A randomized, double-blind, placebo-controlled trial of a Bach Flower Remedy

N.C. Armstrong and E. Ernst

The aim of this randomized, double-blind, placebo-controlled trial was to investigate the efficacy of ‘Five Flower Remedy’ also known as Rescue Remedy, for examination anxiety in healthy university students. One hundred university students were recruited. They were registered at the University of Exeter, aged between 18 to 65 years, and enrolled to undertake university examinations between May to July 1998. Participants were randomized to take one to four doses of either ‘Five Flower Remedy’ or indistinguishable placebo during day 1 to 7 of the experiment. Self-reported anxiety was quantified using the 40-item Spielberger State-Trait Anxiety Inventory (STAI) as the primary outcome measure. Visual analogue scales were used as the secondary outcome measure to allow the determination of daily anxiety scores. Forty five percent of volunteers completed the study (21 subjects in the experimental and 24 in the placebo group). Reporting of anxiety, as measured by the state component of the STAI at enrolment, was not found to differ between verum and placebo and no gender differences were noted. However, self-reports of trait anxiety were found to be lower in men compared to women at enrolment. Eight days before examination, men reported significantly lower levels of state anxiety but no significant differences were subsequently identified the evening before the examination. No significant differences were identified between or within groups for mean values of the daily VAS scores or indeed the linear contrast variable constructed to determine any trends associated with the weekly VAS reports. It is concluded that ‘Five Flower Remedy’ had no specific effects in treating anxiety under these trial conditions.

The Welsh physician and homoeopath Dr Edward Bach (1886–1936) created the treatments which today are known under the name ‘Bach Flower Remedies’ (or Flower Remedies). Essentially Dr Bach believed that most illnesses are caused by negative states of mind (e.g. fear, jealousy, despair). In his native Welsh countryside he identified a total of 38 remedies which, according to his conviction, would alleviate such negative feelings and thus would help to restore health. Dr Bach therefore did not focus on diagnoses but claimed to treat the whole person; two patients with the same disease might well be treated with two different remedies. In this and several other respects, Dr Bach was strongly influenced by the teachings of homoeopathy.

Like homoeopathic remedies, ‘Bach Flower Remedies’ (BFRs) are highly dilute. They are usually produced by dropping fresh flowers into water to produce the ‘mother tincture’ which is subsequently added to brandy (as a preservative) to make up the medicine. BFRs do not contain pharmacological amounts of constituents of the flowers they originate from. Proponents of this treatment claim that their mode of action does not depend on a molecular mechanism comparable to conventional pharmaceuticals. Much like homoeopathic remedies, they are believed to work through the ‘energy’ that is transmitted from the flowers to the remedy (NN 1995). As this ‘energy’ (so far) defines quantification, critics would argue that BFRs are pure placebos, and any perceived clinical effects of BFRs must be caused purely by the placebo or other non-specific effects (Frick 1999). Meanwhile, BFRs have become a thriving business, and many consumers strongly believe in the usefulness of these treatments.

Very few clinical trials exist which aim to resolve the controversy as to whether BFRs are...
associated with specific or non-specific therapeutic effects. The aim of this trial was to test whether ‘Five Flower Remedy’ (also marketed as ‘Rescue Remedy’), one of the most popular BFRs, is efficacious in reducing examination anxiety in university students.

SUBJECTS AND METHODS

Subjects

University students were found through a series of advertising activities and were given a verbal and written explanation of the study. Interested volunteers were screened according to predefined inclusion/exclusion criteria. The inclusion criteria included current registration at the University of Exeter, age between 18 to 65 years, and registration to under-take university examinations between May to July 1998. All volunteers indicated that they experienced some degree of anxiety or ‘exam nerves’ before taking examinations. Exclusion criteria included pregnancy, concurrent treatment for depression, concurrent use of antipsychotic medication, anxiolytic agents, other flower remedies, and/or other psychotherapeutic approaches to anxiety relief.

Sample size calculation

Sample size was calculated using the approximate formula (Senn 1997) for sample size:

$$n = \frac{(Z_{\alpha/2} + Z_{\beta})^2 \sigma^2}{\Delta^2}$$

such that the value of $\alpha$ was set at 0.05, the value $\beta$ selected was 0.2 to allow the detection of any significant differences with 80% power, with $\sigma$ (the standard deviation) estimated as 11.41 (Spielberger 1983) and a clinically relevant difference ($\Delta$) of 6 using data from previous studies (Spielberger 1983). This yielded an approximate sample size of 41 per group.

Randomization and concurrent controls

Volunteers were randomized using block randomization (block size of four) to receive either Five Flower Remedy or an indistinguishable placebo. Both were prepared by Healing Herbs. Five Flower Remedy is the rescue remedy combination according to Dr Bach containing the following flower essences: Prunus cerasifera, Clematis vitalba, Impatiens glandulifera, Helianthemum nummularium and Ornithogalum umbellatum. Verum vials were prepared by adding 2 drops of Five-Flower essence to 10 ml of brandy (40% alcohol by volume). Placebo remedy was prepared in a similar manner with the exclusion of addition of Five-Flower essence. Placebo and verum vials were separately packed in cardboard boxes and posted to the Department of Complementary Medicine, University of Exeter where they were received by the study pharmacist who independently assigned a study number to each bottle.

Dosing regimen

Participants were instructed to take one to four doses of their assigned remedy per day and to follow this regimen between day 1 to 7 of the trial period. One dose (as recommended by the manufacturer) comprised 4 drops of remedy in a small glass of water to be sipped frequently over an unspecified period of time. Benefit is claimed to be derived from small, regular use rather than by the volume of remedy that is taken.

Measurement of anxiety

The 40-item Spielberger State-Trait Anxiety Inventory (STAI) was the primary outcome measure used to quantify anxiety (Spielberger 1983) experienced by the participants. It comprises two separate scales namely, the State anxiety score which aims to quantify the extent to which an emotional state exists at a given moment in time at a particular level of intensity and the Trait anxiety score which refers to relatively stable individual differences in anxiety-proneness. The stronger the anxiety Trait, the more probable that the individual will experience more intense elevations in State anxiety in a threatening situation. This two-component self-evaluation questionnaire was administered three times throughout the investigation: at induction, 7 days before the examination specified as likely to cause most anxiety to the student, the night before examination and after the examination. Visual analogue scales with polar opposite concepts (‘no anxiety’ and ‘worst imaginable anxiety’) were used as the secondary outcome measure to allow the determination of daily anxiety scores.

Adverse effects data and compliance

Participants were asked to indicate the number of doses (0 to a maximum of 4) used per day and were invited to detail any perceived adverse effects during the course of the investigation.

Follow-up

Reminder letters were sent to participants who had not returned their study documents at the expected date of return. One month later, all non-responders were sent a short multiple choice questionnaire in order to remind those who had completed but had still not returned the study
documents to do so and to ascertain the reasons for non-completion.

**Ethical approval and consent**

Ethical approval was obtained from the local ethics committee (South and West Local Research Ethics Committee, Department of Medical Affairs, Royal Devon & Exeter Hospital, Barrack Road, Exeter, EX2 5DW). A clinical trial certificate was deemed unnecessary by the Medicines Control Agency. Written informed consent was obtained from all participants.

**Statistical analysis**

Blinded data analysis was carried out using SPSS version 8.0 for Windows. $P$ values of less than 0.05 were considered statistically significant. Analysis of covariance was used to produce a measure of anxiety which is adjusted by baseline in such a way that the result is uncorrelated with the baseline. The analysis of covariance estimator also has the advantage that its variance is generally lower than that using raw outcomes or change scores. The linear contrast variable was calculated from daily VAS score for anxiety, designed to isolate the slope of any underlying linear trend. Using the following equation to model the linear relationship, $\text{VAS}_{LC} = 3(V_7) + 2(V_6) + (V_5) - (V_3) - 2(V_2) - 3(V_1)$ where $V$ is the VAS score for each day (indicated by the subscript number), $V_4$ having a zero coefficient, then assuming that if $V_t = a + bt + \text{error}$ where $V$ is the VAS score, $a$ is the fixed constant, $b$ is the slope of the linear relationship and $t$ is the time point, then $\text{VAS}_{LC} = 28b + \text{error}$. Values are expressed as $\text{VAS}_{LC}/28$ in order to give units as mm per day (mmd$^{-1}$).

**RESULTS**

Of the 100 subjects who were recruited and randomized, 51 were assigned to Five Flower Remedy (19 men, 32 women) and 49 to placebo (14 men, 35 women). 45% of the participants completed the study; 21 subjects in the experimental group and 24 subjects in the placebo group (Fig. 1). The relative percentages of each gender completing remained comparatively constant within the groups. Of those subjects who did not complete the study, three withdrew, 39 were lost of followup and 13 gave reasons for non-participation (Fig. 1). Eight of those non-completing subjects who gave reasons for non-participation took the assigned remedy for a number of days (Five Flower Remedy 1.80 ± 0.84, placebo 2.00 ± 1.73; two-tailed $t$-test $P = 0.099$) before terminating. Reasons for non-completion in this group included forgetting to take remedy (four subjects), using all the remedy before 7 days (one subject) not finding remedies palatable (one subject), breaking remedy bottle (one subject), perceiving an adverse effect (one subject). Of those who reported that they did not take the remedy at all, 60% (three subjects) reported that they had forgotten to take the remedy, 20% (one subject) reported that the remedy was not palatable and 20% (one subject) perceived the participant burden to be too high.

Table 1 compares the baseline characteristics of participants and non-participants. The average age of those recruits who went on to complete the study was very similar to those

---

**Fig. 1** Trial profile.
who did not complete. Participants reported that they experienced examination nerves to an equal degree between groups (Pearson two-sided chi-square $p = 0.387$). Similar proportions of participants and non-participants were smokers (Pearson two-sided chi-square $p = 0.145$) but of those who did smoke, non-participants reported greater cigarette consumption (two-tailed $t$-test $p = 0.006$). Equal proportions of participants and non-participants reported that they consumed alcohol (Pearson two-sided chi-square $p = 0.94$). However, reported alcohol intake was greater in non-participants (two-tailed $t$-test $p = 0.035$). Reported frequency of exercise session lasting more than 20 min was comparable between groups (Pearson two-sided chi-square $p = 0.163$). Both components of the STAI were similar (state anxiety, participants $43.4 \pm 11.2$ non-participants $43.7 \pm 11.2$; two-tailed $t$-test $p = 0.884$; trait anxiety, participants $41.4 \pm 11.2$ non-participants $41.8 \pm 10.1$; two-tailed $t$-test $p = 0.844$). Concurrent non-parametric tests (Mann–Whitney $U$ tests) were found to agree with all the above parametric tests.

Table 2 compares the baseline characteristics of the experimental and the control groups. Similar proportions of men and women in both groups completed the study; these were broadly representative of the proportions upon recruitment. Within the experimental group, 14 were male (31%) and 31 (69%) were female and within the placebo group, 19 (35%) were male and 36 (65%) were female. No significant differences between or within groups were found for mean age of participants and no interactions between the main effects were identified (two-way ANOVA $p = 0.566$, $p = 0.138$ and $p = 0.456$ respectively). Year of study and reported ‘examination nerves’ seem similar across groups and gender, as do the proportions of smokers and consumers of alcohol. Reported number of exercise sessions per week was not significantly different between or within groups and no interaction between these factors was identified (two-way ANOVA $p = 0.718$, $p = 0.319$ and $p = 0.977$ respectively). Reporting of situational anxiety (State at time 0) was not found to differ significantly according to group allocation or gender of the participant. Again no interaction between the main effects was found (two-way ANOVA $p = 0.936$, $p = 0.110$ and 0.763 respectively). A main effect for gender was identified for reported trait anxiety upon enrolment. When asked to ‘describe how you generally feel’ (trait anxiety score), men consistently scored lower than women (Trait at time 0) with no statistical significance for group allocation or interaction (two-way ANOVA $p = 0.828$, $p = 0.043$ and $p = 0.586$ respectively).

Table 3 compares the treatment outcomes between the experimental and the control groups. Self-reports of dose of remedy taken were not found to significantly differ according to treatment group allocation or gender of the participant. A statistically significant interaction was identified which indicated that the men in the placebo group took significantly fewer doses of their allocated remedy (two-way ANOVA $p = 0.014$). Men were found to report significantly lower levels of anxiety eight days before the examination (State 1 score) identified as most likely to cause increased anxiety as measured by the State component of the STAI (two-way ANOVA $p = 0.001$). No statistically significant

---

### Table 1: Baseline characteristics of participants and those who were either lost to follow-up or withdrew. Categorical variables represent number (%) of responding subjects

<table>
<thead>
<tr>
<th>Baseline measurements</th>
<th>Participants</th>
<th>Non-participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men (n = 14)</td>
<td>Women (n = 31)</td>
</tr>
<tr>
<td></td>
<td>Men (n = 19)</td>
<td>Women (n = 36)</td>
</tr>
<tr>
<td>Mean (SD) no. of exercise sessions</td>
<td>19.7 (0.7)</td>
<td>22.8 (6.7)</td>
</tr>
<tr>
<td>Experience exam. nerves</td>
<td>12 (86)</td>
<td>23 (74)</td>
</tr>
<tr>
<td>Experience exam. nerves</td>
<td>1 (7)</td>
<td>4 (13)</td>
</tr>
<tr>
<td>Experience exam. nerves</td>
<td>1 (7)</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Experience exam. nerves</td>
<td>0 (0)</td>
<td>3 (10)</td>
</tr>
<tr>
<td>Experience exam. nerves</td>
<td>2 (14)</td>
<td>15 (48)</td>
</tr>
<tr>
<td>Experience exam. nerves</td>
<td>1 (7)</td>
<td>12 (42)</td>
</tr>
<tr>
<td>Experience exam. nerves</td>
<td>1 (7)</td>
<td>3 (10)</td>
</tr>
<tr>
<td>Cigarette smoking</td>
<td>10 (71)</td>
<td>20 (65)</td>
</tr>
<tr>
<td>Cigarette smoking</td>
<td>4 (29)</td>
<td>11 (35)</td>
</tr>
<tr>
<td>Mean (SD) age (years)</td>
<td>19.7 (0.7)</td>
<td>22.8 (6.7)</td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td>9.3 (79)</td>
<td>6.9 (5.5)</td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td>0 (0)</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Mean (SD) Units/wk</td>
<td>4 (29)</td>
<td>11 (35)</td>
</tr>
<tr>
<td>Mean (SD) no. of exercise sessions</td>
<td>1 (7)</td>
<td>12 (42)</td>
</tr>
<tr>
<td>Total</td>
<td>1 (7)</td>
<td>12 (42)</td>
</tr>
<tr>
<td>Trait score</td>
<td>39 (29)</td>
<td>45.3 (11.1)</td>
</tr>
<tr>
<td>Trait score</td>
<td>36.6 (11.5)</td>
<td>43.5 (9.4)</td>
</tr>
</tbody>
</table>
differences between the experimental and placebo groups were identified for the State anxiety score measured on the evening before the examination (State 2 score). Analysis of covariance was carried out on the State 2 score using ANOVA. Men were found to report lower anxiety as measured on the VAS scale than those in the placebo group (two-way ANOVA p = 0.011 and p = 0.048 respectively). No statistically significant differences were found for the mean of the daily VAS scores. The linear contrast variable for the weekly VAS reports was not found to be significantly different between or within groups.

Daily mean VAS scores along with the standard error of the means are presented in Table 3 and Figure 2. No significant differences between the men and women in the experimental and placebo groups were found at the beginning of the study period. Men were found to report significantly less anxiety on the VAS scale than those in the placebo group (two-way ANOVA p = 0.048 respectively). On day 3 participants taking the active remedy reported significantly less anxiety on the VAS scale than those in the placebo group (two-way ANOVA p = 0.041). No other statistically significant effects were observed over the study period.

Correlation between anxiety and dose taken shows no clear relationship. This was investigated according to treatment allocation for all
participating students. Three subjects from both the verum and the placebo group reported a total of five different perceived adverse effects. Those reported in the verum group included headaches (two subjects – one withdrew from the study after 3 days as a result) and skin eruptions. The reported adverse effects in subjects taking the placebo included vomiting before the examination, hayfever symptoms and depressive mood (subject withdrew as a result).

DISCUSSION

These data fail to provide convincing evidence to suggest that ‘Five Flower Remedy’ has specific therapeutic effects under the above experimental conditions. Participants taking the verum remedy reported lower anxiety than those taking placebo on day 3 of the study. However, this result needs to be interpreted with great caution as we have no a priori expectation for this finding – it is indeed more likely therefore that this may simply be the result of multiple comparisons. The gender differences as demonstrated in the lower state anxiety score at 8 days before examination (pre-remedy) and the VAS score on days 1 and 2 (after commencement of remedy) perhaps indicate that the men were slower to respond to the phenomenon of heightened feelings of anxiety during the period of undertaking examinations. This would explain the finding that both men and women reported a raised state anxiety score by the evening before the examination and the fact that the daily VAS scores were not significantly different for the remainder of the study period.

The particular model of ‘Five Flower Remedy’ and examination stress was chosen for several reasons. The remedy is popular, was specially created by Dr Bach for anxiety and is believed to work independently of the patient’s state of mind or personality (Chancellor 1971, Vlamis 1990). In books on complementary medicine it is referred to as ‘Dr Bach’s emergency stress formula’ (The Burton Goldberg Group 1994). Proponents point out that it has been successfully used for over 50 years and has a positive, calming effect in acute emergency situations including anxiety [8]. There is also ample anecdotal evidence to suggest that it works for stress-related indications (Chancellor 1971, The Burton Goldberg Group 1994, Vlamis 1990, Weisglas 1979). Moreover, it seemed that examination stress of university students is a model that is readily reproducible and easily studied.

It is on this latter assumption that the present study failed in a major way. We found it not easy to recruit students in sufficient numbers for the trial. More importantly, once we had randomised 100 participants, we were confronted with a drop-out rate much higher than anticipated. This occurred despite intense effort to follow up these individuals. The main reason for the high drop-out rate was complexity of the primary outcome measure: future trials should aim at improving compliance through using a simpler, more user-friendly instrument. Our initial power calculation was thus largely invalidated, and the results of our study are therefore in danger of
falling victim to a type II error. However, the results show no consistent trend in favour of active treatment over placebo, and drop-outs were in most respects similar to complying students (Table 1). Thus a type II error is an unlikely explanation for the negative results of this trial.

The above evaluation of the data refer to per protocol analyses. With an excessively high drop-out rate, one could argue that an intention to treat analysis would have been more appropriate. However, in view of the negative result of the analyses presented, an intention to treat analysis of our data would inevitably have resulted in negative results as well. It was thus decided that a per protocol analysis was a sufficient approach.

To date, only two other controlled trials of BFR have been published. One relates to a US student project which reported positive outcomes in terms of well being in non-clinical subjects treated with BFR (Weisglas 1979). The other trial was a pilot study from Germany (Von Ruhle 1995). Twenty-four women with overdue pregnancies were included in this investigation. Those receiving BFR reported some benefit which, possibly due to the small sample size, did not reach statistical significance. Both of these studies have obvious and serious methodological flaws which our trial attempted to overcome. It is thus the most rigorous test of BFRs published so far.

In the light of the paucity of research into BFRs, most experts and enthusiasts of BFRs agree on the need for more research (Howard 1998, Mantle 1997). The present study is the largest clinical trial of BFRs so far. However, its negative result should not be over-interpreted: we may have produced some evidence to suggest that ‘Five Flower Remedy’ is ineffective in reducing examination stress but we have certainly not shown that all BFRs are ineffective for all medical conditions.

In conclusion, this study provides no evidence for the efficacy of Five Flower Remedies under the given experimental conditions. Further research is needed for testing other BFRs for other indications.

Acknowledgements

The authors are grateful to J. Barnes for her input this study. NCA was supported by the Cecil Pilkington Trust. The study was furthermore supported by a grant from ‘The Twelve Healers Trust’. We thank ‘Healing Herbs’ for supplying both placebo and ‘Five Flower Remedy®’. We are also grateful to Dr Kenneth Read for statistical analyses.

REFERENCES

Fricke U 1999 Die Tops and Flops der Naturmedizin. Bild der Wissenschaft 11: 52–57
Senn S 1997 Statistical issues in drug development. John Wiley & Sons, Chichester
Spielberger CD 1983 State-trait anxiety inventory (Form Y). Mind Garden Inc. Palo Alto, California
The Burton Goldberg Group 1994 Alternative medicine – the definitive guide. Future Medicine Publ., Puyallup Was
Vlamis G 1990 Bach Flower Remedies to the rescue. Healing Arts Press, Rochester Vt
Weisglas MS 1979 Personal growth and conscious evolution through Bach Flower Remedies [dissertation] California Institute of Asian Studies