EUROPEAN JOURNAL OF
PSYCHO-TRAUMATOLOGY
THE OFFICIAL JOURNAL OF THE EUROPEAN SOCIETY FOR TRAUMATIC STRESS STUDIES

Taylor & Francis
Taylor & Francis Group

BASIC RESEARCH ARTICLE

🔓 OPEN ACCESS | Check for updates

A longitudinal evaluation of active outreach after an aeroplane crash: screening for post-traumatic stress disorder and depression and assessment of self-reported treatment needs

Juul Gouweloos-Trines[a,b], Hans Te Brake[a], Marit Sijbrandij [c], Paul A. Boelen [b,d], Chris R. Brewin[e] and Rolf J. Kleber[b,d]

[a]Knowledge Center Impact, Arq Psychotrauma Expert Group, Diemen, The Netherlands; [b]Department of Clinical Psychology, Utrecht University, Utrecht, The Netherlands; [c]Department of Clinical, Neuro- and Developmental Psychology, Free University of Amsterdam, Amsterdam, The Netherlands; [d]Arq Psychotrauma Expert Group, Diemen, The Netherlands; [e]Department of Clinical, Educational and Health Psychology, University College London, London, UK

ABSTRACT

Background: In 2009, an aeroplane crashed near Amsterdam. To remedy unmet mental health needs, active outreach was used to identify victims at risk for post-traumatic stress disorder (PTSD) and depression.

Objective: The active outreach strategy was evaluated by examining the accuracy of screening methods in predicting PTSD and depression, self-reported treatment needs, and the extent to which perceived treatment needs predict trajectories of PTSD.

Method: In 112 adult survivors, semi-structured telephone interviews were held at 2 (T1, $n = 76$), 9 (T2, $n = 77$) and 44 months (T3, $n = 55$) after the crash. The Trauma Screening Questionnaire (TSQ) and the Patient Health Questionnaire-2 (PHQ-2) measured symptoms of PTSD and depression, respectively. At T3, a clinical interview assessed PTSD and depression diagnoses. Based on the TSQ scores at the three time-points, participants were grouped into five 'trajectories': resilient ($n = 38$), chronic ($n = 30$), recovery ($n = 9$), delayed onset ($n = 9$) and relapse ($n = 3$).

Results: The TSQ accurately predicted PTSD at T3 (sensitivity: .75–1.00; specificity: .79–.90). The PHQ-2 showed modest accuracy (sensitivity: .38–.89; specificity: .67–.90). Both measures provided low positive predictive values (TSQ: 0.57; PHQ-2: .50 at T3). A number of participants reported unmet treatment needs (T1: 32.9%; T2: 19.5%; T3: 10.9%). Reporting unmet needs at T2 was more often assigned to a chronic PTSD trajectory compared to reporting no needs ($p < .01$).

Conclusions: The prevalence of unmet needs at 44 months after the crash within a chronic PTSD trajectory indicated that active outreach may be warranted. Nevertheless, although the TSQ was accurate, many participants screening positive did not develop PTSD. This implies that, although active outreach may benefit those with unmet needs, it also has its costs in terms of possible unnecessary clinical assessments.

Una evaluación longitudinal del Contacto Activo luego de un Accidente Aéreo: Tamizaje para TEPT y Depresión y Evaluación de las Necesidades de Tratamiento Auto-reportadas

Antecedentes: En 2009 ocurrió un accidente aéreo cerca de Amsterdam. Para subsanar las necesidades de salud mental insatisfechas, se usó contacto activo para identificar víctimas en riesgo de trastorno de estrés postraumático (TEPT) y depresión.

Objetivo: Evaluar esta estrategia examinando: 1) la precisión de los métodos de tamizaje para predecir TEPT Y depresión, 2) necesidades de tratamiento auto-reportadas, y 3) el grado en que las necesidades de tratamiento percibidas predicen las trayectorias de TEPT.

Métodos: En 112 adultos sobrevivientes, se realizaron entrevistas telefónicas semi-estructuradas a los 2 (T1, n=76), 9 (T2, n=77) y 44 meses (T3, n=55) posteriores al accidente. El Cuestionario de Tamizaje de Trauma (TSQ por sus siglas en inglés) y el Cuestionario de Salud de los Pacientes-2 (PHQ-2 por sus siglas en inglés) midieron síntomas de TEPT y depresión, respectivamente. En T3, se evaluó con una entrevista clínica los diagnósticos de TEPT y depresión. En base a los puntajes de TSQ en los 3 puntos de tiempo, los participantes fueron agrupados en 5 'trayectorias': resiliente (n=38), crónico (n=30), recuperación (n=9), inicio demorado (n=9) y recaída (n=3).

Resultados: la TSQ predijo TEPT en forma precisa al T3 (sensibilidad: .75 a 1.00, especificidad: .79 a .90). El PHQ-2 mostró modesta precisión (sensibilidad: .38 a .89, especificidad: .67 a .90). Ambas medidas entregaron bajos valores predictivos positivos (TSQ: .57, PHQ-2: .50 en T3). Un número de participantes reportaron necesidades de tratamiento insatisfechas (T1: 32.9%, T2: 19.5%, T3: 10.9%). El reportar necesidades insatisfechas en T2 fue más

HIGHLIGHTS
• Reporting unmet needs 2 months after the crash ($n = 25$32.9%) was related to a chronic PTSD trajectory. Furthermore 11% of survivors reported unmet treatment needs 44 months after the crash.
• Using a brief screener (TSQ) participants at risk for PTSD are accurately detected.
• A significant proportion ($n = 21$27.6%) of participants reported no treatment needs most of whom were classified into the resilient trajectory. This suggests that after the crash participants themselves can often provide an accurate estimation of their need for professional care.

CONTACT Juul Gouweloos-Trines ✉ j.gouweloos@impact.arq.org 🌐 Knowledge Center Impact, Arq Psychotrauma Expert Group, Diemen, The Netherlands

frecuentemente asignado a una trayectoria crónica de TEPT comparado a no reportar necesidades (p<.01).

Conclusiones: La prevalencia de necesidades insatisfechas 44 meses después del accidente dentro de una trayectoria crónica de TEPT indicó que el contacto activo puede estar justificado. Sin embargo, aunque la TSQ fue precisa, muchos participantes detectados como positivos no desarrollaron TEPT. Esto implica que, aunque el contacto activo puede beneficiar a aquellos con necesidades insatisfechas, tiene también sus costos en términos de posibles evaluaciones clínicas innecesarias.

飞机坠毁后的积极联系策略的纵向评估：PTSD和抑郁症的筛查以及自述治疗需求的评估

背景：2009年，一架飞机在阿姆斯特丹坠毁。通过积极联系（Active Outreach）策略识别有创伤后应激障碍（PTSD）和抑郁症风险的受害者，以解决其未满足的心理健康护理需求。

目的：通过检查以下方面对该策略进行：1）PTSD和抑郁症的筛查方法的预测准确性，2）自我报告的治疗需求，以及3）感知到的治疗需求对PTSD轨迹的预测程度。

方法：在112名成年幸存者中，在坠机后第2（T1，n = 76），9（T2，n = 77）和44个月（T3，n = 55）进行半结构化电话访谈。创伤筛查问卷（TSQ）和患者健康问卷-2（PHQ-2）分别测量了创伤后应激障碍和抑郁症的症状。在T3，临床访谈评估PTSD和抑郁症诊断。基于3个时间点的TSQ评分，参与者被分为5个症状'轨迹'组：韧性（n = 38），慢性（n = 30），恢复（n = 9），延迟发作（n = 9）和复发（n = 3）。

结果：TSQ准确预测T3时的PTSD（灵敏度：.75至1.00，特异性：.79至.90）。 PHQ-2显示适度的准确度（灵敏度：.38至.89，特异性：.67至.90）。 两种测量均提供低阳性预测值（TSQ：.57，PHQ-2：在T3时为.50）。 一些参与者报告了未满足的治疗需求（T1：32.9%，T2：19.5%，T3：10.9%）。 在T2报告有未满足的需求的参与者，与报告没有需求的相比，更常被分到慢性PTSD的轨迹组（p <.01）。

结论：在事故发生后44个月在慢性创伤后应激障碍轨迹组中普遍存在未满足需求，这表明积极联系策略可能是必要的。但尽管TSQ是准确的，许多参与筛查阳性的患者并未出现创伤后应激障碍。这意味着，虽然积极联系可能会使那些有未满足需求的人受益，但在可能的不需要的临床评估方面付出代价。

1. Introduction

In 2009, an aeroplane of Turkish Airlines crashed near Amsterdam, the Netherlands. Of the 135 passengers, 126 people survived the crash, most of whom had Dutch or Turkish nationality. The regional Community Health Service (CHS) launched an outreach strategy to identify survivors at risk of mental health problems, such as post-traumatic stress disorder (PTSD) and depression, and to support them in finding appropriate psychological care.

Several prospective studies found four distinct trajectories of mental health symptoms after disasters and large-scale incidents: resilient (no or only mild symptoms), recovery (initial high symptoms that decrease over time), chronic (continuously high symptoms) and delayed onset (high symptoms that start at a minimum of 6 months post-event) (Van de Schoot et al., 2018). It is generally acknowledged that many victims of disasters and large-scale incidents follow a resilient trajectory; they are able to cope with the aversive event (Bonanno, Brewin, Kaniasty, & La Greca, 2010; Kleber & Brom, 1992). However, a significant minority can develop severe mental disturbances, of which PTSD, major depressive disorder (MDD) and anxiety disorders are most commonly reported (Norris et al., 2002). In the specific context of an aeroplane crash, a prevalence of 40% for PTSD and 33% for depression was found (Gregg et al.,

1995), which is high compared to other disasters (10–20%) (Bonanno et al., 2010; Norris et al., 2002).

For individuals who need treatment, timely identification is crucial: evidence-based treatment such as trauma-focused cognitive-behavioural therapy may effectively treat acute PTSD and prevent a chronic course of PTSD symptoms (Roberts, Kitchiner, Kenardy, & Bisson, 2009). Although in some countries existing pathways to mental health care via general practitioners have appeared to be sufficient in the post-disaster context (Van der Velden et al., 2006), victims may experience significant barriers to service utilization. The proportion of people with unmet treatment needs after disasters has been found to vary between 14% and 50% (Dyb, Jensen, Glad, Nygaard, & Thoresen, 2014; Stuber, Galea, Boscarino, & Schlesinger, 2006). Detection of PTSD may be hampered, as general practitioners do not always recognize it (Rosenbaum, 2004). In addition, victims experience barriers such as the stigma attached to psychological problems, avoidance of trauma-related memories or the idea that others need help more than they do (Stuber et al., 2006; Weisæth, 2001). These barriers may prevent people with severe levels of distress from finding and using the care and support they need. Therefore, recent guidelines on psychosocial care after disasters or crises stress the importance of an active approach to identify people with high levels of distress and to foster timely

referral to (specialized) mental health care, often referred to as 'proactive outreach' (Bisson et al., 2010; Shultz & Forbes, 2014; Te Brake et al., 2009).

To implement proactive outreach, screening programmes have attempted to detect people with significant levels of distress. Validated, accurate screening instruments include the SPAN (Startle, Physiological arousal, Anger and Numbness), the Impact of Event Scale – Revised (IES-R) and the Trauma Screening Questionnaire (TSQ) for the detection of PTSD (Mouthaan, Sijbrandij, Reitsma, Gersons, & Olff, 2014), and the Posttraumatic Adjustment Scale and the Patient Health Questionnaire-2 (PHQ-2) for measuring depression (Kroenke, Spitzer, & Williams, 2003; O'Donnell et al., 2008). However, research on the performance of screening instruments in the disaster context is scarce, where the logistics of detecting and monitoring victims is often challenging (e.g. the Amsterdam plane crash survivors came from several different countries). In addition, the results of studies on the diagnostic accuracy of screening instruments that may be used as part of active outreach are inconsistent. Although some studies found high levels of both sensitivity and specificity for PTSD screeners (Brewin, 2005), other studies showed that specificity was low 6 months after traumatic injury (Mouthaan et al., 2014) and during the first 18 months after a terrorist attack (Brewin, Fuchkan, Huntley, & Scragg, 2010a). This implies that many healthy people need to be subjected to assessment (i.e. false positives). Furthermore, research has predominantly focused on screening for PTSD, instead of other common disorders such as depression, and on screening for current disorders rather than the risk for future disorders (O'Donnell et al., 2008).

The potential benefit of active outreach depends not only on the existence of accurate instruments to detect those at risk, but also on the number of people having unmet treatment needs, in the short and long term, and on the extent to which people themselves are able to fulfil their needs. As most studies are cross-sectional (e.g. Dyb et al., 2014; Stuber et al., 2006), little is known about the course of treatment needs over time and whether treatment needs in the first months after a disaster predict adverse long-term trajectories of PTSD. Furthermore, the long-term psychological well-being of people without treatment needs has attracted little attention (see for exceptions Van den Berg, Wong, van der Velden, Boshuizen, & Grievink, 2012; Shalev, Ankri, Peleg, Israeli-Shalev, & Freedman, 2011). After traumatic injuries, individuals with acute stress disorder symptoms who declined further assessment and treatment recovered less well than those who accepted treatment (Shalev et al., 2011). In the post-disaster mental health field, knowledge about the long-term health of victims who report not needing any treatment is crucial to understand whether we can rely on people's own estimation of whether they need professional care.

In this longitudinal study, we evaluated an outreach strategy that sought to identify and help those at risk for PTSD and depression after an aeroplane crash. The programme was partly modelled after the Screen and Treat programme developed by Brewin et al. (2010b) after the London bombings. First, people were screened for PTSD and depression. Those who screened positive and/or expressed unmet treatment needs were advised by the CHS to contact their local health-care providers (such as general practitioners) who could provide formal assessment and, when needed, provide or refer to evidence-based treatment. Contact details of local health-care providers specialized in evidence-based trauma treatment were available to support victims in finding appropriate care.

In this study, we first examined whether we were able to screen people effectively for clinically significant PTSD and depression by measuring the performance (sensitivity and specificity) of screening methods 44 months (3.5 years) post-event. Secondly, we examined the proportion of self-reported treatment needs at 2 months, 9 months and 44 months after the crash. Thirdly, we sought to investigate the extent to which potential treatment needs assessed at 2 months after the crash predicted PTSD trajectories.

2. Methods

2.1. Procedure

The regional CHS (GGD Kennemerland) conducted a survey at 2 (T1), 9 (T2) and 44 months (T3) after the crash. The survey consisted of a semi-structured section on treatment needs and psychosocial support and a structured section to screen for PTSD and depression. At T3, a structured clinical interview, the Mini International Neuropsychiatric Interview (MINI) (Sheehan et al., 1997), was added to test the performance of the screening instruments. Instruments were administered by telephone in Turkish, Dutch or English by social workers or psychologists. Time between filling out the screening instrument and conducting the clinical interview was no longer than 1 week. The clinical interviewer was unaware of the results of the screening instruments (i.e. severity of symptoms of PTSD and/or depression) so that this could not influence the diagnostic assessment. A second researcher independently scored the diagnostic decisions of the clinical interview in 30% of the cases by listening in on the interview. Concerning depressive disorders, agreement between raters was 82% (100% after deliberation between researchers on interpreting the severity of symptoms and scoring the diagnostic criteria). Agreement was 100% regarding PTSD.

The regional Medical Research Ethics Committee of Noord Holland gave approval for this study (M012-014).

2.2. Participants

Figure 1 provides a flow diagram presenting the number of participants and reasons for non-response. We excluded children under 14 years. Survey response rates were 62.8%, 63.6% and 45.5% at T1, T2 and T3, respectively. In total, 89 respondents participated at one or more time-points and 35 participated at all three time-points. When asked, reasons given for non-participation were current hospitalization/physical complaints ($n = 5$), already receiving treatment ($n = 2$), not having any psychosocial problems ($n = 5$) and not wanting to talk about the event ($n = 2$). Of all the people participating at T3, 69.1% ($n = 38$) completed the MINI.

Demographics (age, gender, nationality) for the total population, the samples at each time-point and non-responders are provided in Tables 1 and 2. We found no significant differences in age between those who completed the surveys and non-responders [T1: $t(116) = 0.05$, $p = .621$; T2: $t(116) = 1.64$, $p = .106$; T3: $t(116) = 0.92$, $p = .360$]. At T2, participants were more often female [$\chi^2(1) = 4.14$, $p = .042$] and of Dutch origin [$\chi^2(2) = 7.98$, $p = .019$] compared to non-responders at T2. At T1 and T3, there were no such differences regarding gender [T1: $\chi^2(1) = 3.01$, $p = .083$; T3: $\chi^2(1) = 1.96$, $p = .162$] and origin [T1: $\chi^2(2) = 5.93$, $p = .052$; T3: $\chi^2(2) = 3.02$, $p = .221$]. Dropouts at T3 reported higher scores on depressive symptoms at T2 ($n = 31$, $M = 2.52$, $SD = 2.20$) compared to participants at T2 who did not drop out at T3 ($n = 45$, $M = 1.38$, $SD = 1.59$) [$t(74) = 2.62$, $p = .011$, 95% confidence interval (CI) 0.27–2.00]. There were no significant differences in terms of depressive symptoms at T1 [$t(71) = 1.06$, $p = .290$] or in terms of PTSD symptoms at T1 [$t(68) = 0.10$, $p = .92$] and T2 [$t(73) = 1.90$, $p = .602$]. At T3, those refusing to undergo the MINI reported higher symptom levels of PTSD ($n = 16$, $M = 6.20$, $SD = 3.53$) compared to those completing the MINI [$M = 3.52$, $SD = 2.88$]: $t(51) = 2.86$, $p = .006$, 95% CI 0.80–4.56]. There were
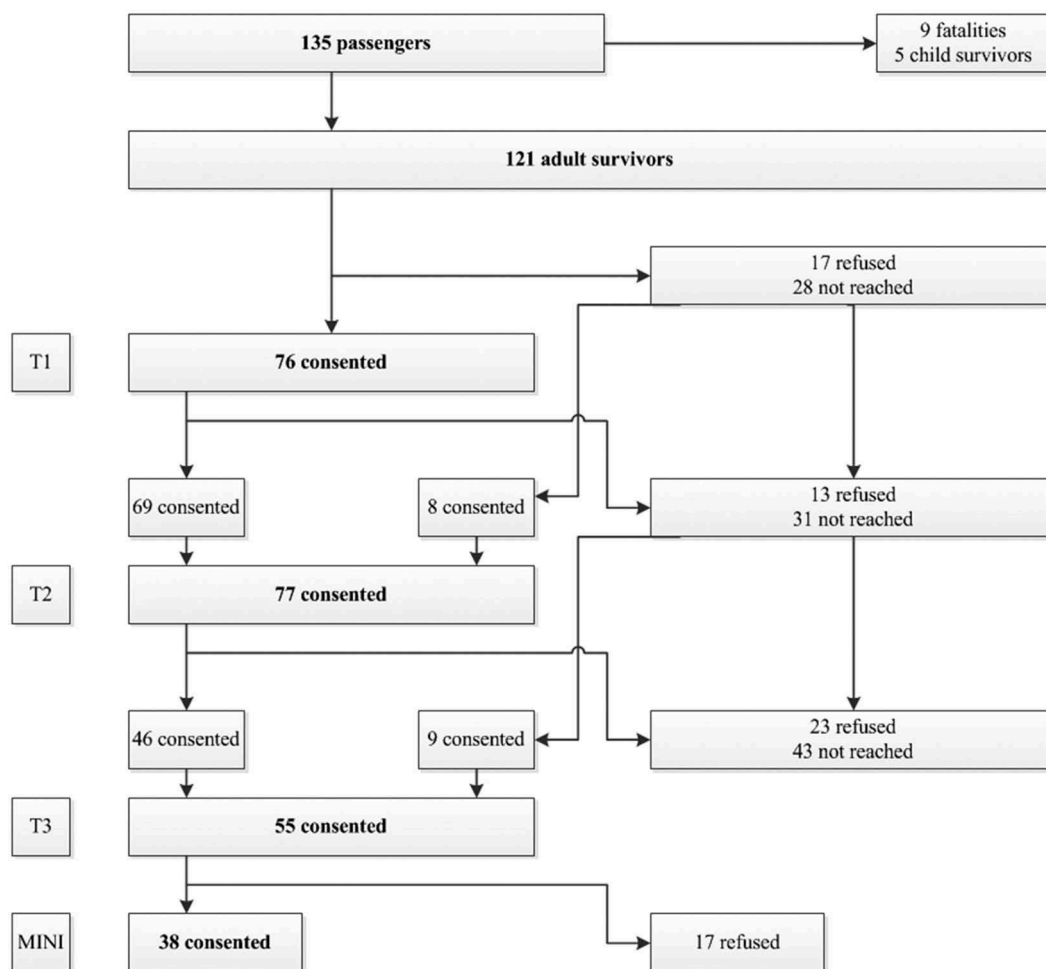

Figure 1. Response and non-response at T1 (2 months after the crash), T2 (9 months after the crash) and T3 (44 months after the crash).

Table 1. Demographics, treatment needs and symptoms of post-traumatic stress disorder (PTSD) and depression in the base population and studied samples.

		Adult survivors	T1: 2 months	T2: 9 months	T3: 44 months	T3: MINI
Total *n*		121	76	77	55	38
Age	*n*	118	73	74	55	38
	Mean (SD)	40.15 (13.20)	40.03 (14.05)	41.68 (13.97)	44.27 (13.13)	42.95 (12.60)
Gender	Men	84 (69.4%)	49 (64.5%)	49 (63.6%)	35 (63.6%)	24 (63.2%)
	Women	36 (29.8%)	27 (35.5%)	28 (36.4%)	20 (36.4%)	14 (36.8%)
Nationality	Dutch	60 (49.6%)	44 (57.9%)	45 (58.4%)	32 (58.2%)	23 (60.5%)
	Turkish	46 (38.0%)	25 (32.9%)	26 (33.8%)	17 (30.9%)	10 (26.3%)
	Iranian	4 (3.3%)	3 (3.9%)	3 (3.9%)	3 (5.5%)	3 (7.9%)
	Other[a]	11 (9.1%)	4 (5.3%)	3 (3.9%)	3 (5.5%)	2 (5.3%)
Treatment needs	*n*		74	76	54	38
	Unmet needs		25 (32.9%)	15 (19.5%)	6 (10.9%)	6 (15.8%)
	Met needs (treatment)		28 (36.8%)	42 (54.5%)	34 (61.8%)	23 (60.5%)
	No need		21 (27.6%)	19 (24.9%)	14 (25.9%)	9 (23.7%)
TSQ score	*n*		70	75	52	37
	Mean (SD)		5.23 (3.46)	4.79 (3.47)	4.27 (3.31)	3.49 (2.91)
PHQ-2 score	*n*		73	76	53	37
	Mean (SD)		1.95 (1.93)	1.84 (1.93)	1.51 (1.41)	1.38 (1.36)

Variables may have different sample sizes and percentages do not always add up to 100% owing to missing values. [a] Other countries were: USA, UK, Bulgaria, Germany, Italy, Sudan, Syria and Taiwan.
MINI, Mini International Neuropsychiatric Interview; TSQ, Trauma Screening Questionnaire; PHQ-2, Patient Health Questionnaire-2.

Table 2. Demographics of non-responders.

		T1: 2 months	T2: 9 months	T3: 44 months	T3: MINI
Total *n*		45	44	66	17
Age	*n*	45	44	63	17
	Mean (SD)	38.76 (12.02)	37.59 (11.51)	42.03 (13.30)	47.24 (14.17)
Gender	Men	35 (77.8%)	35 (79.5%)	49 (74.2%)	11 (64.7%)
	Women	9 (20.0%)	8 (18.2%)	16 (24.2%)	6 (35.3%)
Nationality	Dutch	16 (35.6%)	15 (34.1%)	28 (42.4%)	9 (52.9%)
	Turkish	21 (46.7%)	20 (45.5%)	29 (43.9%)	7 (36.8%)
	Iranian	1 (2.2%)	1 (2.3%)	1 (1.5%)	0
	Other[a]	7 (15.6%)	8 (18.2%)	8 (12.2%)	1 (5.8%)

Variables may have different sample sizes and percentages do not always add up to 100% owing to missing values. [a] Other countries were: USA, UK, Bulgaria, Germany, Italy, Sudan, Syria and Taiwan.
MINI, Mini International Neuropsychiatric Interview.

no such differences in the level of depressive symptoms [$t(52) = 0.810$, $p = .422$] or in the distribution of age [$t(53) = 1.12$, $p = .267$], gender [$t(1) = 0.12$, $p = .912$] or origin [$t(2) = 1.53$, $p = .466$].

2.3. Measures

Symptoms of PTSD during the past week were measured using the TSQ, a 10-item questionnaire with a yes/no response format (Brewin et al., 2002). It consists of five items about re-experiencing and five items about arousal; scores range from 0 (asymptomatic) to 10. At the standard cut-off of six symptoms in any combination, the TSQ showed a sensitivity and specificity of .86 and .93, respectively, in predicting PTSD in rail crash survivors at 6–12 months post-trauma (Brewin et al., 2002). After the 2005 London bombings, sensitivity decreased from .95 within the first month to .63 2 years later, while specificity rose from .26 to .64 (Brewin et al., 2010a).

Symptoms of depression were measured with the PHQ-2, a two-item measure that enquires about the frequency of depressed mood and anhedonia over the past 2 weeks (Kroenke et al., 2003). The PHQ-2 uses a four-option response format (from 'not at all' to 'nearly every day'). The total

score ranges from 0 to 6. A PHQ-2 score ≥3 has been found to agree well with formal diagnosis, with sensitivities of .83–.87 and specificities of .78–.92 being reported in primary care and medical settings (Kroenke et al., 2003; Löwe, Kroenke, & Gräfe, 2005).

Anxiety-, mood- or trauma-related disorders according to the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) were assessed with the MINI version 5.0.0 (Sheehan et al., 1997). Comparison with other 'gold standard' clinical interviews revealed that the validity and reliability of the MINI are high (Lecrubier et al., 1997; Sheehan et al., 1997).

Subjective treatment needs were measured with questions about received psychological treatment since the crash ('Are you currently receiving any psychological treatment?' and 'Did you receive any psychological treatment since the crash?') and unmet treatment needs ('Are you in need of any psychological support?'), using a yes/no response format. Open-ended questions followed up on the kind of treatment they received ('What kind of psychological support do you receive?'). Two researchers (JG and HtB) independently categorized these perceived needs as: (1) unmet treatment needs; (2) met needs, defined as receiving or having received

counselling or treatment since the crash (from a psychologist, psychiatrist or other mental health-care professional); or (3) no need for treatment. Participants who had received psychological treatment but expressed additional treatment needs (six participants) were coded as having unmet treatment needs. Differences in coding between the researchers were discussed and resolved.

2.4. Data analysis

Statistics were computed in SPSS Statistics 23, with two-tailed alpha levels of $< .05$ indicating statistical significance.

To evaluate the accuracy of the TSQ and PHQ-2, sensitivity (the probability that someone with a MINI diagnosis of PTSD or MDD screened positive), specificity (the probability that someone without a MINI diagnosis screened negative), positive predictive value (PPV; the proportion of those who screened positive and met criteria for PTSD or depression) and negative predictive value (NPV; the proportion of individuals who screened negative and did not meet criteria for PTSD or depression) were established. Receiver operating characteristics (ROC) curves were calculated for the screening instruments in relation to the 44 month diagnoses of PTSD and MDD. The area under the curve (AUC) represents the accuracy of the instruments. Missing data were replaced with the scale mean (PHQ-2: one case with one missing item; TSQ: two cases with two and three missing items).

To analyse the treatment needs and course of PTSD, multiple imputation (MI) implemented in SPSS was used to handle missing data on the TSQ. The missing values appeared to be not at random (as described in Section 2.2). MI is the preferred way to handle missing values not at random compared to, for instance, listwise deletion (Newman, 2003). We included the total sum scores of the TSQ for those who had participated at least at one time-point in the imputation model, and used 'fully conditional specifications' and 'predictive mean matching' (TSQ scores were slightly positively skewed) (Van Buuren, 2012) as methods for addressing the missing data. Multiple (500) versions of the data set were produced. When statistical analyses were performed, the parameter estimates for all of the imputed data sets were pooled.

We examined whether treatment needs at T1 predicted PTSD trajectories by first distinguishing PTSD trajectories in this sample. Because of the small sample size, we computed trajectories by first determining whether someone was at risk for PTSD or not at each time-point using validated cut-off scores (scoring 6 or higher on the TSQ indicated being at risk) (Brewin et al., 2002). We then inspected and rated all the possible combinations of being at risk or not for PTSD at the three time-points (see Table 3 for a detailed description of the combinations). We identified five PTSD trajectories in the imputed sample ($N = 89$): resilient ($n = 38$, 42.7%), chronic ($n = 30$, 33.7%), recovery ($n = 9$, 10.1%), delayed onset ($n = 9$, 10.1%) and relapse ($n = 3$, 3.4%) (Table 3). Similar trajectories were found in the non-imputed sample ($N = 38$): resilient ($n = 17$, 44.7%), chronic ($n = 12$, 31.6%), recovery ($n = 5$, 13.2%), delayed onset ($n = 3$, 7.9%) and relapse ($n = 1$, 2.6%). We used Fisher's exact and chi-squared tests to examine whether treatment needs predicted these trajectories.

3. Results

3.1. Accuracy of screening instruments

Of the participants who completed the MINI at T3 ($N = 38$), five (13.2%) met criteria for current PTSD (plane crash as index trauma) and six for MDD (15.8%). PTSD caused by other traumatic experiences or other depressive disorders were not reported. Two participants were diagnosed with comorbid PTSD and MDD. Table 4 presents the TSQ and PHQ-2 scores for individuals participating in the clinical interview.

Table 5 shows that AUCs of the TSQ were adequate at all time-points (varying from .78 to .93), indicating that the TSQ accurately predicted PTSD status. The TSQ yielded good sensitivity and specificity at all time-points. For example, at T3, 80% of the participants with a PTSD diagnosis screened positive on the TSQ (sensitivity; true positives), while 20% went undetected (false negatives). In addition, 90% of the participants without a PTSD diagnosis screened negative (specificity; true negatives), while 10% without the disorder were identified as at risk for PTSD (false positives). Although the NPV was high at

Table 3. Treatment needs at T1 related to post-traumatic stress disorder (PTSD) trajectories.

Trajectory	n (%)	Explanation trajectories T1–T2–T3	Treatment needs at T1, n (%)			Mean TSQ		
			Met needs	No need	Unmet needs	T1	T2	T3
Resilient	38 (42.7)	0–0–0[a]	11 (39.3)	14 (66.7)	5 (20.0)	2.54	1.89	2.06
Chronic	30 (33.7)	1–1–1	10 (35.7)	3 (14.3)	12 (48.0)	8.37	8.27	8.28
Recovery	9 (10.4)	1–0–0/1–1–0	5 (17.9)	0 (0.0)	2 (8.0)	7.46	4.00	3.57
Delayed onset	9 (10.4)	0–1–0/0–1–1[b]	2 (7.1)	3 (14.3)	4 (16.0)	3.78	7.27	4.96
Relapse	3 (3.4)	1–0–1	0 (0.0)	1 (4.7)	2 (8.0)	9.00	3.67	7.45

[a] 0 indicates not at risk for PTSD and 1 indicates at risk for PTSD, measured with the TSQ; a score of 6 or higher implies being at risk for PTSD. [b] 0–0–1 did not exist in this sample.
TSQ, Trauma Screening Questionnaire.

Table 4. Risk groups for post-traumatic stress disorder (PTSD) and depression measured with the Trauma Screening Questionnaire (TSQ) and Patient Health Questionnaire-2 (PHQ-2).

		T1: 2 months	T2: 9 months	T3: 44 months
TSQ[a]	n	28	33	37
	Risk	8	9	7
	No risk	20	24	30
PHQ-2[b]	n	30	33	38
	Risk	6	7	8
	No risk	24	26	30

[a] At risk was defined as a score of 6 or higher. [b] At risk was defined as a score of 3 or higher.

all three time-points (ranging from 95% to 100%), the PPV was low (ranging from 38% to 57%), indicating that given a positive test result, only around half would be diagnosed with PTSD. The use of alternative cut-off points did not enhance the performance of the TSQ in this population.

AUCs of the PHQ-2 were adequate (varying from .72 to .87) (Table 5). The PHQ-2 proved accurate in identifying those without MDD: at all time-points specificity and NPV were good. However, at T1 and T3 sensitivity was low (.60 and .38, respectively), indicating a high number of false negatives (i.e. two out of five participants with an MDD diagnosis at T1, and three out of eight participants with an MDD diagnosis at T3 did not screen positive). PPV was also low at these time-points: 50% of those who screened positive did not have a diagnosis of MDD. At T2, a cut-off point of 2 increased the sensitivity of the PHQ-2 (.83), without losing its specificity (.85). At T1 and T3, a cut-off point of 3 proved optimal in this population compared to other cut-off points.

To check whether the TSQ and PHQ-2 were accurate in detecting any trauma-related psychopathology, we assessed the sensitivity and specificity of both instruments in detecting any anxiety-, mood- or trauma-related disorder measured with the MINI. For both instruments, sensitivity was low, while specificity was high (Appendix A).

3.2. Proportion of subjective treatment needs

Perceived unmet treatment needs were reported by 32.9% of the participants at T1, decreasing to 19.5% at T2 and 10.9% at T3 (Table 1). At each time-point,

approximately a quarter of the participants perceived no treatment needs. The group with met treatment needs since the crash increased from 36.8% at T1 to 62.8% at T3. Of the nine participants with a PTSD and/or MDD diagnosis at T3, six reported having had psychological treatment since the crash, while three reported unmet treatment needs. The course of perceived treatment needs is presented in Figure 2.

3.3. Subjective treatment needs and PTSD trajectories

There was a significant association between potential treatment needs at T1 and the PTSD trajectories (Fisher's exact test, $p = .013$). Of the participants with no treatment needs at T1, most were included in the resilient group (66.7%), while those with unmet treatment needs were most often included in the chronic group (48.0%) (Table 3). Since the recovery, delayed onset and relapse groups were small ($n < 10$), we checked whether the differences between the chronic and resilient groups regarding unmet treatment needs and no needs were significant using a 2×2 contingency table. Those with unmet needs more often showed a chronic trajectory compared to those with no treatment needs, who more often showed a resilient trajectory [$\chi^2(1)$ ($N = 34$) = 9.66, $p = .005$]. This reflects the fact that the odds of belonging to the resilient group were 11.2 times higher when someone expressed no treatment needs at T1 than if someone expressed unmet treatment needs at T1.

4. Discussion

This study evaluated, in a longitudinal design, the implementation of an outreach strategy after an aeroplane crash near Amsterdam in 2009 by examining the accuracy of validated screening instruments for PTSD and MDD, self-reported treatment needs and the extent to which perceived treatment needs predicted trajectories of PTSD. Taking into account the relatively low prevalence of PTSD diagnosis in this study and the uncertainty that arises with few PTSD cases, the TSQ proved stable over time in detecting both PTSD cases and non-cases at 44 months after the

Table 5. Diagnostic performance of the Trauma Screening Questionnaire (TSQ) and Patient Health Questionnaire-2 (PHQ-2) in predicting post-traumatic stress disorder (PTSD) and major depressive disorder (MDD).

Screener	Time-point	Sensitivity	Specificity	PPV	NPV	AUC (95% CI)	n
TSQ[a]	1	.75	.79	.38	.95	.78 (.59–.98)	28
	2	1.00	.86	.56	1.00	.96 (.90–1.00)	33
	3	.80	.90	.57	.97	.93 (.83–1.00)	37
PHQ-2[b]	1	.60	.88	.50	.92	.72 (.45–.96)	30
	2	.89	.67	.92	1.00	.96 (.89–1.00)	33
	3	.38	.90	.50	.84	.87 (.74–.96)	38

[a] A cut-off of 6 or higher was taken to indicate risk for PTSD. [b] A cut-off of 3 or higher was taken to indicate risk for depression.
PPV, positive predictive value; NPV, negative predictive value; AUC, area under the curve.
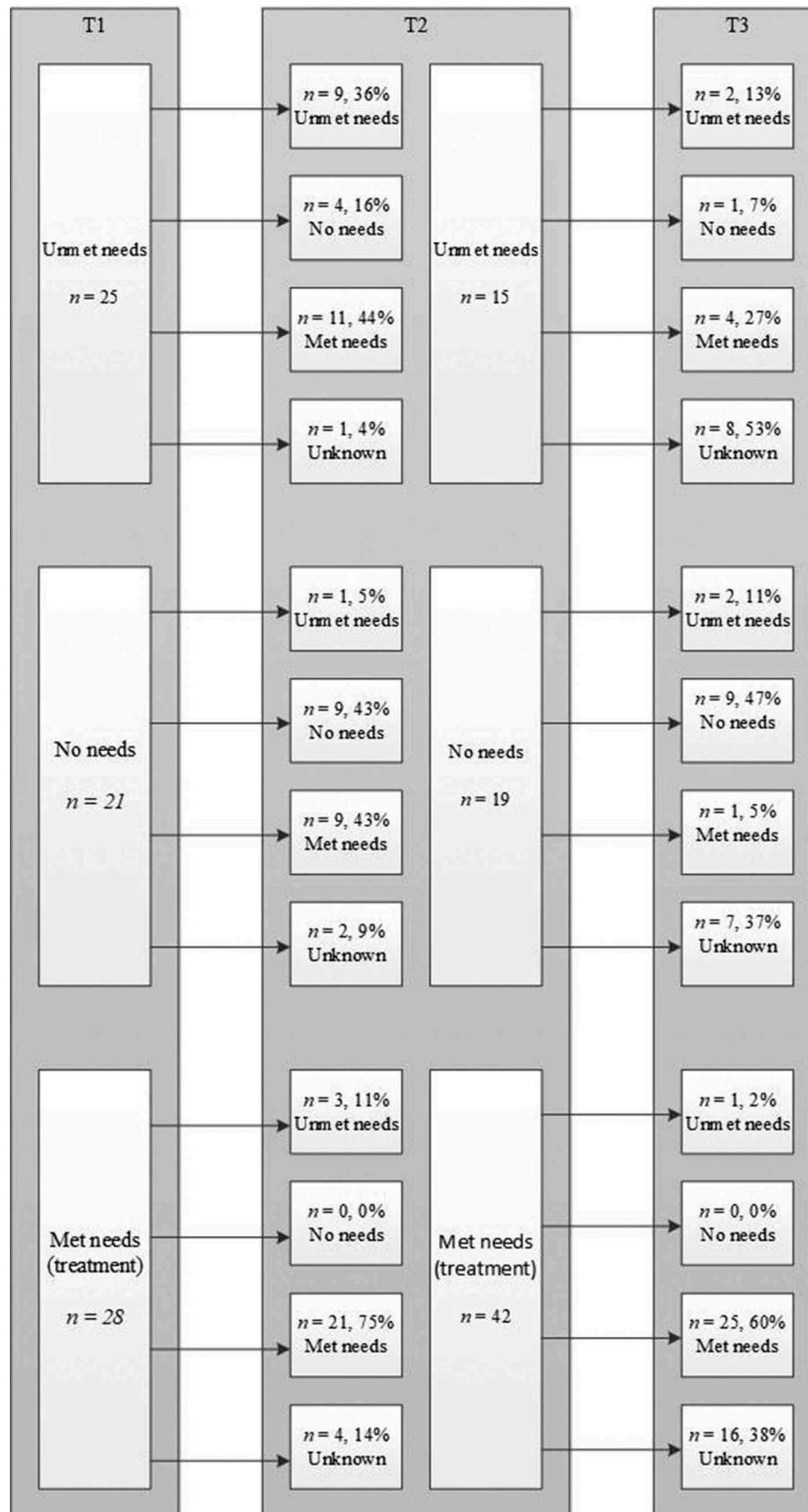
Figure 2. Flow diagram of the course of unmet treatment needs, no need for treatment and met needs (treatment group) between time-points 1 and 2 and between time-points 2 and 3.

crash (i.e. sensitivity and specificity were good). The accuracy of the TSQ was comparable to previous findings among rail-crash survivors (Brewin et al., 2002). It was also comparable to the accuracy of another PTSD screener, the IES-R, which was used 6 years after the 2004 Indian Ocean tsunami disaster (Arnberg, Michel, & Johannesson, 2014). The findings provided less support for the use of the PHQ-2 to detect MDD. Its

performance was inconsistent over time: either a significant proportion of participants with MDD went undetected or a significant proportion without the disease were classified as being at risk. Using an adjusted cut-off score did not improve the stability and overall performance of the PHQ-2. Although the PHQ-2 proved to be an accurate screening instrument in primary care and medical settings (Kroenke et al., 2003; Löwe et al., 2005), an instrument of two items to measure depression may be too brief in a disaster setting including victims with various cultural backgrounds. A study in a Turkish clinical sample showed that somatization symptoms were a predominant mode of expression of depressive symptoms (Uluşahin, Başoğlu, & Paykel, 1994).

The PPVs of both the TSQ and the PHQ-2 were low, indicating that a significant proportion of those at risk will eventually not be diagnosed with PTSD or MDD. In our sample, the highest chance of someone with a positive result on the TSQ developing PTSD was 57%. This is probably due to the low prevalence of PTSD and MDD among this group of victims (Baldessarini, Finklestein, & Arana, 1983). Low PPV is a recurrent problem of screening instruments (Mouthaan et al., 2014), including among disaster victims (Arnberg et al., 2014). It underscores that screening positive should not be taken as indicative of a PTSD or MDD diagnosis, but instead should signal further assessment. A low PPV also highlights the need to carefully weigh up the advantages and disadvantages of screening. A significant proportion of participants who screen positive (and receive an additional assessment) will eventually not develop a clinical disorder. There is also no information about possible negative effects of screening healthy people, meaning that we do not know whether screening carries a risk of medicalization, framing normal behaviour as problematic and interfering with healthy ways of coping with disaster. On the other hand, the decision not to screen carries a risk of untreated PTSD, which has been related to a chronic PTSD course, low quality of life, other clinical disorders such as substance abuse and considerable health-care costs (Bichescu et al., 2005; Priebe et al., 2009).

The discussion about the advantages and disadvantages of the use of screening instruments touches on a larger debate about the degree to which post-disaster strategies for psychosocial care should involve an 'outreach' approach (Dückers & Thormar, 2014; Te Brake & Dückers, 2013; Te Brake et al., 2009). Interventions after a disaster greatly differ in how actively they monitor the victim's health. Some interventions primarily trust the existing health-care system and victims' ability to find the care they need (e.g. Van der Velden, Loon, IJzermans, & Kleber, 2006), while other interventions lean towards actively intervening because of

indications that a proportion of those in need may be overlooked (e.g. Stene, Wentzel-Larsen, & Dyb, 2016). Dückers and Thormar (2014) comment that these differences exist partly because of differences in the availability of existing health-care systems after disasters, and partly because of different attitudes among scholars and policy advisors regarding the extent to which one should rely on victims' own capacity to find treatment.

We found that a significant proportion of participants reported unmet treatment needs, even 3.5 years post-event (33% at 2 months to 11% at 44 months). Unmet needs 2 months after the crash were also more often related to a chronic PTSD trajectory than no treatment needs. These findings are in agreement with many disaster studies showing that existing pathways to mental health care were not always sufficient (Reifels et al., 2013). In countries like the UK (Brewin et al., 2010b) and Norway (Dyb et al., 2014), it was found that outreach generated more referrals of affected individuals compared to normal referral channels. Even when general practitioners were strongly involved, the creation of additional flexible pathways increased victim self-referral, suggesting that a disaster-specific case-management service may indeed enhance service access (Dyb et al., 2014; Reifels et al., 2015). The finding in this study that unmet needs were more often related to a chronic trajectory may even raise the question of whether a strategy in which participants at risk were automatically routed into an assessment and, when needed, evidence-based treatment, such as after the London bombings (Brewin et al., 2010b), would have generated more recoveries.

In line with a more neutral attitude to outreach, this study showed that a substantial number of participants (e.g. 30% at 2 months and 26% at 44 months) indicated that they did not need any treatment. No perceived treatment need 2 months after the crash was also related to a resilient trajectory from 2 to 44 months after the crash, implying that we can trust in victims' self-judgement of their need for care. This, in turn, suggests that many of those without self-reported treatment needs do not require additional screening or assessment. However, we need to be cautious about generalizing this finding to other disasters or crises. Boscarino, Adams, Stuber, and Galea (2005) showed that, among adults with diagnoses of PTSD or MDD after the 2011 New York terrorist attacks, a substantial number (especially African Americans) did not believe that they had a problem that needed treatment. As a group of residents with ethnic minority status exposed to terrorism differs from victims of a plane crash travelling for business reasons or holidays, these contradictory results underscore that, after each disaster, psychosocial responses

need to proactively address the needs and potential barriers to health care experienced by various survivor subgroups (Reifels et al., 2013).

4.1. Strengths and limitations

By evaluating the course of a disaster response in a longitudinal design, this study contributes towards the improvement of guidelines for psychosocial support and their implementation into actual practice. The study nevertheless has a number of limitations, most of which result directly from its unique setting. Given the circumstances in which the research was performed, neither random sampling nor controlled experimentation were possible, precluding any direct conclusions about the effectiveness of 'active outreach' in terms of preventing (exacerbation of) post-traumatic stress and associated problems. Sample size was limited, making it impossible to control for confounding variables (e.g. it was not possible to compare the accuracy of the screening instruments between Dutch and Turkish victims). Non-response was high, especially at follow-up, and individuals with higher symptoms of depression were more likely to drop out. We also found higher levels of PTSD symptoms in participants who were lost to the MINI. These limitations not only point to possible attrition biases in our results, but also have implications for the usefulness of active outreach: a substantial minority of people may not be willing to participate, which has also been found after some other disasters (Dijkema, Grievink, Stellato, Roorda, & Van der Velden, 2005). Another limitation to be considered is that we were not able to assess which kind of treatment people received. Also, other life events, such as sickness or loss of a loved one, were not included in the study, which, especially over the longer term, are likely to have occurred in participants' lives, and may have influenced their perceived needs. Finally, specifically for PTSD, the gold standard clinical interview would have been the Clinician Administered PTSD Scale for DSM-IV (Weathers, Keane, & Davidson, 2001). However, since we were not only interested in PTSD as an outcome measure, we used the MINI (the gold standard for detecting any psychopathology).

To fully understand the potential benefits of active outreach, we recommend future research to examine the impact of screening and clinical assessment on the resilience of disaster victims, to investigate the cultural validity of screeners in a disaster context, and to focus on alternative outcomes that are likely to benefit from active outreach, such as social acknowledgement.

4.2. Conclusion

Although guidelines recommend detecting individuals at risk of mental health problems after crises and disasters, literature on the evaluation and practical implementation of outreach strategies is scarce. This study among 121 adult survivors of an aeroplane crash suggests that, given the risk of chronic PTSD in combination with unmet needs, screening may be warranted. At the same time, as a significant proportion of victims screening positive will not develop a PTSD diagnosis, more research is needed to study the effects of screening healthy people. These conclusions need to be verified in larger samples and other disaster populations.

ORCID

Marit Sijbrandij http://orcid.org/0000-0001-5430-9810
Paul A. Boelen http://orcid.org/0000-0003-4125-4739

References

Arnberg, F. K., Michel, P. O., & Johannesson, K. B. (2014). Properties of Swedish posttraumatic stress measures after a disaster. *Journal of Anxiety Disorders*, 28(4), 402–409.

Baldessarini, R. J., Finklestein, S., & Arana, G. W. (1983). The predictive power of diagnostic tests and the effect of prevalence of illness. *Archives of General Psychiatry*, 40 (5), 569–573.

Bichescu, D., Schauer, M., Saleptsi, E., Neculau, A., Elbert, T., & Neuner, F. (2005). Long-term consequences of traumatic experiences: An assessment of former political detainees in Romania. *Clinical Practice and Epidemiology in Mental Health*, 1(1), 17.

Bisson, J. I., Tavakoly, B., Witteveen, A. B., Ajdukovic, D., Jehel, L., Johansen, V. J., … Sezgin, A. U. (2010). TENTS guidelines: Development of post-disaster psychosocial care guidelines through a Delphi process. *British Journal of Psychiatry*, 196(1), 69–74.

Bonanno, G. A., Brewin, C. R., Kaniasty, K., & La Greca, A. M. (2010). Weighing the costs of disaster: Consequences, risks, and resilience in individuals, families, and communities. *Psychological Science in the Public Interest*, 11(1), 1–49.

Boscarino, J. A., Adams, R. E., Stuber, J., & Galea, S. (2005). Disparities in mental health treatment following the world trade center disaster: Implications for mental

health care and health services research. *Journal of Traumatic Stress*, 18(4), 287–297.

Brewin, C. R. (2005). Systematic review of screening instruments for adults at risk of PTSD. *Journal of Traumatic Stress*, 18(1), 53–62.

Brewin, C. R., Fuchkan, N., Huntley, Z., Robertson, M., Thompson, M., Scragg, P., … Ehlers, A. (2010b). Outreach and screening following the 2005 London bombings: Usage and outcomes. *Psychological Medicine*, 40(12), 2049–2057.

Brewin, C. R., Fuchkan, N., Huntley, Z., & Scragg, P. (2010a). Diagnostic accuracy of the trauma screening questionnaire after the 2005 London bombings. *Journal of Traumatic Stress*, 23(3), 393–398.

Brewin, C. R., Rose, S., Andrews, B., Green, J., Tata, P., McEvedy, C., … Foa, E. B. (2002). Brief screening instrument for post-traumatic stress disorder. *The British Journal of Psychiatry*, 181(2), 158–162.

Dijkema, M. B., Grievink, L., Stellato, R. K., Roorda, J., & Van der Velden, P. G. (2005). Determinants of response in a longitudinal health study following the firework-disaster in Enschede, The Netherlands. *European Journal of Epidemiology*, 20(10), 839–847.

Dückers, M. L., & Thormar, S. B. (2014). Post-disaster psychosocial support and quality improvement: A conceptual framework for understanding and improving the quality of psychosocial support programs. *Nursing & Health Sciences*, 17(2), 159–165.

Dyb, G., Jensen, T., Glad, K. A., Nygaard, E., & Thoresen, S. (2014). Early outreach to survivors of the shootings in Norway on the 22nd of July 2011. *European Journal of Psychotraumatology*, 5(1), 23523.

Gregg, W., Medley, I., Fowler-Dixon, R., Curran, P., Loughrey, G., Bell, P., … Harrison, G. (1995). Psychological consequences of the Kegworth air disaster. *The British Journal of Psychiatry*, 167(6), 812–817.

Kleber, R. J., & Brom, D. (1992). *Coping with trauma: Theory, prevention and treatment*. Abingdon-on-Thames, UK: Taylor & Francis (reprint 2003).

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2003). The patient health questionnaire-2: Validity of a two-item depression screener. *Medical Care*, 41(11), 1284–1292.

Lecrubier, Y., Sheehan, D. V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K. H., … Dunbar, G. C. (1997). The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: Reliability and validity according to the CIDI. *European Psychiatry*, 12(5), 224–231.

Löwe, B., Kroenke, K., & Gräfe, K. (2005). Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *Journal of Psychosomatic Research*, 58(2), 163–171.

Mouthaan, J., Sijbrandij, M., Reitsma, J. B., Gersons, B. P., & Olff, M. (2014). Comparing screening instruments to predict posttraumatic stress disorder. *PloS One*, 9(5), e97183.

Newman, D. A. (2003). Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods*, 6(3), 328–362.

Norris, F. H., Friedman, M. J., Watson, P. J., Byrne, C. M., Diaz, E., & Kaniasty, K. (2002). 60,000 disaster victims speak: Part I. An empirical review of the empirical literature, 1981–2001. *Psychiatry: Interpersonal and Biological Processes*, 65(3), 207–239.

O'Donnell, M. L., Creamer, M. C., Parslow, R., Elliott, P., Holmes, A. C., Ellen, S., … Bryant, R. A. (2008). A predictive screening index for posttraumatic stress

disorder and depression following traumatic injury. *Journal of Consulting and Clinical Psychology*, 76(6), 923.

Priebe, S., Matanov, A., Janković Gavrilović, J., McCrone, P., Ljubotina, D., Knežević, G., … Schützwoh, M. (2009). Consequences of untreated posttraumatic stress disorder following war in former Yugoslavia: Morbidity, subjective quality of life, and care costs. *Croatian Medical Journal*, 50(5), 465–475.

Reifels, L., Bassilios, B., Spittal, M. J., King, K., Fletcher, J., & Pirkis, J. (2015). Patterns and predictors of primary mental health service use following bushfire and flood disasters. *Disaster Medicine and Public Health Preparedness*, 9(3), 275–282.

Reifels, L., Pietrantoni, L., Prati, G., Kim, Y., Kilpatrick, D. G., Dyb, G., … O'Donnell, M. (2013). Lessons learned about psychosocial responses to disaster and mass trauma: An international perspective. *European Journal of Psychotraumatology*, 4(1), 22897.

Roberts, N. P., Kitchiner, N. J., Kenardy, J., & Bisson, J. I. (2009). Systematic review and meta-analysis of multiple-session early interventions following traumatic events. *American Journal of Psychiatry*, 166(3), 293–301.

Rosenbaum, L. (2004). Post-traumatic stress disorder: A challenge for primary care–Misunderstood and incognito. *The British Journal of General Practice*, 54(499), 83.

Shalev, A. Y., Ankri, Y. L., Peleg, T., Israeli-Shalev, Y., & Freedman, S. (2011). Barriers to receiving early care for PTSD: Results from the Jerusalem trauma outreach and prevention study. *Psychiatric Services*, 62(7), 765–773.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Janavs, J., Weiller, E., Keskiner, A., … Dunbar, G. C. (1997). The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. *European Psychiatry*, 12(5), 232–241.

Shultz, J. M., & Forbes, D. (2014). Psychological first aid: Rapid proliferation and the search for evidence. *Disaster Health*, 2(1), 3–12.

Stene, L. E., Wentzel-Larsen, T., & Dyb, G. (2016). Healthcare needs, experiences and satisfaction after terrorism: A longitudinal study of survivors from the Utøya attack. *Frontiers in Psychology*, 7, 1809.

Stuber, J., Galea, S., Boscarino, J. A., & Schlesinger, M. (2006). Was there unmet mental health need after the September 11, 2001 terrorist attacks? *Social Psychiatry and Psychiatric Epidemiology*, 41(3), 230–240.

Te Brake, H., & Dückers, M. (2013). Early psychosocial interventions after disasters, terrorism and other shocking events: Is there a gap between norms and practice in Europe? *European Journal of Psychotraumatology*, 4(1), 19093.

Te Brake, H., Dückers, M., De Vries, M., Van Duin, D., Rooze, M., & Spreeuwenberg, C. (2009). Early psychosocial interventions after disasters, terrorism, and other shocking events: Guideline development. *Nursing & Health Sciences*, 11(4), 336–343.

Uluşahin, A., Başoğlu, M., & Paykel, E. S. (1994). A cross-cultural comparative study of depressive symptoms in British and Turkish clinical samples. *Social Psychiatry and Psychiatric Epidemiology*, 29(1), 31–39.

Van Buuren, S. (2012). *Flexible imputation of missing data*. Boca Raton, FL: CRC Press.

Van de Schoot, R., Sijbrandij, M., Depaoli, S., Winter, S. D., Olff, M., & Van Loey, N. E. (2018). Bayesian PTSD-trajectory analysis with informed priors based on a systematic literature search and expert elicitation. *Multivariate Behavioral Research*, 53(2), 267–291.

Van den Berg, B., Wong, A., van der Velden, P. G., Boshuizen, H. C., & Grievink, L. (2012). Disaster

exposure as a risk factor for mental health problems, eighteen months, four and ten years post-disaster–A longitudinal study. *BMC Psychiatry*, *12*(1), 147.

Van der Velden, P. G., Grievink, L., Kleber, R. J., Drogendijk, A. N., Roskam, A. J. R., Marcelissen, F. G., … Gersons, B. P. (2006). Post-disaster mental health problems and the utilization of mental health services: A four-year longitudinal comparative study. *Administration and Policy in Mental Health and Mental Health Services Research*, *33*(3), 279–288.

Van der Velden, P. G., Loon, P. V., IJzermans, J., & Kleber, R. J. (2006). Psychosociale zorg direct na een ramp [Psychosocial care directly after a disaster]. *De Psycholoog*, *12*, 658–663.

Weathers, F. W., Keane, T. M., & Davidson, J. R. (2001). Clinician-administered PTSD scale: A review of the first ten years of research. *Depression and Anxiety*, *13*(3), 132–156.

Weisæth, L. (2001). Acute posttraumatic stress: Nonacceptance of early intervention. *The Journal of Clinical Psychiatry*, *62*(17), 35–40.

Appendix A. Diagnostic performance of the Trauma Screening Questionnaire (TSQ) and Patient Health Questionnaire-2 (PHQ-2) in predicting any anxiety-, mood- or trauma-related disorder.

Screener	Time-point	Sensitivity	Specificity	Positive predictive power	Negative predictive power	*n*
TSQ[a]	1	.50	.86	.78	.63	28
	2	.53	.94	.90	.67	34
	3	.40	1.00	1.00	.60	38
PHQ-2[b]	1	.27	.87	.67	.54	30
	2	.44	1.00	1.00	.65	33
	3	.30	.89	.75	.53	38

[a] A cut-off of 6 or higher was taken to indicate risk for PTSD. [b] A cut-off of 3 or higher was taken to indicate risk for depression.