PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	A preliminary prospective study of nutritional, psychological and combined therapies for Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) in a private care setting.
AUTHORS	Arroll, Megan; Howard, Alex

VERSION 1 - REVIEW

REVIEWER	Dr Jonathan Price Clinical Tutor in Psychiatry University of Oxford I have no conflicts of interest to declare
REVIEW RETURNED	10-Apr-2012

THE STUDY	I enjoyed reading this manuscript, which is generally very well presented and easy to read and understand. I have the following comments to amplify my 'yes/no' responses above: - study design. The research question is 'to evaluate whether treatments X/Y/Z reduce CFS-ME symptoms over a period of three months, and whether there are differences in treatment effectiveness between X, Y and Z'. The correct study design for this question is the RCT. The authors mention the preliminary nature of this study but, in my view, they fall into the trap of assigning causality to changes in symptoms which are unjustified with this level of evidence. There are several important limitations which very significantly limit the ability to assign causality - non-randomised design; high drop out rate (about 50%); and differential drop-out rate between groups (from 67% in the psychology group to 40% in the nutrition and combined groups). - methods. These ARE adequately described, apart from the description of the nature of the interventions. I am reasonably clear about the focus of each intervention, and some of their specific components, but I would like to know more about: - the intensity of the intervention (number of individual / group sessions, length of individual / group sessions, 'homework' expected if any), therapist involvement (training of therapists, seniority of therapists), mean hours of face to face psychology /nutrition intervention (if available) - main outcome measure. The outcomes measures chosen are appropriate, and well described. However, there does not appear to be a main outcome measure that is seen by the authors as being of key importance to patients. While the study is 'exploratory / preliminary', it would have been helpful to have a pre-determined primary outcome measure.

	- abstract / summary / key messages / limitations accurate. These are, in general, accurate, but I do think that they need to be much more limited in their claims about the effectiveness of these treatment modalities at this early stage in the development of the evidence. I do not agree that this study, for example, justifies the 'key message' 'patient-centred approaches for the management of ME-CFS reduce symptomatology over time', in the absence of a control group; or that 'psychological interventions can help individuals to regain a sense of control over their condition', when we have no data for two thirds of the psychological intervention group, who *might* have worsened, however unlikely this might seem to those of us who are 'prejudiced' in favour of such interventions. The core message, transmitted consistently in abstract, summary and key messages, needs to be 'this study provides early evidence that treatments X, Y and Z may be effective treatments for some people with CFS-ME, but, due to the study's methodological limitations, it is important that this possible treatment effect is investigated further in high quality randomised controlled studies.' - standard of written English acceptable for publication. In general, definitely yes. However, the manuscript needs careful proof-reading for multiple but minor errors of grammar and punctuation, e.g. bottom of page 3 - averse rather than adverse; bottom of page 3- missing comma after medical management; page 5 - 'patient satisfaction of such approaches CAM has been high'; ... ... - I have answered 'yes this is fine' to the final question on supplemental documents.
RESULTS & CONCLUSIONS	The significant limitations of this study (non-randomised design, high drop-out, differential drop-out between groups) need to be made more explicit in abstract, discussion, and key messages, in order for publication to be justified. My concern is that publication of the manuscript as it stands might encourage the view that 'this nutritional therapy' / 'this psychological therapy' / 'this combined therapy' is / are effective treatments for people with CFS-ME. Although this study presents preliminary evidence that this might be the case, as the authors point out in their final sentence (of a very long manuscript), random assignment is needed to provide good quality evidence.
REPORTING & ETHICS	The authors appear to be the CEO and Lead Researcher at a private health facility dedicated to the treatment of CFS-ME. Their commitment to the effective treatment of this disorder, and to research into effective treatments, is highly laudable, but their roles in this organisation and, by implication, their potential to 'benefit' from publication of research supporting their work, needs to be explicit as a potential 'conflict of interest'.

REVIEWER	Dr Esther Crawley Consultant Senior Lecturer School of Social and Community Medicine University of Bristol Bristol BS6 6AJ
REVIEW RETURNED	29-Apr-2012

THE STUDY	1. The study design is not appropriate to answer the research question "evaluate the effectiveness of treatments" because there isn't a control group. 2. There is insufficient detail in the methods. a)For example, it is not clear how many follow up questionnaires were sent out and when. Were they sent out at regular intervals on several occasions or just once? b) How many patients were approached? How many agreed to take part? Without this information, it is hard to know whether the patients recruited are representative of the population of patients attending the clinic. c) There is a high drop out rate. When did the drop outs happen? Immediately or at follow up? d) Page 11, line 33. The authors excluded cases with missing data. Was this missing data at assessment? Which groups were these patients allocated to? Did they have data at outcome? The authors should consider including the data and reporting N at assessment. 3. The statistical methods are not appropriate because the authors need to correct for variations at baseline. 4. The authors should refer to the recently published PACE trial, a large definitive trial investigating CBT, GET and adaptive pacing.
RESULTS & CONCLUSIONS	1. The tables are very unwieldy and need to be greatly improved. It may be easier to present the results when proper statistical analyses has been performed. 2. Throughout the results, the authors report “significant differences” (for example, Page 14, line 34). The authors need to clarify what they mean. I assume they are not talking about clinically significant change but what some authors call “statistically significant differences”. The Authors may wish to read Sterne and Davey-Smith discussion on the problems or reporting statistically significant results (BMJ 2001; 322(7280):226-231.) 3. In the discussion, the authors need to be clearer about the limitations of this study. For example, drop out was not equal across all groups. In the psychology group the drop out was highest with retention rates varying from 33% (14/42), 61% (27/44), 40% (31/52). However this is not discussed. 4. The authors need to take care over some of their conclusions. For example page 16, line 8: "Considering that the options available on the National Health Service, mainly CBT and GET, are often perceived as coping strategies at best, and physically damaging at worst" may be considered inflammatory by many of your readers. In addition, the statements such as " tailored treatments such as described here may be more palatable, and hence effective." (page 16) go beyond the evidence presented in this paper.

VERSION 1 – AUTHOR RESPONSE

Reviewer: Dr Jonathan Price
Clinical Tutor in Psychiatry
University of Oxford

I have no conflicts of interest to declare

I enjoyed reading this manuscript, which is generally very well presented and easy to read and understand.

- study design. The research question is 'to evaluate whether treatments X/Y/Z reduce CFS-ME symptoms over a period of three months, and whether there are differences in treatment effectiveness between X, Y and Z'. The correct study design for this question is the RCT. The authors mention the preliminary nature of this study but, in my view, they fall into the trap of assigning causality to changes in symptoms which are unjustified with this level of evidence. There are several important limitations which very significantly limit the ability to assign causality - non-randomised design; high drop out rate (about 50%); and differential drop-out rate between groups (from 67% in the psychology group to 40% in the nutrition and combined groups).
Response: The limitations have been discussed in more depth to avoid falling into the causality trap and highlighted by Dr Price and the differential drop-out rates have been discussed as per Dr Crawley's point below.
- methods. These ARE adequately described, apart from the description of the nature of the interventions. I am reasonably clear about the focus of each intervention, and some of their specific components, but I would like to know more about: - the intensity of the intervention (number of individual / group sessions, length of individual / group sessions, 'homework' expected if any), therapist involvement (training of therapists, seniority of therapists), mean hours of face to face psychology /nutrition intervention (if available)
Response: Extra detail has been added regarding the different programs.
- main outcome measure. The outcomes measures chosen are appropriate, and well described. However, there does not appear to be a main outcome measure that is seen by the authors as being of key importance to patients. While the study is 'exploratory / preliminary', it would have been helpful to have a pre-determined primary outcome measure.
Response: The outcome measure have now been divided into primary, secondary ME/CFS specific and secondary psychological for clarity. This has led to newly devised tables which we hope will clarify the importance of the functional ability and fatigue scales and also simplify the results section.
- abstract / summary / key messages / limitations accurate. These are, in general, accurate, but I do think that they need to be much more limited in their claims about the effectiveness of these treatment modalities at this early stage in the development of the evidence. I do not agree that this study, for example, justifies the 'key message' 'patient-centred approaches for the management of ME-CFS reduce symptomatology over time', in the absence of a control group; or that 'psychological interventions can help individuals to regain a sense of control over their condition', when we have no data for two thirds of the psychological intervention group, who *might* have worsened, however unlikely this might seem to those of us who are 'prejudiced' in favour of such interventions. The core message, transmitted consistently in abstract, summary and key messages, needs to be 'this study provides early evidence that treatments X, Y and Z may be effective treatments for some people with CFS-ME, but, due to the study's methodological limitations, it is important that this possible treatment effect is investigated further in high quality randomised controlled studies.'
Response: The language throughout the paper has been tempered so that the message reflects the data at hand more accurately and included Dr Price's suggestion for the core message within the 'key message' section of the paper.
- standard of written English acceptable for publication. In general, definitely yes. However, the manuscript needs careful proof-reading for multiple but minor errors of grammar and punctuation, e.g.

bottom of page 3 - averse rather than adverse;
bottom of page 3- missing comma after medical management;
page 5 - 'patient satisfaction of such approaches CAM has been high';
Response: The paper has been proof-read by a professional proof-reader and errors corrected.
As discussed above, the significant limitations of this study (non-randomised design, high drop-out, differential drop-out between groups) need to be made more explicit in abstract, discussion, and key messages, in order for publication to be justified. My concern is that publication of the manuscript as it stands might encourage the view that 'this nutritional therapy' / 'this psychological therapy' / 'this combined therapy' is / are effective treatments for people with CFS-ME. Although this study presents preliminary evidence that this might be the case, as the authors point out in their final sentence (of a very long manuscript), random assignment is needed to provide good quality evidence.
Response: As per point above, the language throughout the paper has been tempered so that the message reflects the data at hand more accurately.
The authors appear to be the CEO and Lead Researcher at a private health facility dedicated to the treatment of CFS-ME. Their commitment to the effective treatment of this disorder, and to research into effective treatments, is highly laudable, but their roles in this organisation and, by implication, their potential to 'benefit' from publication of research supporting their work, needs to be explicit as a potential 'conflict of interest'.
Response: Conflict of interest section has been amended as above.

Reviewer: Dr Esther Crawley
Consultant Senior Lecturer
School of Social and Community Medicine
University of Bristol

1. The study design is not appropriate to answer the research question "evaluate the effectiveness of treatments" because there isn't a control group.
Response: This has been changed to: This study therefore aims to provide preliminary evidence for the utility of three types of patient-centered approaches to the management of ME/CFS over time (baseline and follow-up) offered at a private health-care center in the UK.
2. There is insufficient detail in the methods.
a)For example, it is not clear how many follow up questionnaires were sent out and when. Were they sent out at regular intervals on several occasions or just once?
Response: The following detail has been added: Those that expressed an interest (N = 145) were emailed a spreadsheet that contained the questionnaires and asked to complete it at their convenience. Subsequently, participants were requested to complete the questionnaire pack on a second occasion, three months from the baseline measures.
b) How many patients were approached? How many agreed to take part? Without this information, it is hard to know whether the patients recruited are representative of the population of patients attending the clinic.
Response: Further detail has been added for clarification: Of the 145 individuals who expressed an interest in the study, 142 time-one questionnaires were returned, equating to a 97.9% response rate at baseline (two participants from the psychology group and one from the combined group dropped out at this stage). Therefore, excluding the four cases deleted due to insufficient data, 138 One-hundred and thirty-eight cases were used for baseline analysis; individuals completed the questionnaire battery at time-one (excluding the four deleted cases); 42 participants in the psychology group, 44 in the nutrition group and 52 in the combined group.
c) There is a high drop out rate. When did the drop outs happen? Immediately or at follow up?
Response: As can been seen from the additional information above, the drop-outs occurred at follow-up.
d) Page 11, line 33. The authors excluded cases with missing data. Was this missing data at assessment? Which groups were these patients allocated to? Did they have data at outcome? The

authors should consider including the data and reporting N at assessment.
Response: Four cases were excluded at baseline due to incomplete data as in some cases numerous pages of the questionnaire were missing, i.e. more than 5% of data was missing hence case deletion was conducted to deal with this as advised by Tabachnick and Fidell (Using multivariate statistics. 4th ed. Needham Heights, MA: Allyn & Bacon; 2001.) The following detail has been added to clarify which groups the excluded cases were from: therefore these were excluded from the analysis (one individual from the nutrition group and three from the combined group).
3. The statistical methods are not appropriate because the authors need to correct for variations at baseline.
Response: The statistical test used to investigate differences from time-one to time-two were Wilcoxon Signed Ranks test as the data was not suitable for parametric tests. We accept the point that comparisons across groups should include correction for variations at baseline (here general fatigue, physical fatigue and swollen lymph nodes and glands), which would involve either analysis of covariance (ANCOVA) or multivariate analysis of covariance (MANCOVA), both of which are not suitable for non-parametric data. However, we have now conducted these tests on the outcome variables with controls for baseline variation and the results were non-significant. Therefore, we have removed these findings from the paper and stated that with control for variation at baseline, no significant differences were observed between the three groups.
4. The authors should refer to the recently published PACE trial, a large definitive trial investigating CBT, GET and adaptive pacing.
Response: This study has now been referred to in the introduction.

5. The tables are very unwieldy and need to be greatly improved. It may be easier to present the results when proper statistical analyses has been performed.
Response: The tables have been shortened to take into account the division of outcomes into primary and secondary and some tables have been removed to reflect the reanalysis.
6. Throughout the results, the authors report “significant differences” (for example, Page 14, line 34). The authors need to clarify what they mean. I assume they are not talking about clinically significant change but what some authors call “statistically significant differences”. The Authors may wish to read Sterne and Davey-Smith discussion on the problems or reporting statistically significant results (BMJ 2001; 322(7280):226-231.)
Response: We have now noted that the significant differences are statistical, rather than clinical.
7. In the discussion, the authors need to be clearer about the limitations of this study. For example, drop out was not equal across all groups. In the psychology group the drop out was highest with retention rates varying from 33% (14/42), 61% (27/44), 40% (31/52). However this is not discussed.
Response: This has now been discussed: Also, there was a high drop-out rate from time-one to time-two and this rate differed across groups. The highest drop-out rate was in the psychology group; whilst we cannot be sure why this occurred, it is postulated that the retention was poor in the group as the individuals in the psychology program had more activities to engage in and may have felt overburdened with the research questionnaires in addition to their session and homework (this would not be the case in the combined group as the therapeutic activities are phased-in as mentioned above).
8. The authors need to take care over some of their conclusions. For example page 16, line 8: "Considering that the options available on the National Health Service, mainly CBT and GET, are often perceived as coping strategies at best, and physically damaging at worst" may be considered inflammatory by many of your readers. In addition, the statements such as " tailored treatments such as described here may be more palatable, and hence effective." (page 16) go beyond the evidence presented in this paper.
Response: This statement has been removed.

VERSION 2 – REVIEW

REVIEWER	Dr Esther Crawley

	Consultant Senior Lecturer University of Bristol
REVIEW RETURNED	31-May-2012

THE STUDY	I would be happy to review this paper in detail if the authors can correct the statistical reporting. I am afraid that it is difficult to see what changes they have made because they have not referenced the changes in the authors response with the page number and line change in the new version of the document. In my original review, I suggested that the authors corrected for variations at baseline. If we were doing this study, we would use regression methods to do this. This would enable the authors to look at outcome corrected for baseline and investigate differences between treatments. The authors have chosen to do a Wilcoxen signed rank test. However, their table of results reports the means for time 1 and time 2. Please can the authors explain which statistical test they used for the means in the table. I am concerned that the authors have included the means for time 1 and time 2 even if the p value is generated from a non-parametric test. The authors have stated that they "have conducted tests on the outcome variables with controls for baseline variation and the results were non-significant. Therefore, we have removed these findings from the paper and stated that with control for variation at baseline, no significant differences were observed between the three groups". I cannot find a description that they have done this analyses in the paper, nor can I find the results of this analyses. Please can the authors describe where they have made the changes (page number and line). In the results section of the paper, the authors report an improvement in "physical functioning, role limitations pain etc" (page 16, lines 11 -22). Is this based on mean differences between the groups at time 1 and time 2? Or is it based on percentage change? If it is based on percentage change (as suggested in the methods), is this mean difference between the groups as a whole? Or is this percentage change only for the smaller number of patients included in the analyses. Please can the authors clarify.
RESULTS & CONCLUSIONS	There are several outcome papers that are not included in the discussion.
REPORTING & ETHICS	I think it would be helpful if the authors could clarify their financial interest in the Optimum clinic not just their position within the organisation.
GENERAL COMMENTS	In the first review I suggested that the tables were unwieldy and should be shortened. Reviewer 1 also suggested that the authors should consider what primary outcome(s) were important. The tables are still unwieldy. The authors need to think hard about what is important and present this in simplified tables. Not all the information needs to be in table. Some can be presented as negative findings in the text.

VERSION 2 – AUTHOR RESPONSE

Reviewer: Dr Esther Crawley
Consultant Senior Lecturer
University of Bristol

I would be happy to review this paper in detail if the authors can correct the statistical reporting. I am afraid that it is difficult to see what changes they have made because they have not referenced the changes in the authors response with the page number and line change in the new version of the document.
Response: Please see below response with regard to statistical reporting. At the end of this document, Dr Crawley's previous comments have been attached and page/line numbers added for ease of review.
In my original review, I suggested that the authors corrected for variations at baseline. If we were doing this study, we would use regression methods to do this. This would enable the authors to look at outcome corrected for baseline and investigate differences between treatments. The authors have chosen to do a Wilcoxen signed rank test. However, their table of results reports the means for time 1 and time 2. Please can the authors explain which statistical test they used for the means in the table. I am concerned that the authors have included the means for time 1 and time 2 even if the p value is generated from a non-parametric test.
Response: Firstly, may we clarify that we are looking for two types of difference: between-groups and change over time. However, as the data is not suitable for parametric test, we cannot perform mixed analysis of variance tests. We do not feel that regressions are appropriate here when looking at group differences as we are not attempting to predict outcomes; therefore our first choice for analysis of between-group differences was analysis of variance (ANOVA) for three independent groups. However, as previously mentioned, the data did not satisfy parametric assumptions so this was not possible; hence we conducted Kruskal-Wallis tests. We did take on board Dr Crawley's previous point that correcting for baseline variation would be an appropriate way to analyse the data so we did indeed carry out ANCOVAs which illustrated non-significant results across the three groups; therefore we excluding these findings and reported this on page 20, lines 16-18, to save space (we would be happy to include these analyses as a supplemental note however). The Wilcoxon signed ranks tests refer to the within-participants comparisons, i.e. change over time, not between the three treatment groups. We included means within the tables for illustrative purposes, however we appreciate that Wilcoxon signed ranks tests do not compare means so we would kindly ask the managing editor which measure of dispersion would be more appropriate for the BMJ Open here, means or mean ranks, and we will amend accordingly. We have taken statistical advice from a number of external colleagues and believe that ours methods are correct for the data. We would have preferred to have data that could have been analysed by parametric tools as these are more powerful but we cannot ignore the assumptions of parametric testing. Non-parametric test are by their nature more conservative, therefore our findings are also conservative.
The authors have stated that they "have conducted tests on the outcome variables with controls for baseline variation and the results were non-significant. Therefore, we have removed these findings from the paper and stated that with control for variation at baseline, no significant differences were observed between the three groups". I cannot find a description that they have done this analyses in the paper, nor can I find the results of this analyses. Please can the authors describe where they have made the changes (page number and line).
Response: We did not include these analyses in the last version of the paper due to space (as stated above). It has been challenging to add additional text for the clarification of points without excluding some areas so we decided to remove the details of this analysis. This is noted on page 20, lines 16-18 as above.
In the results section of the paper, the authors report an improvement in "physical functioning, role limitations pain etc" (page 16, lines 11 -22). Is this based on mean differences between the groups at time 1 and time 2? Or is it based on percentage change? If it is based on percentage change (as suggested in the methods), is this mean difference between the groups as a whole? Or is this percentage change only for the smaller number of patients included in the analyses. Please can the authors clarify.
Response: These analyses relate to within-participant differences, as inferred by the sub-title of the section 'comparisons from time-one to time-two', i.e. we are not looking at between-group differences here. However, this sub-title has been changed to 'Comparisons within-groups across time' for clarification, page 15, line 47. The tests used for these analyses are Wilcoxon signed rank tests (i.e. non-parametric version of the within-participants t-test) and the findings are reported in the tables to save space in the main body of the paper. Of the comparisons that were significant, we calculated the

percentage change over time to illustrate change in scores; therefore the analysis was not carried out on change scores. The percentage change refers to the participants that completed both batteries of questionnaires (time-one and time-two) as it is not possible to estimate change on data that does not exist.
There are several outcome papers that are not included in the discussion.
Response: The following statement has been added: These findings appear consistent with outcomes from other psychological interventions 3;4;6.
Page 21, lines 9-11
I think it would be helpful if the authors could clarify their financial interest in the Optimum clinic not just their position within the organisation.
Response: Alex Howard owns 100% shares in the clinic. Megan Arroll is not a shareholder but has an on-going contract with the clinic for research-related activities.
In the first review I suggested that the tables were unwieldy and should be shortened. Reviewer 1 also suggested that the authors should consider what primary outcome(s) were important. The tables are still unwieldy. The authors need to think hard about what is important and present this in simplified tables. Not all the information needs to be in table. Some can be presented as negative findings in the text.
Response: The findings were positioned in the tables due to the word count. We are now over the word limit due to extended clarification of points made in the first round of reviews; therefore we feel that even more text in the main body of the paper would result in an article that would not be reader-friendly. As tables are opened in new window on the BMJ Open webpage, we feel that it would be better to keep the word count as close as possible to the suggested limit and the data in the tables as readers can navigate between these. However, if the managing editor would agree on an extended word count, we would be happy to add the negative findings in the text.

Previous comments with responses and page/line numbers for reference.
From the managing editor
It is very important for the credibility of the article (and our journal) that any potential competing interests are declared. The Optimum Health Clinic's website states: "The Optimum Health Clinic Research Department was established in June 2011 by Alex Howard, with Dr Megan Arroll. The aim of the department is to develop a high quality evidence base for the OHC approach, and to publish the findings in high impact scientific journals. The goal is that this will be a significant step towards government funding being available for treatment at OHC." Therefore the statement that there are no competing interests is arguably not entirely accurate. Please can you make clear your relationships to the OHC in the competing interests statement?
Response: Information added into the competing interests sub-heading: Alex Howard is the founder and CEO of the Optimum Health Clinic and Megan Arroll is the Director of Research at the Optimum Health Clinic, where this study was conducted.
Page 24, lines 40-42
With regard to the title/abstract/methods - please try to frame the research question more clearly in the title, while including the study design. 'Longitudinal' is arguably inaccurate as this applies to studies that are carried out for much longer than three months.
Response: The title has been altered to a ‘preliminary prospective study’.
Page, line: Title not on main document.
Were either of the authors one of the 'practitioners who recommends the best course of action for his/her needs'? Please state in the paper.
Response: The following clarification was added: (please note, this was not either of the authors of the current study).
Page 7, lines 20-22

Reviewer: Dr Jonathan Price
Clinical Tutor in Psychiatry
University of Oxford

I have no conflicts of interest to declare

I enjoyed reading this manuscript, which is generally very well presented and easy to read and understand.

- study design. The research question is 'to evaluate whether treatments X/Y/Z reduce CFS-ME

symptoms over a period of three months, and whether there are differences in treatment effectiveness between X, Y and Z'. The correct study design for this question is the RCT. The authors mention the preliminary nature of this study but, in my view, they fall into the trap of assigning causality to changes in symptoms which are unjustified with this level of evidence. There are several important limitations which very significantly limit the ability to assign causality - non-randomised design; high drop out rate (about 50%); and differential drop-out rate between groups (from 67% in the psychology group to 40% in the nutrition and combined groups).
Response: The limitations have been discussed in more depth to avoid falling into the causality trap and highlighted by Dr Price and the differential drop-out rates have been discussed as per Dr Crawley's point below.
Page 2, lines 11-15 & 46-50
Page 3, lines 12-21
Page 20, lines 45-51
Page 22, lines 20-32
- methods. These ARE adequately described, apart from the description of the nature of the interventions. I am reasonably clear about the focus of each intervention, and some of their specific components, but I would like to know more about: - the intensity of the intervention (number of individual / group sessions, length of individual / group sessions, 'homework' expected if any), therapist involvement (training of therapists, seniority of therapists), mean hours of face to face psychology /nutrition intervention (if available)
Response: Extra detail has been added regarding the different programs.
Page 8, lines 30-51
Page 9, lines 30-46
Page 10, lines 11-17
- main outcome measure. The outcomes measures chosen are appropriate, and well described. However, there does not appear to be a main outcome measure that is seen by the authors as being of key importance to patients. While the study is 'exploratory / preliminary', it would have been helpful to have a pre-determined primary outcome measure.
Response: The outcome measure have now been divided into primary, secondary ME/CFS specific and secondary psychological for clarity. This has led to newly devised tables which we hope will clarify the importance of the functional ability and fatigue scales and also simplify the results section.
Page 10, line 20
Page 11, line 40
Page 12, line 13
- abstract / summary / key messages / limitations accurate. These are, in general, accurate, but I do think that they need to be much more limited in their claims about the effectiveness of these treatment modalities at this early stage in the development of the evidence. I do not agree that this study, for example, justifies the 'key message' 'patient-centred approaches for the management of ME-CFS reduce symptomatology over time', in the absence of a control group; or that 'psychological interventions can help individuals to regain a sense of control over their condition', when we have no data for two thirds of the psychological intervention group, who *might* have worsened, however unlikely this might seem to those of us who are 'prejudiced' in favour of such interventions. The core message, transmitted consistently in abstract, summary and key messages, needs to be 'this study provides early evidence that treatments X, Y and Z may be effective treatments for some people with CFS-ME, but, due to the study's methodological limitations, it is important that this possible treatment effect is investigated further in high quality randomised controlled studies.'
Response: The language throughout the paper has been tempered so that the message reflects the data at hand more accurately and included Dr Price's suggestion for the core message within the 'key message' section of the paper.
Page 1, lines 23-25
Page 2, lines 7-11 & 32-50
- standard of written English acceptable for publication. In general, definitely yes. However, the manuscript needs careful proof-reading for multiple but minor errors of grammar and punctuation, e.g.
bottom of page 3 - averse rather than adverse;
bottom of page 3- missing comma after medical management;
page 5 - 'patient satisfaction of such approaches CAM has been high';
Response: The paper has been proof-read by a professional proof-reader and errors corrected.
As discussed above, the significant limitations of this study (non-randomised design, high drop-out, differential drop-out between groups) need to be made more explicit in abstract, discussion, and key messages, in order for publication to be justified. My concern is that publication of the manuscript as it

stands might encourage the view that 'this nutritional therapy' / 'this psychological therapy' / 'this combined therapy' is / are effective treatments for people with CFS-ME. Although this study presents preliminary evidence that this might be the case, as the authors point out in their final sentence (of a very long manuscript), random assignment is needed to provide good quality evidence.
Response: As per point above, the language throughout the paper has been tempered so that the message reflects the data at hand more accurately.
The authors appear to be the CEO and Lead Researcher at a private health facility dedicated to the treatment of CFS-ME. Their commitment to the effective treatment of this disorder, and to research into effective treatments, is highly laudable, but their roles in this organisation and, by implication, their potential to 'benefit' from publication of research supporting their work, needs to be explicit as a potential 'conflict of interest'.
Response: Conflict of interest section has been amended as above.

Reviewer: Dr Esther Crawley
Consultant Senior Lecturer
School of Social and Community Medicine
University of Bristol
1. The study design is not appropriate to answer the research question "evaluate the effectiveness of treatments" because there isn't a control group.
Response: This has been changed to: This study therefore aims to provide preliminary evidence for the utility of three types of patient-centered approaches to the management of ME/CFS over time (baseline and follow-up) offered at a private health-care center in the UK.
Page 6, lines 43-46
2. There is insufficient detail in the methods.
a)For example, it is not clear how many follow up questionnaires were sent out and when. Were they sent out at regular intervals on several occasions or just once?
Response: The following detail has been added: Those that expressed an interest (N = 145) were emailed a spreadsheet that contained the questionnaires and asked to complete it at their convenience. Subsequently, participants were requested to complete the questionnaire pack on a second occasion, three months from the baseline measures.
Page 7, line 32 & 45-48
b) How many patients were approached? How many agreed to take part? Without this information, it is hard to know whether the patients recruited are representative of the population of patients attending the clinic.
Response: Further detail has been added for clarification: Of the 145 individuals who expressed an interest in the study, 142 time-one questionnaires were returned, equating to a 97.9% response rate at baseline (two participants from the psychology group and one from the combined group dropped out at this stage). Therefore, excluding the four cases deleted due to insufficient data, 138 cases were used for baseline analysis; 42 participants in the psychology group, 44 in the nutrition group and 52 in the combined group.
Page 14, lines 8-20
c) There is a high drop out rate. When did the drop outs happen? Immediately or at follow up?
Response: As can been seen from the additional information above, the drop-outs occurred at follow-up.
d) Page 11, line 33. The authors excluded cases with missing data. Was this missing data at assessment? Which groups were these patients allocated to? Did they have data at outcome? The authors should consider including the data and reporting N at assessment.
Response: Four cases were excluded at baseline due to incomplete data as in some cases numerous pages of the questionnaire were missing, i.e. more than 5% of data was missing hence case deletion was conducted to deal with this as advised by Tabachnick and Fidell (Using multivariate statistics. 4th ed. Needham Heights, MA: Allyn & Bacon; 2001.) The following detail has been added to clarify which groups the excluded cases were from: therefore these were excluded from the analysis (one individual from the nutrition group and three from the combined group).
Page 12, lines 51-53
3. The statistical methods are not appropriate because the authors need to correct for variations at baseline.
Response: The statistical test used to investigate differences from time-one to time-two were Wilcoxon Signed Ranks test as the data was not suitable for parametric tests. We accept the point that comparisons across groups should include correction for variations at baseline (here general fatigue, physical fatigue and swollen lymph nodes and glands), which would involve either analysis of

covariance (ANCOVA) or analysis of covariance (ANCOVA), both of which are not suitable for non-parametric data. However, we have now conducted these tests on the outcome variables with controls for baseline variation and the results were non-significant. Therefore, we have removed these findings from the paper and stated that with control for variation at baseline, no significant differences were observed between the three groups.
Page 20, lines 16-18
4. The authors should refer to the recently published PACE trial, a large definitive trial investigating CBT, GET and adaptive pacing.
Response: This study has now been referred to in the introduction.
Page 4, lines 30-36
5. The tables are very unwieldy and need to be greatly improved. It may be easier to present the results when proper statistical analyses has been performed.
Response: The tables have been shortened to take into account the division of outcomes into primary and secondary and some tables have been removed to reflect the reanalysis.
Pages 50-57
6. Throughout the results, the authors report “significant differences” (for example, Page 14, line 34). The authors need to clarify what they mean. I assume they are not talking about clinically significant change but what some authors call “statistically significant differences”. The Authors may wish to read Sterne and Davey-Smith discussion on the problems or reporting statistically significant results (BMJ 2001; 322(7280):226-231.)
Response: We have now noted that the significant differences are statistical, rather than clinical.
Page 20, line 45
7. In the discussion, the authors need to be clearer about the limitations of this study. For example, drop out was not equal across all groups. In the psychology group the drop out was highest with retention rates varying from 33% (14/42), 61% (27/44), 40% (31/52). However this is not discussed.
Response: This has now been discussed: Also, there was a high drop-out rate from time-one to time-two and this rate differed across groups. The highest drop-out rate was in the psychology group; whilst we cannot be sure why this occurred, it is postulated that the retention was poor in the group as the individuals in the psychology program had more activities to engage in and may have felt overburdened with the research questionnaires in addition to their session and homework (this would not be the case in the combined group as the therapeutic activities are phased-in as mentioned above).
Page 22, lines 30-42
8. The authors need to take care over some of their conclusions. For example page 16, line 8: "Considering that the options available on the National Health Service, mainly CBT and GET, are often perceived as coping strategies at best, and physically damaging at worst" may be considered inflammatory by many of your readers. In addition, the statements such as " tailored treatments such as described here may be more palatable, and hence effective." (page 16) go beyond the evidence presented in this paper.
Response: This statement has been removed.

VERSION 3 - REVIEW

REVIEWER	Dr Esther Crawley Reader in Child Health School of Social and Community Medicine University of Bristol
REVIEW RETURNED	12-Aug-2012

GENERAL COMMENTS	**Major problems** *Aim* The title and results for this paper suggest that the aim of this analyses is to compare effectiveness. However, the authors do not present any comparison in effectiveness only stating in the results

	that there is no difference between groups. The authors might consider reporting the outcome with the comparison between treatments as a secondary analyses. *Statistical analyses* The authors state in the statistical methods that: "Wilcoxon sign-rank tests were employed to look for differences over time (baseline and 3-month follow-up) and if differences were significant, percentage change was calculated". They confirm this in their response to my previous comments where they say: "The tests used for thee analyses are Wilcoxon signed rank tests (ie non-parametc vrsion of the within-aprticipants t-test) and the findings are reported in the tables to save space in the main body of the paper. In the results section they state " (Please see Table 2 for the exact p-value for each repeated measures comparison.). The authors then present their results in three tables. Each table is labelled as "Comparisons across time within the XX outcome measures". However, the table appears to be reporting the mean and SD at baseline and follow up. My concern is that these tables are not reporting the results of the Wilcoxen signed rank test as these should report medians. The authors have not included the number included in each analyses. I am concerned that the authors are reporting the mean of all the patients at baseline as the confidence intervals are tighter than at follow up. The tables should only report the patients at baseline who were included at follow up. *Other problems* 1. Delete "more integrative interventions" in abstract. There is currently no evidence that the Optimum clinics approaches are more integrative. 2. The authors keep referring to "patient-centred treatment". As all treatment offered within the NHS should be patient centred (see NICE guidelines) this is not very descriptive. Please change "patient-centred approaches as the only descriptor (see Abstract line 22, key messages page 2, line 32) to a description of what the optimum clinic offers (as they do line 44 in Key messages) which will be more descriptive such as psychology, nutrition and combined treatment. 3. Strengths and limitations. By stating that they are "filling a gap" in investigating "tailored, multidisciplinary and patient-centered treatments" the authors imply that other treatments are not tailored, multidisciplinary and patient centred. All NICE recommended treatments should be tailored, multidisciplinary and patient centred. I suggest that authors change this to describe what they actually offered. This will also help deal with what appears to be a contradiction in this paper as it appears that the optimum clinic offers their treatment in groups and not individual sessions.

	4. Introduction. Please remove “drop out rates were high in this group” line27 referring to the systematic review. The largest study (PACE) had extremely low drop out rates. If the authors want a detailed discussion they should refer to Hempels detailed review in 2006 which discusses this issue. 5. Page 5 introduction. By discussing NICE strategies at the start of the paragraph, page 19 and then say dietary interventions are recommended for CFS/ME. The authors suggest that dietary strategies are recommended by NICE saying “are thus also recommended for those with ME/CFS”. NICE actually does not recommend dietary strategies for CFS/ME. The reference the authors use is for IBS and is prior to NICE guidance and does not make dietary recommendations for CFS/ME. 6. Objective. Whilst it is true that we need more research in to different treatment approaches for CFS/ME, the reviews of CAMs treatment do not find that there needs to be more research in to “patient-centred” individually tailored treatments. They just state there needs to be more high quality, suitably powered and well conducted trials in to a range of treatments. 7. I note that the editor has pointed out that this study cannot evaluate effectiveness. I agree. Please delete “explore the effectiveness” page 8, line 9 in Study design. 8. The authors reference their own “in prep” paper for one of the questionnaires. The authors should only reference published papers and should describe the methodology of preparing this questionnaire (or leave it out for now). 9. The authors use mean substitution for missing data. Using mean substitution will reduce variation providing artificially low p values. Please either analyse only complete data for each variable or use multiple imputation techniques. In the authors response, they have said they cannot make the tables easier to read because of the word count. I would suggest that the authors consider cutting the last three paragraphs in the introduction as they are not necessary and do not add to this report.

VERSION 3 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:
Reviewer: Dr Esther Crawley
Reader in Child Health
School of Social and Community Medicine
University of Bristol

Major problems
Aim
The title and results for this paper suggest that the aim of this analyses is to compare effectiveness. However, the authors do not present any comparison in effectiveness only stating in the results that there is no difference between groups. The authors might consider reporting the outcome with the comparison between treatments as a secondary analyses.
Response: The present title of this paper is 'A preliminary prospective study of nutritional, psychological and combined therapies for Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) in a private care setting', which does not infer a comparison of effectiveness and was suggested by the editor (we appreciate that previous version of this paper did contain title inferring effectiveness). Reporting of the comparisons across groups has been positioned as secondary analysis; as we are already using the terms 'primary' and 'secondary' with regards to outcomes, we believe it would be confusing to also delineate further in terms of analysis when the comparisons across groups are clearly discussed after comparison over time.
The word 'effectiveness' now only appears once in this paper and is not referring to the present study,
Statistical analyses
The authors state in the statistical methods that: "Wilcoxon sign-rank tests were employed to look for differences over time (baseline and 3-month follow-up) and if differences were significant, percentage change was calculated". They confirm this in their response to my previous comments where they say: "The tests used for thee analyses are Wilcoxon signed rank tests (i.e. non-parametric version of the within-participants t-test) and the findings are reported in the tables to save space in the main body of the paper. In the results section they state " (Please see Table 2 for the exact p-value for each repeated measures comparison.). The authors then present their results in three tables. Each table is labelled as "Comparisons across time within the XX outcome measures". However, the table appears to be reporting the mean and SD at baseline and follow up. My concern is that these tables are not reporting the results of the Wilcoxen signed rank test as these should report medians.
Response: The revised tables now contain medians and percentiles for the non-parametric Wilcoxon Signed Rank Tests.
Pages 28-40
The authors have not included the number included in each analyses. I am concerned that the authors are reporting the mean of all the patients at baseline as the confidence intervals are tighter than at follow up. The tables should only report the patients at baseline who were included at follow up.
Response: The revised tables only include data for participants included in the follow-up and a column stating the total number in each analysis has been added.
Pages 28-40

Other problems
1. Delete "more integrative interventions" in abstract. There is currently no evidence that the Optimum clinics approaches are more integrative.
'More integrative interventions' has been replaced by 'alternative techniques'.
Page 1, line 17
2. The authors keep referring to "patient-centred treatment". As all treatment offered within the NHS should be patient centred (see NICE guidelines) this is not very descriptive. Please change "patient-centred approaches as the only descriptor (see Abstract line 22, key messages page 2, line 32) to a description of what the optimum clinic offers (as they do line 44 in Key messages) which will be more descriptive such as psychology, nutrition and combined treatment.
Response: The term 'patient-centered' has been removed or changed to 'psychological, nutritional and combined' within the paper.
Page 1, line 22; page 2, lines 8, 22, 30, 53; page 6, lines 28, 44, 49; page 20, line 24
3. Strengths and limitations. By stating that they are "filling a gap" in investigating "tailored, multidisciplinary and patient-centered treatments" the authors imply that other treatments are not tailored, multidisciplinary and patient centred. All NICE recommended treatments should be tailored, multidisciplinary and patient centred. I suggest that authors change this to describe what they actually offered. This will also help deal with what appears to be a contradiction in this paper as it appears that the optimum clinic offers their treatment in groups and not individual sessions.
Response: As per comment 2 above. However, as can be seen on pages 8-9, the treatments offered by the clinic are a combination of group work and individual session in the psychology programme and all of the nutrition session are one-to-one.
4. Introduction. Please remove "drop out rates were high in this group" line27 referring to the

systematic review. The largest study (PACE) had extremely low drop out rates. If the authors want a detailed discussion they should refer to Hempels detailed review in 2006 which discusses this issue.
Response: In the Edmonds et al. (2004) review of GET as a treatment for CFS, it is stated that 'Physical functioning was significantly improved with exercise therapy group (SMD -0.64, CIs -0.96 to -0.33) but there were more dropouts with exercise therapy (RR 1.73, CIs 0.92 to 3.24).' under abstract, main findings. Therefore, we have retained this sentence but changed the wording to 'However, drop-out rates were higher in the GET groups than control groups suggesting that individuals with ME/CFS are averse to this type of therapy.' In addition we have noted that the PACE trial had low drop-out rates. We hope that this provides a more balanced view.
Page 4, lines 24-26, 30
5. Page 5 introduction. By discussing NICE strategies at the start of the paragraph, page 19 and then say dietary interventions are recommended for CFS/ME. The authors suggest that dietary strategies are recommended by NICE saying "are thus also recommended for those with ME/CFS". NICE actually does not recommend dietary strategies for CFS/ME. The reference the authors use is for IBS and is prior to NICE guidance and does not make dietary recommendations for CFS/ME.
Response: The previous sentence to the one mentioned above stated that: 'Dietary management may also reduce symptomatology for those with concurrent irritable bowel syndrome (IBS).', therefore we were discussing individuals with both IBS and ME/CFS. However, we accept that NICE does not recommend dietary strategies (although the guidelines do note that some individuals find dietary supplementation helpful); therefore we have changed the sentence to 'Dietary management may also reduce symptomatology for those with concurrent irritable bowel syndrome (IBS) 19, although this is not currently recommended by NICE.'
Page 5, lines 20-24
6. Objective. Whilst it is true that we need more research in to different treatment approaches for CFS/ME, the reviews of CAMs treatment do not find that there needs to be more research in to "patient-centred" individually tailored treatments. They just state there needs to be more high quality, suitably powered and well conducted trials in to a range of treatments.
Response: As per comment 2 above, the term 'patient-centered' has been removed. However, the Porter et al. (2010) paper states 'Individualized treatment plans that involve several pharmacological agents and natural remedies appear promising as well.' Therefore, we have retained the following sentence with reference to the Porter et al. paper: 'A recent review of CAM techniques 31 highlight the need for further exploration of individually tailored interventions for the alleviation of the condition's often debilitating and intrusive symptomatology.'
Page 6, lines 42-7
7. I note that the editor has pointed out that this study cannot evaluate effectiveness. I agree. Please delete "explore the effectiveness" page 8, line 9 in Study design.
Response: This sentence has been changed to 'This preliminary prospective study aimed to investigate whether psychological, nutritional and combined approaches to the treatment of ME/CFS influenced symptom report measures over a 3-month time period and whether there were significant differences in these changes between group.'
Page 7, lines 12-18
8. The authors reference their own "in prep" paper for one of the questionnaires. The authors should only reference published papers and should describe the methodology of preparing this questionnaire (or leave it out for now).
Response: It is not unusual to reference an 'in-prep' questionnaire in the literature; for instance Moss-Morris and colleague have referenced their Cognitive and Behavioural Responses to Symptoms Questionnaire in numerous publications as 'in-prep' without discussing the methodology of preparing the questionnaire (e.g. Laura Dennison, Rona Moss-Morris, Eli Silber, Ian Galea, Trudie Chalder (2010). Cognitive and behavioural correlates of different domains of psychological adjustment in early-stage multiple sclerosis, Journal of Psychosomatic Research, Volume 69, Issue 4, Pages 353–361). However, we have excluded this measure from the present paper in light of Dr Crawley's concerns.
Page 1, line 42, 48; page 2, line; page 12, lines 12-21; page 19, lines 47-49; page 26, lines 42-43
9. The authors use mean substitution for missing data. Using mean substitution will reduce variation providing artificially low p values. Please either analyse only complete data for each variable or use multiple imputation techniques.
Response: Mean substitution is an appropriate method of dealing with missing data if there is less than 5% missing data in the variable. (Please see Tabachnick and Fidell, Using multivariate statistics. 4th ed. Needham Heights, MA: Allyn & Bacon; 2001, for a discussion of this.) However, in light of Dr Crawley's concerns, we have re-investigated the missing data (which was not in every variable and

consisted of very few cases in variables were it appeared, hence less than 5%) and re-analysed the data.

Page 12, lines 30-34; pages 28-40

In the authors response, they have said they cannot make the tables easier to read because of the word count. I would suggest that the authors consider cutting the last three paragraphs in the introduction as they are not necessary and do not add to this report.

Response: The managing editor has stated that word count is not an issue with BMJ Open; therefore we have retained the paragraphs mentioned above but simplified the tables. We sincerely hope that these are now clear and easier to read.