	0	1	2	3	
c	Your spouse/partner	0	1	2	3	N/A
d	Your friends and relatives	0	1	2	3	

Please indicate how true each of the following statements is of your immediate supervisor.

		Not at all true	Not too true	Some-what true	Very True
4.	My supervisor is competent in doing (his/her) job.	0	1	2	3
5.	My supervisor is very concerned about the welfare of those under him/her.	0	1	2	3
6.	My supervisor goes out of his/her way to praise good work	0	1	2	3

House, J. & Wells, J.S. (1978). Occupational stress, social support, and health. In A. McLean, G. Black, & M. Colligan (Eds.), Reducing Occupational Stress: Proceedings of a Conference (pp. 8–29). Washington, DC, US: National Institute of Occupational Health and Safety.

General Self-Efficacy Scale

	Not at all true	Hardly true	Moderately true	Exactly true
1. I can always manage to solve difficult problems if I try hard enough	☐	☐	☐	☐
2. If someone opposes me, I can find the means and ways to get what I want.	☐	☐	☐	☐
3. It is easy for me to stick to my aims and accomplish my goals.	☐	☐	☐	☐
4. I am confident that I could deal efficiently with unexpected events.	☐	☐	☐	☐
5. Thanks to my resourcefulness, I know how to handle unforeseen situations.	☐	☐	☐	☐
6. I can solve most problems if I invest the necessary effort.	☐	☐	☐	☐
7. I can remain calm when facing difficulties because I can rely on my coping abilities.	☐	☐	☐	☐
8. When I am confronted with a problem, I can usually find several solutions.	☐	☐	☐	☐
9. If I am in trouble, I can usually think of a solution	☐	☐	☐	☐
10. I can usually handle whatever comes my way.	☐	☐	☐	☐

Schwarzer, R., & Jerusalem, M. (1995). Measures in health psychology: A user’s portfolio. In & M. J. J. Weinman, S. Wright (Ed.) (pp. 35–37). NFER-NELSON.

Work-Family Conflict Scale

Below are five statements with which you may agree or disagree. Using the 1 – 7 scale below, indicate your agreement with each item by circling the appropriate number. The words “work” and “job” refer to all work-related activities that you do as part of your paid employment. The word “family” refers to the following family roles that pertain to you, including being a parent, being a spouse/partner, and overall homelife.
1 = strongly disagree; 2 = disagree; 3 = slightly disagree;
4 = neither agree nor disagree; 5 = slightly agree; 6 = agree; 7 = strongly agree

1. The demands of my work interfere with my home and family life.

1 2 3 4 5 6 7

2. The amount of time my job takes up makes it difficult to fulfil my family responsibilities.

1 2 3 4 5 6 7

3. Things I want to do at home do not get done because of the demands my job puts on me.

1 2 3 4 5 6 7

4. My job produces strain that makes it difficult to fulfil family duties.

1 2 3 4 5 6 7

5. Due to work-related duties, I have to make changes to my plans for family activities.

1 2 3 4 5 6 7

Netemeyer, R. G., Boles, J. S., & Mcmurrian, R. (1996). Development and Validation of Work -- Family Conflict and Family -- Work Conflict Scales. Journal of Applied Psychology, 81(4), 400–410

Job Diagnostic Survey Autonomy subscale

1 = very inaccurate, 2 = mostly inaccurate, 3 = slightly inaccurate, 4 = uncertain, 5 = slightly accurate, 6 = mostly accurate, 7 = very accurate

1. I have almost complete responsibility for deciding how and when the work is to be done.
2. I have very little freedom in deciding how the work is to be done.
3. My job does not allow me an opportunity to use discretion or participate in decision making
4. My job gives me considerable freedom in doing the work

Hackman, J. R., & Oldham, G. R. (1974). The Job Diagnostic Survey: An instrument for the diagnosis of jobs and the evaluation of job redesign projects (Tech. Rep. No.4). Yale University

Job Discrepancy and Satisfaction Scale

Instructions: Please rate how much you agree with each of the following statements. Choose a score from 1 to 4 in reference to the scale below:

1 – Not at all satisfying; 2 – Somewhat satisfying; 3 – Moderately satisfying; 4 – Very satisfying

1 ______ How does the type of work that you currently do compare to what you think it should be?
2 ______ How does the amount of pay that you currently receive compare to what you think it should be?
3 ______ How do the number of opportunities for promotion that you currently have compare to what you think they should be?
4 ______ How does the quality of supervision that you currently receive compare to what you think it should be?
5 ______ How does the quality of colleagues and people you currently work with compare to what you think it should be?
6 ______ How do the working conditions in your job compare to what you think they should be?
7 ______ How does the amount of autonomy or personal freedom that you have compare to what you think it should be?
8 ______ How does your overall satisfaction with your current job compare to what you think it should be?

Source:

Nagy, M.S. (2002). Using a single-item approach to measure facet job satisfaction. *Journal of Occupational and Organizational Psychology*, 75, 77-86.

Overcommitment subscale

The following items refer to your present occupation. For each of the following statements, please indicate whether you strongly agree, agree, disagree or strongly disagree.

		Strongly disagree	*Disagree*	*Agree*	*Strongly agree*
OC1	I get easily overwhelmed by time pressures at work.	□	□	□	□
OC2	As soon as I get up in the morning I start thinking about work problems.	□	□	□	□
OC3	When I get home, I can easily relax and 'switch off' work.	□	□	□	□
OC4	People close to me say I sacrifice too much for my job.	□	□	□	□
OC5	Work rarely lets me go, it is still on my mind when I go to bed.	□	□	□	□
OC6	If I postpone something that I was supposed to do today I'll have trouble sleeping at night.	□	□	□	□

Siegrist, J., Li, J., & Montano, D. (2014). Psychometric properties of the effort-reward imbalance questionnaire. Duesseldorf University, 1–14.

Personality BFI-10 Measure

Instruction: How well do the following statements describe your personality?

I see myself as someone who …	Disagree strongly	Disagree a little	Neither agree nor disagree	Agree a little	Agree strongly
… is reserved	(1)	(2)	(3)	(4)	(5)
… is generally trusting	(1)	(2)	(3)	(4)	(5)
… tends to be lazy	(1)	(2)	(3)	(4)	(5)
… is relaxed, handles stress well	(1)	(2)	(3)	(4)	(5)
… has few artistic interests	(1)	(2)	(3)	(4)	(5)
… is outgoing, sociable	(1)	(2)	(3)	(4)	(5)
… tends to find fault with others	(1)	(2)	(3)	(4)	(5)
… does a thorough job	(1)	(2)	(3)	(4)	(5)
… gets nervous easily	(1)	(2)	(3)	(4)	(5)
… has an active imagination	(1)	(2)	(3)	(4)	(5)

An additional Agreeableness item will be added based on the advice of the authors to increase the validity of this subscale and include the item *"…is considerate and kind to almost everyone*".

Rammstedt, B. & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41*, 203-212.

Appendix C – Invitation to Participate Email

Subject: Mental Health Staff Burnout Study

Dear Mental Health Colleague

You are receiving this email as your NHS Trust is supporting a study of professional burnout in mental health clinical professionals. The study aims to develop a better understanding how different pressures and resources, both organisational and personal, impact the risk of exhaustion and disengagement at work, otherwise known as 'burnout'. The hope is that the results will feed into the development of tools which enable staff to self-monitor potential difficulties with their professional role as well as provide information on suitable interventions to support staff.

If you would like to find out more about this study and what it involves or take part in this study, please click on the link below.

Mental Health Staff Burnout Study

Many thanks and best wishes

Ben Davis, PhD
Trainee Clinical Psychologist
Clinical Psychology Unit
Department of Psychology
University of Sheffield
Email bdavis1@sheffield.ac.uk

Appendix D – Study Information Sheet

The University Of Sheffield. Department of Psychology

Development of a Predictive Model of Mental Health Staff Burnout

Researchers

Lead Researcher: Ben Davis (bdavis1@sheffield.ac.uk)

Supervised by: Dr Jaime Delgadillo (j.delgadillo@sheffield.ac.uk) and Professor Michael Barkham (m.barkham@sheffield.ac.uk)

Study Information

You are being invited to participate in this research project. This information sheet explains why the research is being done and what it will involve, to help you decide whether you would like to take part. Please take the time to read the following information carefully.

Why have I been chosen?

We are contacting all staff who engage in clinical work with patients at the NHS Trust where you work.

What is the study about?

We aim to develop a model of the most important factors which predict exhaustion and disengagement, or burnout, for mental health staff who engage in clinical work with patients. Using information about work pressures (e.g. workload) and work resources (e.g. supervision) as well as personal attributes (e.g. experience; social support), we aim to establish which of these demands and resources (or lack of) are most indicative of risk of burnout. The findings may help develop interventions to help staff at risk of burnout.

What will taking part involve?

The first part of the study involves completion of an online questionnaire asking about your workload and work environment. There are also questions about you which relate to how you feel about your role and aspects of your life which may affect the impact your job has on you. This will take about 10-15 minutes to complete.

The second part of the study involves completing a short questionnaire every month about how you are currently feeling about your job and this will take less than 5 minutes to complete.

Are there any disadvantages or risks of taking part?

Some of the questions ask you how you respond in certain situations and how you feel about your job. This might make you feel embarrassed and/or uncomfortable or may make you think more about how you feel about your current role. If this does happen you can choose to complete the survey later or withdraw from the study altogether. If taking part in this study does raise awareness of difficulties in your role then you may find it helpful to speak to work colleagues or your line manager or contact workplace wellbeing.

Are there any benefits in taking part?

Reflecting on the demands of your role and what helps you cope may have some benefits however this is not the intention of the study. We hope that the results will feed into processes to support staff who may be at risk of burnout or help inform interventions for staff who show high levels of burnout. If you would like to see a simple summary of your own burnout data when all data has been collected prior to analysis, then you can email the lead investigator to receive this if you are happy for this to be sent via email.

Who is organising and funding the study?

The University of Sheffield is organising and funding this study.

Who has ethically approved this study?

This study has been ethically approved by the NHS Research and Ethics Committee.

Legal statement under the General Data Protection Regulation (GDPR)
The University of Sheffield is the sponsor for this study based in England. We will be using information from you in order to undertake this study and will act as the data controller for this study. This means that we are responsible for looking after your information and using it properly. The University of Sheffield will keep identifiable information about you until you have completed the study in 2019. Anonymised information about you will be kept for 5 years after the study has finished until 2024.
Your rights to access, change or move your information are limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate. If you withdraw from the study after all your data has been collected, we will keep the information about you that we have already obtained. To safeguard your rights, we will not keep any identifiable information.

You can find out more about how we use your information at https://www.sheffield.ac.uk/govern/data-protection/privacy/general.

What will be done with the data and results?
To take part in the study you will be asked to use your work email address as a username, and if you consent to take part we will use it to send you links for the second part of the study each month. All data will be anonymized and held securely, and your Trust will only have access to anonymous analysed data which would be available on publication.

The results from this study will be written up and submitted as a thesis for the clinical psychology doctorate at the University of Sheffield. Additionally, the results will be disseminated through publishing in a peer-reviewed journal. No participants will be identifiable in any publications as data will be pooled from all participants.

What if I wish to complain about the way the study has been carried out?
If you would like to make a complaint about this project, in the first instance you should contact the lead researcher. If you do not feel satisfied that your complaint has been dealt with appropriately you can contact the lead

researcher's supervisor. If you feel that your complaint has not been handled to your satisfaction following this, you can contact. Prof Glenn Waller, Head of Department at g.waller@sheffield.ac.uk

What next?

If you have any questions or would like an electronic or paper copy of this information please email bdavis1@sheffield.ac.uk. If you would like to take part, then please click the link below to consent to the study and then complete the first part of the survey. As previously stated, you can withdraw at any time.

Consent to take part in the study

Appendix E – Consent Form

I confirm that I have read and understood the information sheet explaining the research project led by Ben Davis entitled **Development of a Prognostic Model of Mental Health Staff Burnout Using Machine Learning**:

○ **I confirm**

I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without there being any negative consequences. I understand that it will not be possible to remove my anonymized data from the study once data collection is completed and data analysis has begun:

○ **I understand**

I understand that my data will be kept strictly confidential. I give permission for members of the research team to have access to my anonymised data. I understand that my name will not be identified or identifiable in any reports that result from the research:

○ **I understand**

I give my full consent to take part in the study:

○ **Yes** ○ **No**

Appendix F – NHS Ethical and HRA Approval

NHS
Health Research
Authority

East of England - Cambridge Central Research Ethics Committee

Royal Standard Place
Nottingham
NG1 6FS

Please note: This is the favourable opinion of the REC only and does not allow you to start your study at NHS sites in England until you receive HRA Approval

14 February 2019

Dr Jaime Delgadillo
Clinical Psychology Unit
Cathedral Court,
Vicar Lane Sheffield
S1 2LT

Dear Dr Delgadillo

Study title: **Development of a Prognostic Model of Mental Health Staff Burnout Using Machine Learning**
REC reference: **19/EE/0054**
IRAS project ID: **256342**

The Proportionate Review Sub-committee of the East of England - Cambridge Central Research Ethics Committee reviewed the above application on 06 February 2019.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all studies that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact hra.studyregistration@nhs.net outlining the reasons for your request. Under very limited circumstances (e.g. for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the study.

Ethical opinion

On behalf of the Committee, the sub-committee gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Conditions of the favourable opinion

The REC favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission must be obtained from each host organisation prior to the start of the study at the site concerned.

Truncated …

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see “Conditions of the favourable opinion”).

Extract of the meeting minutes

Ethical issues raised, noted and resolved in discussion:

The PR Sub-Committee agreed that this was a well presented study with no material ethical issues.

Approved documents

Truncated…

With the Committee’s best wishes for the success of this project.**19/EE/0054**
Please quote this number on all correspondence

Yours sincerely

pp.

Dr Gusztav Belteki
Chair

Email: NRESCommittee.EastofEngland-CambridgeCentral@nhs.net

Enclosures: *List of names and professions of members who took part in the review*

Copy to: *Dr Benjamin Davis*

Ymchwil Iechyd a Gofal Cymru
Health and Care Research Wales

NHS
Health Research Authority

Dr Jaime Delgadillo
Clinical Psychology Unit
Cathedral Court,
Vicar Lane Sheffield
S1 2LT

Email: hra.approval@nhs.net
Research-permissions@wales.nhs.uk

20 February 2019

Dear Dr Delgadillo

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

Study title: Development of a Prognostic Model of Mental Health Staff Burnout Using Machine Learning
IRAS project ID: **256342**
REC reference: **19/EE/0054**
Sponsor **University of Sheffield**

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.

How should I continue to work with participating NHS organisations in England and Wales? You should now provide a copy of this letter to all participating NHS organisations in England and Wales, as well as any documentation that has been updated as a result of the assessment. Participating NHS organisations in England and Wales **will not** be required to formally confirm capacity and capability before you may commence research activity at site. As such, you may commence the research at each organisation 35 days following sponsor provision to the site of the local information pack, so long as:

- You have contacted participating NHS organisations (see below for details)
- The NHS organisation has not provided a reason as to why they cannot participate
- The NHS organisation has not requested additional time to confirm.

You may start the research prior to the above deadline if the site positively confirms that the research may proceed.

Truncated …

Who should I contact for further information?
Please do not hesitate to contact me for assistance with this application. My contact details are below. Your IRAS project ID is **256342**. Please quote this on all correspondence.
Yours sincerely

Joanna Ho
Assessor

Email: hra.approval@nhs.net

Copy to: *Amrit Sinha, Sponsor Representative, University of Sheffield*
Dr Benjamin Davis, Student, University of Sheffield

Appendix G – Power calculations

A priori calculations

$f^2 = \frac{R^2}{1-R^2}$ (J. Cohen, 1992). Previous studies showed R^2 of >=0.25 gives estimated large effect size f^2=0.33. Instead cautious estimation of effect size of f^2=0.15 (moderate) was used. For α=.05, variables=36, power =80%, minimum sample size was estimated at 203. Accounting for 80%:20% train-test paradigm gives sample of 254 (203/.8). Accounting for 10-fold cross-validation gives 283 (254/.9). Accounting for 60% attrition gives 708 (283/.4). All values rounded up.

Post hoc calculations

Multiple linear regression models of all factors or reduced variable set factors (OLBI T1 scores excluded) for OLBI at T7, resulted in large adjusted R^2 values (>=0.38) This equates to a large effect size (f^2=0.61). Pseudo adjusted R^2 for logistic regression models were almost identical (0.39). For actual sample size of 151, training samples were n=101. Post hoc power calculations using G*power, α=.05, f^2=0.61, variables=36, sample size=101 gave power as 97.9%