**A pilot evaluation of a high-intensity treatment pathway for prolonged grief reactions in a Devon NHS Talking Therapies Service (Goff et al): Supporting Online Materials)**

Detailed overview of measures

Reflecting standard NHS-TT practice, levels of anxiety, depression and associated functional impairment were assessed at each session attended with the standard NHS-TT minimum data set.

The Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001) indexed depression severity, with each of nine (self-report) items describing common depression symptoms. How frequently participants have experienced each symptom over the past two weeks is rated on a scale from 0 (not at all) to 3 (nearly every day). The scores are summed, leading to a range from 0-27, with higher scores indicating greater severity of depression. A score of > 10 indicates clinically significant depression and a change of > 6 points indicates reliable improvement/deterioration. Studies indicate satisfactory internal consistency (Cronbach’s α of 0.86 – 0.89; Kroenke et al., 2001).

The Generalized Anxiety Disorder Scale (GAD-7; Spitzer et al., 2006) indexed symptoms of anxiety, with each of seven (self-report) items describing common symptoms of anxiety (and rated on the same frequency scale as the PHQ-9). Items are summed, with the scale ranging from 0-21 (higher scores indicating greater anxiety severity). Scores > 8 indicate clinically significant anxiety and a change of > 4 points indicates reliable improvement/deterioration. The measure has satisfactory internal consistency (α = 0.92; Spitzer et al., 2006).

The Work and Social Adjustment Scale (WSAS; Mundt et al., 2002) consists of five (self-report) items rating degree of impairment linked to psychopathology in work, home management, social leisure, private leisure, and relationship domains. Each item is rated on a scale of impairment from 0 (not at all impaired) to 8 (very severely impaired). No time period over which to rate the item is specified. Items are summed, leading to a scale range of 0 to 40, with higher scores indicating greater functional impairment. Scores < 10 indicate little or no significant impairment and are associated with subclinical populations, scores between 10 and 19 indicate significant impairment but less severe clinical symptomatology, and scores > 20 indicate more severe impairment associated with moderately severe or worse clinical symptomatology. The measure has adequate internal consistency when used in an NHS-TT setting (α = .82; Zahra et al., 2014). There is no agreed threshold for reliable change. Therefore, we estimated the criterion for reliable improvement/deterioration based on the reliability and intake standard deviation of a similar NHS-TT sample reported in Zahra and colleagues (2014). This resulted in a change of >11 points being set as the threshold.

The NHS-TT minimum dataset was optionally supplemented with the Brief Grief Questionnaire (BGQ; Shear et al., 2006) at assessment and final treatment session. This ask five questions about grief reactions to bereavement (e.g., “How much does your grief still interfere with your life?”), each rated on a scale from 0 (not at all) to 2(a lot). No time period over which to rate the item is specified. Scores are summed, with higher scores indicating a more marked grief reaction. The scale has adequate internal reliability in a primary care setting (α = .82; Patel et al., 2019). Patel et al (2019) demonstrated that scores of 4 or greater on the BGQ correspond with a score of 30 or more on the longer Inventory of Complicated Grief (Prigerson et al. 1995), which in turn is highly predictive of meeting ICD-11 diagnostic criteria for PGD (Mauro, unpublished data provided to DSM-5 committee). Therefore, in the present study we use a BGQ score of > 4 to assess pathway eligibility (and a score of < 4 was used to indicate clinical remission post treatment). Further, we divide PGD into lower severity (BGQ 4-7) and higher severity (BGQ > 8) categories. There is no agreed threshold for reliable change on the BGQ, so we estimated this based on the reliability reported in Patel and colleagues (2019) and the intake standard deviation of the present sample who had been through the NHS-TT grief pathway. This resulted in a change of > 3 points being set as the threshold. The BGQ was an optional not mandated measure due to operational challenges adding it to the electronic outcome measure capture system used in service. The company who run the system were only able to add it as ‘clinician delivered’ measure rather than ‘web form deployable’ measure, meaning it could not routinely be sent out to clients to complete online with the other minimum data set measure and instead clinicians had to manually administer and enter scores. This significantly increases the workforce burden and session time associated with completing the measure, so for pragmatic reasons this was an optional not mandatory measure.

References

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606-613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>

Mundt, J. C., Marks, I. M., Shear, K. M., & Griest, J. H. (2002). The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *British Journal of Psychiatry, 180*, 461-464. <https://doi.org/10.1192/bjp.180.5.461>

Patel, S. R., Cole, A., Little, V., Skritskaya, N. A., Lever, E., Dixon, L. B., & Shear, K. M. (2019). Acceptability, feasibility and outcome of a screening programme for complicated grief in integrated primary and behavioural health care clinics. *Family Practice, 36*(2), 125-131. <https://doi.org/10.1093/fampra/cmy050>

Shear, K. M., Jackson, C. T., Essock, S. M., Donahue, S. A., & Felton, C. J. (2006). Screening for complicated grief among Project Liberty service recipients 18 months after September 11, 2001. *Psychiatric Services, 57*(9), 1291-1297. <https://doi.org/10.1176/ps.2006.57.9.1291>

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*(10), 1092-1097. <https://doi.org/10.1001/archinte.166.10.1092>

Zahra, D., Qureshi, A., Henley, W., Taylor, R., Quinn, C., Pooler, J., Hardy, G., Newbold, A., & Byng, R. (2014). The work and social adjustment scale: reliability, sensitivity and value. *International Journal of Psychiatry in Clinical Practice, 18*(2), 131-138. <https://doi.org/10.3109/13651501.2014.894072>

Supplementary Tables

Table S1

Therapist characteristics (N=17)

|  |  |
| --- | --- |
| Variable | Summary |
| Age | 44.75 (9.81), 32 to 61 |
| Gender | 13 female, 4 male |
| Years experience as a high intensity therapist | 10.0 (5.83), 2 to 21 |
| Year mental health experience prior to high intensity training | 11.19 (6.17), 4 to 27 |
| Primary clinical credentials | All high intensity trained; in addition 5 were occupational therapists; 6 were nurses; one was a social worker. |
| Clients treated in audit | 5.71 (5.02), 1 to 20 |

Note: for age, years experience, and clients treated, data are mean (SD), range.

Table S2

Clinical and demographic characteristics of clients in hgiehr severity PGD subgroup versus remaining clients

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Higher severity PGD subgroup | Remaining participants | Comparison |
| *N* | 40 | 40 | - |
| Age | 50.25 (14.43) | 52.60 (14.96) | t<1 |
| Female Gender | 30/40 (75%) | 30/40 (75%) | χ2<1 |
| Taking psychotropic medication | 17/40 (43%) | 14/40 (35%) | χ2<1 |
| Time since loss (months)a | 69.90 (100.78) | 92.77 (133.28) | t<1 |
| PHQ-9 depression | 15.58 (4.08) | 14.80 (4.44) | t<1 |
| GAD-7 anxiety | 14.65 (4.48) | 11.23 (4.81) | t=3.30, p=.001 |
| WSAS functioning | 20.73 (8.06) | 17.85 (7.73) | t=1.63, p=.11 |
| BGQ grief b | 8.85 (0.80) | 5.94 (1.39) | t=10.06, p<.001 |
| PHQ-9 caseness | 36/40 (90%) | 37/40 (93%) | χ2<1 |
| GAD-7 caseness | 38/40 (95%) | 30/40 (75%) | χ2=6.28, p=.01 |
| WSAS caseness | 37/40 (93%) | 36/40 (90%) | χ2<1 |
| PTSD status | 8/40 (20%) | 7/40 (18%) | χ2<1 |
| Sessions attended | 11.3 (5.55) | 9.20 (5.94) | t=1.69, p=.09 |
| Min adequate dose | 36/40 (90%) | 30/40 (75%) | χ2=3.12, p=.08 |

*Note:* continuous variables are mean (one SD) values; categorical variables are number (%) values; higher severity sub sample = those with BGQ data and scoring at least 8 on it; continuous variables compared with independent sample t-tests; categorical variables compare with chi-squared tests; a time since death not recorded for one client; b BGQ completed by 40/40 participants in higher severity PGD subgroup, but only 18/40 in remaining subgroup.

Table S3

Clinical Outcome Data for Participants with and without PTSD

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | PHQ-9 depression | GAD-7 anxiety | WSAS functioning | BGQ grief |
| *PTSD subsample* |  |  |  |  |
| *N* with complete dataa | 14 | 14 | 14 | 6 |
| First session | 15.50 (4.03) | 13.57 (3.88) | 20.64 (8.04) | 8.67 (1.37) |
| Last session | 6.29 (4.29) | 6.57 (5.19) | 13.07 (9.03) | 4.00 (2.61) |
| Change | 9.21 (4.98) | 7.00 (6.37) | 7.57 (7.77) | 4.67 (2.42) |
| t-test | 6.02, p<.001 | 4.11, p<.01 | 3.65, p<.01 | 4.72, p<.01 |
| Cohen’s d | 1.86 (0.96, 2.72) | 1.10 (0.41, 1.76) | 0.97 (0.32, 1.60) | 1.93 (0.50, 3.31) |
| Reliable improvement | 11/14 (79%) | 10/14 (71%) | 6/14 (43%) | 5/6 (80%) |
| Reliable deterioration | 0/14 (0%) | 0/14(0%) | 0/14(0%) | 0/6(0%) |
| Recovery | 11/14 (79%) | 9/14 (64%) | 4/14 (29%) | 3/6 (50%) |
| Response (50% reduction in symptoms) | 11/14 (79%) | 8/14 (57%) | 7/14 (50%) | 3/6 (50%) |
| *Remaining participants* |  |  |  |  |
| *N* with complete datab | 65 | 65 | 65 | 42 |
| First session | 15.17 (4.34) | 12.94 (5.06) | 19.11 (8.01) | 7.81 (1.81) |
| Last session | 6.32 (5.82) | 5.98 (5.56) | 9.46 (9.21) | 3.05 (2.48) |
| Change | 8.85 (5.69) | 6.95 (5.48) | 9.65 (8.91) | 4.76 (2.85) |
| t-test | 12.54, p<.001 | 10.24, p<.001 | 8.73, p<.001 | 10.82, p<.001 |
| Cohen’s d | 1.56 (1.19, 1.92) | 1.27 (0.94, 1.60) | 1.08 (0.77, 1.39) | 1.67 (1.20, 2.14) |
| Reliable improvement | 48/65 (74%) | 49/65 (75%) | 26/65 (40%) | 32/42 (76%) |
| Reliable deterioration | 1/65 (2%) | 2/65 (3%) | 1/65 (2%) | 1/42 (2%) |
| Recovery | 50/65 (77%) | 49/65 (75%) | 40/65 (62%) | 27/42 (64%) |
| Response (50% reduction in symptoms) | 43/65 (66%) | 43/65 (66%) | 35/65 (54%) | 27/42 (64%) |

*Note:* continuous variables are mean (one SD) values; categorical variables are number (%) values; Cohen’s d effect sizes have 95% confidence interval in parentheses; a in the PTSD subsample, only 6/14 individuals had first and last treatment session BGQ scores; b in the remaining subsample of 66, one individual only attended a single session, so had no post data for PHQ-9, GAD-7 and WSAS and 42 participants had first and last session BGQ scores. A series of chi-squared tests (for binary variables) and repeated measures ANOVAs examined if there were any significant differences between the subgroups. The only significant difference was that WSAS recovery rates were lower for the PTSD sample than the remaining sample, p=.02.