Does emotional disclosure have any effects? A systematic review of the literature with meta-analyses

Catherine Meads, Arie Nouwen
University of Birmingham

Objectives: Emotional disclosure has been widely publicized as having beneficial effects on physical and psychological health. A full systematic review was undertaken, with standard health technology appraisal methods, with the aim to assess the effects of emotional disclosure on healthy participants and those with pre-existing morbidity, particularly on longer-term physical health, performance, and psychological outcomes.
Methods: Randomized controlled trials of emotional disclosure were obtained from database searches (Medline (1966–2003), Embase (1980–2003), Cochrane Library (2002, issue 4), Web of Science (1981–2003), Cinahl (1982–2003), and Theses (March 2003), Internet sites (including Professor J.W. Pennebaker's home pages), and personal contacts. Quality was assessed qualitatively and by Jadad score. Meta-analysis was conducted, using Revman 4.1 software, where more than two trials reported the same outcome.
Results: Sixty-one trials were found meeting the inclusion criteria. Most had less than 100 participants and the median Jadad score was 0. A wide variety of physical, physiological, immunological, performance, and psychological outcomes were measured, but fewer were reported. There was no clear improvement for emotional disclosure compared with controls in objectively measured physical health and most other outcomes assessed.
Conclusions: The opinion that this intervention is beneficial needs to be reassessed in light of the totality of evidence available.

Keywords: Systematic review, Meta-analysis, Emotional disclosure, Health technology assessment

Emotional disclosure is a technique whereby people are encouraged to write about a traumatic, stressful, or upsetting event, usually from their recent or distant past and encouraged to go into as much detail about their feelings surrounding the event as possible. They write (or talk into a tape recorder) in private for 15–30 minutes typically for 3–4 days within a relatively short time period such as consecutive days or within 2 weeks. The intervention has become established in the research literature and a relatively standard form was developed that is used in several randomized, controlled trials (RCTs), with the control group being no treatment, waiting list, or written control. It is seen as similar to psychotherapy or counseling in that people confront a distressing memory, label it,

and discuss its causes and consequences (54). Nondisclosure of important psychological experiences is believed to be a form of inhibition acting like a long-term low-level stressor, which could cause or exacerbate psychosomatic processes leading to illness or ineffective functioning (56). The underlying assumption is that the disclosure of emotional-laden material lowers stress levels (5) and subsequently enhances physical health (80). Emotional disclosure is also hypothesized to help people reorganize their thoughts and feelings about traumatic experiences, allowing better understanding of the traumatic event(s) in their lives (62). However, in contrast to traditional psychotherapy, emotional disclosure lacks an interpersonal element (16). Some query the benefit

of revisiting painful emotions without forming new insights (41;81), whereas others query whether this intervention has any effects at all (25).

A meta-analysis on this intervention was published in 1998 (74), combining the results of thirteen emotional disclosure trials. It looked at categories of reported health, psychological well-being, physiological functioning, general functioning, and health behaviors and found that health was enhanced in the first four categories, but health behaviors were not influenced. It reported an overall effect size of d = 0.47 and an illness rate reduction from 61 percent in the control group to 38 percent in the experimental group (74). Since this meta-analysis was conducted, many more RCTs have been published, therefore, it is now difficult to gain an accurate picture of the beneficial effects of emotional disclosure. The aim of this systematic review is to assess the effects of the standard emotional disclosure intervention on healthy participants, people under psychological stress and people with pre-existing physical morbidity, with the focus on longer-term physical health, performance, and psychological outcomes.

METHODS

A scoping review of the published literature was made to develop an effective search strategy. The subsequent inclusion and exclusion criteria used are summarized in Table 1.

A wide variety of sources were searched for published and gray literature, including the following:

- Bibliographic databases: Cochrane Central Register of Controlled Trials and Database of Abstracts of Reviews of Effectiveness (2002, Issue 4), Medline (Ovid; 1966–Feb 2003), Embase (1980–Feb 2003) Cinahl (1982–Feb 2003), Science Citation Index (Web of Science) (1981–Feb 2003), and ISSI (Mar 2003), FRANCIS (Mar 2003), Index to Theses (Mar 2003) and UMI Proquest digital dissertations (Mar 2003) databases
- Citation search on Pennebaker J.W. in BIDS ISI (Mar 2003)
- General search of Internet sites using Google (July 2002) and Scirus (Aug 2002) search engines using the search term emotional disclosure. The first 100 references on each were checked.
- The list of emotional disclosure trials on J.W. Pennebaker's Web site (July 2002; 55)
- Hand search of relevant journals from 2002–2003: Behavior Therapy, Behavior Research and Therapy, Health Psychology, Psychology and Health, Psychology, Health and Medicine, Psychosomatic Medicine and the Abstracts from the 61st Annual Scientific Meeting, American Psychosomatic Society, Phoenix, AZ, March 5–8, 2003.
- Contact with some emotional disclosure researchers and other interested academics
- Citations checked in reviews and trials identified by the searches.

For Medline, Embase, and Cinahl, search terms included emotions, catharsis, health status, emotional disclosure, emotional expression, rehearsal, self-disclosure, writing, and journal; for the other databases, emotion or emotion and disclosure.

One reviewer (C.M.), using explicit predetermined criteria, made inclusion and exclusion decisions independently of trial results and extracted effectiveness and quality assessment data onto predefined forms. A second reviewer (A.N.) extracted effectiveness data from six trials. The results were compared, and no discrepancies were found. If trials had more than one intervention group, the one that was most similar to the basic emotional disclosure intervention as in the single intervention group trials was used, for example real trauma rather than imaginary trauma (24). Written control was used in preference to nonwritten control in cases where both were present.

Quality of trials was assessed qualitatively, by Jadad score (30) and the presence or absence of a CONSORT-style flow diagram or information to construct one (48). Publication bias was assessed by funnel plot using the most frequently reported outcome of objective health center or general practitioner (GP) visits (health center visit, HCV; see Table 2 for a list of acronyms used in this report) plotted against study size. No attempt was made to impute missing data from statistics such as p values or Cohen's d. A level of statistical significance of $p < .05$ was used throughout. Meta-analysis using Review Manager software version 4.1 and Metaview 4.1 was carried out if more than two trials reported the same outcomes. Weighted mean difference (WMD) was used where variables were assessed with the same measure, such as HCV or the Impact of Events Scale (IES). Standardized mean difference (SMD) was used where there were different measures for the same outcome (e.g., symptoms of depression or anxiety). Meta-analysis for fixed effects used the inverse variance method and DerSimonian and Laird for random effects models (7).

RESULTS

Number of Studies Identified

Database searches found 1,194 references of which 347 were duplicates. A first trawl gave 148 potentially relevant studies. Sixty-one were included, seventy-two excluded, and brief details obtained on fifteen that may have been included if more details had been available. The main reasons for exclusion were verbal emotional disclosure in front of a listener and being nonrandomized studies. Six RCTs were excluded because they presented combined intervention and control results only or subgroup analyses only.

The sixty-one RCTs included were separated into three main categories:

- Pre-existing physical conditions such as headaches (9), fibromyalgia (22), rheumatoid arthritis (31;47), with asthma (77),

Table 1. Inclusion Criteria

Criteria	Included	Excluded
Study design	a. RCTs only for longer-term outcomes (follow-up 1 category below). b. RCTs or randomized crossover trials for immediate physiological or immunological outcomes.	Non-randomized studies.(Within-subject or crossover trials allowed for physiological and immune outcomes only).
Population	Any (i.e., healthy or with physical illness or psychological problem or both).	Actors.
Intervention	Emotional disclosure, which can be written or verbal. Any element of written included. If verbal only, must be done without a listener present, i.e., into a tape recorder or similar. No time limit on the length of disclosure.	Verbal emotional disclosure in the presence of a listener (e.g., counselor, psychotherapist, therapist or doctor). Counseling or psychotherapy, expressive dance, film, hypnosis.
Control	a. Either written or verbal nonemotional or fact-based activity for the same time in same modality (written, verbal) as intervention. If verbal, must be done without a listener present as above. b. Nonintervention control (do nothing). c. Waiting list control.	RCTs where the intention or expectation is that the control group may have an effect, e.g., RCTs with one written and one verbal disclosure group only, control groups as expressive dance, positive event disclosure, or relaxation therapy.
Outcome measures	Objective or subjectively measured health center visits or other physical health outcomes, psychological health, performance, physiological or immune system outcomes.	Trials reporting psychological outcomes only during the intervention period only.
Follow-up A (physical and psychological health, performance measures)	Minimum 1 week following the end of the intervention phase of the trial. No maximum time limit for follow-up.	Studies presenting baseline characteristics only, with no follow-up reported Studies. presenting combined intervention and control group results only. Studies presenting no results comparing intervention and control groups. Studies presenting subgroup analyses only.
Follow-up B (physiological and immune measures)	During intervention or any follow-up length.	Follow-up longer than 1 week for crossover trials.

RCT, Randomized controlled trial.

"symptomatic," (42;43), HIV (44), cancer (10;11;69;82;87), and rehabilitation after anterior cruciate ligament reconstruction (85). – thirteen trials

Table 2. Acronyms Used

CAT	College Adjustment Test
GPA	Grade point average
HCV	Health center visits
IES, IES-R	Impact of Events Scale, Impact of Events Scale–revised
MAACL-R	Multiple Affect Adjective Checklist–Revised
PILL	Pennebaker's Inventory of Limbic Languidness
RCT	Randomized controlled trial
SCL-90, SCL-90-R	Symptom Checklist-90, Symptom Check list-90–revised
SD	Standard deviation
SDS	Zung Self-rating Depression Scale
SMD	Standardized mean difference
SMU-HQ	Southern Methodist University–Health Questionnaire
WMD	Weighted mean difference

- Psychosocial stressors such as having a baby in an intensive care unit (2), child sexual abuse (3), posttraumatic stress disorder (21;73), frequent clinic visitors (20), severe trauma (24;70;71), lost loved one (38;66), taking professional exams (39), relationship breakup (40), high distress after hurricane flooding (51), prisoners (68), unemployed (79), and bereaved (84). – eighteen trials
- Healthy volunteers – Twenty-eight RCTs (4;6;8;13–15;17;23;28;32–34;36;37;45;49;50;57–59;61;63–65;67;72;76;86; including one on children, 67) and two randomized crossover trials (8;60)

Five RCTs (4;13;15;35) and the randomized crossover trials (8;60) measured physiological or immunological outcomes only and will not be discussed further here (46).

For the included RCTs, the time of intervention varied between one episode of 20 minutes and five episodes of 45 minutes, with a median of 60 minutes total writing time. Follow-up time varied between 17 days and 15 months but was mostly between 6 and 12 months. Intervention and control group numbers varied between 5 and 79. Numbers

in each group were not reported in seven RCTs (2;36;37; 39;40;42;76).

Quality of Included Randomized Controlled Trials

The median Jadad score was 0. The method of randomization was given in six RCTs (17;31;59;77;82;85), plus one used minimization (10;11). Allocation concealment was mentioned in five (3;69;72;77;82), and some element of blinding in seventeen (3;20;21;24;32;33;36;38;43;57;58;61;69;72; 77;82). It was frequently unclear who was being blinded (investigators or outcome assessors—participants could not be blinded) and how successful the blinding was. There was a power calculation in four RCTs (24;29;77;82) and explicit intention to treat analysis in one (77). This RCT also gave the only CONSORT-style diagram. The attempt to construct a CONSORT-style diagram from details given in the other RCTs had varying amounts of success. In each group, the number of participants who were randomized was present in thirty RCTs (3;9–11;23;31;36;38;43;44;47;50;51;58;66– 72;76;77;79;82;84–87), those who received allocated intervention in thirty-one (10;11;14;20;21;23;24;29;31–33;36; 38;45;49;57;58;61;64;67;69;71;72;77;79;82;84;85;87), and those who were followed up for any of the follow-up measures was reported in thirty-three (3;6;10;11;17;20;23; 29;31–34;36;38;43;44;49;51;57–59;65–69;72;76;77;79;82; 84–87). The percentage of participants lost to follow-up could be calculated in a different thirty-three RCTs (6;9– 11;14;17;22;23;31;36;40;43;44;51;58;66–70;72;76;77;79; 82;84;86;87), of these, fifteen RCTs had 20 percent or more participants lost to follow-up (9;14;22;38;51;59;65;66; 69;70;72;76;84;86;87) and one had 59 percent (72). Six of the RCTs had considerable imbalance in losses to follow-up (44;51;72;76;84;87), usually losing more intervention group participants than controls. The funnel plot did have some evidence of asymmetry, which may suggest some publication bias, but there were too few data points to have any degree of certainty.

A wide variety of outcomes were measured, but fewer were reported. As a sample, Table 3 shows these measures for the 13 RCTs in people with pre-existing physical conditions.

Findings from Included Studies

The RCT findings are ordered by outcome measure, then for each outcome by the three categories of participants.

Objective Health Measures

In three RCTs with participants with rheumatoid arthritis, two found no significant differences (31;47), the third found a significant improvement in disease state for the intervention group compared with control in rheumatoid arthritis and asthma (77). The RCT measuring disease stage in prostate cancer did not report the results (69). The RCT of reha-

bilitation after anterior cruciate ligament injury (17) found a significant improvement for the intervention group for the number of step ups the person could manage but not for range of motion (extension and flexion) and a measure of how much the participant used a relevant muscle for knee stability (85). One RCT on healthy participants measured illness-related absences, but these were not reported separately from total absences, including annual leave (17).

Health Center Visits

Sixteen RCTs measured objective HCV (20;23;24;29;32– 34;37;42;49;57–59;61;68;84), and it was fully reported in 11 (i.e., gave means and standard deviations (SDs) or medians and ranges; see Figure 1).

The summary WMD for the ten RCTs giving means and SDs was −0.09 (95 percent confidence interval [CI], 4−0.19 to +0.02) fixed effects and −0.06 (−0.26 to +0.13) random effects, suggesting that there is no difference in objectively measured HCV overall for the intervention group compared with control at follow-up. There was significant heterogeneity, which may be partly explained by the single RCT in people with pre-existing physical morbidity finding significantly more HCVs for the intervention group whereas in the other two subgroups the trend was in the opposite direction. In healthy volunteers, eleven RCTs measured but only five gave means and SDs (see Figure 1). The summary WMD in this subgroup was −0.11 (95 percent CI, −0.22 to 0.00) fixed and random effects, suggesting that there may be a decrease in HCV for the intervention groups. Of the six RCTs not giving sufficient details for the meta-analysis, three showed significantly fewer HCV (32;58;59), one more (34), and two no significant differences (49;57). The RCT showing more service use (34) reported medians and ranges and had considerable baseline imbalance, as the intervention group used twice as many services as control. This finding was not apparent in the other RCTs. For the four RCTs that reported means but no SDs (49;57–59), mean SDs from the fully reported RCTs (of SD = 0.5) were inserted for intervention and control groups. The effect of this was to change the summary WMD to −0.2 (95 percent CI, −0.29 to −0.12) fixed effects and −0.21 (95 percent CI, −0.35 to −0.07) random effects, suggesting that, if results for these RCTs had been available, the meta-analysis may have shown a clearer trend for the intervention group for healthy participants.

Subjective Health Measures

Sixteen RCTs measured subjective HCV (3;6;14;20;22;23; 37;38;45;49;65;66;69;72;82) and ten gave sufficient results for meta-analysis (3;6;20;37;38;65;66;69;82). The summary WMD was −0.95 (95 percent CI, −1.11 to −0.78) fixed and −0.55 (−1.13 to +0.03) random effects. The results suggest fewer HCV for the intervention, but there was considerable heterogeneity. Also, examining the RCTs that gave insufficient detail for meta-analysis, five showed no

Table 3. Outcome Measures of the Pre-existing Physical Condition RCTs

First author	Outcomes measured (outcomes in brackets not reported)
D'Souza et al. (9)	Headache pain, headache frequency, McGill Pain Inventory short form, Migraine Disability Assessment Scale, days using pain medications in previous month, (SCL-90-R somatization scale).
Gillis et al. (22)	Self-report HCV, Fibromyalgia Impact Questionnaire, Fatigue Severity Scale, medication, sleep quality, Positive And Negative Affect Schedule-X, support.
Kelley et al. (31)	American Rheumatism Association Joint Condition including joint count, grip strength and walking time. Arthritis Impact Measurement Scales-2 including physical dysfunction, pain, affective disturbance.
Lumley et al. (42)	Objective HCV, 'health,' physical symptoms, health interference with daily functioning.
Lumley & Provenzano (43)	Credit hours attempted and earned, GPA.
Mann (44)	Compliance, treatment side effects, Life Orientation Test.
Meyer et al. (47)	Joint status, walking speed, grip strength, over the counter medications, Arthritis Impact Measurement Scales-2 physical functioning, pain, fatigue, emotional functioning.
de Moor (10;11)	(Brief Symptom Inventory) Centre for Epidemiological Studies-Depression Scale, IES, Profile of Mood States, Pittsburgh Sleep Quality Index, Perceived Stress Scale.
Rosenberg et al. (69)	Self-report HCV, health care use, use of medicines, pain, (disease stage, health behaviors, physical symptoms), Psychological symptoms, (Functional Assessment of Cancer Therapy Scale, Profile of Mood States, Short form-36, SCL-90-R).
Smyth et al. (77)	Forced expiratory volume at 1 minute, physician's global assessment of rheumatoid arthritis.
Stanton et al. (82)	Self-report HCV, perceived somatic symptoms, Coping questionnaire, Functional Assessment of Cancer Therapy Scale, IES, Profile of Mood States.
Strough (85)	Range of motion, strength, biofeedback, subjective knee rating, Profile of Mood States, rumination, self-efficacy.
Walker et al. (87)	(Side effect severity), IES, Positive And Negative Affect Schedule.

For abbreviations, see Table 2.

significant difference between intervention and control groups (14;22;45;49;72) and one was not reported (23). This finding suggests that the ones that did give summary statistics may have been a biased sample.

A wide variety of subjective health measures were used but fewer reported and most showed no significant differences between intervention and control groups, for example medication compliance in HIV (44), Southern Methodist University Health Questionnaire (88) (SMU-HQ (23;79) and health behaviors (29;57;58;61;64;69). Six RCTs (3;21;37;57;68;79) used the Pennebaker Inventory of Limbic Languidness (PILL) questionnaire (52) but one did not report their results (79). Meta-analysis of the four RCTs with sufficient information, (3;21;37;68) using WMD gave 4.97 (2.16 to 7.78) fixed and 3.27 (−3.43 to 9.96) random effects. This finding suggests that, at follow-up, there is more reported illness for the intervention group. Regarding pain, of the five RCTs to measure this, two reported less pain for the intervention group (9;69), and three RCTs showed no significant difference (22;31;47).

For RCTs in people with pre-existing physical morbidity, twenty-seven outcomes were measured and twenty-one reported. Of the reported outcomes, twelve showed no difference between intervention and control groups or no relevant statistical test was given. There were improvements for the intervention groups compared with controls for sleep quality (22), fibromyalgia impact (22), rheumatoid arthritis physical dysfunction (31), generalized physical symptoms (42), health interference with daily functioning (42), and perceived somatic symptoms (82). None of the RCTs demonstrated worse subjective physical health. For psychosocial stress RCTs, one split SMU-HQ physical symptom scales and reported results and statistical tests on each subscale separately (23). The remaining five outcomes showed no significant differences. In healthy volunteers, of twenty-two outcomes measured, four were not reported and thirteen showed either no difference between intervention and control groups or no relevant statistical test was reported. The remaining five outcomes provided conflicting results. One RCT showed fewer physical symptoms for the intervention group (67), whereas another showed more symptoms and more days off due to illness (72), a third showed more symptom severity (86), and the fourth RCT showed less activity restriction from illness for the intervention groups (78).
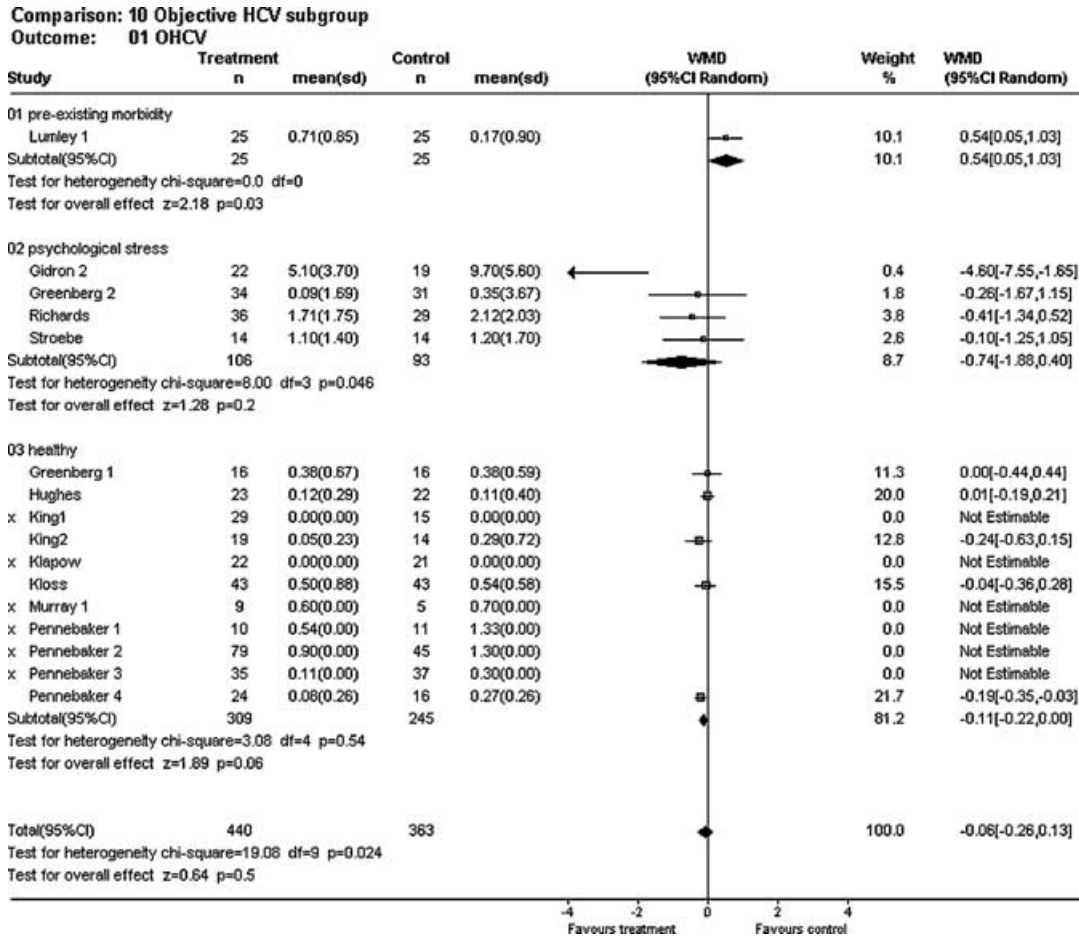
Comparison: 10 Objective HCV subgroup
Outcome: 01 OHCV

Study	Treatment n	mean(sd)	Control n	mean(sd)	WMD (95%CI Random)	Weight %	WMD (95%CI Random)
01 pre-existing morbidity							
Lumley 1	25	0.71(0.85)	25	0.17(0.90)		10.1	0.54[0.05,1.03]
Subtotal(95%CI)	25		25			10.1	0.54[0.05,1.03]
Test for heterogeneity chi-square=0.0 df=0							
Test for overall effect z=2.18 p=0.03							
02 psychological stress							
Gidron 2	22	5.10(3.70)	19	9.70(5.60)		0.4	-4.60[-7.55,-1.85]
Greenberg 2	34	0.09(1.69)	31	0.35(3.67)		1.8	-0.26[-1.67,1.15]
Richards	36	1.71(1.75)	29	2.12(2.03)		3.8	-0.41[-1.34,0.52]
Stroebe	14	1.10(1.40)	14	1.20(1.70)		2.6	-0.10[-1.25,1.05]
Subtotal(95%CI)	106		93			8.7	-0.74[-1.88,0.40]
Test for heterogeneity chi-square=8.00 df=3 p=0.046							
Test for overall effect z=1.28 p=0.2							
03 healthy							
Greenberg 1	16	0.38(0.67)	16	0.38(0.59)		11.3	0.00[-0.44,0.44]
Hughes	23	0.12(0.29)	22	0.11(0.40)		20.0	0.01[-0.19,0.21]
x King1	29	0.00(0.00)	15	0.00(0.00)		0.0	Not Estimable
King2	19	0.05(0.23)	14	0.29(0.72)		12.8	-0.24[-0.63,0.15]
x Klapow	22	0.00(0.00)	21	0.00(0.00)		0.0	Not Estimable
Kloss	43	0.50(0.88)	43	0.54(0.58)		15.5	-0.04[-0.36,0.28]
x Murray 1	9	0.60(0.00)	5	0.70(0.00)		0.0	Not Estimable
x Pennebaker 1	10	0.54(0.00)	11	1.33(0.00)		0.0	Not Estimable
x Pennebaker 2	79	0.90(0.00)	45	1.30(0.00)		0.0	Not Estimable
x Pennebaker 3	35	0.11(0.00)	37	0.30(0.00)		0.0	Not Estimable
Pennebaker 4	24	0.08(0.26)	16	0.27(0.26)		21.7	-0.19[-0.35,-0.03]
Subtotal(95%CI)	309		245			81.2	-0.11[-0.22,0.00]
Test for heterogeneity chi-square=3.08 df=4 p=0.54							
Test for overall effect z=1.89 p=0.06							
Total(95%CI)	440		363			100.0	-0.06[-0.26,0.13]
Test for heterogeneity chi-square=19.08 df=9 p=0.024							
Test for overall effect z=0.64 p=0.5							

-4   -2   0   2   4
Favours treatment   Favours control

Figure 1. Forest plot of objective health center visits.

Performance

These scores were mostly measured only in healthy volunteers, and the types of outcome used reflected that most participants were college students. Six RCTs reported grade point average (GPA; 6;36;43;58;59), two RCTs showed higher scores for the intervention group (6;43), and four showed no significant differences. Full results were available for two RCTs (6;43), but two others gave means but not SDs (58;59). For these, mean SDs from the fully reported RCTs (of SD = 0.8) were inserted for intervention and control groups. This had very little effect on the overall result, suggesting that if results for these RCTs had been available, the meta-analysis would have shown a clear difference in GPA (see Figure 2).

Two RCTs to record Scholastic Aptitude Tests did not report the results (58;59). Absences from school or work were not different in the two RCTs to measure this (17;67). The one RCT to measure subsequent employment in a group of unemployed participants found that more obtained a job in the intervention group (79). It was stopped when this was found to have occurred but job seeking behaviors were no different between the two groups. In three other RCTs,

working memory and thought generation were found to be no different between intervention and control groups at followup (36;59). One RCT measured total charges paid for medical treatment (in US$) and found a very wide range in both intervention and control groups (34).

Psychological Outcomes

These have been listed in seven categories — mood or affect, anxiety, depression or emotional distress, IES (27), College Adjustment Test (CAT; 58), SCL-90 and SCL-90-R (12), and miscellaneous.

Twenty-three RCTs measured mood or affect (3;6;10; 11;17;21–24;31;40;51;58;61;64–66;69–72;76;82;85;87) using a variety of different measures, many reporting positive and negative affect separately, giving thirty-one outcomes. Five RCTs reported positive mood (21;23;66;72;87) and six negative mood (3;6;21;23;51;76;87) in sufficient detail for meta-analysis. The SMD for positive mood was 0.56 (0.22 to 0.91) for both fixed and random effects and for negative mood was 0.37 (0.12 to 0.62) fixed and 0.51 (0.01 to 1.01) random effects, with considerably more
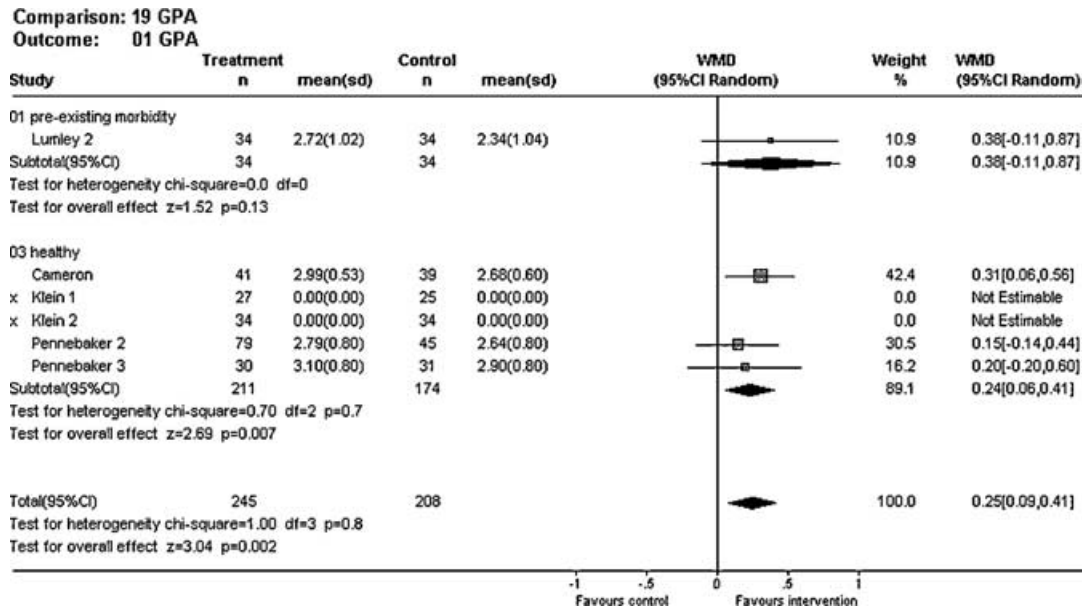
Figure 2. Grade point average (GPA) with mean SDs inserted.

heterogeneity between the RCTs. This finding suggests increases in positive and negative mood at follow-up for the intervention group.

Eight RCTs measured anxiety (10;11;37;57;66–68;71; 72) and six reported sufficient detail for meta-analysis (10;11;37;66;68;71;72). The SMD was 0.16 (−0.39 to +0.18) fixed and −0.40 (−0.97 to +0.17) random effects, showing no difference in anxiety. Twenty-one RCTs measured depression or emotional distress in various ways (3;10;11;14;21;24;34;37;38;40;45;47;50;61;64–67;70;71; 82), and ten had sufficient detail for meta-analysis (3;10;11;21;24;37;38;40;65;66;71). The SMD was 0.22 (+0.05 to +0.40) fixed and 0.21 (−0.13 to +0.55) random effects with little heterogeneity. The results suggest that there may be an increase in depression for the intervention group. One RCT measured depression in two ways, MAACL-R (89) and Self-Rating Depression Scale (90), one showing less depression and one showing more. (66) The SMD changed very little if one or the other was used and did not alter the conclusions.

Fifteen RCTs measured IES (2;10;11;21;24;36;38–40; 51;66;70;71;75;76;82;84;87), and nine gave results separately for the two subscales of avoidance and intrusion in sufficient detail for meta-analysis (10;11;21;24;38;51;66;71; 76;84;87; see Figure 3). For IES avoidance, the WMD was −0.06 (−1.13 to +1.00) fixed and 0.08 (−1.54 to +1.70) random effects and for IES intrusion 0.17 (−0.76 to +1.10) fixed and −0.04 (−2.31 to +2.22) random effects, suggesting that the intervention overall has no effect on avoidance or intrusion. Both Forest plots showed some heterogeneity, particularly IES intrusion where one RCT demonstrated statistically significantly more intrusion (87), whereas another showed the opposite (71).

For CAT, of the four RCTs in healthy volunteers to report this outcome (6;29;36;58), none found significant differences between the two groups. Eight RCTs measured SCL-90 or SCL-90-R (2;3;9;20;24;69–71). One found lower scores on the SCL-90-R for the intervention group for the total questionnaire score (71), and one a significant difference between intervention and control groups but failed to report the direction (2). The others did not report the results or found no significant differences. Two RCTs demonstrated less somatization for the intervention group (20;71), and one did not report (9). A wide variety of other psychological outcomes were measured, fewer were reported, and most showed no significant differences. The significantly different results were less sleep disturbance (10;11), better rehabilitation efficacy (85), less grief and more grief recovery (38), less grief recovery (66), better psychological well-being (32), and a group by time interaction on posttraumatic growth (86).

DISCUSSION

The results of this systematic review show no clear beneficial effect of emotional disclosure compared with controls for most physical and psychological outcomes. With the increasing publication of journal articles in health psychology, it is very difficult to view the totality of the evidence without a summary being available. A systematic review such as this tries to fill the gap in as impartial a way as possible. However, systematic reviews are secondary research and are particularly susceptible to publication and selective reporting biases. They are also limited by the quality of the reports of the included RCTs. Although the comprehensive search strategy helped minimize this bias and the funnel plot showed little evidence of asymmetry, that 15 potentially includable
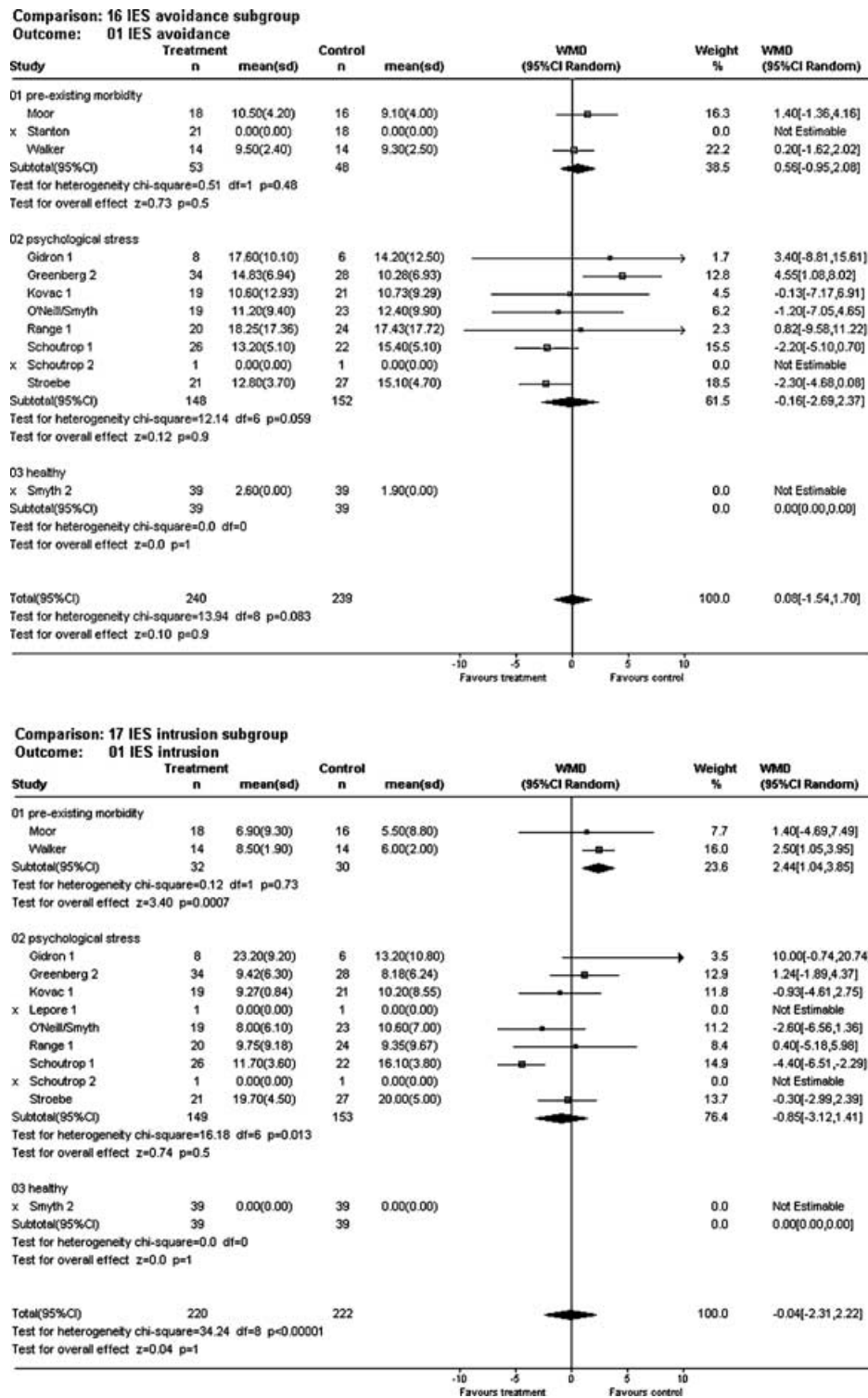
Figure 3. Forest plots of Impact of Events Scale (IES) avoidance and intrusion.

study references were found suggests that some publication bias may have occurred. These could represent the tip of an iceberg of a considerable volume of unpublished research. Where researchers have conducted small RCTs on physically healthy students and where the results have shown no significant differences in the primary outcome measure(s), it seems likely that there will be no great imperative to publish.

Related to the above is the major constraint of lack of complete information on published RCTs. It is noticeable

that approximately a third of outcomes were measured but not reported. It may be safe to assume that, if they had had statistically significant results then many would have been reported. Therefore, this assumption suggests that some reporting bias is operating. Of the outcomes that were reported, many are just by the statistical tests done on the summary results, with no summary measures such as means and SDs given. This finding meant that the statistics and conclusions arising could not be checked and that many results could not be entered into meta-analysis. As a result, it is difficult to gain a true picture of the trend of results. This difficulty is particularly apparent when such a variety of outcome measures have been used to explore the effects of the intervention.

The quality of the RCTs was assessed using CONSORT criteria and the Jadad scale. This strategy could be seen as a weakness in that most included RCTs were published in psychology not medical journals. Standards of reporting are different in each discipline, psychology journals generally adhering to the Publication Manual of the American Psychological Association (1). However, this manual includes the following guidance:

The sample should be adequately described: (i) Give the total number of subjects and the number assigned to each experimental condition; (ii) If any did not complete the experiment, state how many and explain why they did not continue; (iii) Describe randomization; (iv) Mention all relevant results; (v) Be sure to include descriptive statistics (e.g., means or medians); and (vi) Where means are reported, always include an associated measure of variability, such as SDs, variances, or mean square errors. If authors of RCTs and editors of the relevant journals had followed this guidance, then far more information would have been available for the systematic review and the effectiveness of the emotional disclosure intervention would have been clearer.

One of the main differences between this systematic review and the previous meta-analysis is our decision to disaggregate categories of reported health, psychological well-being, and general functioning into component outcomes. The reason was to investigate the impact of emotional disclosure on a wide variety of different dependent variables. In medical RCTs, composite outcomes give greater precision but at a cost of the greater difficulty in interpretation of the results (18). This finding is also likely to be true in systematic reviews. There is a tendency to assume that, if the composite outcome was found to be statistically significantly different between the intervention and control groups, then all of the component outcomes will also vary in the same way. However, this systematic review has shown that outcomes in each of the composites vary independently, one from another. Another disadvantage of using composite categories is the problem of allowing more than one effect size per study and nonindependence. If a single effect size from each RCT is used, the rules by which the outcome measure is chosen for each RCT needs to be made explicit. Several RCTs in the previous meta-analysis reported multiple outcomes and did

not state the primary outcome. This finding means that one of several results could have been chosen for the meta-analysis. As RCTs tend to highlight their most positive findings, one of the more positive results may have been used, which would not mirror the true spread of results. This finding may give a more optimistic overall effect size. Alternatively, if all of the outcomes were averaged within each RCT and that figure were used to derive the overall effect size, then some RCTs would have contributed a single outcome, whereas others an averaged outcome. Therefore, the more fully reported RCTs with a spread of outcomes would contribute a lower effect size than the RCTs that presented their most positive results only.

The disadvantage of using separate outcomes is that far fewer RCTs measure each variable. In this systematic review, an added disadvantage is that even fewer RCTs reported their results fully. In the absence of adequate information for meta-analysis for some of the outcomes, a vote-counting approach has been used instead. We acknowledge that this is not ideal, because it ignores sample size, effect size, and the variance of results. However, with the increasing number of emotional disclosure RCTs being published, this should gradually become less of a problem.

At the moment, it remains unclear as to whether the equivocal results in this systematic review are due to small sample sizes and lack of power in the RCTs, poor quality of reporting, or because the emotional disclosure intervention actually has little effect. Accordingly, the trend of results provides a mixed picture. There is no clear balance in favor of the emotional disclosure intervention for many of the outcomes measured. This finding is not what one would expect from reading the reviews and editorials on emotional disclosure (16;53;80). It may be that the way the RCTs have been reported has resulted in a more positive picture of the effects of this intervention than is actually the case. This possibility is all the more worrying as this intervention has been recommended in a therapeutic setting (5;62;83), when the benefits have not been established, and has been evaluated for use in this setting (19;26). On the other hand there is little evidence from the RCTs reviewed that this intervention does any harm. This systematic review is not suggesting that all emotional disclosure has no or very little effect. It is suggesting that the current evidence available has not clearly demonstrated the effectiveness of this brief emotional disclosure intervention.

CONTACT INFORMATION

Catherine Meads, MBChB, MSc (c.a.meads@bham.ac.uk), Research Officer, Department of Public Health and Epidemiology, **Arie Nouwen**, PhD (a.nouwen@bham.ac.uk), Clinical Lecturer, School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

REFERENCES

1. American Psychological Association. *Publication manual of the American Psychological Association*. 4th ed. Washington, DC: American Psychological Association; 1994.

2. Barry LM. *The benefits of journal writing: Reducing maternal psychological distress levels after the neonatal intensive care unit*. PhD Thesis. Santa Barbara, CA: University of California; 2000.

3. Batten SV, Follette VM, Hall ML, Palm KM. Physical and psychological effects of written disclosure among sexual abuse survivors. *Behav Ther*. 2002;33:107-122.

4. Booth RJ, Petrie KJ, Pennebaker JW. Changes in circulating lymphocyte numbers following emotional disclosure: Evidence of buffering? *Stress Med*. 1997;13:23-29.

5. Bootzin RR. Examining the theory and clinical utility of writing about emotional experiences. *Psychol Sci*. 1997;8:167-169.

6. Cameron LD, Nicholls G. Expression of stressful experiences through writing: Effects of a self-regulation manipulation for pessimists and optimists. *Health Psychol*. 1998;17:84-92.

7. Clarke M, Oxman A, eds. *Cochrane Reviewers' Handbook 4.2.0*. Oxford: The Cochrane Library, Update Software; 2003.

8. Czajka JA. *Behavioral inhibition and short-term physiological responses*. MA Thesis. Southern Methodist University, Dallas, TX, USA; 1987.

9. D'Souza P, Lumley MA, Kraft C, et al. *Emotional disclosure and relaxation training for migraine and tension headaches: A randomised trial*. Phoenix, AZ. Presentation to 61st Annual Scientific Meeting of the American Psychosomatic Society; 2003.

10. de Moor C, Sterner J, Hall ML, et al. A pilot study of the effects of expressive writing on psychological and behavioural adjustment in patients enrolled in a phase II trial of vaccine therapy for metastatic renal cell carcinoma. *Health Psychology*. 2002;21:615-619.

11. de Moor C, Warneke C, Sterner J, et al. Emotional expression writing program for cancer patients. *Psychosom Med*. 2001;63:1143.

12. Derogatis LR, Rickels K, Rock AF. The SCL-90 and the MMPI: A step in the validation of a new self-report scale. *Br J Psychiatry*. 1976;128:280-289.

13. Dickerson SS, Kemeny ME, Aziz N, Kim KH, Fahey JL. Immunological effects of induced shame and guilt. *Psychosom Med*. 2001;63:159.

14. Donnelly DA, Murray EJ. Cognitive and emotional changes in written essays and therapy interviews. *J Soc Clin Psychol*. 1991;10:334-350.

15. Esterling BA, Antoni MH, Fletcher MA, Margulies S, Schneiderman N. Emotional disclosure through writing or speaking modulates latent Epstein-Barr virus antibody titres. *J Consult Clin Psychol*. 1994;62:130-140.

16. Esterling BA, L'Abate L, Murray EJ, Pennebaker JW. Empirical foundations for writing in prevention and psychotherapy: Mental and physical health outcomes. *Clin Psychol Rev*. 1999;19:79-96.

17. Francis ME, Pennebaker JW. Putting stress into words: The impact of writing on physiological, absentee and self-reported emotional well-being measures. *Stress Manag*. 1992;6:280-287.

18. Freemantle N, Calvert M, Woods J, Eastaugh J, Griffin C. Composite outcomes in clinical trials. Greater precision but with greater uncertainty? *JAMA*. 2003;289:2554-2559.

19. Gallagher P, Maclachlan M. Evaluating a written emotional disclosure homework intervention for lower-limb amputees. *Arch Phys Med Rehabil*. 2002;83:1464-1466.

20. Gidron Y, Duncan E, Lazar A, et al. Effects of guided written disclosure of stressful experiences on clinic visits and symptoms in frequent clinic attenders. *Fam Pract*. 2002;19:161-166.

21. Gidron Y, Peri T, Connolly JF, Shalev AY. Written disclosure in posttraumatic stress disorder: Is it beneficial for the patient? *J Nerv Ment Dis*. 1996;184:505-507.

22. Gillis ME, Lumley MA, Koch H, et al. Written emotional disclosure in fibromyalgia: Effects on sleep quality and fatigue. *Sleep*. 2002;25:538.

23. Greenberg MA, Stone AA. Emotional disclosure about traumas and its relation to health: Effects of previous disclosure and trauma severity. *J Pers Soc Psychol*. 1992;63:75-84.

24. Greenberg MA, Wortman CB, Stone AA. Emotional expression and physical health: Revising traumatic memories or fostering self regulation? *J Pers Soc Psychol*. 1996;71:588-602.

25. Greenhalgh T. Writing as therapy, effects on immune mediated illness need substantiation in independent studies. *BMJ*. 1999;319:270-271.

26. Hannay D, Bolton G. Therapeutic writing in primary care: A feasibility study. *Prim Care Psychiatry*. 1999;5:157-160.

27. Horowitz MJ, Wilner N, Alvarez W. Impact of events scale: A measure of subjective stress. *Psychosom Med*. 1979;41:209-218.

28. Hughes CF. *Effects of expressing negative and positive emotions and insight on health and adjustment to college*. PhD Thesis. Southern Methodist University, Dallas, TX, USA; 1993.

29. Hughes CF, Uhlmann C, Pennebaker JW. The body's response to processing emotional trauma: Linking verbal text with autonomic activity. *J Personality*. 1994;62:565-585.

30. Jadad A, Moore R, Carroll D. Assessing the quality of reports of randomised clinical trials: Is blinding necessary? *Control Clin Trials*. 1996;17:1-12.

31. Kelley JE, Lumley MA, Leisen JC. Health effects of emotional disclosure in rheumatoid arthritis patients. *Health Psychol*. 1997;16:331-340.

32. King LA. The health benefits of writing about life goals. *Pers Soc Psychol Bull*. 2001;27:798-807.

33. King LA, Miner KN. Writing about the perceived benefits of traumatic events: Implications for physical health. *Pers Soc Psychol Bull*. 2000;26:220-230.

34. Klapow JC, Schmidt SM, Taylor LA, et al. Symptom management in older primary care patients: Feasibility of an experimental, written self-disclosure protocol. *Ann Intern Med*. 2001;134:905-911.

35. Klein DJ, Cacioppo JT. The effects of emotional disclosure and traumatic life event history on blood pressure and heart rate in college-aged females. *Psychophysiology*. 1996;33:S51.

36. Klein K, Boals A. Expressive writing can increase working memory capacity. *J Exp Psychol Gen*. 2001;130:520-533.

37. Kloss JD, Lisman SA. An exposure based examination of the effects of written emotional expression. *Br J Health Psychol*. 2002;7:31-46.

38. Kovac SH, Range L. Writing projects: Lessening undergraduates' unique suicidal bereavement. *Suicide Life Threat Behav*. 2000;30:50-60.

39. Lepore SJ. Expressive writing moderates the relation between intrusive thoughts and depressive symptoms. *J Pers Soc Psychol*. 1997;73:1030-1037.

40. Lepore SJ, Greenberg MA. Mending broken hearts: Effects of expressive writing on mood, cognitive processing, social adjustment and health following a relationship breakup. *Psychol Health*. 2002;17:547-560.

41. Littrell J. Is the Re-experience of painful emotion therapeutic? *Clin Psychol Rev*. 1998;18:71-102.

42. Lumley M, Leegstra S, Provenzano K, Warren V. The health effects of writing about emotional stress on physically symptomatic young adults. *Psychosom Med*. 1999;61:84-130:89.

43. Lumley MA, Provenzano K. Stress management through written emotional disclosure improves academic performance among college students with physical symptoms. *J Educ Psychol*. 2003;95:641-649.

44. Mann T. Effects of future writing and optimism on health behaviors in HIV-infected women. *Ann Behav Med*. 2001;23:26-33.

45. Marlo H, Wagner MK. Expression of negative and positive events through writing: Implications for psychotherapy and health. *Psychol Health*. 1999;14:193-215.

46. Meads CA, Lyons A, Carroll D. *The impact of the emotional disclosure intervention on physical and psychological health — a systematic review*. Birmingham, UK: WMTAC, Department of Public Health and Epidemiology, University of Birmingham; 2003. Report number 43.

47. Meyer T, Lumley MA, Markowitz A, et al. *Written and verbal emotional disclosure about stress in patients with rheumatoid arthritis*. Phoenix, AZ. Presentation to 61st Annual Scientific Meeting of the American Psychosomatic Society, March 5-8, 2003.

48. Moher D, Schulz KF, Altman DG. The CONSORT Statement: Revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet*. 2001;357:1191-1194.

49. Murray EJ, Lamnin AD, Carver CC. Emotional expression in written essays and psychotherapy. *J Soc Clin Psychol*. 1989;8:414-429.

50. Murray EJ, Segal DL. Emotional processing in vocal and written expression of feelings about traumatic experiences. *J Trauma Stress*. 1994;7:391-405.

51. O'Neill HK, Smyth JM. *Effects of written disclosure on post-disaster psychological adjustment and symptomatology*. Available at: http://www.colorado.edu/hazards/qr/qr138/qr138.htm. Accessed: November 7, 2002.

52. Pennebaker JW. *The psychology of physical symptoms*. New York: Springer; 1982.

53. Pennebaker JW. Putting stress into words: Health, linguistic, and therapeutic implications. *Behav Res Ther*. 1993;31:539-548.

54. Pennebaker JW. Writing About emotional experiences as a therapeutic process. *Psychol Sci*. 1997;8:162-166.

55. Pennebaker JW. *Reference list of writing/disclosure studies*. Available at: http://homepage.psy.utexas.edu/homepage/faculty/Pennebaker/Reprints/writingrefs.html. Accessed: 2002.

56. Pennebaker JW, Barger SD, Tiebout J. Disclosure of traumas and health among holocaust survivors. *Psychosom Med*. 1989;51:577-589.

57. Pennebaker JW, Beall SK. Confronting a traumatic event: Toward an understanding of inhibition and disease. *J Abnorm Psychol*. 1986;95:274-281.

58. Pennebaker JW, Colder M, Sharp LK. Accelerating the coping process. *J Pers Soc Psychol*. 1990;58:528-537.

59. Pennebaker JW, Francis ME. Cognitive, emotional and language processes in disclosure. *Cogn Emotion*. 1996;10:601-626.

60. Pennebaker JW, Hughes CF, O'Heeron RE. The psychophysiology of confession: Linking inhibitory and psychosomatic processes. *J Pers Soc Psychol*. 1987;52:781-793.

61. Pennebaker JW, Kiecolt-Glaser JK, Glaser R. Disclosure of traumas and immune function: Health implications for psychotherapy. *J Consult Clin Psychol*. 1988;56:239-245.

62. Pennebaker JW, Seagal JD. Forming a story: The health benefits of narrative. *J Clin Psychol*. 1999;55:1243-1254.

63. Petrie KJ, Booth RJ, Pennebaker JW. The immunological effects of thought suppression. *J Pers Soc Psychol*. 1998;75:1264-1272.

64. Petrie KJ, Booth RJ, Pennebaker JW, Davison KP, Thomas MG. Disclosure of trauma and immune response to a hepatitis B vaccination program. *J Consult Clin Psychol*. 1995;63:787-792.

65. Range L, Kovac SH. Can writing autobiographical essays lessen suicidal thinking? *Arch Suicide Res*. 2002;6:373-382.

66. Range L, Kovac SH, Marion MS. Does writing about the bereavement lessen grief following sudden, unintentional death? *Death Studies*. 2000;24:115-134.

67. Reynolds M, Brewin CR, Saxton M. Emotional disclosure in school children. *J Child Psychol Psychiatry*. 2000;41:151-159.

68. Richards JM, Beal WE, Seagal JD, Pennebaker JW. Effects of disclosure of traumatic events on illness behavior among psychiatric prison inmates. *J Abnorm Psychol*. 2000;109:156-160.

69. Rosenberg HJ, Rosenberg SD, Ernstoff MS, et al. Expressive disclosure and health outcomes in a prostate cancer population. *Int J Psychiatr Med*. 2002;32:37-53.

70. Schoutrop M, Brosschot JF, Lange AJ. Writing assignments after trauma: Decreased re-experiencing and within/across session physiological habituation. *Psychosom Med*. 1999;61:84-130:95.

71. Schoutrop M, Lange A, Hanewald G, Davidovich U. Structured writing and processing major stressful events: A controlled trial. *Psychother Psychosom*. 2002;71:151-157.

72. Sheffield D, Duncan E, Thomson K, Johal SS. Written emotional expression and well-being: Result from a home-based study. *Aust J Disaster Trauma Studies*. 2002;1:1-13.

73. Sloan DM, Marx BP, Soler-Baillo J. Physiological correlates of emotional processing through written disclosure. *Psychophysiology*. 2002;39:S77.

74. Smyth JM. Written emotional expression: Effect sizes, outcome types and moderating variables. *J Consult Clin Psychol*. 1998;66:174-184.

75. Smyth JM, Anderson C, Hockemeyer J, Stone AA. Does emotional non-expressiveness or avoidance interfere with writing about stressful life events? An analysis in patients with chronic illness. *Psychol Health*. 2002;17:561-569.

76. Smyth JM, Hockemeyer J, Anderson C, et al. Structured writing about a natural disaster buffers the effect of intrusive thoughts on negative affect and physical symptoms. *Aust J Disaster Trauma Studies*. 2002;1:1-10.

77. Smyth JM, Stone AA, Hurewitz A, Kaell A. Effects of writing about stressful experiences on symptom reduction in patients with asthma or rheumatoid arthritis. *JAMA*. 1999;281:1304-1309.

78. Smyth JM, True N, Souto J. Effects of writing about traumatic experiences: The necessity for narrative structuring. *J Soc Clin Psychol*. 2001;20:161-172.

79. Spera SP, Buhrfeind ED, Pennebaker JW. Expressive writing and coping with job loss. *Acad Manag J*. 1994;37:722-733.

80. Spiegel D. Healing words: Emotional expression and disease outcome. *JAMA*. 1999;281:1328-1329.

81. Spiegel D. *Minding the body*. Available at: http://whyfiles.org/039emotion/spiegel2.htm. Accessed: November 7, 2002.

82. Stanton AL, Danoff-Burg S, Sworowski LA, et al. Randomized, controlled trial of written emotional expression and benefit finding in breast cancer patients. *J Clin Oncol*. 2002;20:4160-4168.

83. Stephens C. Health benefits of the disclosure of emotions about traumatic experiences: What is the evidence and potential for therapeutic benefits? *Aust J Disaster Trauma Studies*. 2002;1:1-6.

84. Stroebe M, Stroebe W, Schut H, Zech E, van den Bout J. Does disclosure of emotions facilitate recovery from bereavement? Evidence from two prospective studies. *J Consult Clin Psychol*. 2002;70:169-178.

85. Strough HC. *The effects of disclosive writing on psychological responses and subjective and objective outcomes following anterior cruciate ligament (ACL) rehabilitation*. PhD Thesis. West Lafayette, IN: Purdue University; 1998.

86. Ullrich PM, Lutgendorf SK. Journaling about stressful events: Effects of cognitive processing and emotional expression. *Ann Behav Med*. 2002;24:244-250.

87. Walker BL, Nail LM, Croyle RT. Does emotional expression make a difference in reactions to breast cancer? *Oncol Nurs Forum*. 1999;26:1025-1032.

88. Watson D, Pennebaker JW. Health complaints, stress and distress: Exploring the central role of negative affectivity. *Psychol Rev*. 1989;96:234-254.

89. Zuckerman M, Lubin B. *Manual for the MAACL-R: The multiple affect adjective checklist — revised*. San Diego: Education and Industrial Testing Service; 1985.

90. 90. Zung WWK. A self-rating depression scale. *Arch Gen Med*. 1965;12:63-70.