

## **Author's response to reviews**

**Title:** Associations between Diet and Disease Activity in Ulcerative Colitis Patients using a Novel Method of Data Analysis

### **Authors:**

Elizabeth A Magee ([E.Magee@dundee.ac.uk](mailto:E.Magee@dundee.ac.uk))  
Laurie M Edmond ([L.M.Edmond@dundee.ac.uk](mailto:L.M.Edmond@dundee.ac.uk))  
Shiona M Tasker ([s.m.tasker@dundee.ac.uk](mailto:s.m.tasker@dundee.ac.uk))  
Chris Kong ([s.c.kong@dundee.ac.uk](mailto:s.c.kong@dundee.ac.uk))  
Richard Curno ([r.curno@dundee.ac.uk](mailto:r.curno@dundee.ac.uk))  
John H Cummings ([J.H.Cummings@dundee.ac.uk](mailto:J.H.Cummings@dundee.ac.uk))

**Version: 2 Date:** 11 January 2005

**Author's response to reviews:** see over

#### Reviewer 1 Comments:

I think the authors put a lot of time and effort in this study. It also takes some effort from the reader to understand how the Food Sigmoidoscopy Score (FSS) was calculated. The proposed score is a new way to study the effect of diet on disease activity. It will need to be verified in more studies to prove its reliability.

**We agree.**

This new approach should be given a chance to be known. For this reason I believe this article should be published. Future studies will be the only way to tell how reliable this method will be.

I don't have any specific changes that I recommend. I just suggest the following:

1- Adequately label the figures and tables to fit with the text.

**The mismatch was introduced during the download but will be rectified.**

2- Regarding figure 3 referred to as table 2 in the text, it is better to show the actual calculation of the sig. score amount / diary energy e.g. for patient 3 the value would be  $750 \times 3.5/2087 = 1.258$ , so it will be easier for the reader to understand. This can be added separately at the end of the table.

**Additional text has been incorporated into the table legend as suggested.**

3- Regarding table 3, it is better to add the unit to the average portion which is g (gram) at the top of column 6.

**Average portion unit introduced.**

#### Reviewer 2 Comments:

1 Abstract. This accurately reflects the data and avoids over-speculation

2. Introduction. Again this is fine, not overly long and relevant.

3. Methods. This describes succinctly the methods and subjects. The study validity essentially revolves around the accuracy of the method of assessing the rectal mucosa and the food assessment. I would like to see more justification of the validity of the sigmoidoscopic score. This is notoriously unreliable and there is actually scant evidence in the literature that the inter-observer reliability makes it a worthwhile scoring system. In particular the authors have used the rectum as the scoring point but about 10-15% of UC patients may have rectal sparing, even more if they are using rectal drugs. The CAI is also not 1000% valid but considerably more data exist about its utility and reproducibility. The lack of a widely accepted activity measure for colitis is hampering studies of the disease so the authors are not alone in choosing an inexact method. There was poor correlation between CAI and sigmoidoscopy score and the distribution of scores was very strange with concentration on particular values (eg 1.5). It would be interesting to know whether there was any difference in the pattern of scoring between the two sigmoidoscopy scorers. I have major doubts that a scoring system which has 12 (13) possible scores for rectal inflammation is of sufficient validity.

We agree that sigmoidoscopic scoring is imperfect and we looked at several before choosing that of Baron – which is a validated scoring system and hopefully more widely accepted.

The two clinicians worked closely together and often discussed cases where there was any doubt. The two patterns of scoring were not formally tested.

The distribution of the sigmoidoscopy scores is interesting and points to some of the scoring parameters being linked.

The use of rectal therapies will affect the sigmoidoscopic grading. The use of rectal preparations was recorded at every clinic visit. However in the final analysis there did not seem to be any relationship between mucosal disease activity and use of rectal therapy, so no correction factor was introduced. Clinical activity index data also exist and it is reasonable to ask what would happen if the calculations were performed using this value instead of sigmoidoscopy scores. This calculation was performed on a subset of data earlier on in the study. The correlation ( $r^2$ ) for these 2 (food) scores was 0.13, which compares to the correlation between the raw CAI and SS data of 0.25. This leads to the question - why choose sigmoidoscopy scores over the clinical activity index? The answer was simply a value judgement. We considered the sigmoidoscopy score to be more objective and free from patient bias, while the clinical activity index is more likely to be affected by patient values and perceptions of their disease.

If this paper is mainly about finding a new method for analyzing dietary data in colitis, I also think that it would be useful to show the model with a CAI-diet score done in exactly the same way as the fss score. For future studies and for serial studies it can be better to have a non-invasive method of evaluating colitis activity to enhance participation rates.

This is discussed above. Non invasive methods rely on patient judgement unfortunately.

With respect to the dietary data, I think it would be very useful to see the analysis by dietary constituents that are alluded to (by WISP) even if they are negative. In addition, it would be very useful to see the analysis by food groups, such as in the paper cited by Jowett et al. This information should be in the public domain to allow for meta-analyses and comparisons between studies, even if, and especially if, it is negative. An example otherwise of positive publication bias.

The WISP analysis is a huge amount of data and it is planned to include this in another paper. We also have second diaries for this population at the six month time point and a lot of additional data.

It is perhaps worth mentioning that we would interpret the results obtained by Jowett in terms of sulphite being a risk factor for UC and not total sulphur.

In the methods, the authors have corrected dietary data for calorie intake but not for smoking, disease extent and sex which might be related to dietary intake and activity. Could this be done?

Disease extent will be related to the macroscopic appearance of the rectal mucosa (at a distance 5-10 cm from the anal verge). Correcting dietary data for calorie intake takes into account individual food requirements (e.g. due to age, sex, activity levels etc.). Smoking is a recognised factor in UC but as few of our volunteers smoked this correction was not performed.

The bowel habit diary contributes little and should be dispensed with.

OK

Results. There are two pages of results and 6 of discussion. I feel that this is rather imbalanced and would like to see more results presented (as suggested above) and the discussion trimmed a bit. In terms of the findings, the authors have concentrated on one factor, namely sulphite, in the foods with a high FSS and not really searched for other explanations.

This is not correct. The risk factor we have concentrated on is thiamin. The relevance of sulphite to the discussion is simply because it influences thiamin status. We also discuss resistant starch intake. However we are advocating a particular diet (for whatever reasons) rather than promoting the taking of supplements.

The only statistical justification for this is the mann whitney test performed on the rankings of fss for selected high sulphite foods. It would be interesting to know the opinion of a statistician as to the validity of this and the strength of this finding. Some justification for this test should be mentioned in the methods. It appears that the mann whitney u test was performed using the top 7 sulphite containing foods in the diet. Why 7 and not 5 or 10? Was this a prior hypothesis or were several models tried?

Again the question here is why a particular approach was chosen over another. There are a number of ways to test the significance of the sulphite containing food order. One way is to plot the sulphite content (for those foods containing sulphite) of an average portion of the food against its sigmoidoscopy score and perform regression analysis (dose response type graph). To get round the problem of the widespread and inconsistent sulphiting protocols and the instability of sulphite (making it difficult to determine how much sulphite was actually added to each food) we used the permitted sulphite levels (i.e. an external data set) to represent the sulphite content. This method was used in a previously published abstract on a slightly different (overlapping) data set and also gave a significant result.

The decision as to where to make the cut for the Mann-Whitney test was based on the distribution of the sulphite amounts per portion for the foods and the likelihood of the foods actually being sulphited. The point we chose appeared the most obvious cut point. An alternative would have been to make no cut at all but as this test gives equal weight to each food, the position of jam and marmalade would then be as important as the position of say, white wine. White wine (always sulphited in our experience) contains maybe 50 times more sulphite per portion than jam and marmalade (often not sulphited). This test is thus an approximation but does draw attention to the distribution of sulphite containing foods.

Combining the above two above approaches demonstrates that sulphite-containing foods have high FSS values and that the sulphite containing foods that have the highest FSS values contain the most sulphite. A number of different approaches were tested but for brevity only a single approach was documented.

I am also particularly drawn to the fact that of the top 10 foods, 5 are not listed as being high sulphite. The place of these 5 is not mentioned. This suggests to me that the authors may be giving undue emphasis to data that fits their hypothesis and ignoring other data. For example, can they give an explanation for why lamb is in the worst ten and pork in the best 10? Why do 'natural' foods like kiwi, carrots and grapes feature in the worst ten but highly altered foods like pizza and breakfast cereals in the low risk ten and why the difference between cerals and cereal bars? The low risk foods do appear to be mostly genuinely low sulphite except the lowest ranked one (crackers) which if it is classified as a dry biscuit is mentioned in the table of high sulphite foods. I believe also that salads have sulphites added as preservatives and lettuce features in the low risk foods category. Further analysis of food groups might help to elucidate the relationships between these factors.

We have spent the last few months trying to interpret the food orders presented here. The best way of doing it is to look not only at the absolute position of the food but also its position in its class. For example taking meats (and fish); plotting the thiamin content of a typical portion against the sigmoidoscopy score yields a striking correlation. This explains the respective positions of pork (1.31 mg B1 per portion) and lamb (0.22 mg B1 per portion). Grapes and kiwi fruit are not in the foods to avoid list because their position probably has more to do with being consumed by sick people than with making people sick. Cereal bars can contain sulphited fruit but probably reflect a

convenience diet (convenience food generally having a higher probability of being sulphited). Cereals are generally fortified and it is conceivable that milk is beneficial to UC patients. We could speculate a reason for the position of every food but the judgements we have made are reflected in the advice table.

Conclusions. This is really a preliminary paper describing the method and as the title says trying to establish this. The conclusions are wide ranging and to my mind rather too firm in their stating. Patients have already been presented with a table of high sulphur foods (Roediger WE. Decreased sulphur aminoacid intake in ulcerative colitis. Lancet 1998; 351: 1555) which does not appear to be referenced in this paper. The table of high and low risk foods seems to me very preliminary and should be further established before being presented in this manner. As a clinician dealing with people with colitis, I feel that a proliferation of poorly evidenced diets is going to make it very difficult to be clear with patients and if it turns out to be wrong and patients do not get benefit, then when we have better data patients will have lost faith. The conclusions and discussion should focus more on the uncertainties in the data and the need for an intervention trial to confirm these findings.

The omission of the above Lancet letter is an oversight that has been rectified. Roediger's paper was the justification for the study from which this baseline data was extracted.

We agree about the uncertainties and the need for more work and confirmation (2<sup>nd</sup> to last paragraph of the original discussion).

Figures and tables.

Table 1 legend appears to match fig 2. This table of the results of previous studies should be abandoned. Many of the studies are of poor quality and design and the design is not apparent in the table so the reader does not know how strongly to weigh the evidence. The higher quality studies are also not all included.

We agree with the above except deleting the table as their inclusion is more to set this work in its historic context rather than to critically review them.

The figure 6 of foods to be avoided or consumed is in my mind premature and should be removed. Patients will have easy access to this paper through an electronic journal and may place undue reliance on the findings of this preliminary study.

The problem with removing the dietary advice table is that patients (and clinicians) will possibly then make their own judgements about diet based on the FSS table. We believe that there are good reasons for the inclusion or exclusion of the foods shown in this table (further discussed above) and hence that it has some value. We have however removed some of the minor components (< 100 g) from this table as there is more uncertainty with these products due to limited data.

Overall therefore I feel that further analysis and more conclusion would strengthen this paper.