Cognitive-Behavioral Group Therapy for Irritable Bowel Syndrome: Effects and Long-Term Follow-Up

A. M. van Dulmen, MA, J. F. M. Pennis, MD, PhD, and G. Bleijenberg, PhD

Little is known about the effectiveness of cognitive-behavioral treatment for patients with irritable bowel syndrome on a group basis. Previous studies have used only small samples, and studies with long term follow-up are lacking. The aim of the present study was to investigate: a) the effectiveness of a cognitive-behavioral group treatment compared with a waiting list control condition in alleviating abdominal complaints and b) the long-term effectiveness of cognitive-behavioral group treatment. In study 1, we performed a controlled study with 25 patients in the group treatment condition and 20 patients in the waiting list control condition. Treatment consisted of eight 2-hour group sessions over a period of 3 months. In study 2, all patients were treated and followed up for an average of 2.25 years (range 6 months-4 years) after the completion of the group treatment. The abdominal complaints of the patients who underwent treatment were found to improve significantly more than the complaints of the patients awaiting treatment. Moreover, in agreement with the purpose of the therapy, the number of successful coping strategies was found to increase more and patients' avoidance behavior was found to decrease more in the treatment group than in the waiting list control group. The positive changes appeared to persist during follow-up. Cognitive-behavioral group treatment is effective in alleviating irritable bowel syndrome, in stimulating coping strategies, and in reducing avoidance behavior. At long-term follow-up, the abdominal complaints, the number of successful coping strategies, and the avoidance behavior were still improved compared with the pretreatment assessment.

Key words: cognitive-behavioral group therapy, controlled study, coping, irritable bowel syndrome, long-term follow-up.

INTRODUCTION

Abdominal complaints have a high prevalence in the general population and are associated with a high rate of medical visits (1). Sixteen percent of the patients attending a general practitioner with abdominal complaints are referred for extensive physical examination, most often to internists (2). In the majority of referrals, no somatic explanation can be found (3, 4). These so-called functional abdominal complaints can also be referred to as irritable bowel syndrome, of which two definitions are known based on broad (5) and restrictive criteria (6), respectively. Irritable bowel syndrome appears to have a poor prognosis (7–9). Standard medical treatments, such as medication and dietary advice, are rather ineffective in improving this prognosis (10, 11).

Recently, we found, at an outpatient internal medicine clinic, that, during medical consultations, doctors are able to influence complaint-related cognitions and that these achieved changes are related to improvement of irritable bowel symptoms at follow-up (9, 12). Thus, at least some patients with irritable bowel syndrome can be treated successfully during medical consultations with minimal psychological intervention. Patients who did not improve or improved less after the consultations appeared to report more somatic attributions and catastrophizing cognitions as well as less psychological attributions and self-efficacy expectations (9). For these patients with severe and chronic complaints, referral to a behavior therapist may be indicated.

Psychotherapeutic treatment programs administered individually have shown to be effective in alleviating irritable bowel syndrome (13–20). However, given the large numbers of patients with irritable bowel syndrome and the general concern with cost containment in health care, it seems worthwhile to establish the effects of cognitive-behavioral group treatment. The first uncontrolled study in which the effectiveness of behavioral group therapy was evaluated was reported by Wise et al. in 1982 (21). After group treatment, one-third of a group of 20 treated patients reported a reduction of their abdominal
complaints. Positive results were also reported by Blanchard and Schwarz who evaluated the short term effects of cognitive-behavioral group treatment on three groups with a total of 14 patients (22). This group treatment was an adaptation of their behavioral treatment program administered individually (16). These uncontrolled studies suggest that cognitive-behavioral group treatment is effective in alleviating abdominal complaints. Long term effects of cognitive-behavioral group treatment are not known. The aim of the present study was to investigate the effectiveness of group treatment compared with a waiting list control group (study 1) as well as the long term effects of this cognitive-behavioral group treatment (study 2).

STUDY 1

Methods

Subjects. Forty-seven outpatients with refractory irritable bowel syndrome were recruited by the Outpatient University Clinic for Internal Medicine for cognitive-behavioral group treatment. Irritable bowel syndrome was broadly defined as abdominal pain with or without disordered defecation with a duration of at least 3 months in the absence of any recognized gastrointestinal pathology (5); the abdominal pain-predominant patients were considered to have refractory irritable bowel syndrome (13, 23). Patients were allocated to the treatment condition, which was presented as “a course in coping with abdominal complaints” until the first group was full, and then the next patients were put on the waiting list until a new group began. Ultimately, seven groups were completed. To restrict the duration of the waiting period, each group consisted partly of patients from the waiting list and partly of recently referred patients. The treatment condition consisted of 27 patients, 16 women and 11 men, with a mean age of 44 (SD = 11) years and a mean duration of abdominal complaints of 5.0 (SD = 4.2) years. The waiting list control group consisted of 20 patients, 8 women and 12 men, with a mean age of 48 (SD = 14) years and a mean duration of complaints of 5.3 (SD = 4.0) years. The mean duration of the waiting period was 3.5 months (SD = 1.5), which was about the same duration as the group treatment. During treatment and waiting period, no concomitant somatic treatment was given. Two patients in the treatment condition returned incomplete questionnaires, so the effects of the group treatment could only be evaluated in 25 patients.

Group Treatment. All patients underwent the same treatment, which consisted of eight 2-hour group sessions over a period of 3 months. An experienced psychotherapist and a junior psychologist conducted all sessions, aimed at modifying maladaptive cognitions and behavior and stimulating different and more effective ways of coping with the complaints. The group treatment was based on principles from cognitive-behavioral treatment (24, 25) and combined:

1. Patient education about the role of cognitions, behavior, emotions, and environment in relation to the abdominal complaints. Attention was given to correction of unjustified or dysfunctional attributions, to events or circumstances preceding or following the presence of complaints, and to reactions, in thinking, feeling, or behavior, to the presence of complaints.
2. Homework. Patients were stimulated to change their complaint-related cognitions and behavior and to try out new ways of thinking and behaving between sessions. The patients described their experiences in a diary.
3. Group conversation. During the sessions, much time was spent talking about the educational part and the patients’ experiences with the homework. Attention was given to mutual recognition of the problems the participants had in coping with their pain.
4. Training in progressive muscle relaxation. Each session was concluded by a muscle relaxation exercise (26). Patients were instructed how to use these exercises at home and in daily life. Later on, coping imagery exercises were added.

The information that was presented during each session, the relaxation instructions, and the homework were also given in written form as a course book. Themes of the sessions were: 1. General introduction 2. Cognitions and abdominal pain 3. Changing cognitions 4. Discovering antecedent cognitions and behavior 5. Changing behavior 6. Role of environment 7. Changing complaint-related lifestyle 8. Coping in the future.

Assessment. Patients completed the following questionnaires at first and second assessment: patients in the treatment group before and after the completion of the treatment and patients in the waiting list control group before and after the waiting period:

Diary. The Daily Abdominal Complaint Score (DAC) was measured using a prescheduled diary in which patients rated their abdominal pain through self-observation four times daily during 2 weeks on a scale from 0 (no pain) to 4 (serious interfering pain). Thus the DAC score could vary between 0 and 16 (27, 28). This score is reliable and able to measure change (28).

Daily duration of the abdominal pain was also measured by means of the diary in which patients reported four times daily how long they had experienced abdominal pain since the last measurement. Scores ranged from 0 (0 hours), 1 (less than 2 hours), 3 (more than 2 hours but less than 4 hours), to 5 (more than 4 hours). Scores at breakfast were doubled because at that time patients often reported more pain than during other times of the day. Thus the DAC score could vary between 0 and 28 (27, 28).

Daily avoidance behavior was measured using the above mentioned diary in which patients reported four times daily whether or not they had avoided certain activities because of the abdominal pain since their last observation. An individual avoidance score was presented as a percentage, indicating the number of times the patient reported avoidance, divided by the total number of observations during the 2 weeks.

Finally, in the diary, patients reported four times daily whether they had suffered from the following gastrointestinal complaints: flatulence, belching, nausea, heartburn, abdominal rumbling. Details of the defecation were also noted: difficult or painful evacuation, a feeling of incomplete evacuation, and the type of defecation (hard, well formed, or pulpy stools).

Abdominal Complaint Inventory. In the abdominal complaint inventory, biographical data, history, and details of the complaints as well as coping strategies and use of medication for abdominal complaints were listed.

The severity score of the abdominal complaints (range 0–9) was determined by summing the reported frequency of the complaints (0–3), interference with daily activities (0–3), and avoidance behavior (0–3). A score of 0 refers to patients who experienced abdominal complaints less than once a month without significant interference, or who did not feel they avoided a certain activity.
being bothered by these complaints, 9 to patients who reported daily complaints and much interference and avoidance behavior (7, 9).

The number of coping strategies patients used to cope with their complaints successfully was counted. For that purpose, patients were asked to indicate how many of 14 different ways of behaving or thinking (eg, relaxation excercises; avoiding certain foods or drinks; thinking that the complaints are bearable) most often resulted in experiencing less complaints. This list of coping strategies was adapted from the Pain Coping Inventory, a reliable 33-item scale referring to the patient's use of cognitive and instrumental strategies when in pain (29).

Psychological Well Being. The Symptom Check List (SCL-90) was used as a parameter for psychological well being; the total score was used (30). A high score reflects a decreased level of psychological well being.

Outcome Measures. Following its purpose, treatment was evaluated in terms of changes in Daily Abdominal Complaint score (DAC), duration of the pain, daily avoidance, severity score, number of successful coping strategies, and psychological well being. Moreover, DAC scores were used to assess the outcome of treatment on an individual subject basis by calculating the percentage change for each patient. A patient was defined as clinically improved when the DAC score had decreased 50% or more (31).

Statistics. Differences between subgroups of patients were investigated using the Mann-Whitney U and χ² tests. For each patient, a change score was calculated by subtracting scores at second assessment from baseline scores. Using these change scores, it was possible to compare the amount of changes in treated patients and controls using Mann-Whitney U and independent t tests (for the number of coping strategies). Statistical significance for all tests was set at the 5% level.

Results

Treatment Group vs. Waiting List Control Group at Baseline. No significant differences emerged among the baseline scores on demographic and symptom measures between treatment group (N = 27) and control group (N = 20). Moreover, no baseline differences emerged among the number of successful coping strategies, daily avoidance behavior, and the level of psychological well being. So both groups appeared to be well matched.

Changes Between First and Second Assessment. Table 1 presents the means on the DAC score, daily duration, daily avoidance, severity score, number of successful coping strategies, and psychological well being at first and second assessment for treated patients and controls. Patients' DAC score, daily duration, daily avoidance, and number of successful coping strategies appeared to improve significantly more during the group treatment than during the waiting period. No significant differences were found comparing improvement in the level of psychological well being or in any of the secondary gastrointestinal complaints.

The mean decrease in the DAC score for all patients in the treatment group was 37% (SD = 40); all controls showed a mean increase of 22% (SD = 80). This difference was statistically significant (p = .005). Forty-four percent of the patients in the treatment group and 11% of the controls showed a decrease in the DAC score of more than 50% (p = .02). Moreover, in the treatment and control groups, no differences were found in the mean change in the DAC score between female and male patients.

Treatment of Waiting List Controls. Eighteen of the 20 patients from the waiting list control group subsequently underwent group treatment. During this group treatment, the DAC score of these patients decreased, with an average of 23% (SD = 40), not significantly different from the original treatment group (p = .26).

Two patients refused the offered treatment. The baseline scores from these two patients were within 0.6 SD from the mean of the treated patients, except for psychological well being, which was found to be much lower in the two refusers.

Improvement in Psychological Well Being. Although patients' psychological well being did not appear to improve as a result of the group treatment,
improvement in psychological well being was found to correlate significantly with reductions in the daily abdominal complaint score ($r = .46$, $p = .002$), daily duration ($r = .35$, $p = .01$), and severity score ($r = .46$, $p = .001$) during treatment ($N = 43$).

**STUDY 2**

Methods

Subjects. All patients from study 1 ultimately received the same treatment, so we report the results of a long term follow-up study of both groups of patients together. Patients were asked to complete questionnaires at several occasions after completing treatment, namely at 6, 12, 24, 36, and 48 months posttreatment. Because not all patients had completed treatment at least 48 months ago, follow-up points differed. Follow-up results were available for 32 patients from the 45 patients who underwent treatment; six patients completed treatment but did not return the follow-up questionnaire; seven patients returned an incomplete follow-up questionnaire. The follow-up group consisted of 16 women and 16 men. The mean length of time between completion of the group treatment and follow-up assessment was 2.25 years (SD = 1.12) with a range from 6 months to 4 years. Twenty-seven patients completed the prescheduled diary during 2 weeks at follow-up.

Assessment at Follow-Up. At follow-up, patients completed the same instruments as described in study 1. Data collected at the end of the waiting period were used as pretreatment measurements for the follow-up study of former controls.

Statistics. Changes in scores on complaints (DAC, duration, and severity), daily avoidance, and psychological well being between pretreatment and follow-up were measured using Wilcoxon matched-pairs signed-ranks test. Changes in the number of coping strategies were measured using a paired t test. Change scores were calculated by subtracting the scores at follow-up from the scores at pretreatment and were used to compare the amount of change in former controls and treated patients, in patients with a short and a long follow-up, and in female and male patients, by means of Mann-Whitney U test. Differences between subgroups of patients were investigated using Mann-Whitney U and $\chi^2$ tests. Spearman rank correlation coefficients were used to investigate the relationships between improvement in the DAC score and the use of medication and health care services at follow-up. Statistical significance was set at the 5% level.

**Results**

Follow-Up Group vs. Missings. At baseline, the 32 patients in the follow-up group and the 13 patients who did not return the follow-up questionnaire or filled it incompletely were comparable with respect to sex, age, duration of complaints before referral, and expectations regarding the treatment. Furthermore, no differences were found between both groups in the DAC score, daily duration of the abdominal pain, daily avoidance, severity score, number of successful coping strategies, or psychological well being immediately after treatment. Thus, not returning the follow-up questionnaires did not appear to be related to differences in the treatment outcome. The five patients who did not complete the prescheduled diary at follow-up did not differ with respect to any of the above mentioned variables from the patients who completed the diary at follow-up.

Concomitant Treatment During Follow-Up. During follow-up, three patients from the original treatment condition and two patients from the waiting list received either individual cognitive-behavioral psychotherapy or psychiatric treatment for problems that came up during the group therapy (posttraumatic stress disorders and obsessive-compulsive disorder, respectively). These five patients did not appear to improve more or less on any of the outcome measures from those not receiving concomitant psychotherapeutic treatment.

Changes Between Pretreatment and Follow-Up. Table 2 shows the means on abdominal complaints, daily avoidance, number of successful coping strategies, and psychological well being at pretreatment and at follow-up. The means on all outcome measures at follow-up were improved compared with the pretreatment means and reached conventional levels of significance ($p < .05$) on the DAC score, daily duration, daily avoidance, and the number of successful coping strategies. Moreover, at follow-up, patients appeared to be bothered significantly less by the secondary gastrointestinal complaints flatulence ($p = .01$), abdominal rumbling ($p = .003$), and difficult or painful defecation ($p = .01$) as noted in the diary compared with pretreatment. Between pretreatment and follow-up, the DAC score decreased with an average of 23% (SD = 58). The more the DAC score decreased, the less patients reported using medication for abdominal complaints ($r = .51$).

| TABLE 2. Mean Scores on Complaints, Daily Avoidance, Number of Successful Coping Strategies, and Psychological Well Being at Pretreatment and Follow-Up ($N = 32$) with Level of Significance of the Difference Between Pretreatment and Follow-Up |
|--------------------------------------------------|------------------|------------------|------------------|
| Diary†                                           | Pretreatment     | Follow-Up        | $p$              |
| DAC score                                        | (M (SD))         | (M (SD))         |                  |
| Daily duration                                   | 5.60 (3.25)      | 3.91 (3.36)      | .005             |
| Daily avoidance                                  | 11.47 (7.75)     | 8.72 (8.37)      | .039             |
| Severity score                                   | 21.22 (30.20)    | 13.11 (27.86)    | .035             |
| Number of successful coping strategies           | 5.06 (1.99)      | 4.34 (2.56)      | .091             |
| Psychological well being                         | 149 (45.90)      | 145 (56.42)      | .39              |

†The prescheduled diary was completed by 27 patients.
Using the criterion for clinical improvement of ≥50% decrease, in 37% of the patients, the DAC score appeared to be clinically improved. In addition, the following gastrointestinal complaints were clinically improved in at least half of the patients who were bothered by that complaint: flatulence (reported by 23 patients of which 57% clinically improved), belching (reported by 16 patients of which 63% clinically improved), nausea (reported by 16 patients of which 50% clinically improved), heartburn (reported by 11 patients of which 73% clinically improved), and abdominal rumbling (reported by 23 patients of which 65% clinically improved). No differences emerged in comparing the follow-up results of former controls and first-treated patients and in the comparison of the results of female and male patients.

**Short vs. Long Term Follow-Up.** As there is a considerable range in the duration of the follow-up, we investigated whether the treatment outcome differed between patients with a shorter and a longer duration of follow-up. Comparing patients with ≤1 year \((N = 9)\) and >1 year \((N = 23)\) follow-up revealed no differences in any of the outcome measures nor did comparing patients with ≤2 years \((N = 17)\) and >2 years \((N = 15)\) follow-up.

**DISCUSSION**

The results of the present study indicate that a cognitive-behavioral group treatment is more effective in alleviating irritable bowel syndrome than a waiting period without intervention. Cognitive-behavioral group treatment not only improved the daily observed intensity and duration of the complaints but also reduced patients' daily avoidance behavior and increased the number of successful cognitive and behavioral coping strategies. It may be assumed that this is a direct result of the treatment, which aimed at restructuring patients' complaint-related cognitions and behavior.

The improvement found at posttreatment appeared to be maintained at follow-up; an average of more than 2 years after completion of the group treatment, the abdominal complaints were still improved compared with the pretreatment levels. Moreover, patients still used more successful coping strategies than before treatment. We could not find any difference between a short and a long term follow-up, so it may be concluded that our brief group treatment has long term effects. Analyzing the specific nature of the coping strategies revealed that doing relaxation exercises was not the only successful strategy, as was suggested in an earlier study (32). More positive cognitions and emotions, such as "worrying less," "getting angry less frequently," and "thinking that the pain will decrease," appeared to have a favorable influence on the abdominal complaints, too. Apparently, comparable with earlier findings (12, 20), changes in abdominal complaints are associated with changes in thoughts and behaviors.

Besides the favorable treatment results mentioned above, our study has some methodological assets, too. First, in this study, patients rated their abdominal complaints four times daily, whereas in other studies complaints were either rated only once-per-day (13–20, 22, 32–34) or were not assessed at all by self-observation (21). Moreover, patients in this study not only rated their abdominal pain and type of defecation through daily self-observation but also five other gastrointestinal complaints associated with irritable bowel syndrome. Most abdominal symptoms occur erratically throughout the day (35), so we believe that assessing complaints four times daily is more valid than assessing complaints once-per-day. A consequence of this assessment difference is that the clinical outcome of our study cannot be compared with that found by others. Second, in contrast to other studies, the present study evaluated a rather large patient sample \((N = 43)\) using a controlled design and a long term follow-up. Other studies evaluating group therapy for patients with IBS either were uncontrolled, used self-selected patients or only small samples, or only investigated short term effects (21, 22).

The present study fulfilled a number of Klein's stringent criteria necessary for a satisfactory treatment trial in irritable bowel syndrome (10). The study used clear entry criteria and demographic features to establish the generalizability of the results, adequate sample sizes, valid statistical techniques, a treatment trial of sufficient length (3 months), baseline comparisons between conditions, and relevant outcome measures. Nevertheless, this study had methodological limitations. The first shortcoming is the not strict random allocation. A period of allocation to the experimental group was followed by a period of allocation to the waiting list and so on. A consequence of a complete random design would have been that the patients in the waiting list had to wait a long time before treatment could be offered to them. Despite this lack of proper randomization, our experimental and control group
COGNITIVE-BEHAVIORAL GROUP THERAPY FOR IBS

appeared to be similar with respect to baseline measures and as a consequence were considered to be well matched. Therefore, we believe that this shortcoming has not distorted our findings.

Furthermore, an attention placebo control group would have been a better comparison. In psychotherapy research, however, it is very difficult to provide proper attention placebo treatment. Also, the therapist can seldom or never be blinded. We saw improvements in the number of coping strategies and in the avoidance behavior corresponding with the aim of the treatment, so it is less probable that these positive results arose just because patients were given attention, regardless of its nature.

Moreover, we investigated the short term effects of cognitive-behavioral group treatment in a controlled trial. In investigating the long term effects, no waiting list control group was used because waiting to start treatment for an average of 2 years was believed to be unethical. Patients could have improved regardless of the followed treatment. However, the prognosis of irritable bowel syndrome is known to be poor (7–9). In this light, our results can considered to be positive. It would increase the impact of our findings if future studies could prove the long term superiority of cognitive-behavioral group treatment over the natural course or usual care for patients with IBS.

One could furthermore argue that we should have adjusted our a-level for the number of comparisons (36, 37). Yet, following the purpose of the treatment offered, we expected to find differences in a restricted number of six outcome measures. Even if we would have adjusted our a-level, significant improvements in Daily Abdominal Complaint score and daily avoidance would continue to exist.

Finally, in the present study, IBS was defined broadly, based on the Rome criteria (5), as abdominal pain with or without disordered defecation not explained by any structural or biochemical abnormalities; all patients in this study appeared to suffer from abdominal pain with or without disordered defecation. In the more restrictive definition of IBS (6), abdominal pain and disordered defecation must both be present. For that reason, in studies on patients with restrictive IBS treatment, outcome is evaluated on the basis of changes in pain and constipation and/or diarrhea (22). In agreement with our definition of IBS, we calculated a reduction score based on changes in abdominal pain only, without taking into account possible disordered defecation.

The clinical outcome of cognitive-behavioral group therapy might improve further by prolonging the treatment, assessing long term effects using a control group, or altering the size of the groups. The group treatment we offered consisted of only eight therapy sessions, less than the number of sessions offered by others (22). Although our short term treatment appeared to have long term effects, it is still possible that more sessions would increase the effectiveness of the treatment.

As a result of the broad criteria we used in defining irritable bowel syndrome, our results, strictly speaking, can be generalized reliably only to the subgroup of broadly defined abdominal pain-predominant refractory outpatients with irritable bowel syndrome. Yet, gastrointestinal complaints associated with restrictively defined irritable bowel syndrome were also found to improve during follow-up, so the present results can possibly be generalized to this group of patients, too.

Remarkably, patients’ psychological well being did not appear to improve as a result of the group treatment. However, improvement in psychological well being appeared to be related to improvement in abdominal complaints. Results of a cross-lagged panel analysis (38) confirmed that the level of patients’ psychological well being is a consequence of the daily observed intensity of the abdominal complaints rather than a precursor, consistent with the results of an earlier study that found that physical health had a positive effect on mental health (38, 39).

Short term cognitive-behavioral group treatment appears to have a long term effectiveness in alleviating abdominal complaints of referred patients with severe and chronic IBS. Therefore, doctors are recommended to refer their patients with refractory IBS at an early stage to cognitive-behavioral therapy.

This study was funded by the Centre for Women’s Studies, University of Nijmegen, The Netherlands.

REFERENCES


29. Kraaimaat FW, Schevikhoven REO van: Inventarisatielijst Pijngedrag (Pain-Coping Inventory). Utrecht, The Netherlands, Department of Clinical Psychology, Academic Hospital, 1984