Brief screening instrument for post-traumatic stress disorder

CHRIS R. BREWIN, SUZANNA ROSE, BERNICE ANDREWS, JOHN GREEN, PHILIP TATA, CHRIS McEVEDY, STUART TURNER and EDNA B. FOA

Background Brief screening instruments appear to be a viable way of detecting post-traumatic stress disorder (PTSD) but none has yet been adequately validated.

Aims To test and cross-validate a brief instrument that is simple to administer and score.

Method Forty-one survivors of a rail crash were administered a questionnaire, followed by a structured clinical interview 1 week later.

Results Excellent prediction of a PTSD diagnosis was provided by respondents endorsing at least six re-experiencing or arousal symptoms, in any combination. The findings were replicated on data from a previous study of 157 crime victims.

Conclusions Performance of the new measure was equivalent to agreement achieved between two full clinical interviews.

Declaration of interest The study was funded by the NHSE London Regional Office (project no. RDC01702). The views and opinions expressed herein do not necessarily reflect those of the NHSE (LRO) or the Department of Health.

How can those survivors of traumatic events likely to develop post-traumatic stress disorder (PTSD) be most efficiently identified? Recently studies have investigated the performance of 4-, 6- and 12-item screening instruments requiring respondents to rate the frequency and/or severity of some of the 17 major symptoms contributing to a DSM–III–R (American Psychiatric Association, 1987) or DSM–IV (American Psychiatric Association, 1994) diagnosis (Meltzer-Brody *et al*, 1999; Fullerton *et al*, 2000). These have shown promising results equivalent to longer 17-item measures, although none has yet been validated on independent samples. Brewin *et al* (1999) found that victims of violent crime at high risk of developing PTSD 6 months later could be identified by their reports at 3 weeks post-crime of at least three re-experiencing or arousal symptoms. Our aim in this study was to test a brief 10-symptom screening instrument with survivors of a rail crash and then to validate the findings on our crime victim data.

METHOD

Sample 1: rail crash survivors

Participants

Participants had all been passengers on one of two trains that crashed into one another at Ladbroke Grove, London, on 5 October 1999. There were high levels of injury and loss of life, both from the impact and from smoke inhalation. There were three groups of respondents: 18 patients treated at St Mary's Hospital, Paddington; 15 patients treated at the Royal Berkshire Hospital, Reading; and 8 members of a survivors' group set up after the crash. The sample consisted of 21 men and 20 women with a mean age of 38.3 years (s.d.=10.3 years).

Measures

Screening questionnaire. This was designed for trauma victims in general and consisted initially of a single sheet of 16 items. Of these, five were re-experiencing items and five were arousal items taken from the PTSD Symptom Scale–Self Report version (PSS–SR; Foa *et al*, 1993). As in our previous work, the threshold for a positive response was designed to correspond to a rating of 2 on the 0–3 scale employed by the original PSS–SR. The PSS–SR instructions were amended as follows: 'Please consider the following reactions which sometimes occur after a traumatic event. This questionnaire is concerned with your personal reactions to the traumatic event which happened to you. Please indicate whether or not you have experienced any of the following at least twice in the past week.' Respondents ticked either 'Yes' (scored 1) or 'No' (scored 0). A further three items enquired about negative emotions and there were three filler items, but analyses involving these items are not reported. The final 10-item version of the Trauma Screening Questionnaire (TSQ) is given in the Appendix.

Clinician-Administered PTSD Scale (CAPS–1; Blake et al, 1995). This is a well-validated structured clinical interview designed to elicit the frequency and severity of symptoms and to assign a DSM–IV diagnosis. Interrater and test–retest reliability are good (Blake *et al*, 1995). In this study a subsample of 28 CAPS interviews were subjected to independent blind rating, which produced 100% agreement on the presence or absence of a PTSD diagnosis with the interviewer rating.

Procedure

In the course of routine clinical follow-up following their involvement in the train crash, teams from the Brent, Kensington, Chelsea & Westminster Mental Health Trust and the Royal Berkshire Hospital contacted patients by letter, inviting them to take part in a study of the care received following major disasters. The 41 respondents agreeing to take part (18 out of 44 contacted from St Mary's; 15 out of 25 contacted from the Royal Berkshire Hospital; no response data available from the survivors' group) were asked to describe their current reactions to the accident and its aftermath by completing the screening questionnaire. They then gave consent for a second interview conducted approximately 1 week later, during which the CAPS was administered. All interviews, which took

https://doi.org/10.1192/bjp.181.2.158 Published online by Cambridge University Press

place between May and November 2000, were conducted by telephone and were tape-recorded. Patients meeting the criteria for PTSD at the second interview were informed of their diagnostic status and treatment options were discussed with them.

Analysis

The performance of the screening instrument was assessed by reference to two standard criteria: sensitivity (i.e. the probability that someone with a PTSD diagnosis will have tested positive) and specificity (i.e. the probability that someone without a PTSD diagnosis will have tested negative). These criteria are independent of the prevalence of the disorder in the population, and so can be compared readily across studies. In practice, the researcher or clinician generally wants to know the answer to two slightly different questions that are sensitive to population prevalence. What is the probability that someone with a positive test will report a diagnosis of PTSD? What is the probability that someone with a negative test will not receive a PTSD diagnosis? The answers to these questions are given by the positive and negative predictive power of the screening test, respectively. The performance of the test was also expressed in terms of the percentage of cases correctly classified as having or not having PTSD (overall efficiency).

Sample 2: crime victims

Participants

We recruited victims of violent crime from police and hospital sources. The 157 participants who entered the study (118 men, 39 women) had an average age of 35 years (s.d.=13 years). Full details of the sample are given in Rose *et al* (1999).

Measures

Post-traumatic stress symptoms and diagnoses were assessed using the PSS–SR (Foa *et al*, 1993). As before, symptoms were counted as present if they were rated at least 2 on the 0–3 scale. A diagnosis of PTSD was assigned if the DSM–III–R criteria of at least one re-experiencing symptom, three avoidance/numbing symptoms and two arousal symptoms were met. Diagnoses based on the PSS–SR have been shown to be highly concordant with diagnoses based on structured interviews (Foa *et al*, 1993).

Procedure

Participants completed the PSS–SR during interviews that in all cases were conducted within 1 month of the crime (mean 21 days post-crime, range 9–31 days, s.d.=5.6 days). Thus, participants did not fulfil the duration criterion for a DSM–III–R diagnosis of PTSD, although they fulfilled all other criteria.

RESULTS

Sample 1: rail crash survivors

A total of 14 out of the 41 respondents received a CAPS diagnosis of PTSD, which is a prevalence rate of 34%. First we confirmed that a threshold of around three or four re-experiencing and arousal symptoms offered optimum predictive power relative to other possible cut-offs.

Performance at these two cut-offs is presented in Table 1, which shows that either of these thresholds offers overall efficiency of around 80%. Although both thresholds perform similarly, arguably the threshold of three re-experiencing symptoms offers the best balance of sensitivity and specificity. Using a threshold of four re-experiencing symptoms would improve specificity but at a cost of weaker sensitivity. A threshold of three arousal symptoms would again offer relatively more sensitivity, but the threshold of four arousal symptoms offers relatively better specificity and optimum overall efficiency. Table 1 also shows the diagnostic efficiency of requiring respondents to endorse at least six out of the ten re-experiencing or arousal symptoms in any combination. This cut-off maximised overall efficiency and led to a substantial increase in sensitivity and specificity, with excellent positive and negative predictive power.

Sample 2: crime victims

Forty-two respondents received a PSS–SR diagnosis of PTSD, which is a prevalence rate of 26.8%. Performance at the cut-offs of three re-experiencing symptoms and four arousal symptoms again maximised overall efficiency, which ranged from 88% to 92% (Table 2). As with the rail crash survivors, we investigated what was the optimum criterion for endorsing any combination of the ten re-experiencing and arousal symptoms. Once again the optimum cut-off was six symptoms, yielding a comparable level of overall efficiency and excellent positive and negative predictive value.

DISCUSSION

Performance of the Trauma Screening Questionnaire

There are a number of well-established risk factors for PTSD, such as female gender, previous trauma and previous psychiatric disorder, but few of these risk factors are consistent across different types of study and none of them accounts for a sufficient amount of the variance to be practically useful as predictors (Brewin *et al*, 2000). Most potential screening instruments have

Table 1 Sensitivity, specificity and power to predict post-traumatic stress disorder of different symptom combinations in the rail crash sample

Screening criterion	Number meeting criterion	Sensitivity	Specificity	Positive predictive power	Negative predictive power	Overall efficiency
At least three re-experiencing symptoms	15	0.79	0.85	0.73	0.88	0.83
At least four re-experiencing symptoms	9	0.57	0.96	0.89	0.81	0.83
At least three arousal symptoms	21	0.93	0.70	0.62	0.95	0.78
At least four arousal symptoms	11	0.64	0.93	0.82	0.83	0.83
At least six re-experiencing or arousal symptoms in any combination	14	0.86	0.93	0.86	0.93	0.90

https://doi.org/10.1192/bjp.181.2.158 Published online by Cambridge University Press

Table 2 Sensitivity, specificity and power to predict post-traumatic stress disorder of different symptom combinations in the crime victims sample

Screening criterion	Number meeting criterion	Sensitivity	Specificity	Positive predictive power	Negative predictive power	Overall efficiency
At least three re-experiencing symptoms	39	0.74	0.93	0.79	0.91	0.88
At least four re-experiencing symptoms	24	0.50	0.97	0.87	0.84	0.85
At least three arousal symptoms	44	0.86	0.93	0.82	0.95	0.91
At least four arousal symptoms	37	0.79	0.97	0.89	0.92	0.92
At least six re-experiencing or arousal symptoms in any combination	35	0.76	0.97	0.91	0.92	0.92

therefore focused on symptom patterns to identify and predict cases of PTSD. This study represents the first attempt to cross-validate a brief, symptom-based screening instrument for PTSD in independent samples. In both samples a threshold of around three to four re-experiencing or arousal symptoms provided a reasonably sensitive and specific measure of the presence of PTSD. The level of prediction obtained among the rail crash survivors (83% overall efficiency) when symptoms in each cluster were counted separately was lower than we achieved in our previous study, where the overall efficiency was 85–92%. However, by adopting the strategy of allowing respondents to endorse any combination of six or more re-experiencing and arousal symptoms, efficiency was increased to 90%. Using this criterion, equivalent levels of screening performance were obtained despite the samples differing in the type of trauma, the prevalence of PTSD and the time elapsed since the trauma.

This is a striking result for a number of reasons. One reason is that with the rail crash survivors we were capitalising on *post hoc* analyses of the data designed to yield optimum prediction, whereas in re-analysing the crime victim data we were replicating previously established cut-off scores. The second reason is that with the crime victims the items used for prediction were also used in the calculation of PTSD, whereas among the rail crash survivors prediction was kept separate from the diagnosis of PTSD. Finally, the crime victims study utilised a questionnaire assessment of PTSD status, whereas with the rail crash survivors we used a structured clinical interview. Despite these promising findings, it will be important to establish the criterion validity of the instrument in other settings and with other traumas in order to overcome any possible sources of bias in the two samples used.

Comparison with existing instruments

Most existing instruments involve the use of rating scales and decision rules, and contain 17 items or more (see Brewin *et al*, 2002, for a review). They can be scored in two ways, either requiring respondents to meet the diagnostic criteria for PTSD (endorsing at least one re-experiencing symptom, three avoidance/numbing symptoms and two arousal symptoms) or to exceed a cut-off score. Data regarding the performance of such measures have been published by Blanchard *et al* (1996) using the PTSD Checklist, by Foa *et al* (1997) using the Post-traumatic Diagnostic Scale, by Davidson *et al* (1997) using the Davidson Trauma Scale and by Başoğlu *et al* (2001) using the Traumatic Stress Symptom Checklist. The performance of briefer screening instruments has been reported by Meltzer-Brody *et al* (1999) using the 4-item SPAN, and by Fullerton *et al* (2000) using the BPTSD-12 and BPTSD-6. However, the performance of some of these brief measures is probably inflated by the use of *post hoc* cut-off scores and none has yet been adequately validated. When we required the endorsement of at least six re-experiencing or arousal items in any combination, the overall efficiency of the screening instrument in this study was superior to all these measures, of whatever length. Performance was equivalent to that obtained from a comparison of diagnoses yielded by the two most highly regarded interview assessments currently available for PTSD: the Structured Clinical Interview for DSM–IV (SCID; First *et al*, 1996) PTSD module and the CAPS. In a sample of 123 combat veterans, a CAPS total score of 65 was found to have a sensitivity of 0.84 and a specificity of 0.95 relative to a SCID diagnosis (Blake *et al*, 1995).

General considerations in screening for PTSD

It is quite possible that other combinations of symptoms would be as effective as using the re-experiencing and arousal items. Previously it has been claimed that the avoidance and numbing symptom cluster is likely to be most efficient for screening purposes, because it is less common to reach the threshold for these symptoms than it is for the re-experiencing and arousal symptom clusters (e.g. North *et al*, 1999). Our data show that the greater predictive power of the avoidance and numbing cluster is almost certainly due to the fact that more symptoms are required to meet the criterion. If equivalent numbers of re-experiencing or arousal symptoms are required, levels of prediction appear to be just as good. Avoidance and numbing symptoms were not included in our instrument for several reasons. First, there are more of these items, so the length of the instrument would be increased; second, some of the items (e.g. the amnesia and foreshortened future items) are not always well comprehended by respondents.

To be useful, screening instruments ideally should be short and contain the minimum number of items necessary for accurate case identification. They should be simple and preferably not require respondents to ponder over large numbers of alternative scale points. They should be written in a language that is easy to understand. Their purpose should be plain and they should be acceptable to respondents. For ease of administration, self-report questionnaires would appear to be the most flexible solution. If they are to be scored by non-specialists, which would widen their applicability, simple decision rules for determining who passes and fails the screen would be at a premium. Also highly

https://doi.org/10.1192/bjp.181.2.158 Published online by Cambridge University Press

desirable for successful instruments is that they be accurate at detecting both current PTSD and the risk of future PTSD, and that they should work well with different traumas, with different periods of time elapsed post-trauma and with varying prevalence of PTSD.

Our instrument appears to meet most of these criteria. All the items are simple and easy to understand. The use of a clear frequency threshold allied to a Yes/No response format also simplifies matters for respondents, whereas other measures require them to make ratings on four- or five-point scales. Moreover, having a single symptom scale makes our measure extremely practical for use by other health professionals, who may not be familiar with the disorder and with the structure of PTSD symptom clusters. We have shown that among crime and disaster victims excellent levels of prediction can be obtained with as few as ten items, and that enquiring about more PTSD symptoms has little additional value for screening purposes. It should be noted that all these data were collected, on average, 3 weeks post-trauma or later, and our experience is that screening usually should be delayed until this time, because during the initial post-trauma period natural recovery processes are in operation (Brewin, 2001). The next step is to implement the use of the instrument in primary health care or hospital settings in order to demonstrate that it is effective in improving the identification and treatment rates for cases of PTSD.

CLINICAL IMPLICATIONS

- Current post-traumatic stress disorder can be predicted effectively with as few as ten Yes/No questions about re-experiencing and arousal symptoms.
- A useful threshold is to ask whether symptoms have occurred at least twice in the past week.
- To allow for natural recovery processes, screening probably should not be attempted before 3–4 weeks post-trauma.

LIMITATIONS

- The findings should be replicated on larger samples.
- The findings should be replicated on samples experiencing different kinds of trauma.
- It is possible that other combinations of symptoms will prove to be as effective as using the re-experiencing and arousal items.

CHRIS R. BREWIN, PhD, Subdepartment of Clinical Health Psychology, University College London; SUZANNA ROSE, PhD, West Berkshire Traumatic Stress Service; BERNICE ANDREWS, PhD, Royal Holloway, University of London; JOHN GREEN, PhD, PHILIP TATA, MPhil, CHRIS McEVEDY, MRCPsych, Brent, Kensington, Chelsea & Westminster Mental Health Trust, London; STUART TURNER, MD, Camden & Islington Community Health Services NHS Trust, London; EDNA B. FOA, PhD, University of Pennsylvania, USA

Correspondence: Chris R. Brewin, Subdepartment of Clinical Health Psychology, University College London, Gower Street, London WC1E 6BT, UK

(First received 5 July 2001, final revision 2 January 2002, accepted 17 January 2002)

APPENDIX

Trauma Screening Questionnaire (TSQ)

Your own reactions now to the traumatic event

Please consider the following reactions which sometimes occur after a traumatic event. This questionnaire is concerned with your personal reactions to the traumatic event which happened to you. Please indicate (Yes/No) whether or not you have experienced any of the following at least twice in the past week.

1. Upsetting thoughts or memories about the event that have come into your mind against your will
2. Upsetting dreams about the event
3. Acting or feeling as though the event were happening again
4. Feeling upset by reminders of the event
5. Bodily reactions (such as fast heartbeat, stomach churning, sweatiness, dizziness) when reminded of the event
6. Difficulty falling or staying asleep
7. Irritability or outbursts of anger
8. Difficulty concentrating
9. Heightened awareness of potential dangers to yourself and others
10. Being jumpy or being startled at something unexpected

REFERENCES

American Psychiatric Association (1987) *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn, revised) (DSM–III–R). Washington, DC: APA.

___ (1994) *Diagnostic and Statistical Manual of Mental Disorders* (4th edn) (DSM–IV). Washington, DC: APA.

Başoğlu, M., Şalcioğlu, E., Livanou, M., *et al* (2001) A study of the validity of a screening instrument for traumatic stress in earthquake survivors in Turkey. *Journal of Traumatic Stress*, **14**, 491–509.

Blake, D. D., Weathers, F. W., Nagy, L. M., *et al* (1995) The development of a clinician-administered PTSD scale. *Journal of Traumatic Stress*, **8**, 75–90.

Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., *et al* (1996) Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy*, **34**, 669–673.

Brewin, C. R. (2001) Cognitive and emotional reactions to traumatic events: implications for short-term intervention. *Advances in Mind–Body Medicine*, **17**, 163–168.

___, Andrews, B., Rose, S., *et al* (1999) Acute stress disorder and posttraumatic stress disorder in victims of violent crime. *American Journal of Psychiatry*, **156**, 360–366.

___, ___ & Valentine, J. D. (2000) A meta-analysis of risk factors for posttraumatic stress disorder in adults exposed to trauma. *Journal of Consulting and Clinical Psychology*, **68**, 748–766.

___, Rose, S. & Andrews, B. (2002) Screening for posttraumatic stress disorder in civilian populations. In *Reconstructing Early Intervention After Trauma* (eds R. Ørner & U. Schnyder). Oxford: Oxford University Press, in press.

Davidson, J. R. T., Book, S. W., Colket, J. T., *et al* (1997) Assessment of a new self-rating scale for post-traumatic stress disorder. *Psychological Medicine*, **27**, 143–160.

First, M. B., Spitzer, R. L., Gibbon, M., *et al* (1996) *Structured Clinical Interview for DSM–IV Axis I Disorders*. New York: New York State Psychiatric Institute.

https://doi.org/10.1192/bjp.181.2.158 Published online by Cambridge University Press

Foa, E. B., Riggs, D. S., Dancu, C. V., *et al* (1993) Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *Journal of Traumatic Stress*, **6**, 459–473.

—, Cashman, L., Jaycox, L., *et al* (1997) The validation of a self-report measure of posttraumatic stress disorder: the Posttraumatic Diagnostic Scale. *Psychological Assessment*, **9**, 445–451.

Fullerton, C. S., Ursano, R. J., Epstein, R. S., *et al* (2000) Measurement of posttraumatic stress disorder in community samples. *Nordic Journal of Psychiatry*, **54**, 5–12.

Meltzer-Brody, S., Churchill, E. & Davidson, J. R. T. (1999) Derivation of the SPAN, a brief diagnostic screening test for post-traumatic stress disorder. *Psychiatric Research*, **88**, 63–70.

North, C. S., Nixon, S.J., Shariat, S., *et al* (1999) Psychiatric disorders among survivors of the Oklahoma City bombing. *Journal of the American Medical Association*, **282**, 755–762.

Rose, S., Brewin, C. R., Andrews, B., *et al* (1999) A randomised trial of psychological debriefing for victims of violent crime. *Psychological Medicine*, **29**, 793–799.

https://doi.org/10.1192/bjp.181.2.158 Published online by Cambridge University Press