1. FIRST CONTROLLED TRIAL RESULTS

Summarised in Perrin RN, Edwards J and Hartley P An evaluation of the effectiveness of osteopathic treatment on symptoms associated with Myalgic Encephalomyelitis. A preliminary report. *Journal of Medical Engineering and Technology*, 22(1), 1-13.1998

Statistical methods

Symptom picture

Firstly, for each questionnaire, separate two-sample t-tests were performed on the difference of mean scores from start to months 3, 6, 9 and 12. The most important statistic after each 3 months is the p-value. Those tests with $p<0.05$ indicate significant differences between symptom reduction of the control group and the patient group at the 5% level of significance.

Secondly the estimated differences in mean reductions between the two groups were calculated, with corresponding 95% confidence intervals. These provide an alternative way of presenting the information instead of hypothesis tests. An interval that excludes zero corresponds with a significant test. Each interval is a range of values inside which the difference in the mean score of symptom reduction (controls minus patients) lies with 95% confidence (significant if the interval excludes zero). The confidence interval also provides information about the size of the difference (note that the estimated difference lies in the middle of the interval).

Finally, for each questionnaire, a one-sample t-test was performed for patients only, to analyse any difference from month 12 until a year follow up. Ho = increase in mean score indicating a worsening of symptoms over the follow-up period.

Muscle fatigue

The mean results of the knee extensor fatigability tests were statistically analysed using a paired t-test comparing torque X time before and after the year-long study.

Results

Self report questionnaires

The tables and bar charts below represent the mean score obtained from all the patients for each of the questionnaires at intervals during the 12 month test period, and subsequently at a one year follow up. The score values are expressed as the mean from the group of subjects as a percentage of the maximum severity of the symptom

Mean value results

In the table below (Table 1), the columns represent the mean value of each questionnaire result where 100% = the most severe case possible and 0% = symptom free.

Beneath the means are the p-values calculated using two-sample t-tests on the reductions in the scores from the start to 3, 6, 9, and 12 months respectively, comparing the mean reduction of symptoms for both groups. Those tests with $p<0.05$ indicate significant differences between the mean reductions of the control group and the patient group at the 5% level of significance. Thus if $p>0.05$ then it would signify that there was no statistical difference between the mean improvement of the patient group (1) and the control group (2). A p-value of <0.05 in all the questionnaires would show a significant improvement in the overall health of the patient group compared to that of the control.

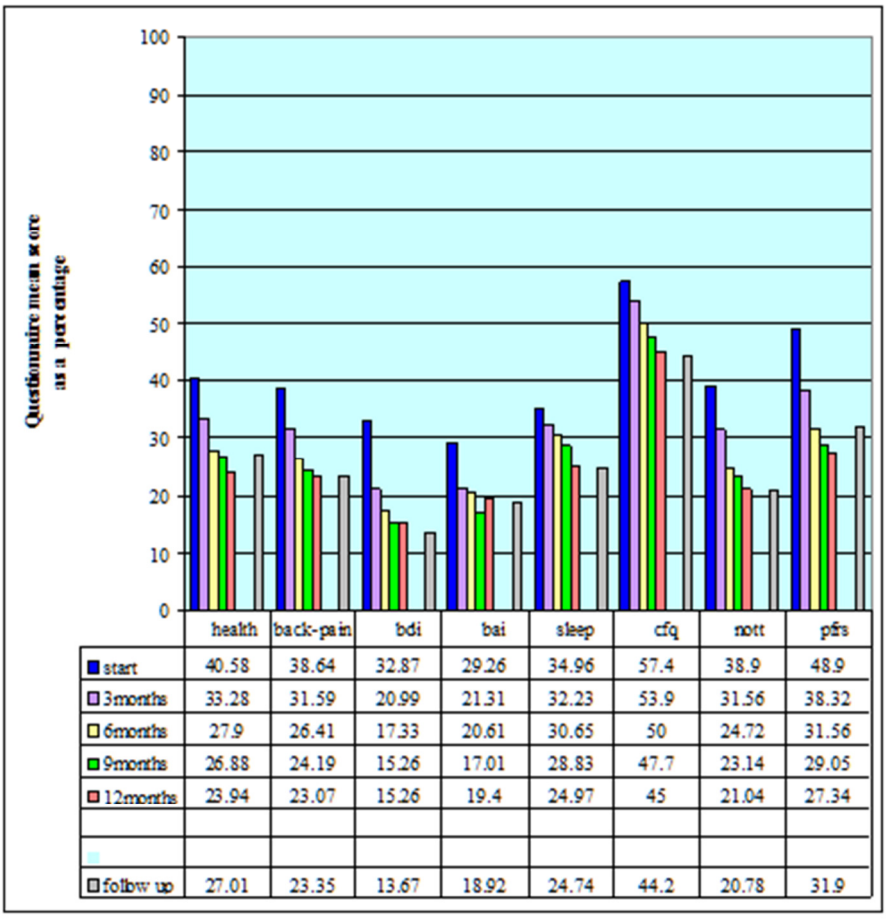
Note that the final row of Table 1 only pertains to the patient group, and shows the mean scores in the follow up a year after the end of the trial. The p-values in this row were calculated using a one-sample t-test. The scores at the end of the follow-up year were compared with the scores at 12 months. If there is no significant deterioration in the health of the patient group the p-values will all exceed 0.05.

The results are further illustrated graphically in the two bar charts Figs. 1 and 2. A reduction of symptoms is represented by a diminution of the height of the corresponding bar.

	health %	**back-pain %**	**Bdi %**	**Bai %**	**sleep %**	**cfq %**	**Nott %**	**Pfrs %**
Start								
Group 1	40.58	38.64	32.87	29.26	34.96	57.40	38.9	48.90
Group 2	33.92	26.97	25.44	24.80	42.68	61.40	31.1	44.95
3 months								
1.	33.28	31.59	20.99	21.31	32.23	53.9	31.56	38.32
2.	31.49	27.11	24.49	25.28	40.75	58.60	34.98	42.78
p-value	0.180	0.151	0.023	0.023	0.706	0.853	0.122	0.032
6 months								
1.	27.9	26.41	17.33	20.61	30.65	50.00	24.72	31.56
2.	34.05	28.36	21.31	28.46	40.18	61.5	37.35	44.80
p-value	0.002	0.005	0.033	0.001	0.535	0.004	0.001	0.000
9 months								
1.	26.88	24.19	15.26	17.01	28.83	47.70	23.14	29.05
2.	33.66	28.36	20.99	25.76	46.08	60.20	37.08	44.10
p-value	0.000	0.000	0.003	0.000	0.005	0.002	0.000	0.000
12 months								
1.	23.94	23.07	15.26	19.40	24.97	45.00	21.04	27.34
2.	36.22	29.19	22.66	27.51	41.54	61.70	39.45	47.43
p-value	0.000	0.000	0.000	0.000	0.006	0.000	0.000	0.000
Follow up (Group 1 only)	27.01	23.35	13.67	8.92	24.74	44.2	20.78	31.9
p-value	0.678	0.529	0.298	0.069	0.850	0.837	0.545	0.194

Table 1 Phase 1: Means for symptom severity showing statistical significance

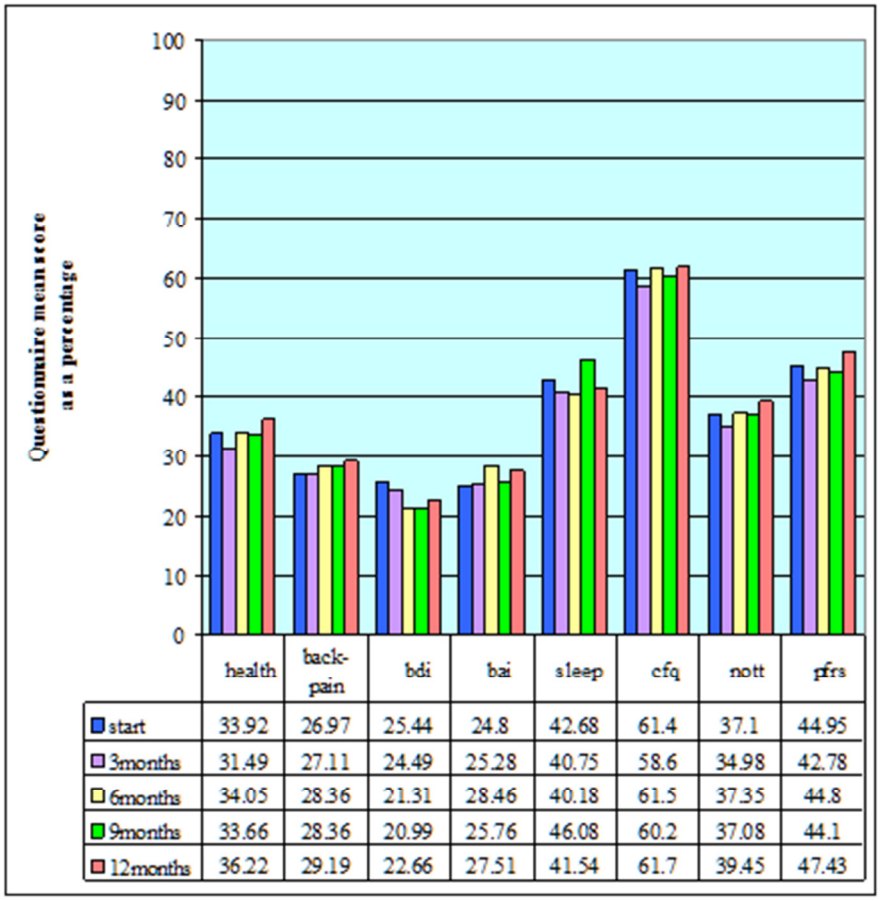
The columns represent the mean value of each questionnaire result where 100% = the most severe case possible and 0% = symptom free. The p-values were calculated using two-sample t-tests on the reductions in the scores from the start to 3, 6, 9, and 12 months respectively, comparing the mean percentage change in both groups. The p-values in the final row were calculated using a one-sample t-test, comparing the final scores with those at the end of the follow-up year.



	health	back-pain	bdi	bai	sleep	cfq	nott	pfrs
start	40.58	38.64	32.87	29.26	34.96	57.4	38.9	48.9
3months	33.28	31.59	20.99	21.31	32.23	53.9	31.56	38.32
6months	27.9	26.41	17.33	20.61	30.65	50	24.72	31.56
9months	26.88	24.19	15.26	17.01	28.83	47.7	23.14	29.05
12months	23.94	23.07	15.26	19.4	24.97	45	21.04	27.34
follow up	27.01	23.35	13.67	18.92	24.74	44.2	20.78	31.9

Figure 1 Bar chart of patient group's mean scores (phase 1)

The bar chart shows the changes in the questionnaire scores in the treated group over the course of the first clinical trial and the year follow-up. The questionnaires were as follows: Column 1. Health = A general health questionnaire examining most common symptoms of CFS/ME; Column 2. examined back pain; Column 3. bdi = The revised Beck depression inventory; Column 4. bai = The Beck anxiety inventory; Column 5 = The Morgan-Gledhill sleep questionnaire; Column cfq = Broadbent's cognitive function questionnaire; Column 7. = The Nottingham Health Questionnaire which concentrates on symptoms of fatigue; Column 8. pfrs= The Profile of Fatigue Related States. As all the questionnaires have different scales and scoring systems, the results were converted to a percentage of the maximum score possible.



	health	back-pain	bdi	bai	sleep	cfq	nott	pfrs
start	33.92	26.97	25.44	24.8	42.68	61.4	37.1	44.95
3months	31.49	27.11	24.49	25.28	40.75	58.6	34.98	42.78
6months	34.05	28.36	21.31	28.46	40.18	61.5	37.35	44.8
9months	33.66	28.36	20.99	25.76	46.08	60.2	37.08	44.1
12months	36.22	29.19	22.66	27.51	41.54	61.7	39.45	47.43

Figure 2 Bar chart of control group's mean scores (Phase 1)

The bar chart shows the changes in the questionnaire scores in the untreated control group over the course of the first clinical trial. As in Fig. 1 the results were converted to a percentage of the maximum score possible for each questionnaire.

Muscle fatigue

The individual results of the isometric tests are listed in Table 2. Torque X time, the impulse torque, measured in Newton metre seconds (Nms) for each member of the patient group before and after the year's treatment was recorded. The difference and percentage change in impulse torque was calculated and is inversely proportional to the post exercise fatigue, thus the greater the percentage change the less the fatigue. This test was only carried out on the patient group as the control group members of phase 1 were completely anonymous to the author having been independently selected by Action for M.E. Also the control group members were CFS/ME sufferers living throughout the UK. The task of bringing them to Salford for the tests to be carried out would have proved logistically too difficult and possibly stressful and potentially harmful to the subjects.

There are only five negative values with only one, RP13, recording a statistically significant lower value of impulse torque after the year. All the other patients showed an increase in the impulse torque with some over 100% improved on their first score.

The mean results of the knee extensor fatiguability tests for all subjects in the patient group were statistically analysed using a paired t-test, comparing the impulse torque before and after the year-long study.

The mean value at the beginning was 1791Nms and at the completion had risen significantly to 2266 Nms. The results of the paired t-test on the above data are shown in Table 3, where the null hypothesis was that the mean impulse torque /Nms of right knee during first 30 seconds of final push was the same for tests 1 and 2.

Patient Code No.	Test 1 (year start)/Nms	Test 2 (year end) /Nms	Difference Between	% Change
RP01	606	2211	1605	264.85
RP02	1077	1767	690	64.07
RP03	512	1255	743	145.12
RP04	716	1783	1067	149.02
RP05	732	1182	450	61.48
RP06	1264	1559	295	23.34
RP07	1982	2641	659	33.25
RP11	583	711	128	21.96
RP12	1691	2971	1280	75.69
RP13	1917	1366	-551	-28.74
RP14	5033	5014	-19	-0.38
RP15	2484	2945	461	18.55
RP16	2653	2990	337	12.7
RP17	1807	2604	797	44.11
RP19	245	704	459	187.35
RP20	518	826	308	59.46
RP21	659	1331	672	101.97
RP23	4776	4670	-106	-2.22
RP24	475	1179	704	148.21
RP25	1440	1777	337	23.4
RP26	5140	4658	-482	-9.38
RP27	277	1343	1066	384.84
RP28	1793	2118	325	18.13
RP29	1109	1320	211	19.03
RP30	1594	1866	272	17.06
RP31	700	863	163	23.29
RP32	1695	3221	1526	90.03
RP33	3002	2927	-75	-2.5
RP34	3698	3742	44	1.19
RP35	1727	2658	931	53.91
RP36	2757	3476	719	26.08
RP38	1364	1998	634	46.48
RP40	3093	3103	10	0.32
MEAN	1791.48	2266.03	474.55	26.49

Table 2 Phase 1: Impulse torque /Nms of right knee during first 30 seconds of final push

Column 3 = impulse torque at the start of the research. Column 4 = impulse torque at the end of the year. Column 6 = column 4 - 3. The final column shows the percentage change in the impulse torque. A minus sign indicates a reduction in the impulse torque.

	Number of subjects	Mean Impulse Torque/ Nms	Standard Deviation
Test 1	33	1791	1,353
Test 2	33	2266	1,170
		Mean Difference = 475	% Mean Difference 26.49

Table 3 Mean difference of impulse torque in patient group (phase 1)

The table shows the mean difference in impulse torque, pre and post treatment over a year, calculated from the results in Table 2.

Estimated difference in means m2- m1 is 474.5 with 95% confidence interval (296.4, 652.7). m1 = m2 has a two-sided test statistic of 5.43 on 32 degrees of freedom, resulting in a p-value of 0.000 to three decimal places. Hence, the null hypothesis H0 was rejected at the 5% (even at the 0.1%) level demonstrating that the means are significantly different.

2. SECOND CONTROLLED TRIAL RESULTS

Further analysis of the changes in muscle fatigue following the treatment were carried out following a second clinical trial published in the International Journal of Osteopathic Medicine (2011)