Social Science & Medicine 285 (2021) 114268

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Mental and physical health of international humanitarian aid workers on short-term assignments: Findings from a prospective cohort study

K. De Jong [a,*], S.E. Martinmäki [b], H. Te Brake [b], J.F.G. Haagen [b], R.J. Kleber [c,d]

[a] Médecins Sans Frontières (MSF), the Netherlands
[b] ARQ Centre of Expertise on the Impact of Disasters and Crises, the Netherlands
[c] Utrecht University, the Netherlands
[d] ARQ National Psychotrauma Centre, the Netherlands

ARTICLE INFO

Keywords:
Humanitarian staff mental health
Stress
Post-traumatic
Conflict
Well-being
Burnout

ABSTRACT

Research findings show humanitarian work impacts one's health. We conducted a prospective observational study among 618 international humanitarian aid workers (iHAWs)' recruited from 76 countries to investigate health changes and ill-health risk factors after mostly short-term (<1 year) medical emergency assignments. The aid workers were assigned to 27 countries. Data collected between 2017 and 2020.

We also compared a gold-standard clinical interview with self-report questionnaires to assess whether self-report scores overestimate the prevalence of clinical anxiety, depression and PTSD. Analyses consisted of repeated measures ANOVAs and adjusted odds ratios, using pre-assignment (T1), post-assignment (T2) and two-month follow-up data (T3). Humanitarian workers experienced on average, 2.6 experienced and witnessed potential traumatic events, and 4.8 male and 5.6 female assignment-related stressors. Self-report health indicators demonstrated a significant increase in emotional exhaustion, loss of vitality, decreased social functioning and emotional well-being between T1 and T2, all of which improved between T2 and T3. PTSD, depression, experienced role limitations, physical functioning, pain, and general health – remained stable. Anxiety levels decreased significantly between T1 and T2. The presence of DSM-5 disorders anxiety (6.6 %), depression (1.3 %) and PTSD (0.3 %) was low compared to norm populations, except for alcohol-use disorder (13 %). None of the reported T2 risk factors was significant at T3. Compared to the clinical interview, self-report cut-off thresholds inflated the presence of a potential anxiety disorder (3×), PTSD (8×) and depression (25×). Humanitarian work is highly stressful but most iHAWs remained healthy. Looking into how iHAWs stay healthy may be a more useful way forward.

1. Introduction

Currently, over 200 million people need humanitarian assistance, mostly due to conflicts and disasters. Due to growing needs record funding levels (US$27.3) were reported in 2017 (Overseas Development Assistance, 2019). Aid workers are a diverse group of international staff, professional consultants, and locally contracted national staff working (Stoddard et al., 2019).

IHAWs tend to have higher stress levels compared to general population (Jachens et al., 2019; Young and Pakenham, 2020). This may be the result of potentially psycho-traumatic events, such as attacks on aid workers. Despite a stabilizing trend in the previous decade, the number of attacks on aid workers is on the rise, with causalities exceeding all past years in 2019 (Stoddard, 2020). Most attacks occurred in six countries (Syria, South Sudan, DRC, Afghanistan, CAR, Mali, Yemen). Males, compared to females, were three times more exposed to attacks but sexual harassment and violence was reported mainly by females (Gritti, 2015; Stoddard et al., 2019). Males, compared to females, were three times more exposed to attacks but sexual harassment and violence was reported mainly by females. Most aid workers themselves however do not consider trauma as a key stressor (Young et al., 2018).

Workers are also confronted with chronic, assignment-related stress such as the overwhelming needs of the beneficiaries, and lack of resources (Holtz et al., 2002). This can evoke moral dilemmas and feelings of helplessness among iHAWs (Eriksson et al., 2009). Organisational stressors such as high deployment frequency, work conflicts, poor

* Corresponding author. MSF, Plantage Middenlaan 14, 1018DD, Amsterdam, the Netherlands.
E-mail address: Kaz.de.Jong@amsterdam.msf.org (K. De Jong).

https://doi.org/10.1016/j.socscimed.2021.114268
Received 5 October 2020; Received in revised form 7 July 2021; Accepted 23 July 2021
Available online 4 August 2021

management, a lack of management support, a lack of reciprocity, perceived inequity at work, and heavy workload are likely more important causes of stress for iHAWs (Cardozo et al., 2005; Dubravka et al., 2016; Eriksson et al., 2009).

Above mentioned (organisational) stressors have been identified as risk factors for the physical and mental health of iHAWs (Young et al., 2018). IHAWs reported elevated levels of mental distress from post-traumatic stress disorder (PTSD), anxiety, depression, alcohol misuse and burnout compared to the general population (Connorton et al., 2012; Young and Pakenham, 2020). These findings support the prevailing perspective that iHAWs' exposure to extreme and chronic stress is high and gives rise to health problems that impede the delivery of humanitarian assistance (Connorton et al., 2012; Lopes Cardozo et al., 2012). For instance, through staff turnover, institutional knowledge loss, and increased healthcare costs (Korff et al., 2015). A recent special reported warns against a growing epidemic of aid workers with psychological trauma (Macpherson and Burkle, 2021).

The reported elevated levels of mental distress and the deterioration in quality of life are, however, in contrast to reported stable or improving health outcomes of professionals confronted with potentially traumatic assignment experiences and distressing work-related situations (Kim et al., 2017; Koen et al., 2011). These discrepancies are possibly due to methodological limitations in previous aid work study designs. Firstly, only a few studies systematically addressed the full scope of the humanitarian work context, such as organisational and environmental stressors (Brooks et al., 2015). Secondly, except for one longitudinal study (Lopes Cardozo et al., 2012), aid worker health studies were cross-sectional, not allowing for conclusions on causality (Galea et al., 2008). Thirdly, other methodological shortcomings included using exclusively self-report questionnaires without gold standard clinical interview, anecdotal reports, both inflating the risk of biased results (Limb, 2011), and the use of small sample sizes, reducing the statistical power to undesirable levels. Self-report instruments can overestimate the prevalence and severity of mental health disorders (Charlson et al., 2019).

The present prospective study aims to clarify discrepancies between iHAWs' health studies by reporting on a structured clinical interview and a comprehensive set of self-reporting health indicators, focusing on various mental and physical health indicators, and quality of life, as well as traumatic, organisational and environmental stressors. We expect, despite high levels of stress, that the iHAWs' health remains stable on all indicators (Kim et al., 2017; Koen et al., 2011).

Other objectives were to establish the prevalence of DSM-5 disorders after a humanitarian assignment and to determine potential demographic and assignment-related risk factors for ill-health. Lastly, we assessed whether self-report questionnaires for common disorders inflated actual pathology in comparisons to a gold standard clinical interview as has been found in psychopathology research (Charlson et al., 2019).

2. Method

2.1. Participants

The sample consisted of 609 iHAWs of Médecins Sans Frontières (MSF) Operational Centre Amsterdam. International humanitarian aid workers are defined as staff not from the country within which they are working (Egeland et al., 2011, p. 59). They were aged between 24 and 76, mostly female, European and with a university degree. Most iHAWs had prior iHAW experience (78 %). The participants originated from 76 different countries: 24 countries in Europe, 21 countries in Africa, 4 countries in North America, 21 countries in Asia, 2 countries in Oceania, and 4 countries from South America. Some (16 %) had experiences as paid aid workers in their home countries. Deployment was mostly short-term (1< year), emergency focussed in 27 different countries. Assignments were in high security settings: Syria and/or Iraq (n = 100, 20 %), Democratic Republic of Congo (n = 62, 12 %), Bangladesh (n = 59, 12 %), South Sudan (n = 47, 9 %), and Nigeria (n = 42, 8 %) . See Appendix 1 for an overview of all aid assignment destinations. Complete and detailed demographic information about the participants are presented in Table 1.

We expected a low prevalence of post-assignment mental health issues and aimed for a sample size that could detect small effect sizes (f = 0.10). With 15 planned repeated measures ANOVAs, the required Bonferroni alpha correction, and an expected pre-post-follow-up dropout of 26 % (Lopes Cardozo et al., 2012), a sample of at least 508 participants would provide adequate power. This study received ethical approval

Table 1
Participant information (n = 609).

Pre-assignment	***N* (%)**
Age	
In years *(M, SD)*	40.5 (10.8)
Biological sex	
Female	343 (56.3)
Male	266 (43.7)
Continent of origin	
Africa	62 (10.6)
Asia	80 (13.7)
Europe	301 (51.6)
N. America	111 (19.0)
S. America	14 (2.4)
Oceania	15 (2.6)
Education	
Secondary or high school	10 (1.8)
Higher vocational training/technical training	40 (7.1)
University degree: Bachelors or Masters	373 (66.1)
Postgraduate degree	141 (25.0)
Relationship status	
Single, never married	249 (41.9)
Married	130 (21.9)
Committed relationship but not married	131 (22.1)
Separated	31 (5.2)
Divorced	47 (7.9)
Widowed	6 (1.0)
Assignment function	
Coordinator	179 (31.8)
Activity manager & clinical medical specialist	371 (63.9)
Supervisor & specialist	26 (4.5)
Other	7 (1.2)
Prior assignment experience	
First-timer	110 (21.9)
Veteran	392 (78.1)
Number of assignments *(M, SD)*	4.7 (5.7)
Previously worked as national staff	
Any experience	73 (15.5)
No experience	397 (84.5)
In years *(M, SD)*	5.0 (3.7)
Early departure	
Early departure	80 (16.4)
Post-assignment	
Assignment duration	
<3 months	148 (29.5)
3–6 months	167 (33.3)
6–9 months	100 (19.9)
>9 months	87 (17.3)
In months *(M, SD)*	6.4 (3.9)
Return information	
As planned	287 (57.3)
Completed after extended	135 (26.9)
Needed to evacuate	8 (1.6)
Project/position closed	4 (0.8)
Early return	61 (12.2)
Other	6 (1.2)

Note. National staff refers to locally recruited staff members from the aid-recipient country (Stoddard et al., 2009). Early departure: departure within two months after previous assignment. First-timer: no prior NGO experience. Assignment experience in number of assignments, national staff experience in years: based on the sample with any experience.

from the internal Ethics Review Board of Médecins Sans Frontières on February 24, 2017 (ID 1642).

2.2. Study design

The current study was a prospective observational study with three measurement occasions: pre-assignment, post-assignment and a two-month follow-up.

2.3. Procedure

All iHAWs with field-based contracts and start- and end-of-assignment dates between December 2017 and February 2019 were eligible: data collection ended February 2020. Office staff going on brief field visits were not considered eligible. An independent non-MSF researcher informed and recruited participants during their face-to-face pre-assignment MSF office briefings or via a video call. All participants signed an informed consent.

The participants completed the questionnaires on an online survey platform, either at the office or remotely. The pre-assignment measurement (T1) took place 0–14 days before travelling to the assignment area, the post-assignment measurement (T2) immediately during debriefing, with a maximum of four weeks after returning. Participants debriefing face-to-face were asked to participate in a clinical interview. Two trained non-MSF psychologists (S.M, J.H.) conducted these interviews. The two-month follow-up measurement (T3) was done remotely on the aforementioned online survey platform. T3 took place earlier (4–8 weeks after T2) if the participant was due to leave for a new assignment within two months. In case a participant was already on a new assignment the follow-up measurement was stopped; this occurred 8 times (1.3 %) at post-assignment and 46 times (7.6 %) at follow-up. The data collection ended prematurely on March 12, 2020 due to Dutch government COVID-19 measures, logistical issues (borders closing) and to avoid confounding the study results.

Participants' scores were monitored for severe and acute suicidal ideation after completing every measurement; no cases were detected.

2.4. Measures

All measures, including those outside the scope of the current study, are presented in Appendix 2. Demographic information was collected at T1, assignment characteristics at T2 and assignment-related healthcare services utilisation at T3. We provide descriptions for current study measures below.

Stressor measures. The *Life Events Checklist for DSM-5* (LEC-5) (Weathers, 2013) screens for self-reported potentially traumatic events (PTEs). It was used at pre-assignment (T1) to determine the extent of lifetime exposure while at post-mission (T2) it assessed exposure to PTEs during the assignment. The psychometrics for the LEC-5 are not available. Given the minimal revisions between the LEC-5 and the original, psychometrically adequate, version of the LEC (Gray et al., 2004), few psychometric differences are expected. The LEC internal consistency score was not checked because it is unlikely that respondents experience the same patterns or clusters of events.

The MSF designed *Humanitarian Field Stressor List (HFSL)* assessed the number and severity of 39 potential humanitarian field stressors. The items were rated on a six-point scale ranging from 0 ('none/not applicable') to 5 ('high'). Six dimensions were assessed: field conditions, cultural stressors, work-related stressors, team stressors, code of conduct and experienced traumatic experiences.

Health outcome measures. The health indicators were assessed at all measurement moments. The *Hopkins Symptom Checklist* (HSCL-25) (Parloff et al., 1954), a 25-item self-report questionnaire (four-point scale ranging from 1 ("not at all") to 4 ("often")), was used to assess symptoms of anxiety (10 items) and depression (15 items) during the past week. Subscale scores were calculated by summing up the standardised items scores. A cut-off score of 1.75 was used to screen for elevated symptoms of depression or anxiety (Derogatis et al., 1974). HSCL-25 has adequate psychometric properties (Mollica et al., 1987), can be used cross-culturally (Tinghög and Carstensen, 2010), and in non-psychiatric populations (Winokur et al., 1984). The internal consistency in the current sample was high for both depression ($\alpha = 0.90$) and anxiety ($\alpha = 0.87$) subscales.

The *Post-Traumatic Check List DSM-5* (PCL-5) (Blevins et al., 2015) measured PTSD DSM-5 symptoms using 20 self-report items rated on a six-point scale (0 ('not at all') to 5 ('extremely')). Items rated as 2 ('moderately') or higher were treated as a PTSD symptom endorsement. A provisional PTSD diagnosis was established if participants reported at least: 1 Criterion-B item, 1 Criterion-C item, 2 Criterion-D item, and 2 Criterion-E item symptom endorsements, and endorsed at least one PTE on their LEC-5 score. Internal consistency, test-retest reliability, convergent and divergent validity are good (Blevins et al., 2015). The PCL-5 detects clinical change over time (Wortmann et al., 2016). In the current sample, the scale had a good internal consistency ($\alpha = 0.89$).

The Maslach Burnout Inventory (MBI-HSS) (Maslach et al., 1981) included 16 items (0 ('never') to 6 ('always')), clustered around three general burn-out scales: emotional exhaustion, depersonalisation and personal accomplishment. The MBI has a good construct, factorial and congruent validity when compared to another known burnout measure (Schaufeli and Dierendonck, 1993). Internal consistency of the subscales was good in previous studies (Schutte et al., 2000) and acceptable to good in the current sample ($0.67 < \alpha < 0.84$).

The *RAND-36* (Hays and Morales, 2001) had 36 items (2–6 response categories) assessing eight dimensions of health: physical functioning, role limitations caused by physical health problems and/or emotional problems, social functioning, emotional well-being, vitality (energy/fatigue), pain, and general health perceptions. Each scored item was transformed to a 0–100 range (highest possible health-related quality of life: 100, lowest possible: 0). The internal consistency and convergent validity of the instrument are high (Van der Zee et al., 1996). In the current sample the internal consistency of the scales was considered good to acceptable ($0.82 < \alpha <. 71$), except for the general health perceptions scale that was considered poor ($\alpha = 0.51$).

We applied the semi-structured *International Neuropsychiatric Interview* (M.I.N.I. 7.0.2) (T2) (Sheehan et al., 1998) to identify a range of DSM-5 diagnoses commonly seen in clinical settings: a major depressive disorder, anxiety disorders (panic disorder, agoraphobia, social anxiety, obsessive-compulsive disorder or generalised anxiety), post-traumatic stress disorder, eating disorders (anorexia nervosa, bulimia nervosa or binge-eating disorder), or alcohol or substance use disorder. The M.I.N.I. has demonstrated good to very good reliability (Sheehan et al., 1998), sensitivity and specificity scores (Hergueta and Weiller, 2013). Predictive values of 0.85 or higher have been found across all diagnoses (Sheehan et al., 1998). One PhD and one master-level psychologist were trained as M.I.N.I interviewers and received supervision. Inter-rater reliability in the current sample was excellent, with Cohen's Kappa's between 0.96 and 1 for all the diagnoses. To determine whether or not screeners inflate actual pathology, we compared the clinical interview (M.I.N.I.) disorder prevalence rates of PTSD, major depression, and any anxiety disorder with the abovementioned clinical threshold scores for PTSD, depression and anxiety on their associated screeners (PCL-5 and HSCL-25). The clinical interview being the interpretation of a trained professional is regarded the gold standard for a diagnosis in mental health care settings.

2.5. Statistical analyses

We examined differences in baseline characteristics of 'participants' vs. 'not-informed iHAWs' (those without office briefings), 'participants' vs. 'decliners' and the primary outcome health indicators for the subgroups biological sex, level of education, previous national and international staff experience. Comparison was done using $\chi 2$ tests with

continuity correction (or Fisher's exact test when appropriate) and independent *t*-test comparisons or analyses of variance (ANOVA) with Scheffe post-hoc comparisons for continuous data.

To test the likelihood of longitudinal score changes over time we used repeated measures (RM) ANOVAs on the health outcome indicators. Post-hoc pairwise comparison analyses between two measurement moments (T1-T2, T1-T3, T2-T3) were performed on significant outcomes to determine between which measurement moments health changed.

Assumptions required to perform a reliable RM ANOVA were checked for each indicator: the Greenhouse-Geisser correction demonstrated near-perfect sphericity ($\varepsilon \geq 0.95$) for all health indicators. Z-score tests combined with visual inspections of change score histograms and Q-Q plots demonstrated approximate normal distributions all but one indicators.

RM ANOVAs only used data of participants that completed all three measurements. To detect potential non-response bias in the RM ANOVA we performed sensitivity analyses by repeating the RM ANOVAs using multiple imputed (MI) data for missing scale score values. If the RM ANOVA findings with imputed data correspond with prior RM ANOVA results, it strengthens the reliability of our findings. We created 20 imputation data sets using partitioned predictive mean matching (Vink et al., 2015).

Correlational analyses were executed between demographic, assignment-specific risk-factor variables and post-assignment and follow-up clinical threshold scores (i.e., ill-health) on the health indicators. All risk-factor variables with at least a moderate correlation ($r > .30$) with the health indicators were analysed using adjusted odds ratios (AOR).

SPSS Statistics version 25 (IBM Corp, Armonk, NY, USA) was used for all statistical analyses and the imputation of missing data. We applied a two-sided 5 % level of significance for all tests of statistical hypotheses. A Cohen's d effect size was calculated for significant differences on mean score of variables between two groups (Sullivan and Feinn, 2012). A Cohen's d of 0.2 was considered small, 0.5 medium, and 0.8 large. Partial eta squared (η^2) effect sizes were reported for RM ANOVA analyses. A η^2 of < 0.01 is considered small, 0.06 medium, and ≥ 0.14 large (Levine and Hullett, 2002).

3. Results

3.1. Participants

Recruitment. Within the data collection period, 1391 iHAWs departed on an assignment, 794 (58 %) were briefed in the Amsterdam office. Eligible staff members (n = 689) were informed about the study, 618 (89 %) agreed to participate, and 609 (88 %) commenced in the present study.

Study flow. Ninety-six percent (n = 594) of those who agreed to participate in the study completed pre-assignment measures before arriving at the project site. Participant retention at post-assignment measure was good (82 %, n = 502). The average time between return from the assignment and completing the post-assignment measures was 7.0 days (*SD* = 9.4). Retention rates dropped to 61 % (n = 373) at two-month follow-up. With regard to the post-assignment clinical interview, 320 (63 %) of the 509 participants that returned via the Amsterdam office completed the interview. Fig. 1 provides details on the recruitment and study flow.

3.2. Stressors

Pre-assignment lifetime experienced PTEs. Almost all participants (96 %, n = 564) reported one or more lifetime experienced events (average 2.5 events, SD = 2.0). Most common experienced events were physical assault (37 %), transport accidents (34 %), and unwanted or uncomfortable sexual experiences – excluding sexual assault (31 %). Males most often reported experienced confrontations with non-sexual violence, such as physical assault (40 %), assault with weapon (30 %), and exposure to combat or warzone (30 %). Females reported substantial higher rates of sexual violence (19 %), and other unwanted or uncomfortable sexual experiences (46 %).

Assignment-related PTEs. Three quarters (76 %) of the participants reported exposure to a PTE during field assignment (experienced themselves, witnessed, and part-of-the-job). Participants reported on average 0.74 (SD = 1.33) experienced assignment-related PTEs. Most frequently reported experiences: combat or warzone exposure (13 %), physical assault (8.9 %), transport accident (8.9 %), or a natural disaster (8.3 %). If witnessing a PTE was included the mean number of events increased to 2.55 (SD = 2.79). Participants witnessed severe human suffering most often (44 %). Table 2 provides an overview of all experienced lifetime and assignment-related PTEs split by gender. Compared to males, females reported substantial higher assignment-related rates of experienced sexual assault (4 vs 0 %) and other unwanted or uncomfortable sexual experiences (11 vs 2 %). There were no gender differences in the number of experienced PTEs ($p > .05$).

Assignment-related stressors. The most frequently reported stressors were the climate (17 %), unclear (organisational) communication in the project (16 %), workload (16 %), travel to location or assignment destination (16 %), and the security context of the country (15 %). Female participants reported on average 5.6 (SD = 6.0) different sources causing 'significant' levels of stress (score 4 or higher). Male participants reported on average 4.8 (SD = 5.3) sources of significant stress. There were no significant gender differences in the number (T-test; $p = .15$) of experienced significant sources of stress. Chi-square analyses to assess whether females and males experienced different levels of stress for each source of stress. There were no significant differences on most stressors. Males more often reported experiencing significant to high levels of stress (all comparisons ($p < .05$) from being separated from family and friends (22 vs 14 %), and regarding the security and safety conditions (20 vs 12 %), compared to females. Females more often reported (all comparisons $p < .05$) experiencing significant to high levels of stress from unclear communication within the project (31 vs 19 %) and within the team, (26 vs 15 %), the lack of technical support (21 vs 10 %), the country management team (19 vs 12 %), feeling powerlessness or hopelessness (12 vs 7 %), regarding the composition of the team (9 vs 4 %), and from witnessed sexual harassment or violence toward colleagues (5 vs 2 %), compared to males. A gender stratified overview of all environmental, cultural and organisational stressors is presented in the Supplementary Materials (Appendix 3).

3.3. Health & health changes

Table 3 shows RM ANOVA changes on all health indicators including the portion of participants scoring above clinical threshold on indicators with established cutoff scores. Risk factor analyses estimated the likelihood of demographic and assignment-related variables to predict T2 and T3 clinical health levels.

Anxiety and depression (HSCL-25). Anxiety changed significantly ($p = .00$; medium effect size: $\eta^2 = 0.036$). The mean pre-assignment scores were significantly higher (M = 1.50) than post-assignment (M = 1.39) and follow-up scores (M = 1.39). A quarter (26 %, n = 144) of the participants reported clinical levels of pre-assignment anxiety. These rates dropped to 20 % (n = 97) at T2 and 21 % (n = 75) at T3. Being female, a greater number of years of previous national staff experience, a greater number of experienced and witnessed assignment-related PTEs increased the risk for suspected T2 clinical anxiety. None of the variables were risk factors at T3.

Mean depression severity levels remained stable ($p = .59$) over time with scores M = 1.59 (T1), M = 1.57 (T2), and M = 1.59 (T3). A third (30 %, n = 168) of the participants reported clinical levels of pre-assignment depression. These rates increased slightly to 32 % (n =

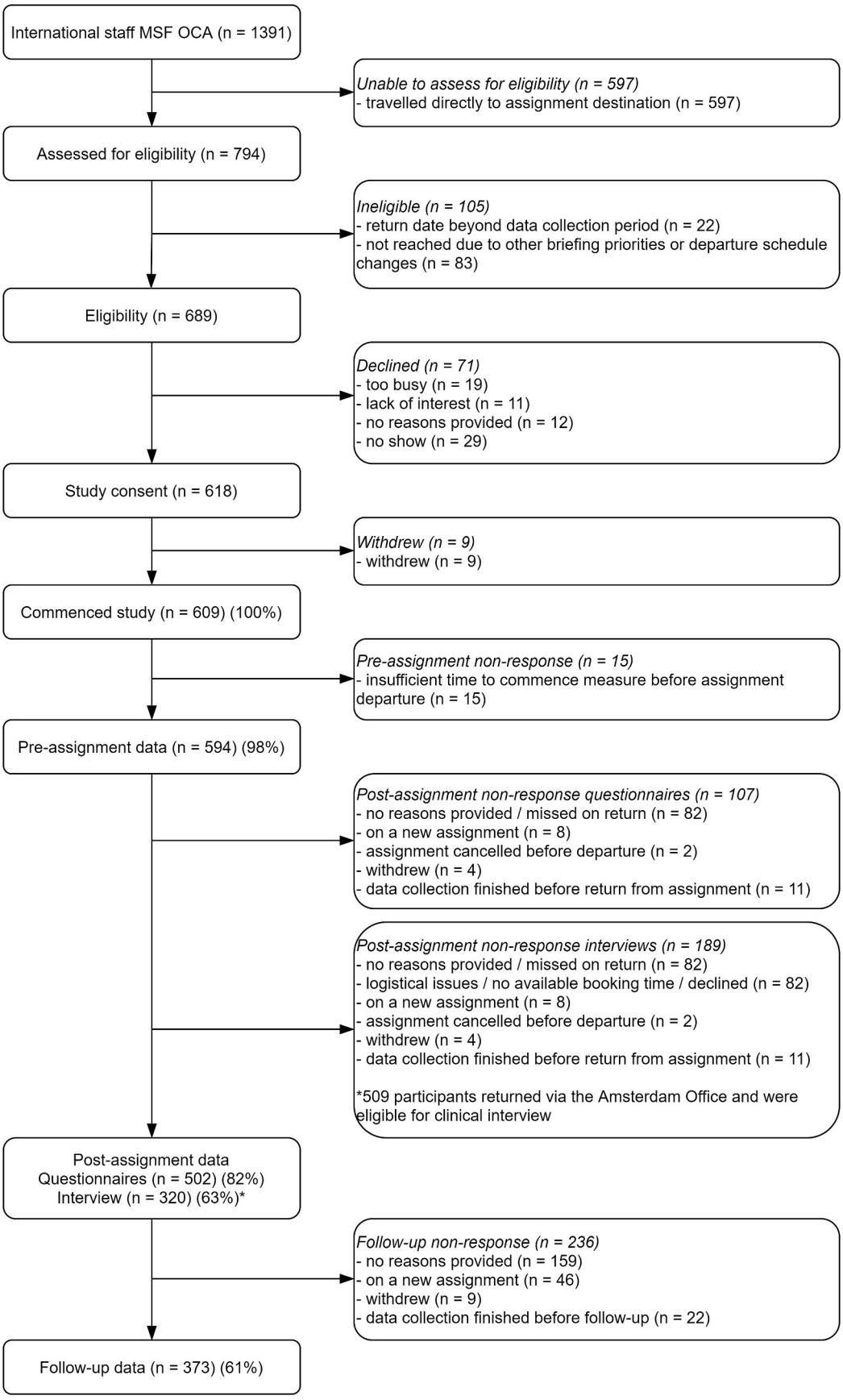

Fig. 1. Participant flow chart.

Table 2
Experienced shocking events (LEC) stratified for gender.

	Experienced Pre-assignment lifetime-events		Experienced Assignment-related		Witnessed Assignment-related		Part of job Assignment-related	
	Female	Male	Female	Male	Female	Male	Female	Male
	N(%)	*N(%)*	*N(%)*	*N(%)*	*N(%)*	*N(%)*	*N(%)*	*N(%)*
Natural disaster	75 (22.4)	55 (21.7)	24 (8.2)	16 (8.5)	20 (6.9)	33 (17.6)	19 (6.5)	22 (11.7)
Fire explosion	42 (12.5)	37 (14.6)	7 (2.4)	9 (4.8)	39 (13.4)	47 (25.0)	30 (10.3)	26 (13.8)
Transport accident	112 (33.4)	90 (35.4)	26 (8.9)	17 (9.0)	41 (14.1)	39 (20.7)	42 (14.4)	36 (19.1)
Serious accident at home, work or recreational	45 (13.4)	49 (19.3)	11 (3.8)	11 (5.9)	22 (7.6)	32 (17.0)	39 (13.4)	31 (16.5)
Exposure to toxic substance	11 (3.3)	15 (5.9)	4 (1.4)	6 (3.2)	10 (3.4)	14 (7.4)	17 (5.8)	30 (16.0)
Physical assault	118 (35.2)	102 (40.2)	24 (8.2)	17 (9.0)	30 (10.3)	32 (17.0)	67 (23.0)	29 (15.4)
Assault with a weapon	64 (19.1)	77 (30.3)	14 (4.8)	14 (7.4)	22 (7.6)	23 (12.2)	62 (21.3)	31 (16.5)
Sexual assault	62 (18.6)	2 (0.8)	11 (3.8)	0 (0)	8 (2.7)	5 (2.7)	76 (26.1)	26 (13.8)
Other unwanted or uncomfortable sexual experiences	154 (46.2)	30 (11.8)	32 (11.0)	4 (2.1)	12 (4.1)	6 (3.2)	52 (17.9)	16 (8.5)
Combat or warzone exposure	66 (19.8)	77 (30.3)	40 (13.7)	37 (19.7)	23 (7.9)	30 (16.0)	78 (26.8)	50 (26.6)
Captivity	9 (2.7)	8 (3.1)	2 (0.7)	1 (0.5)	15 (5.2)	8 (4.3)	34 (11.7)	17 (9.0)
Life-threatening illness or injury	36 (10.8)	33 (13.0)	6 (2.1)	6 (3.2)	59 (20.3)	39 (20.7)	99 (34.0)	48 (25.5)
Severe human suffering	12 (3.6)	12 (4.7)	4 (1.4)	2 (1.1)	129 (44.3)	80 (42.6)	146 (50.2)	76 (40.4)
Sudden violent death	10 (3.0)	8 (3.1)	2 (0.7)	0 (0)	24 (8.2)	19 (10.1)	66 (22.7)	44 (23.4)
Sudden accidental death	9 (2.7)	7 (2.8)	1 (0.3)	1 (0.5)	36 (12.4)	32 (17.0)	86 (29.6)	37 (19.7)
Serious injury, harm or death caused to someone	9 (2.7)	8 (3.1)	2 (0.7)	0 (0)	10 (3.4)	9 (4.8)	19 (6.5)	13 (6.9)

Table 3
Health changes over time – repeated measures ANOVAs & risk factor analyses – adjusted odds ratios.

Domain	+/−	Pre-assignment	Post-assignment	Follow-up	$F_{2,280}$	p	η^2	Sig. pairwise comparisons	Post-assignment risk factors	AOR
		M(SD)	M(SD)	M(SD)						
Anxiety										
Mean	–	1.50 (.48)	1.39 (.44)	1.39 (.44)	11.91	.00	.036	T1>T2*; T1>T3*	Female National staff exp. Number of PTEs	1.88 1.18 1.13
Depression										
Mean	–	1.59 (.51)	1.57 (.53)	1.59 (.57)	.52	.59			Female National staff exp. Number of PTEs	1.72 1.27 1.10
PTSD*										
Sum	–	8.89 (8.57)	7.84 (8.38)	8.19 (9.06)	2.09	.13				
Burnout										
Sum									Age Assignment length	0.69 0.78
Emotional exhaustion	–	1.66 (.89)	1.81 (1.11)	1.78 (1.02)	5.74	.00	.017	T1<T2*; T1<T3*		
Depersonalisation	–	1.15 (.88)	1.17 (1.03)	1.23 (1.00)	1.87	.16				
Personal accomplishments	+	4.80 (.82)	4.77 (.86)	4.67 (.85)	7.03	.00	.020	T1>T3*; T2>T3*		
Quality of life										
Physical functioning	+	95.0 (9.6)	94.1 (12.0)	95.3 (10.2)	2.19	.11				
Social functioning	+	86.9 (16.2)	83.6 (19.6)	85.1 (19.6)	4.26	.02	.014	T1>T2*		
Role limitations physical	+	90.5 (22.0)	88.5 (24.5)	86.6 (27.8)	2.58	.08				
Role limitations emotional	+	86.7 (27.3)	86.9 (26.4)	83.3 (30.8)	2.17	.12				
Emotional wellbeing	+	79.7 (13.0)	77.9 (15.5)	78.8 (14.4)	3.18	.04	.010	T1>T2*		
Vitality	+	71.5 (15.7)	62.8 (21.5)	68.3 (19.0)	37.5	.00	.108	T1>T2*; T1>T3*; T2<T3*		
Pain	+	87.2 (14.5)	86.7 (16.0)	87.5 (16.2)	.367	.69				
General health perception	+	67.1 (12.4)	65.9 (12.0)	66.8 (12.7)	1.88	.15				

Note. '+' = higher scores represent better health. '-' = higher scores represent poorer health. Excluded from PTSD analysis: 7 (no pre-assignment or post-assignment LEC experiences). η^2 = eta squared effect size. Statistically significant post-hoc pairwise findings were determined using $p < .05$. *is post-hoc Bonferroni corrected significance. Adjusted odds ratio (AOR) were calculated for above clinical cut-off (screener) scores of anxiety, depression, PTSD and burnout at T2. None of the risk factors were significant at T3.

158) at T2 and 34 % (n = 118) at T3. Being female, a greater number of years of national staff experience, a greater number of experienced and witnessed assignment-related PTEs increased the risk for suspected T2 depression. None of the variables were risk factors at T3.

PTSD (PCL-5). Mean PSTD severity levels remained stable ($p = .13$) over time with scores M = 8.89 (T1), M = 7.84 (T2), and M = 8.19 (T3). A small portion (2.6 %, n = 15) of the participants reported clinical pre-assignment levels for PTSD. These rates remained stable at T2 (2.4 %, n

= 12), and increased at T3 (3.9 %, n = 14). None of the variables were risk factors for suspected PTSD at T2 and T3.

Burnout (MBI). Burnout scale 'feelings of emotional exhaustion' changed significantly ($p = .00$; small effect size: $\eta^2 = 0.017$) between T1 (M = 1.66) and T2 (M = 1.81). T3 scores (M = 1.78) also differed significantly from T1 scores. The burnout scale 'personal accomplishments' showed a significant change ($p = .00$; small effect size: $\eta^2 = 0.020$). Pairwise comparisons indicated that scores were at their lowest at T3 (M = 4.67) compared to T1 (M = 4.80) and T2 (M = 4.77). No health changes were found in the scale 'depersonalisation' ($p = .16$) with scores M = 1.15 (T1), M = 1.17 (T2), and M = 1.23 (T3). Based on all three scales, a small portion (4.1 %, n = 24) of the participants reported pre-assignment levels for suspected burnout. These rates increased slightly to 5.5 % (n = 27) at T2 and 6.0 % (n = 22) at T3. Older participants and participants with greater assignment length were less likely at risk for suspected T2 burnout. None of the variables were risk factors at T3.

Quality of life (RAND-36). Levels of vitality changed significantly ($p = .00$; large effect size: $\eta^2 = 0.108$). Post-hoc pairwise comparisons showed in a significant drop in vitality levels between T1 (M = 71.5) and T2 (M = 62.8), and a significant increase in vitality levels at T3 (M = 68.3). The vitality levels did not return to baseline demonstrated by a significant difference between T1 and T3 scores ($p = .00$). Significant changes in the levels of social functioning were found ($p = .014$; small effect size: $\eta^2 = 0.014$) between T1 (M = 86.9) and T2 (M = 83.6). Similarly, emotional wellbeing also demonstrated significant changes ($p = .04$; $\eta^2 = 0.010$) between T1 (M = 79.7) and T2 (77.9). The remaining RAND-36 health indicators (physical functioning, physical or emotional role limitations, pain and general health) demonstrated stable health levels over time ($p > .05$).

Clinical interview (T2). A structured clinical interview (N = 320) assessed the presence of psychological disorders at T2 (Table 4). Approximately one in five participants were indicative of current psychological disorders (19 %, n = 60). Alcohol use disorder (AUD) (12 %, n = 38) was reported most often. In most AUD cases the disorder was considered mild (n = 33). The remaining cases showed moderate (n = 4) or severe levels (n = 1). None of the demographic and assignment-related variables were risk factors for a likely clinical disorder based on the interview.

Inflation of self-report scores (T2). We examined differences in the proportions of participants who reported above clinical threshold scores on the questionnaires and the clinical interview. Compared to the gold standard clinical the questionnaires overestimated three times (20 vs 6.6 %) the presence of a potential anxiety disorder, eight times (2.4 vs 0.3 %) the presence of PTSD, and twenty-five times (32 vs 1.3 %) the presence of a major depression.

Table 4
Prevalence of current post-assignment (T2) disorders according to the M.I.N.I. clinical interview stratified for gender.

	Total	Females	Males
	N (%)	*N* (%)	*N* (%)
Separate disorders			
Alcohol use	38 (11.9)	23 (12.0)	13 (10.8)
Substance use	9 (2.8)	2 (1.0)	7 (5.8)
Major depressive disorder	4 (1.3)	1 (0.5)	3 (2.5)
Obsessive compulsive disorder	7 (2.2)	4 (2.1)	3 (2.5)
Panic	3 (0.9)	2 (1.0)	1 (0.8)
Agoraphobia	3 (0.9)	3 (1.6)	0 (0)
Generalised anxiety	3 (0.9)	3 (1.6)	0 (0)
PTSD	1 (0.3)	1 (0.5)	0 (0)
Social anxiety	0 (0)	0 (0)	0 (0)
Anorexia nervosa	0 (0)	0 (0)	0 (0)
Bulimia nervosa	0 (0)	0 (0)	0 (0)
Binge eating	0 (0)	0 (0)	0 (0)
Pooled disorders			
Any anxiety disorders	21 (6.6)	16 (8.3)	5 (4.2)
Any current disorders	60 (18.8)	33 (17.2)	25 (20.8)

Note. Early remission not taken into account for alcohol. PTSD = 'post-traumatic stress disorder'. Total sample n = 320 (Female n = 192. Male n = 120. N = 8 gender specification missing).

Healthcare utilisation (T3). Three hundred and forty-six (346) participants completed the healthcare utilisation follow-up measure of whom 45 (13 %) utilised assignment-related physical health services and 24 (6.9 %) assignment-related mental health services. Another 16 (4.6 %) utilised physical and 8 (2.3 %) mental health services for issues that might or might not be assignment-related, whilst 72 (21 %) utilised physical and 8 (2.3 %) mental health services for issues unrelated to the assignment.

Two hundred and eighteen (n = 218) participants completed both the follow-up measure and post-assignment M.I.N.I. clinical interview, of whom 38 participants had an indication of a current DSM-5 disorder. Ten participants (26 %) with such a reported disorder utilised mental health services. The remaining 28 participants (74 %) with a current mental health disorder did not utilise mental health services at follow-up.

There were substantial differences between the self-report measures and clinical interview regarding the suspected presence of mental health disorders. The self-report rates for suspected clinical pathology were 32.4 % (n = 158) for depression, 20 % for anxiety-based pathologies (n = 97), and 2.4 % (n = 12) for PTSD. In contrast, the clinical interview rates were 1.3 % (n = 4) for depression, 6.6 % (n = 21) for anxiety-based pathologies, and 0.3 % (n = 1) for PTSD.

3.4. Sample representativeness

The representativeness of the sample was determined by comparing demographic information about the current sample with data from the total MSF population, performing decliner analyses, non-responder analyses and sensitivity analyses. For a detailed description see also Appendix 4.

Population comparisons. Five hundred and ninety-seven (597) iHAWs were not informed about the present study because they did not visit the MSF Amsterdam office for a briefing. Compared to them, our study participants were significantly ($p < .05$) more often female, younger, more often in supervisory/specialist and coordination assignment positions, and less likely assigned as activity managers/clinical medical specialists. All effect size differences were small.

Decliner analyses. Study participants were significantly more often female but did not differ in age and roles in the field.

Baseline subgroup differences. Females reported higher levels of anxiety, depression and emotional exhaustion. They also reported lower levels of mental wellbeing and vitality, but better physical functioning. The effect size differences between males and females were small. Participants with national staff experiences reported higher levels of PTSD, lower levels of physical function and felt less emotionally exhausted. The effect size differences were small, with the exception of a medium effect size on physical functioning. There were no baseline differences regarding prior assignment experiences or education.

Non-responder analyses. Compared to baseline, post-assignment non-responders were more often male and reported lower on the burnout 'depersonalisation' scale. Compared to baseline, follow-up non-responders were more often assignment first-timers. There were no baseline differences on the health-indicators between responders and non-responders.

Sensitivity analyses. Missing data were likely missing completely at random (Little's MCAR test ($\chi^2 = 206.5$ (194), $p = .26$). After reanalysis of the RM ANOVA health change scores using multiple imputations for any missing health indicator data results were comparable to the reported RM ANOVAs. It is unlikely that participant dropout influenced the present results.

4. Discussion

Humanitarian aid workers are considered a scientifically overlooked population with (post-traumatic) mental health problems reaching epidemic proportions (Macpherson and Burkle, 2021). Our current findings provide an objective assessment of the extent of the health problems in international humanitarian aid workers – a particular group of aid workers.'The current research confirms that humanitarian assignments are highly stressful. Three-quarters of the iHAWs reported assignment-related exposure to PTEs in addition to a number of organisational and environmental stressors. Males and females reported a similar number of field-related (environmental, cultural and organisational) sources of significant stress, though there are qualitative differences in the sources of stress. Women more often reported sexual assault and other unwanted experiences and witnessing sexual harassment and violence towards colleagues, emphasizing an important gender risk distinction within the humanitarian aid worker community. Women also more often reported experiencing interactional sources of stress (lack of technical support, management issues, team composition, and poor communication within team and project). This may be due to different gender styles in communicating and conflict management. Males tend to be more competing (dominating) with less concern for others, whereas females tend to focus on noncompeting (integrating, obliging, avoiding and compromising) strategies to resolve work conflicts (Rahim and Katz, 2019). It might also explain why women more often reported feelings of powerlessness and hopelessness. Alternatively, women were more often exposed to sexual threats and violence in workplace interactions, which may then lead them to report more interactional stress rather than differences in conflict management style. They may just feel less safe at work, increasing feelings of powerlessness and hopelessness.

Despite the stressful conditions most participants remained healthy. Health can be defined as 'the ability to adapt and to self-manage in the face of social, physical, and emotional challenges' (Huber et al., 2011). Five present findings support our assumptions that iHAWs, on short-term emergency aid assignments, remain predominantly healthy. We found (1) few overall negative health changes on a broad set of health indicators, (2) some health improvements, (3) lower prevalence rates compared to general population norm scores for psychological disorders, (4) post-assignment significant risk factors were no longer significant at follow-up, and (5) as expected, the clinical interview outcomes compared to the self-reporting questionnaires showed substantial inflation of the latter. The outcomes are discussed below.

First, on a broad set of health indicators, only a few negative pre-to post-assignment health changes were detected based on sample average scores with almost exclusively small effect sizes. Emotional wellbeing, social functioning, emotional exhaustion and a loss of vitality deteriorated at post-assignment. Indicative for the recuperation of the iHAWs is the improvement of all negative health changes in the months following their return.

Second, iHAWs reported pre-to post-assignment health improvements. In particular, their anxiety levels were significantly lower. This may be a product of timing of the measurement, a general satisfaction from the assignment, or both. With regard to timing, the current study pre-assignment levels of anxiety were elevated compared to previous longitudinal findings (Lopes Cardozo et al., 2012). Cardozo and colleagues collected pre-assignment data between days and weeks before assignment departure (Lopes Cardozo, personal communication). Our participants were assessed between leaving home and going straight to the assignment. The uncertainties about future threats, leaving their significant others, about their assignment briefings may result in anxiety (Grupe and Nitschke, 2013). At post-assignment, the rewarding nature of humanitarian work (Kim et al., 2017), and prospect of reuniting with family and friends could also explain the decrease of anxiety.

Third, the post-assignment clinical interview iHAWs prevalence rates compared to general population were lower for anxiety (6.6 % versus 6.9 %) and major depressive disorder (1.3 % versus 5.4 %) (Steel et al., 2014). IHAWs also reported similar or higher levels of current health-related quality of life compared to general reference populations (Jenkinson et al., 1999; Roser et al., 2019). Better health among iHAWs may be due to self-selection, and/or human resources department strong selection criteria on iHAWs.

Fourth, although risk factors, such as gender and previous national staff experience, were identified upon assignment return, none of these predicted ill-health at two-months follow-up. These findings contrast with findings of risk factors negatively affecting the health of iHAWs (Brooks et al., 2015; Gritti, 2015). The effect size of predictive risk factors was small at post-assignment. The lack of significance at follow-up may be attributable to a lack of power to detect small magnitude effects due to T3 study attrition. Both findings (small significant T2 effect size and lack of T3 significance) indicate that these risk factors play a negligible role regarding negative health changes. Altogether, the above findings demonstrate the overall capacity of iHAWs, short-term emergency aid assignments, to stay healthy and manage their highly stressful assignment environment.

Our findings confirmed our assumption that self-report questionnaires tend to substantially overestimate mental health pathologies. Using existing general or professional population cut-off scores, our participants were three times more likely to suffer from an anxiety disorder, eight times more likely to qualify for a PTSD diagnosis, and 25 times more likely to experience a depressive disorder, compared to their clinical interview outcomes. However, screeners reporting above clinical threshold levels may still be indicative of mental health difficulties as discussed in the next paragraph.

There are also some negative health outcomes that require discussion: (1) PTSD, anxiety and burnout – showed high RM ANOVA standard deviations, (2) the negative self-reported health changes for social functioning, emotional wellbeing, emotional exhaustion and vitality require attention, (3) higher than reference norm levels for alcohol use disorder and (4) post-assignment health services indicate negatively changing health conditions among the iHAWs. We discuss each of these results in the following paragraphs.

RM ANOVA mean trajectories with high standard deviations suggest the presence of unidentified iHAW sub-populations with different health trajectories. Findings in military populations describe five different pre-assignment, post-assignment, and follow-up deployment trajectories (Van der Wal et al., 2019). It is possible that similar trajectories are also present among iHAWs.

The negative overall health changes in terms of emotional exhaustion and vitality should not be overlooked. These related concepts are considered important features in common mental health disorders (Tuithof et al., 2017). Emotional exhaustion is often considered the core dimension of burnout (Te Brake et al., 2008). A decrease in these aspects may signify the psychological 'wear' or strain of some key elements of working as iHAWS: a high workload, hostile environments, witnessing severe human suffering, unmet medical needs, and impossible moral dilemmas. Fortunately, the magnitude of the effect change for emotional exhaustion is small. It is larger for vitality, but vitality also demonstrates a strong recovery process between post-assignment and follow-up.

Some individuals experienced stressful or shocking events. This may explain the small increase in the portion of participants scoring above the clinical threshold on the screeners' depression, PTSD symptoms and burnout between pre-assignment and follow-up (approximately 2.0 %).

The iHAWs' 12-month prevalence rate of alcohol use disorders (AUD) was substantially higher (12 %) compared to the global 12-month alcohol and other substance use disorder rate (4 %) (Steel et al., 2014). Most iHAWs reported mild AUD (88 %) which is considered a form of self-medication to cope with different stressors and experiences {Citation}(Biron et al., 2011). Alcohol use is unhealthy and is as such an ineffective coping mechanism (Griswold et al., 2018). AUD is associated with increased mortality rates and burden of illness (Carvalho et al., 2019). Our results are supported by previous findings of increased levels of alcohol use among iHAWs (Cardozo et al., 2005; Dubravka et al.,

2016). Higher alcohol consumption may be part of an international humanitarian lifestyle. Alcohol consumption is a colossal global health issue that knows no safe level of consumption (Burton and Sheron, 2018).

Lastly, health service use acts as an indicator of the population's health (Jordans et al., 2019). Almost one in ten of the iHAWs utilised (possible) assignment-related mental healthcare services at follow-up. With regard to mental health service utilisation, there are no comparison data for iHAWs. These rates are comparable to military deployment-related populations (Hom et al., 2017; Sareen et al., 2007). Apparently, most iHAWs contacting mental health services sought support for psychological distress and/or to prevent more serious mental health problems, rather than seeking treatment for current psychological disorders. Those with current psychological disorders that did not seek any support may experience barriers to care (e.g. stigma, career concerns) that prevent them from utilising healthcare services (Hom et al., 2017).

4.1. Strengths and limitations

The present study has a number of strengths. The longitudinal design and large sample size addressed important methodological limitations of prior studies (Connorton et al., 2012). It retained a high number of participants between pre-assignment and post-assignment measurements. The inclusion of the clinical interview, rated by trained psychologists with high inter-rater reliability, made this the first research on iHAWs with diagnosed DSM-5 mental health disorders.

There are also a number of limitations. Small but significant differences compared to iHAWs not briefed were found on gender (more females), and profession (more often specialist/supervisory staff). These differences were significant but not practically relevant (Sullivan and Feinn, 2012). The list of assignment-related stressors is based on iHAWs' field experiences and staff health observations. However, the instrument is not validated. The follow-up measure was relatively soon after the post-assignment measure for operational reasons (staff returning to the field). Hence, the identification of health changes in the longer term remains limited. Anxiety, depression and PTSD symptoms may take time to develop. For example, a quarter of PTSD cases in general have a delayed-onset, preceded by subclinical PTSD symptoms (Smid et al., 2009). These findings may also occur within iHAWs but are beyond the scope of the current study. Some participants returned to the field before the follow-up could be administered, which reduced the follow-up retention rate. Note that the sensitivity analyses detected no bias in the results due to study attrition. Adapting to stressful and demanding circumstances may change how iHAWs perceive concepts such as anxiety and stress. Determining the potential impact of a response shift was beyond the scope of this article. Research into the effect of a response shift showed a small effect in the general population (Schwartz et al., 2006).

Our research focused on a population of a large, international humanitarian organization with an emphasis on relatively short-term emergency assignments (Redfield, 2012). International humanitarian aid workers tend to be highly educated and well-trained. It is unclear to what degree our results are generalizable to other aid workers, such as national staff, iHAWs of small aid organizations, consultants, and different types of humanitarian aid workers (emergency, development). Considering that most participants participated in relatively short-term emergency assignments, it is uncertain whether our findings generalize to long-term aid assignments (>1 year). Long-term missions were recently associated with better mental health outcomes and a higher level of wellbeing (Young and Pakenham, 2020). Our risk factor analyses showed no indications that certain characteristics such as education, gender etc. were predictive of ill-health but we cannot exclude cultural and organisational differences from impacting study outcomes. We did not study national staff. This group runs a greater risk to be killed, wounded or kidnapped (Stoddard et al., 2020), and reported greater psychological distress compared to iHAWs (Cardozo et al., 2005). Considering that national staff makes up for 90 % of the aid workers (Stoddard et al., 2009), further investigation into the distinction between international and national staff is warranted (Shevchenko and Fox, 2008).

4.2. Implications

IHAWs on short-term emergency aid assignments experience a range of stressors partly preventable via organisational changes (e.g., in project communication, workload and travel arrangements). Reducing these results stressors likely enhances the overall health of all staff (Michie and Williams, 2003). This may also help eliminate the need for alcohol as a coping mechanism (Burton and Sheron, 2018). Furthermore, creating a culture of normalcy and mutual (social) support for coping with PTEs, feelings of (moral) distress, professional inadequacy and loneliness is potentially an effective way to help iHAWs in a healthy, non-alcohol-related manner. Besides creating a supportive environment, alcohol control policies can also help reduce the overall amount of consumption (Carvalho et al., 2019).

Management and safety specialists need to be aware of specific gender related sources of potentially traumatic stress. The implementation of organisational policies that minimize the threat of sexual assault and harassment to iHAWs, female iHAWs in particular, is essential. It warrants explicit attention because of its potential long-term harm to health and associated stigma (Stoddard et al., 2019). Furthermore, awareness on gender differences in team and management-related sources of stress is warranted. Women more often experience significant team and management-related stress. Managers and supervisors can encourage aid workers to use effective context appropriate conflict-resolution strategies to decrease organisational stress (Rahim and Katz, 2019). Providing a socially supportive organisational environment helps to decrease a sense of helplessness and promote health and wellbeing in humanitarian aid workers (Aldamman et al., 2019).

Humanitarian organizations should rethink their screening procedures. Development of iHAW-specific norm populations and clinical cut-off scores based on optimal sensitivity and specificity values is essential. This will help to avoid inflating the presence of clinical health issues that mask the distinction between (transient) distress and pathology. Applying screeners with non-iHAW norm scores may have strengthened the perception among the general public that humanitarian assignments are damaging to one's health.

Screeners should focus on health (changes) and act as a dialogue tool for discussing health implications: how to improve one's health or remain healthy in future aid assignments. Vitality, emotional exhaustion and high alcohol use are particularly useful post-assignment iHAW-specific health screeners. PTSD is another important health indicator to monitor because of the repeated exposure to assignment-related PTEs. Clinical interviews of DSM-5 disorders after positive self-report screening provide the next step in the process of early identification of healthcare needs. Although, our clinical interviews showed comparatively low levels of clinical pathology, elevated distress levels must always be taken serious. Distress still indicates potential substantial suffering. Organizations need to acknowledge this suffering and provide (preventative) distress-related health services.

Third, a strategy of watchful waiting is recommended to assess post-assignment health. Increases in post-assignment or follow-up depression, burnout and PTSD prevalence rates, the nature of delayed-onset PTSD, frequency of assignment-related health services utilisation, and the absence of use of health services among those with probable disorders, imply that assignment-related health issues can manifest beyond the actual return date and may not recover spontaneously. A strategy of watchful waiting helps to detect but also helps to avoid overtreatment of staff at risk. Overtreatment is unnecessary, expensive, medicalises health and inhibits natural recovery processes (The PLOS Medicine Editors, 2013). Our study already demonstrated that health improved on

several health indicators two months after assignment return. We suggest a watchful waiting period of 2 months at least.

4.3. Epilogue: paradigm shift?

To improve the overall health of iHAWs on short-term emergency aid assignments, applying a pathogenic perspective is not productive, except for a disorder-affected minority. Our findings put forward a paradigm shift in the analysis of iHAW health. Most research focuses on the question: 'How ill or pathogenic are iHAWs and what makes iHAWs sick?' Rephrasing the question to 'What keeps and makes iHAWs healthy?' will give a different, more comprehensive and more useful perspective on the improvement of iHAWs' health (Antonovsky, 1979). It will open new avenues of scientific interest to explore how 'doing good' by delivering aid acts as a meaning-making mechanism to cope with stress.

Declaration of competing interest

Kaz de Jong is employed currently by Médecins Sans Frontières. Saara Martinmäki has no conflicts of interest. Hans te Brake has no conflicts of interest. Joris Haagen has no conflicts of interest. Rolf Kleber has no conflicts of interest.

Acknowledgments

We acknowledge and thank with gratitude: Dr. D. Sheehan for providing the M.I.N.I. free of charge for the duration of the study; Dr. I. van Vliet for the skilled training and supervision necessary for the correct use of the M.I.N.I.; Prof. Dr. I. Komproe of Utrecht University for the extensive theoretical, methodological, statistical and editorial advice provided; all participating Masters students – Tim Breda, Mina Belojevic, Livia Rebstock, Helena Carter, Louise O'Meara, Elpida Karazisi, Samira Lintzen and Noa van Zwieten – for their collaboration in the data collection; and Natasha Lewer for her final editing check on the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2021.114268.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Médecins Sans Frontières.

Role of funding resource

MSF's Management Team funded the study into the health of its international staff. It was not involved in: study design; collection, analysis, and interpretation of data; writing of the articles; and decisions to submit papers for publication. ARQ Research and Policy Institute Impact independently carried out the collection of the data. The interpretation, (anonymous) analyses and writing of the article were done by all members of the research team. The MSF principal investigator was not involved in the recruitment, retention or follow-up procedures of the study participants, was unaware of the individuals who participated in the study and only had access to anonymized study data. Two scientific experts (professors) from Utrecht University supervised the quality and independence of the scientific process.

References

Aldamman, K., Tamrakar, T., Dinesen, C., Wiedemann, N., Murphy, J., Hansen, M., Elsiddig Badr, E., Reid, T., Vallières, F., 2019. Caring for the mental health of humanitarian volunteers in traumatic contexts: the importance of organisational support. Eur. J. Psychotraumatol. 10 (1), 1694811. https://doi.org/10.1080/20008198.2019.1694811.

Antonovsky, A., 1979. Health, Stress, and Coping: New Perspectives on Mental and Physical Well-Being. Jossey-Bass.

Biron, M., Bamberger, P.A., Noyman, T., 2011. Work-related risk factors and employee substance use: insights from a sample of Israeli blue-collar workers. J. Occup. Health Psychol. 16 (2), 247–263. https://doi.org/10.1037/a0022708.

Blevins, C.A., Weathers, F.W., Davis, M.T., Witte, T.K., Domino, J.L., 2015. The posttraumatic stress disorder checklist for DSM-5 (PCL-5): development and initial psychometric evaluation. J. Trauma Stress 28, 489–498. https://doi.org/10.1002/jts.22059.

Brooks, S.K., Dunn, R., Sage, C.A.M., Amlôt, R., Greenberg, N., Rubin, G.J., 2015. Risk and resilience factors affecting the psychological wellbeing of individuals deployed in humanitarian relief roles after a disaster. J. Ment. Health 24 (6), 385–413. https://doi.org/10.3109/09638237.2015.1057334.

Burton, R., Sheron, N., 2018. No level of alcohol consumption improves health. Lancet 392 (10152), 987–988. https://doi.org/10.1016/S0140-6736(18)31571-X.

Cardozo, B.L., Holtz, T.H., Kaiser, R., Gotway, C.A., Ghitis, F., Toomey, E., Salama, P., 2005. The mental health of expatriate and Kosovar Albanian humanitarian aid workers. Disasters 29 (2), 152–170. https://doi.org/10.1111/j.0361-3666.2005.00278.x.

Carvalho, A.F., Heilig, M., Perez, A., Probst, C., Rehm, J., 2019. Alcohol use disorders. Lancet 394 (10200), 781–792. https://doi.org/10.1016/S0140-6736(19)31775-1.

Charlson, F., van Ommeren, M., Flaxman, A., Cornett, J., Whiteford, H., Saxena, S., 2019. New WHO prevalence estimates of mental disorders in conflict settings: a systematic review and meta-analysis. Lancet 394 (10194), 240–248. https://doi.org/10.1016/S0140-6736(19)30934-1.

Connorton, E., Perry, M.J., Hemenway, D., Miller, M., 2012. Humanitarian relief workers and trauma-related mental illness. Epidemiol. Rev. 34 (1), 145–155. https://doi.org/10.1093/epirev/mxr026.

Derogatis, L.R., Lipman, R.S., Rickels, K., Uhlenhuth, E.H., Covi, L., 1974. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. Behav. Sci. 19 (1), 1–15. https://doi.org/10.1002/bs.3830190102.

Dubravka, S., Thomas, R., Jachens, L., Mihalca, L., 2016. Staff Wellbeing and Mental Health within UNHCR. United Nations High Commissioner for Refugees, Geneva, Switzerland.

Egeland, J., Harmer, A., Stoddard, A., 2011. To Stay and Deliver: Good Practice for Humanitarians in Complex Security Environments. OCHA, United Nations.

Eriksson, C.B., Bjorck, J.P., Larson, L.C., Walling, S.M., Trice, G.A., Fawcett, J., Abernethy, A.D., Foy, D.W., 2009. Social support, organisational support, and religious support in relation to burnout in expatriate humanitarian aid workers. *Mental Health*. Religion & Culture 12 (7), 671–686. https://doi.org/10.1080/13674670903029146.

Galea, S., Maxwell, A.R., Norris, F., 2008. Sampling and design challenges in studying the mental health consequences of disasters. Int. J. Methods Psychiatr. Res. 17 (Suppl. 2), S21–S28. https://doi.org/10.1002/mpr.267.

Gray, M.J., Litz, B.T., Hsu, J.L., Lombardo, T.W., 2004. Psychometric properties of the life events checklist. Assessment 11 (4), 330–341. https://doi.org/10.1177/1073191104269954.

Griswold, M.G., Fullman, N., Hawley, C., Arian, N., Zimsen, S.R.M., Tymeson, H.D., Venkateswaran, V., Tapp, A.D., Forouzanfar, M.H., Salama, J.S., Abate, K.H., Abate, D., Abay, S.M., Abbafati, C., Abdulkader, R.S., Abebe, Z., Aboyans, V., Abrar, M.M., Acharya, P., Gakidou, E., 2018. Alcohol use and burden for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 392 (10152), 1015–1035. https://doi.org/10.1016/S0140-6736(18)31310-2.

Gritti, A., 2015. Building aid workers' resilience: why a gendered approach is needed. Gend. Dev. 23, 449–462. https://doi.org/10.1080/13552074.2015.1095542.

Grupe, D.W., Nitschke, J.B., 2013. Uncertainty and anticipation in anxiety. Nat. Rev. Neurosci. 14 (7), 488–501. https://doi.org/10.1038/nrn3524.

Hays, R.D., Morales, L.S., 2001. The RAND-36 measure of health-related quality of life. Ann. Med. 33 (5), 350–357. https://doi.org/10.3109/07853890109002089.

Hergueta, T., Weiller, E., 2013. Evaluating depressive symptoms in hypomanic and manic episodes using a structured diagnostic tool: validation of a new Mini International Neuropsychiatric Interview (M.I.N.I.) module for the DSM-5 'With Mixed Features' specifier. Int. J. Behav. Dev. 1 (1), 21. https://doi.org/10.1186/2194-7511-1-21.

Holtz, T.H., Salama, P., Cardozo, B.L., Gotway, C.A., 2002. Mental health status of human rights workers, Kosovo, June 2000. J. Trauma Stress 15, 389–395. https://doi.org/10.1023/A:1020133308188.

Hom, M.A., Stanley, I.H., Schneider, M.E., Joiner, T.E., 2017. A systematic review of help-seeking and mental health service utilization among military service members. Clin. Psychol. Rev. 53, 59–78. https://doi.org/10.1016/j.cpr.2017.01.008.

Huber, M., Knottnerus, J.A., Green, L., van der Horst, H., Jadad, A.R., Kromhout, D., Leonard, B., Lorig, K., Loureiro, M.I., van der Meer, J.W.M., Schnabel, P., Smith, R., van Weel, C., Smid, H., 2011. How should we define health? Br. Med. J. 343, d4163. https://doi.org/10.1136/bmj.d4163.

Jachens, L., Houdmont, J., Thomas, R., 2019. Effort-reward imbalance and burnout among humanitarian aid workers. Disasters 43 (1), 67–87. https://doi.org/10.1111/disa.12288.

Jenkinson, C., Stewart-Brown, S., Petersen, S., Paice, C., 1999. Assessment of the SF-36 version 2 in the United Kingdom. J. Epidemiol. Community Health 53 (1), 46–50.
Jordans, M.J., Luitel, N.P., Kohrt, B.A., Rathod, S.D., Garman, E.C., De Silva, M., Komproe, I.H., Patel, V., Lund, C., 2019. Community-, facility-, and individual-level outcomes of a district mental healthcare plan in a low-resource setting in Nepal: a population-based evaluation. PLoS Med. 16 (2) https://doi.org/10.1371/journal.pmed.1002748.
Kim, Y., Patel, N., Diehl, G., Richard, P., 2017. The association between service members' participation in humanitarian aid and disaster relief and mental health symptoms and treatments. Mil. Med. 182 (9), e1849–e1855. https://doi.org/10.7205/MILMED-D-17-00023.
Koen, M.P., van Eeden, C., Wissing, M.P., 2011. The prevalence of resilience in a group of professional nurses. Health SA Gesondheid 16 (1), 1–11. https://doi.org/10.4102/hsag.v16i1.576.
Korff, V.P., Balbo, N., Mills, M., Heyse, L., Wittek, R., 2015. The impact of humanitarian context conditions and individual characteristics on aid worker retention. Disasters 39 (3), 522–545. https://doi.org/10.1111/disa.12119.
Levine, T.R., Hullett, C.R., 2002. Eta squared, partial eta squared, and misreporting of effect size in communication research. Hum. Commun. Res. 28 (4), 612–625. https://doi.org/10.1111/j.1468-2958.2002.tb00828.x.
Limb, C.J., 2011. The need for evidence in an anecdotal world. Trends Amplif. 15 (1), 3–4. https://doi.org/10.1177/1084713811425751.
Lopes Cardozo, B., Gotway Crawford, C., Eriksson, C., Zhu, J., Sabin, M., Ager, A., Foy, D., Snider, L., Scholte, W., Kaiser, R., Olff, M., Rijnen, B., Simon, W., 2012. Psychological distress, depression, anxiety, and burnout among international humanitarian aid workers: a longitudinal study. PloS One 7 (9), e44948. https://doi.org/10.1371/journal.pone.0044948.
Maslach, C., Jackson, S.E., Leiter, M.P., 1981. Maslach Burnout Inventory: MBI. Consulting Psychologists Press, New York.
Macpherson, R., Burkle, F., 2021. Humanitarian aid workers: the forgotten first responders. Prehospital Disaster Med. 36 (1), 111–114. https://doi.org/10.1017/S1049023X20001326.
Michie, S., Williams, S., 2003. Reducing work related psychological ill health and sickness absence: a systematic literature review. Occup. Environ. Med. 60 (1), 3–9. https://doi.org/10.1136/oem.60.1.3.
Mollica, R.F., Wyshak, G., de Marneffe, D., Khuon, F., Lavelle, J., 1987. Indochinese versions of the Hopkins Symptom Checklist-25: a screening instrument for the psychiatric care of refugees. Am. J. Psychiatr. 144 (4), 497–500. https://doi.org/10.1176/ajp.144.4.497.
Overseas Development Assistance, 2019. Global Humanitarian Assistance Report 2018. Development Initiatives LTD, Bristol, United Kingdom. http://devinit.org/post/global-humanitarian-assistance-report-2018/#.
Parloff, M.B., Kelman, H.C., Frank, J.D., 1954. Comfort, effectiveness, and self-awareness as criteria of improvement in psychotherapy. Am. J. Psychiatr. 111 (5), 343–352. https://doi.org/10.1176/ajp.111.5.343.
Rahim, M.A., Katz, J.P., 2019. Forty years of conflict: the effects of gender and generation on conflict-management strategies. Int. J. Conflict Manag. 31 (1), 1–16. https://doi.org/10.1108/IJCMA-03-2019-0045.
Redfield, P., 2012. The unbearable lightness of ex-pats: double Binds of humanitarian mobility. Cult. Anthropol. 27 (2), 358–382. https://doi.org/10.1111/j.1548-1360.2012.01147.x.
Roser, K., Mader, L., Baenziger, J., Sommer, G., Kuehni, C.E., Michel, G., 2019. Health-related quality of life in Switzerland: normative data for the SF-36v2 questionnaire. Qual. Life Res. 28 (7), 1963–1977. https://doi.org/10.1007/s11136-019-02161-5.
Sareen, J., Cox, B.J., Afifi, T.O., Stein, M.B., Belik, S.-L., Meadows, G., Asmundson, G.J. G., 2007. Combat and peacekeeping operations in relation to prevalence of mental disorders and perceived need for mental health care: findings from a large representative sample of military personnel. Arch. Gen. Psychiatr. 64 (7), 843–852. https://doi.org/10.1001/archpsyc.64.7.843.
Schaufeli, W.B., Dierendonck, D.V., 1993. The construct validity of two burnout measures. J. Organ. Behav. 14 (7), 631–647. https://doi.org/10.1002/job.4030140703.
Schutte, N., Toppinen, S., Kalimo, R., Schaufeli, W., 2000. The factorial validity of the Maslach Burnout Inventory-General Survey (MBI-GS) across occupational groups and nations. J. Occup. Organ. Psychol. 73 (1), 53–66. https://doi.org/10.1348/096317900166877.
Schwartz, C.E., Bode, R., Repucci, N., Becker, J., Sprangers, M.A.G., Fayers, P.M., 2006. The clinical significance of adaptation to changing health: a meta-analysis of response shift. Qual. Life Res.: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation 15 (9), 1533–1550. https://doi.org/10.1007/s11136-006-0025-9.
Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J. Clin. Psychiatr. 59 (Suppl. 20), 22–33 quiz 34-57.
Shevchenko, O., Fox, R.C., 2008. "Nationals" and "expatriates": challenges of fulfilling "sans frontières"("without borders") ideals in international humanitarian action. Health and Human Rights 109–122. https://doi.org/10.2307/20460093.
Smid, G.E., Mooren, T.T.M., van der Mast, R.C., Gersons, B.P.R., Kleber, R.J., 2009. Delayed posttraumatic stress disorder: systematic review, meta-analysis, and meta-regression analysis of prospective studies. J. Clin. Psychiatr. 70 (11), 1572–1582. https://doi.org/10.4088/JCP.08r04484.
Steel, Z., Marnane, C., Iranpour, C., Chey, T., Jackson, J.W., Patel, V., Silove, D., 2014. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980–2013. Int. J. Epidemiol. 43 (2), 476–493. https://doi.org/10.1093/ije/dyu038.
Stoddard, A., Harmer, A., DiDomenico, V., 2009. Providing aid in insecure environments: 2009 Update. Trends in violence against aid workers and the operational response 34. https://www.eldis.org/document/A43064.
Stoddard, A., Harvey, P., Czwarno, M., Breckenridge, M., 2019. Aid Worker Security Report 2019. Speakable: Addressing Sexual Violence and Gender-Based Risk in Humanitarian Aid. Humanitarian Outcomes.
Stoddard, A., Harvey, P., Czwarno, M., Breckenridge, M.J., 2020. Aid Worker Security Report 2020: contending with threats to humanitarian health workers in the age of epidemics. Humanitarian Outcomes. https://www.humanitarianoutcomes.org/AWSR2020.
Sullivan, G.M., Feinn, R., 2012. Using effect size-or why the p value is not enough. Journal of Graduate Medical Education 4 (3), 279–282. https://doi.org/10.4300/JGME-D-12-00156.1.
Te Brake, H., Smits, N., Wicherts, J.M., Gorter, R.C., Hoogstraten, J., 2008. Burnout development among dentists: a longitudinal study. Eur. J. Oral Sci. 116 (6), 545–551. https://doi.org/10.1111/j.1600-0722.2008.00584.x.
The PLOS Medicine, 2013. The paradox of mental health: over-treatment and under-recognition. PLoS Med. 10 (5) https://doi.org/10.1371/journal.pmed.1001456.
Tinghög, P., Carstensen, J., 2010. Cross-cultural equivalence of HSCL-25 and WHO (ten) wellbeing index: findings from a population-based survey of immigrants and non-immigrants in Sweden. J. Community Mental Health 46, 65–76. https://doi.org/10.1007/s10597-009-9227-2.
Tuithof, M., ten Have, M., Beekman, A., van Dorsselaer, S., Kleinjan, M., Schaufeli, W., de Graaf, R., 2017. The interplay between emotional exhaustion, common mental disorders, functioning and health care use in the working population. J. Psychosom. Res. 100, 8–14. https://doi.org/10.1016/j.jpsychores.2017.06.018.
Van der Wal, S.J., Gorter, R., Reijnen, A., Geuze, E., Vermetten, E., 2019. Cohort profile: the prospective research in stress-related military operations (PRISMO) study in the Dutch Armed Forces. BMJ Open 9 (3). https://doi.org/10.1136/bmjopen-2018-026670.
Van der Zee, K.I., Sanderman, R., Heyink, J.W., de Haes, H., 1996. Psychometric qualities of the RAND 36-Item Health Survey 1.0: a multidimensional measure of general health status. Int. J. Behav. Med. 3 (2), 104. https://doi.org/10.1207/s15327558ijbm0302_2.
Vink, G., Lazendic, G., Van Buuren, S., 2015. Partitioned predictive mean matching as a multilevel imputation technique. Psychological Test and Assessment Modeling 57 (4), 577–594.
Weathers, F.W., Blake, D.D., Schnurr, P.P., Kaloupek, D.G., Marx, B.P., Keane, T.M., 2013. The Life Events Checklist for DSM-5 (LEC-5). Instrument available from the National Center for PTSD at. www.ptsd.va.gov.
Winokur, A., Winokur, D.F., Rickels, K., Cox, D.S., 1984. Symptoms of emotional distress in a family planning service: stability over a four-week period. Br. J. Psychiatr. 144, 395–399. https://doi.org/10.1192/bjp.144.4.395.
Wortmann, J.H., Jordan, A.H., Weathers, F.W., Resick, P.A., Dondanville, K.A., Hall-Clark, B., Foa, E.B., Young-McCaughan, S., Yarvis, J.S., Hembree, E.A., Mintz, J., Peterson, A.L., Litz, B.T., 2016. Psychometric analysis of the PTSD Checklist-5 (PCL-5) among treatment-seeking military service members. Psychol. Assess. 28 (11), 1392–1403. https://doi.org/10.1037/pas0000260.
Young, T., Pakenham, K., 2020. Risk and protective factors for aid worker mental health: the effects of job context, working conditions and demographics. Disasters. https://doi.org/10.1111/disa.12440.
Young, T., Pakenham, K., Norwood, M., 2018. Thematic analysis of aid workers' stressors and coping strategies: work, psychological, lifestyle and social dimensions. J. Inter. Humanitarian Action 3 (1), 19. https://doi.org/10.1186/s41018-018-0046-3.